Goal-setting reloaded: The influence of minimal and maximal goal standards on task satisfaction and goal striving after performance feedback[☆]

Steffen R. Giessner[a,*], Daan Stam[a], Rudolf Kerschreiter[b], Danny Verboon[a], Ibrahim Salama[a]

[a] *Rotterdam School of Management, Erasmus University, the Netherlands*
[b] *Freie Universität Berlin, Germany*

ABSTRACT

When striving to meet goals, individuals monitor their progress towards achieving them. The discrepancy between their current performance and their goal determines task (dis)satisfaction, and thus whether they will make greater effort. We propose and test a theoretical extension of goal-setting theory, namely that different types of goal standards (minimal or maximal) fundamentally change this monitoring process. Through four experiments we demonstrate that with maximal goals ("ideal" standards), individuals experience greater task satisfaction the nearer their current performance comes to the goal. In contrast, with minimal goals ("at least" standards), their satisfaction level remains low, regardless of how close their performance is to the goal. When goals are exceeded, the reverse applies: with maximal goals, satisfaction remains high regardless of the level of overperformance, while with minimal goals, satisfaction is determined by the level of overperformance. We also demonstrate that task satisfaction levels influence subsequent decisions on goal striving.

1. Introduction

Goals have an enormous impact on how people perform in organizations. Consequently, vast amounts of research have been done on how people set and strive for goals (Carver & Scheier, 2000). In this paper we focus on how people monitor progress towards (and beyond) their goals. Previous research has established that individuals continuously compare their current performance to goal standards (i.e., they monitor goal progress by assessing the so-called goal-performance discrepancy), which gives them a sense of how satisfied they are with their performance on the task (Locke, 1969). This task performance satisfaction[1] is not only important in and of itself but also represents an essential part of the greater construct of job satisfaction (Spector, 1997). Furthermore, it forms the basis for adjusting individual goal-setting (Elicker, Lord, Ash, Kohari, Hruska, et al., 2010; Locke, Cartledge, & Knerr, 1970) and determines the individual's motivation to continue with this behavior (Bandura & Cervone, 1983; Cervone, Jiwani, & Wood, 1991; Locke & Latham, 2002). For example, sales employees who are dissatisfied with their sales of a certain product might as a result reduce their sales target (i.e., the goal) for that product in the following period. Or if those selling a product have a certain selling price they want to

achieve, they might not negotiate any further once they agreed with the buyer on a selling price they are satisfied with. It is therefore of the utmost importance to gain a comprehensive understanding of how people monitor progress towards goals or, more specifically, how differences between current performance and goal standards affect task satisfaction and other subsequent outcomes.

Early research by Locke and colleagues on monitoring goal progress assumed that there was a *linear* relationship between goal–performance discrepancy and task satisfaction (Kernan & Lord, 1991; Locke, 1969). More recent research suggests that the *value function* of prospect theory (Kahneman & Tversky, 1979) may describe this relationship more accurately (Heath, Larrick, & Wu, 1999). Irrespective of their differences, underlying all of these studies is the notion that one universal model can account for all monitoring of progress towards goals. However, predicting task satisfaction with goal achievement using one universal function assumes that all goals are similar in nature and that there are no major defining differences between them. We posit that this assumption may not be realistic.

Indeed, prior research in the area of self-regulation has focused on the self-regulatory functions of goals and their role in determining valence judgments (Brendl & Higgins, 1996). Based on earlier research by

Gould (1939), Brendl and Higgins (1996) differentiated two goal standards, minimal and maximal, that are fundamentally different (see also Giessner & van Knippenberg, 2008). *Minimal goal standards* set a base reference point that needs at least to be achieved; here the differentiation is thus between a negative valence area (i.e., all performance below the reference point is judged to be negative) and a non-negative area (where performance is above the reference point). In contrast, *maximal goal standards* provide a reference point that can ideally be achieved; the division is thus between a positive valence region (i.e., all performance above the reference point is judged to be positive) and a non-positive region (where performance is below the reference point). We argue that the differences between minimal and maximal goal standards have important consequences for the relationship between goal–performance discrepancies and task satisfaction. In other words, in contrast to prior theories (Heath et al., 1999; Kernan & Lord, 1991; Locke & Latham, 2002) we argue that the relationship between goal–performance discrepancies and task satisfaction does not follow a universal value function, but that its shape is dependent on the nature of the goal standard.

We test this idea with four studies that use a range of experimental paradigms and samples. This research contributes to the literature on goal-setting and goal regulation in three important ways. First, we develop a new theoretical model that can be used to examine the effects of goal achievement on task satisfaction for different types of goal standard (minimal and maximal). This refines prior models that explain the relationship between performance–goal discrepancy and task satisfaction and provides a solid basis for further research on goal-setting and self-regulation. Second, we show that these levels of satisfaction can affect subsequent decision-making, such as accepting an offer in a simulated negotiation. This is important, because it provides evidence that goal–performance discrepancies and the resulting task satisfaction can affect subsequent behavior. Third, we discuss the methodological consequences of these findings for theorizing and analyzing the link between performance and satisfaction. Here we argue that understanding this link requires a consideration of goal standards (in addition to the goal level) and may suggest the use of non-linear modeling, two elements that have a substantial influence on how future studies should be designed.

2. Goals and satisfaction with goal achievement

Satisfaction has been defined as a "pleasurable emotional state" (Locke, 1969) and as an attitude (Brief, 1998), resulting in some controversy about the meaning of the construct itself (Weiss, 2002). In our current research we define it as an attitude, in line with Weiss (2002) (see also Bandura & Cervone, 1983). More precisely, as we focus on task satisfaction, we define it as "a positive (or negative) evaluative judgment one makes about one's task achievement" (Weiss, 2002, p. 175). Task satisfaction is an important construct for organizations. Although in organizational behavior the satisfaction construct has been used most commonly in the context of job evaluations, task satisfaction can be seen as sub-facet of job satisfaction. Furthermore, previous research in the goal-setting literature has shown that increased dissatisfaction with an outcome leads either to an adjustment of the goal itself (Elicker et al., 2010; Locke et al., 1970; Podsakoff & Farh, 1989) or to an increased motivation to engage further in the task or behavior (Bandura & Cervone, 1983; Cervone et al., 1991; Locke & Latham, 2002). In simpler terms, feeling dissatisfied with one's performance on a task is more likely to make one adjust one's goal (i.e., to make it either more difficult or easier to achieve) and/or to further perform in order to achieve one's goal than one is inclined to go on without making any adjustments. For instance, not having yet achieved a specific sales target should result in dissatisfaction, which will either increase commitment to achieving it or lead to a lower sales target the following year. Another example would be negotiating the price when buying a house: The more dissatisfied you are with your achievement (due to a larger discrepancy in

the vendor's final price compared to the amount you had in mind as your target price), the less likely you are to agree to the deal.

Consequently, satisfaction (either in the form of task satisfaction or more general job satisfaction) has been a central variable in Locke and Latham's (2002) goal-setting framework. In this paper we focus on how performance (reflected in goal–performance discrepancies) affects task satisfaction and how this then affects decision-making whether to progress with the task.

So how do we evaluate our performance achievements (in terms of satisfaction)? Let us assume that Daan and Rudi are both teaching an MBA course. At the end of the course both receive their student course evaluations (rated on a scale from 1 = bad to 5 = excellent). Daan receives a 4.3 and Rudi receives a 4.0. Who is more satisfied with his evaluation? This question cannot be answered just by knowing the evaluation results. Previous research has established that the link between performance and satisfaction can only be understood if alongside performance one also considers the goal or, more precisely, the goal-performance discrepancy (Carver & Scheier, 1982; Kernan & Lord, 1991; Lord & Hanges, 1987). In other words, what predicts satisfaction levels is not performance per se but rather the discrepancy between the performance and one's own goal (i.e., the goal achievement level). In the example above, if Daan's goal was to receive a course evaluation of 4.5 but Rudi was aiming for a 4.1, Rudi should be more satisfied than Daan with his performance, given that his score was closer to his goal than Daan's.

Early research on goal-setting theory established a negative relationship between levels of goal achievement and feelings of satisfaction or dissatisfaction (Kernan & Lord, 1991; Locke, 1969). Importantly, the literature on goal-setting theory suggests there is a linear relationship between goal–performance discrepancies and task satisfaction such that smaller negative discrepancies (in the case of failure to reach the goal standard) and larger positive discrepancies (in the case of success in reaching the standard) would linearly cause more task satisfaction.

Extending this view, Heath et al. (1999) suggested that the value function of prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) would be a more parsimonious theoretical framework for explaining the relationship between goal–performance discrepancies and evaluative judgment. The value function assumes a non-linear function that distinguishes between positive value (gains) and negative value (losses). The reference point in the value function is the mid-point between gains and losses. If goals are reference points in the value function, then they become the critical points that distinguish between successes and failures. This is not a novel idea in itself. There are several theories on goal-setting that propose this dichotomy (e.g., Brendl & Higgins, 1996; Lewin, Dembo, Festinger, & Sears, 1944), and it is also intuitively appealing. More important, however, is their assumption that if goals act as reference points in the value function of outcomes, then they may well inherit a key property of the value function: *diminished sensitivity*. This principle relates to actual deviation from the goal and whether the end-result is above the goal (i.e., success) or below the goal (i.e., failure). Given the S-shape of the value function, successes should follow a concave curve and failures a convex curve. The concave curve for successes suggests that larger positive discrepancies (successes) from the reference point provide diminishing returns in terms of value: Whereas a small positive discrepancy provides a certain value, a discrepancy that is twice as large will provide less than twice that value. The convex curve for failures suggests that the same logic applies to failures: The larger the negative discrepancy (the failure) from the reference point the less additional reduction in value it offers. Furthermore, given the specific form of the value function for gains and losses, equally sized positive and negative discrepancies from the reference point have differently sized effects on value: a negative discrepancy (failure) has a stronger effect on value than a positive discrepancy.

3. Goal standards and the value function

While there seems to be evidence for the application of the value function as described in prospect theory, recent advances in the goal-setting literature suggest that the way that goals are communicated changes the effects that they have on performance. For instance, Drach-Zahavy and Erez (2002) showed that describing a goal as challenging resulted in a higher level of performance than if the same goal was described as a threatening one. Also, learning versus performance goals shift attention to either developing task-related abilities (i.e., learning goal) or increasing one's motivation to apply that knowledge (i.e., performance goal) (Chen & Latham, 2014; Latham & Brown, 2006; Seijts, Latham, Tasa, & Latham, 2004).

We propose that, in much the same way that these goal narratives have consequences for performance, specific goal standards may also be able to change the evaluative judgments associated with goal–performance discrepancies. We define a *goal standard* as "a criterion or rule established by experience, desires, or authority for the measure of quantity and extent, or quality and value" (Higgins, 1990, p. 302). The valence framework developed by Brendl and Higgins (1996) seems especially suited to understanding how goal standards affect the evaluative value functions of goal–performance discrepancies (see also Austin & Vancouver, 1996; Vallacher & Wegner, 1989). This motivational and cognitive framework is based on previous research on aspiration levels (e.g., Lewin et al., 1944) and self-regulation (e.g., Higgins, 1997). While it focuses on goals and goal-achievement levels to explain an individual's evaluative state, this research has developed separately from the goal-setting literature. Nevertheless, we propose that these frameworks can also be leveraged to describe how specific goals, such as getting published in a top-tier journal (for an academic) or achieving a top sales performance (for a salesperson), affect subsequent evaluations.

We argue that an important extension to previous theory on the relationship between goal–performance discrepancies and satisfaction is offered by *goal standards* of specific goals. Drawing on Gould's (1939) qualitative research on aspiration levels, Brendl and Higgins (1996) differentiate between minimal and maximal goal standards. Minimal goal standards represent "oughts that a person must attain or standards that must be met" (Idson, Liberman, & Higgins, 2000, p. 254), whereas maximal goal standards represent ideals that one hopes to achieve (see also Higgins, Shaw, & Friedman, 1997). Referring back to our previous example of Daan and Rudi, Rudi might have had a minimal goal and wanted to achieve a course evaluation score of at least 4.1. Daan, however, might have had a maximal goal and aimed to ideally achieve a score of 4.5. Brendl and Higgins argue that minimal standards define a region of failure in which goal–performance discrepancies are negative. Therefore, these are perceived as "the lowest goal whose end state will still produce satisfaction" (Higgins et al., 1997, p. 104). In contrast, maximal goals define a positive region in which success implies a clear positive evaluation. Consequently, we argue that these goal standards imply an important qualification of the meaning of the reference points. Minimal standards provide a psychological separation between a negative valence region and a non-negative region. In other words, failure to achieve a minimal goal would be experienced as a "categorical" negative, whereas achieving these goals is not necessarily clearly positive. In the case of Rudi, not achieving the minimal level of 4.1 for the evaluation would lead to dissatisfaction, irrespective of whether the score he is given is 3.5 or 4.0. However, if he exceeds his goal of 4.1, he will feel this to be satisfactory, but the higher the score is above that level, the greater his satisfaction will be. So there is a positive relationship between the degree of overachievement and the level of satisfaction. In contrast, maximal standards separate a positive valence region from a non-positive one. Thus, meeting or going beyond a maximal goal is a "categorical" positive, while failing to meet such a goal is not necessarily negative. In the example of Daan, he should feel any score above 4.5 to be highly satisfactory, irrespective of the degree

to which he exceeds his goal. If he does not achieve his goal of 4.5, however, it should not be perceived as negative per se, and the degree of satisfaction he feels should depend on the size of the goal–performance discrepancy.

Interestingly, whereas the goal-setting literature would suggest that Rudi should be more satisfied with his score of 4.0 than Daan with his 4.3, given that their respective goals were 4.1 and 4.5, if one also takes into account the goal standards of Rudi and Daan, this puts quite a different perspective on these results and suggests Rudi may be more dissatisfied than Daan with not having met his minimal goal standard.

The differentiation between minimal and maximal goals (i.e., hoped-for goals) was also used in early research on goal-setting theory (Locke, 1969; Locke & Bryan, 1968, 1969; Wood & Locke, 1990). However, this research focused on how reliable these self-set goals are in predicting student grades. Thus, although the distinction between different types of goal standards has some grounding in the goal-setting literature as well, previous theory and research did not consider what implications these goals might have for the relationship between goal–performance discrepancies and satisfaction. We integrate research on goal-setting theory (Locke & Latham, 2002) and the regulatory function of goal standards (Brendl & Higgins, 1996) and develop a new theoretical framework on how minimal and maximal goal standards change the relationship between goal–performance discrepancies and task satisfaction. In contrast to previous research, we argue that different goal standards produce unique and qualitatively different value functions, rather than one universal value function.

4. Minimal and maximal goal standards

Although no research has yet examined how minimal and maximal goal standards affect evaluative reactions to goal–performance discrepancies, some related research may provide relevant insights. For instance, Giessner and van Knippenberg (2008) conducted a study in a leadership context in which they show that leaders who were more trusted received more support from followers if they failed to achieve a maximal goal standard, but less support if they failed to achieve a minimal standard. They argued that minimal goals lead to a general negative perception of leader failure. Consequently, other variables, such as how much one has trust in the leader, should also matter less in one's evaluation of the leader if she/he fails to achieve a minimal goal. In contrast, when maximal goal standards are applied, the evaluation of the leader will typically be done on a continuous scale. Therefore, the degree of trust one has in the leader can have an impact on the leader evaluation, despite the failure. This research thus provides some support for the assumption that, for failures at least, minimal goals lead to a categorically negative evaluation whereas maximal goals allow more variation in the evaluations. However, this previous research investigated neither performance–goal discrepancies nor self-evaluations of performance.

Another series of studies by Kessler et al. (2010) applied minimal and maximal goal standards to predict how observers evaluate norm violations (e.g., a police officer using torture on a suspect). They theorized that minimal standards set an absolute cut-off point (e.g., torture is always condemned with no exception), implying that all norm violations, irrespective of the degree of severity (e.g., any form of torture, including slapping or punching), lead to punishment. However, if the goal standard is maximal (e.g., there may be situations in which torture might be tolerated), they assumed a linear function. Thus, the degree of norm violation is significant in terms of the level of punishment. This research is more closely related to our research aim since it examines behavior–goal discrepancies. At the same time, Kessler and colleagues were concerned with how norm deviations were evaluated by others (i.e., in terms of punishment) rather than with self-evaluations of norm deviations. Moreover, positive behavior–goal discrepancies (i.e., exceeding a goal equals success) are not possible with norms – there is no possibility of 'over-performance' with a norm of no torture, for example

– so this leaves unanswered the question of how performance–goal discrepancies relate to task satisfaction when a goal is exceeded.

In sum, this prior research considered situations that involved either failure or a negative deviation from a norm. In both cases, minimal goal standards are assumed to provide a reference point that creates a categorical value function (i.e., fulfilling or not fulfilling the norm), whereas with maximal goal standards the reference point is assumed to have a continuous, linear value function (i.e., the degree of deviation from the norm determines level of punishment). In the current research, we deviate from this assumption for three reasons. First, former research defined the minimal and maximal goal standards in terms of their effects (i.e., the value function as being either categorical or continuous). Our definition, however, follows Brendl and Higgins (1996) and does not include the expected effects on the value function. Second, previous research did not involve a context in which task performance could be evaluated on a continuous scale or in which positive goal–performance discrepancies (i.e., exceeding a goals) were possible. Compared to previous studies, we focus only on task goals and task performance that can be evaluated on a continuous scale and that can give rise to demonstrable areas of success (e.g., if the goal is to collect 500 signatures, one can exceed this by collecting 600). We argue that the goal standards of mere goals can change the psychological value function of goal–performance discrepancies. Finally, since we also consider positive goal–performance discrepancies, we theorize about the effect that goal standards might have on the overall value function.

Based on the goal standard definitions provided by Brendl and Higgins (1996), we argue that failing to achieve a minimal goal standard results in negative evaluations – independent of the degree of the performance–goal discrepancy. This is because this standard distinguishes a negative from a non-negative evaluation area. Thus, irrespective of the goal–performance discrepancy, individuals experience a high degree of dissatisfaction when they fail to achieve minimal goal standard. In contrast, failure to achieve a maximal goal standard results in non-positive evaluations. In this case, one can feel still satisfied in case one is close to the goal. Therefore, we predict that when maximal goal standards are not met (i.e., failure), a more continuous relationship between goal–performance discrepancy and satisfaction exist (cf. Giessner & van Knippenberg, 2008; Kessler et al., 2010). However, with success, we expect that these effects will be reversed. Thus, achieving a maximal goal standard implies a strongly positive evaluation, because achieving an ideal goal should be positive in itself (Brendl & Higgins, 1996; Idson et al., 2000). Consequently, when an individual achieves a maximal goal, goal–performance discrepancies should have less impact on his or her subsequent satisfaction. Achieving a minimal goal, however, is non-negative, and evaluations should not automatically become positive in this situation. Thus, we argue that a better performance should also lead to greater task satisfaction – resulting in more positive evaluation, depending on the goal–performance discrepancy. We therefore expect there to be a linear relationship under these conditions. Fig. 1 summarizes our assumptions, and provides a graphical display of the value function based on minimal and maximal goal standards. Like Heath et al. (1999), we focus on specific points of difference between minimal and maximal standards to derive our hypotheses and test these in our experiments.

Following these assumptions, we expect that when goal standards are maximal (versus minimal), satisfaction levels will generally be higher. This is because goal–performance discrepancies are evaluated on a more continuous basis when failures occur, but when there is success, satisfaction levels are generally expected to be very positive.

Hypothesis 1: Overall, task satisfaction levels are higher when individuals are seeking to meet a maximal compared to minimal goal standards.

We also argue that the relationship between goal–performance discrepancies and satisfaction levels vary depending on the outcome in
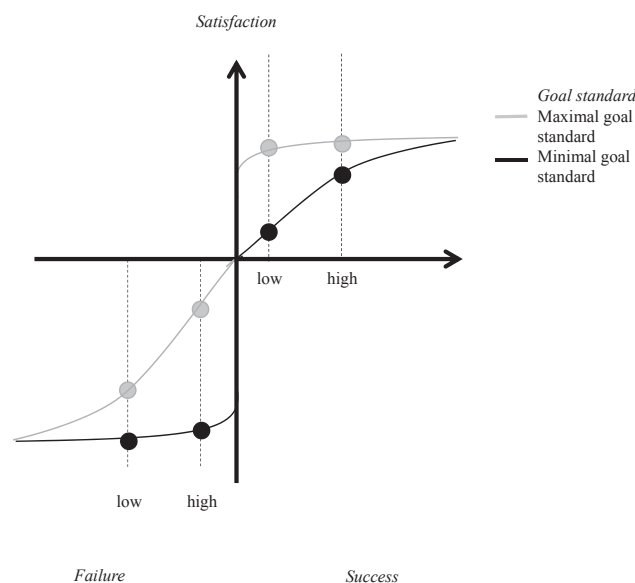

Fig. 1. Goal standards and the value function. The small circles on the dashed lines indicate patterns of the relationship between goal–performance discrepancies and *satisfaction* levels, depending on *goal standards* (minimal or maximal), *performance* (success or failure), and *achievement level* (low or high). We use these patterns to test this model in our studies.

terms of success or failure and the actual level of performance (i.e., achievement level) in either success or failure. The dotted lines in Fig. 1 represent different instances of goal–performance discrepancies and show how the value function changes feelings of task satisfaction when there is a minimal or a maximal goal standard. Based on our previous reasoning, we thus predict that goal standards, performance (i.e., defined as success or failure), and goal achievement levels (i.e., defined as level of performance within failure or success) have a three-way interaction effect on satisfaction.

Hypothesis 2[2]: Goal standards, performance, and level of achievement have a three-way interaction effect on the level of task satisfaction such that:
When individuals experience failure (i.e., do not meet their goal), the level of achievement has less impact on their task satisfaction if a minimal (vs. maximal) goal standard is being used as the reference point. When individuals experience success (i.e., meet or exceed their goal), the level of achievement has less impact on their task satisfaction if a maximal (vs. minimal) goal standard is being used as the reference point.

5. Overview of studies

We report four studies in this paper. All data, syntax, and materials are available at https://osf.io/8gpjr/. These studies differ in their methodology (e.g. the degree to which we use within- and between-participants designs), but all are experiments in which participants undertake a specific task and receive feedback on it. While Studies 1 to 3 test our primary dependent variable, task satisfaction, Study 4 (which engages participants in a simulated negotiation) also measures behavioral consequences in terms of decision-making in a negotiation. We note that, while for Study 1 we aimed for a sample size of more than 50 for each condition, Studies 2 to 4 are all pre-registered and thus we set a-priori sample sizes.

---

[2] Note that, in our pre-registered studies, we refer to sub-hypotheses as 2a and 2b and explain how to test these, too. We provide all these tests in the manuscript but primarily interpret the overall three-way interaction and the overall pattern of results. We thank a reviewer for advising us to do this.

6. Study 1

6.1. Methods

6.1.1. Participants and design

This study was conducted with 300 participants with an average age of 24.64 years (SD = 4.97); 148 were female. Most had either a bachelor's degree (n = 135) or a master's degree (n = 96) as their highest educational qualification. Respondents came from more than 35 countries, with the majority coming from Germany (n = 83) or the Netherlands (n = 66). A mixed design with one within-participants factor (*achievement level*: low/high) and two between-participants factors (*goal standard*: minimal/maximal; *performance*: success/failure) was used. Participants were randomly assigned to the between-participants conditions. The order of the within-participants factor was randomized.

6.1.2. Procedure and measures

The data were collected online and participants were invited via a snowball method, using personal networks. The study was described as being a pilot test for a new short measure of intelligence. Participants were told that they would be answering only a few test questions. Their intelligence score would be ostensibly based on their personal information (e.g., age, gender, education), the number of correct answers, and the time taken to finish the test. Before the study started, participants gave informed consent and answered some demographic questions (i.e., age, gender, education, and mother tongue).

We manipulated the *goal standard* and informed participants in the maximal condition that the *ideal* was to achieve an intelligence score of 350. In the minimal condition, participants were told that they had to obtain a score of *at least* 350. There were two reasons why we chose this wording in order to manipulate maximal and minimal goals. First, minimal is defined as "the least possible" (Merriam-Webster, n.d.) and ideal reflects "an ultimate object or aim of endeavor: goal" (Merriam-Webster, n.d.). Thus, these wordings match our definition of maximal and minimal goals. Second, these wordings have been used in previous experimental research (Giessner & van Knippenberg, 2008).

Participants had to perform two rounds of the ostensible intelligence test – both consisting out of four typical intelligence test questions on pattern recognition and logic. After they worked through the first round of mathematical problems, we gave participants the scores for this first round performance. Here, we manipulated the *performance* and *achievement level*. Participants in the *success* condition received a score of either 380 (*low achievement*) or 440 (*high achievement*). Participants in the *failure* condition received a score of either 260 (*low achievement*) or 320 (*high achievement*).

After each round, participants were asked to rate how satisfied they were with their performance on an 11-point visual scale, which uses a series of smileys (cf. Davies & Brember, 1994). Note that we use a single-item measure of task satisfaction in all of our studies, as has been done in previous research (e.g., Bandura & Cervone, 1983; Heath et al., 1999; Locke, 1969; Mento, Locke, & Klein, 1992). Single-item measures of job satisfaction (Wanous, Reichers, & Hudy, 1997) and sub-facets of job satisfaction (Nagy, 2002) have been shown to be valid and robust if they have strong face validity.

A second round of four mathematical problem then followed. While participants remained in the same goal standard and performance condition as in the previous round, in this round they received the other achievement score (i.e., those who been given 'high achievement' now received 'low achievement' and vice versa). Preliminary analyses with an order factor did not show any interactive effects. Therefore, we do report any results without an order factor.

At the end of the study, an attention item was used to check whether participants had understood the goal standard manipulation correctly. Four different descriptions were offered and participants were asked to select the correct one (i.e., the answers gave the two different standards

[maximal and minimal] and two different goal levels [350 or 400]). A second question asked participants whether they had put in maximum effort while working through the mathematical problems. This question also had four answer options ranging from 1 (Yes, I answered all questions with full effort) to 4 (No, I did not answer any questions with full effort). Finally, the participants were thoroughly debriefed and thanked.

6.2. Results

6.2.1. Attention check and effort

The attention check question on the goal standards was answered correctly by 90% of participants, with mistakes randomly distributed between the two goal standard conditions (n = 30; χ^2 [1] = 2.17, p = .14). Overall, participants seemed to put in a relatively high level of effort while working through the mathematical problems (M = 1.81; SD = 1.12).

6.2.2. Task satisfaction

A three-way ANOVA yielded a significant main effect for performance, $F(1, 296)$ = 60.99, p < .001, η_p^2 = .17, CI [0.11, 0.23], and for achievement level, $F(1, 296)$ = 198.50, p < .001, η_p^2 = .40, CI [0.33, 0.46]. Participants were more satisfied with success (M = 6.83, SD = 2.59) than with failure (M = 5.01, SD = 2.61), and with higher levels of achievement (M = 6.77 SD = 3.00) than with lower levels (M = 5.03, SD = 2.89). We predicted that, overall, participants with maximal goal standards would show more satisfaction than those with minimal standards (Hypothesis 1). This hypothesis was supported with a significant main effect of the goal standard, $F(1, 296)$ = 80.35, p < .001, η_p^2 = .21, CI [0.15, 0.28]. In line with our hypothesis, participants in the maximal goal condition showed more satisfaction than those in the minimal condition (maximal goal: M = 6.87, SD = 2.46; minimal goal: M = 4.67, SD = 2.62).

Hypothesis 2 predicts a three-way interaction effect. While none of the two-way interactions reached significance, the predicted three-way interaction was significant $F(1, 296)$ = 18.19, p < .001, η_p^2 = .06, CI [0.02, 0.11]. In order to understand whether the pattern of means supports the pattern predicted in Hypothesis 2, we performed simple interaction analyses in the performance conditions and thereafter simple main effect analyses to better understand the nature of the interaction (see Fig. 2 for means, SD and CIs).

We predicted that when individuals experience failure, the level of achievement has less impact on their task satisfaction if a minimal (vs. maximal) goal standard is being used as the reference point. In the failure condition, we found a significant main effect for achievement level, $F(1, 151)$ = 120.88, p < .001, η_p^2 = .44, CI [0.35, 0.52], and for goal standard, $F(1, 151)$ = 49.11, p < .001, η_p^2 = .25, CI [0.15, 0.33], which was qualified by an interaction effect, $F(1, 151)$ = 18.39, p < .001, η_p^2 = .11, CI [0.04, 0.19]. Fig. 2 illustrates that achievement levels had a weaker effect on satisfaction in the minimal goal condition, $F(1, 151)$ = 18.30, p < .001, η_p^2 = 0.11, CI [0.04, 0.19], than in the maximal goal condition, $F(1, 151)$ = 151.47, p < .001, η_p^2 = 0.50, CI [0.41, 0.57]. This pattern of results is in line with Hypothesis 2.

For the success condition, we that the level of achievement has less impact on task satisfaction if a maximal (vs. minimal) goal standard is being used as the reference point. In the success condition, we found a significant main effect for achievement level, $F(1, 145)$ = 80.06, p < .001, η_p^2 = .36, CI [0.25, 0.44], and for goal standard, $F(1, 145)$ = 32.53 p < .001, η_p^2 = .18, CI [0.10, 0.27]. The two-way interaction yielded a marginal effect, $F(1, 145)$ = 3.12, p = .08, η_p^2 = .02, CI [0.00, 0.07]. Achievement levels had somewhat stronger effect on satisfaction in the minimal condition, $F(1, 145)$ = 57.01, p < .001, η_p^2 = 0.28, CI [0.18, 0.37], than in the maximal condition, $F(1, 145)$ = 25.96, p < .001, η_p^2 = 0.15, CI [0.07, 0.24]. Again, the pattern of results is in line with Hypothesis 2. Together these patterns with the significant three-way interactive effect supports Hypothesis 2.
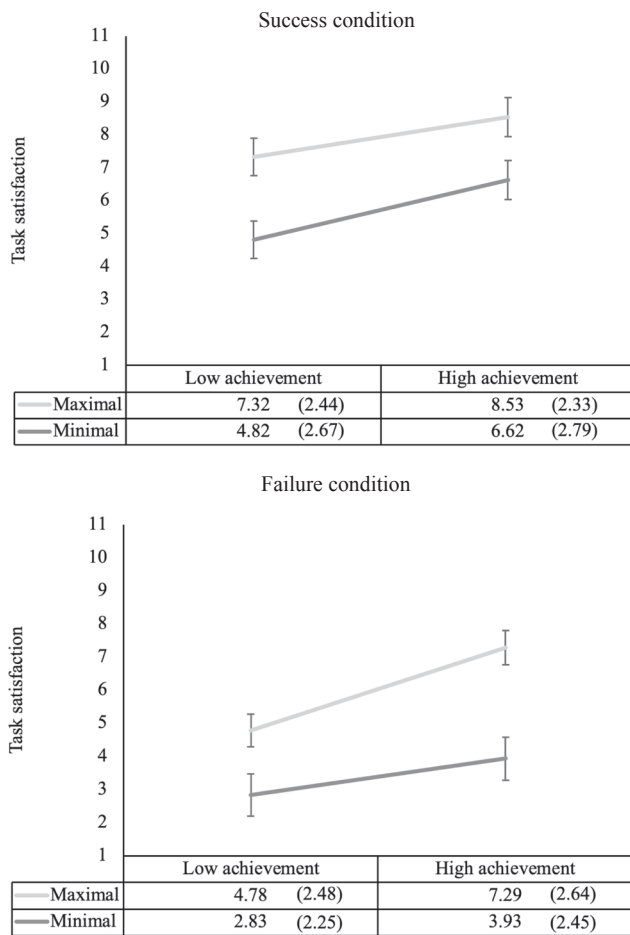
Fig. 2. Mean satisfaction scores in Study 1. Error bars denote 95% confidence intervals around the mean. The table shows means and standard deviations (in brackets).

6.3. Discussion

We found support for both of our hypotheses: a main effect of goal standard on task satisfaction and a three-way interaction effect between goal standard, performance, and achievement level, with a pattern of results that is in line with our predictions in Hypothesis 2. Thus, Study 1 provides support for our prediction that goal standards influence the relationship between the level of achievement (i.e., goal–performance discrepancies) and satisfaction levels.

7. Study 2

Given the initial support for the predicted three-way interaction in Study 1, we decided to test the robustness of the findings on a different task and with a different sample to establish generalizability. Study 2 was pre-registered on osf.io/r3p9w prior to data collection[3]. All materials, data, and syntax can be found on https://osf.io/zmsqg/. We made an a-priori power calculation partly based on Study 1. Because we used a different task and our sampling was via an online panel, we used a smaller effect size of f = 0.20, an alpha error of 0.05 and a power of 0.95 with a 1 degree of freedom for the three-way interaction effect

[3] Please note that during the review process it became clear that our a-priori choice of exclusion criteria might have been unreasonable. We therefore report the results here without these a-priori exclusion criteria. We do report, however, the results with the pre-registered a-priori criteria in the supplement. Note that the results do not actually differ significantly.

with eight groups. We estimated a sample size of 327 using G-Power (Faul, Erdfelder, Buchner, & Lang, 2009). Given that we expected to have to exclude some participants, we aimed for a total sample size of 500.

7.1. Methods

7.1.1. Participants and design

A sample of 501 individuals took part in this study[4]. Participants were on average 37.36 years old (*SD* = 12.02); 367 were female. Respondents came from the UK and the US. We used a mixed design, with one within-participants factor (*achievement level*: low/high) and two between-participants factors (*goal standard*: maximal/minimal; *performance*: success/failure). Participants were randomly assigned to the between-participants conditions. The order of the within-participant factor was randomized.

7.1.2. Procedure and measures

The data were collected using the online panel Prolific Academics (https://app.prolific.ac). We used pre-screening, choosing participants from the UK and the US, who were fluent in English, did not have dyslexia, and had an approval rate of at least 90 percent. The study was described as a pilot test for a new ability test that was relevant to work performance and career success. Participants were informed that they would be expected to complete two test rounds, and would be given a personal test score (PTS) that would ostensibly be based on personal information, the number of correct and incorrect answers, and any answers that were missed out. Before the start of the study participants gave informed consent and answered some demographic questions about their age and gender.

Participants were then informed that they would be doing an adjusted version of the d2-test (Brickenkamp, 1962). The test shows a series of d's and p's with one or two dashes above and/or below (i.e., a total of four dashes possible). The task is to mark all the d's that have two dashes within a time limit of one minute and not to mark any of the other d's or p's. As in Study 2, we manipulated the *goal standard* and informed participants in the maximal condition that the *ideal* was to achieve a PTS of 350. In the minimal condition, participants were told that they had to obtain a PTS of *at least* 350. They then started working through the first round of the d2-test.

The *performance* and *achievement level* manipulations and the measurement of satisfaction followed the procedure as in Study 1. Preliminary analyses with an order factor did not show any interactive effects in the analyses with regard to our primary hypotheses and are therefore not reported here. At the end, the same attention check was undertaken as in Study 1. Finally, after a debriefing, the participants were thanked.

7.2. Results

7.2.1. Attention check

The attention check question on the goal standards was answered correctly by 89% of participants (incorrect: *n* = 55), with more mistakes being made by participants in the minimal goal standard condition (*n* = 41), χ^2 (1) = 14.77, *p* < .001. In round 1, the d2-test performance (correct answers minus incorrect answers) was 26.57 (*SD* = 11.72). In round 2, the average performance was 31.19 (*SD* = 12.74).

7.2.2. Task satisfaction

A three-way mixed ANOVA yielded a significant main effect for both performance, $F(1, 497) = 181.86$, $p < .001$, $\eta_p^2 = .27$, CI [0.22, 0.32],

[4] The difference between the planned sample size and the actual sample size (delta n = 1) is due to the sampling approach in Prolific.
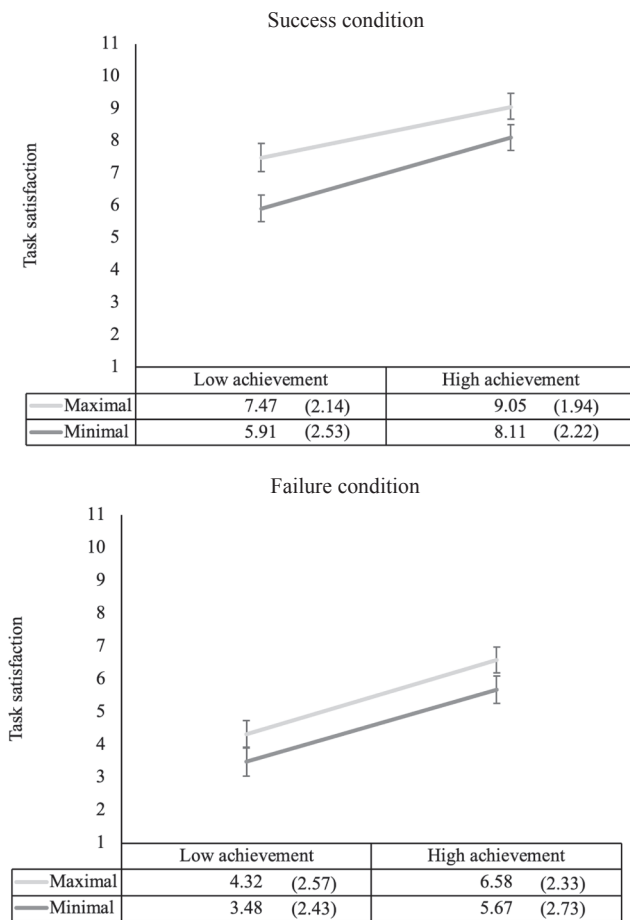
Fig. 3. Mean satisfaction scores in Study 2. Error bars denote 95% confidence intervals around the mean. The table shows means and standard deviations (in brackets).

and for achievement level, $F(1, 497) = 567.97$, $p < .001$, $\eta_p^2 = .53$, CI [0.49, 0.57]. Participants were more satisfied with success ($M = 7.59$, $SD = 2.12$) than with failure ($M = 5,04$, $SD = 2.12$), and with higher levels of achievement ($M = 7.34$, $SD = 2.64$) than with lower levels ($M = 5.28$, $SD = 2.85$).

We predicted that, overall, participants with maximal goal standards would show more satisfaction than those with minimal standards (Hypothesis 1). This hypothesis was supported with a significant main effect of goal standard, $F(1, 497) = 29.84$, $p < .001$, $\eta_p^2 = .06$, CI [0.03, 0.09]. In line with our hypothesis, participants in the maximal goal condition indicated a greater level of satisfaction ($M = 6.76$, $SD = 2.46$) than those in the minimal condition ($M = 5.86$, $SD = 2.61$).

None of the two-way interactions reached significance. We again found support for the three-way interaction as predicted in Hypothesis 2, $F(1, 497) = 4.01$, $p = .046$, $\eta_p^2 = .01$, CI [0.0001, 0.03]. To check whether the pattern of results supports our Hypothesis 2 we performed simple interaction and main effect analyses (see Fig. 3 for means, SD and CIs). In the failure condition, we found a significant main effect for achievement level, $F(1, 249) = 302.30$, $p < .001$, $\eta_p^2 = .55$, CI [0.48, 0.60], and for goal standard, $F(1, 249) = 8.94$, $p = .003$, $\eta_p^2 = .04$, CI [0.007, 0.08]. The interaction, however, was non-significant, $F(1, 249) = 0.09$, $p = .77$, $\eta_p^2 < .001$, $\eta_p^2 < .001$, CI [0.00, 0.01] (see Fig. 3). Nevertheless, the effect was in the expected direction, because achievement levels had a stronger effect on satisfaction in the maximal condition, $F(1, 249) = 166.35$, $p < .001$, $\eta_p^2 = 0.40$, CI [0.32, 0.47], than in the minimal condition, $F(1, 249) = 137.75$, $p < .001$, $\eta_p^2 = 0.36$, CI [0.28, 0.42]. This pattern provides some support for

Hypothesis 2 (within the failure condition). In the success condition, we found a significant main effect for achievement level, $F(1, 194) = 266.11$, $p < .001$, $\eta_p^2 = .52$, CI [0.45, 0.57], and for goal standard, $F(1, 248) = 23.69$, $p < .001$, $\eta_p^2 = .09$, CI [0.01, 0.15]. The two-way interaction was also significant, $F(1, 248) = 7.04$, $p = .008$, $\eta_p^2 = .03$, CI [0.004, 0.07]. Achievement levels had a stronger effect on satisfaction in the minimal condition, $F(1, 248) = 192.16$, $p < .001$, $\eta_p^2 = 0.44$, CI [0.36, 0.50], than in the maximal condition, $F(1, 248) = 87.68$, $p < .001$, $\eta_p^2 = 0.26$, CI [0.19, 0.33]. Thus, overall, these results support Hypothesis 2, with the interactive effects between achievement level and goal standard being stronger in the success condition.

7.3. Discussion

Again, we found support for our hypotheses in this study: There is a main effect of goal standard on task satisfaction and a three-way interaction effect between goal standard, performance, and achievement level, and the pattern of results is in line with our predictions in Hypothesis 2. We note that while the two-way interaction effect in the failure condition did not reach significance, the overall pattern is in line with our predictions in Hypothesis 2. Thus, the findings of Study 2 support the robustness of the findings from Study 1 and attest to their generalizability as we replicated the predicted three-way interaction effect using a different task and with a different sample of participants.

8. Study 3

One weakness of our previous studies was that the goal was relatively abstract. In addition, we may assume that a maximal goal would typically be perceived as more challenging to achieve than a minimal goal. While in practice, this might not be always the case (e.g., think about the tenure criteria at some business schools, which may be framed as either an ideal or a minimal standard, even though tenure is generally rather difficult to attain), it raises the question of whether difficulty per se drives the effects or whether these goal standards themselves have a specific psychological meaning as reference standards that affect the proposed value function. To address these points, we conducted another study in which participants first do a d2-test and then receive their performance score. They are thus given a personal reference point by which to measure their own performance level before we manipulate goal standard and achievement level. In this way, we set a concrete goal standard that is consistent in terms of objective difficulty for all participants.

In this study we only test for the failure condition. The reason for this is practical: The calculations of sample size (based on some pilot tests) indicated that we would require relatively large samples (close to 1,000) to test the full design and this was not feasible (in the time frame) for a controlled laboratory experiment at the university in question. We therefore investigated only the failure condition, since this was the condition for which we found the weakest evidence in Study 2, making it the most pressing condition to study further. In addition, as negative effects are generally stronger than positive ones (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Tversky & Kahnemann, 1992), and our own studies (with the exception of Study 2) showed consistently stronger effects for failure (this includes the pilot studies for Study 2), we reduced the experimental design in this study. Thus, we test only a sub-hypothesis of Hypothesis 2, namely that effect within the failure condition, and run the study in an experimental laboratory at a Dutch business school.

The study was pre-registered on https://osf.io/nfkjd.[5] All materials, data, and syntax can be found on https://osf.io/yjktw/. We made an a-priori power calculation using G-Power (Faul et al., 2009) (effect size:

_____
[5] See comment 3.

0.13; ∝ error probability: 0.05; power: 0.80; numerator df: 1; number of groups: 4). This resulted in a desired sample of 408 participants. Given that we expected to have to exclude some participants, we aimed for a total sample size of 450. In case the sample size ended up below 400 after participants had been excluded, we aimed to recruit an additional 50 participants (i.e., we actually had to do this).

8.1. Methods

8.1.1. Participants and design

In total 504[6] undergraduate business administration students took part in this study. Participants were on average 20.06 years old ($SD = 1.79$); 229 were female. Most of the participants were Dutch ($n = 269$). We used a 2 (*goal framing*: minimal vs. maximal) by 2 (*achievement level*: low vs. high) between-subjects design. Participants were randomly assigned to conditions.

8.1.2. Procedure and measures

Undergraduate business students were recruited for this laboratory study and ethical approval was given by the school. Students took part in exchange for extra course credits. The study was part of a set of three unrelated studies and was run as the first of these. After explaining the purpose of the study (i.e., testing concentration, which is assumed to be predictive of career success) and the procedure (i.e., two rounds of testing), we asked first whether any students had dyslexia (yes/no). After explaining the d2-test (i.e., asking participants to choose the correct answers as fast as possible), participants did a first round of the d2-test (30 lines each with eight options). They then received a point score that indicated how they had performed (i.e., correct answers, minus any incorrect answers). Next, we manipulated the *goal standard*. All participants were asked to improve their performance by 10 points in a second round of the test. This goal was either stated as maximal (i.e., they should *ideally* improve their performance by 10 points) or minimal (i.e., they should improve their performance by *at least* 10 points). Participants then did another d2-test, and in this second round, *achievement level* was the manipulated in the feedback. Participants received feedback indicating either that they were close to the goal (i.e., eight points above the previous performance level) or that they were some distance from it (i.e., only one point above the previous performance level). Next, they rated how satisfied they were, using a sliding scale ranging from 0 to 100. Afterwards, we used an attention check item that asked them which goal standard description they had received (minimal vs. maximal). We also asked whether they had had any technical problems. Finally, we asked for demographic information (i.e., gender, age, nationality, and education). At the end, all participants were debriefed and thanked.

8.2. Results

8.2.1. Attention checks

The attention check question on the goal standards was answered correctly by 85% of participants (incorrect: $n = 76$), and more mistakes were made by those in the maximal goal standard condition ($n = 58$; $\chi^2 (1) = 25.17$, $p < .001$). We had 14 participants with dyslexia.

8.2.2. Performance levels before the experimental manipulations

After the first round of the d2-test, the average performance was 69.27 ($SD = 14.51$). There was no difference between participants who were allocated to the maximal goal standard and those who were

allocated to the minimal goal standard, $t(502) = -0.74$, $p = .46$. Hence, the difficulty level (i.e., achieving an additional 10 points) was the same in both conditions.

8.2.3. Task satisfaction

A two-way ANOVA yielded a significant main effect for achievement level, $F(1, 500) = 45.00$, $p < .001$, $\eta_p^2 = .08$, CI [0.05, 0.12]. Participants were more satisfied with higher levels of achievement ($M = 67.70$, $SD = 19.08$) than with lower levels ($M = 54.92$, $SD = 23.83$). We predicted that, overall, participants with maximal goal standards would show more satisfaction than those with minimal standards (Hypothesis 1). This hypothesis was not supported, because the main effect for goal standard was found to be non-significant, $F(1, 500) = 0.29$, p = .59, η2p = .001, CI [0.00, 0.01]. The mean scores nevertheless suggest that participants in the maximal goal condition felt more satisfaction ($M = 61.80$, $SD = 22.27$) than those in minimal condition ($M = 60.82$, $SD = 22.75$). Obviously, the size of the difference is very small.

In Hypothesis 2 (failure condition), we predicted that when individuals experience failure, the level of achievement has less impact on their task satisfaction if a minimal (vs. maximal) goal standard is being used as the reference point. We indeed found support for this, because the predicted two-way interaction was significant, $F(1, 500) = 9.15$, $p = .003$, $\eta_p^2 = .02$, CI [0.004, 0.04] (see Fig. 4 for means, SD and CIs). Simple main effect analyses within the goal standard conditions show that achievement level has a stronger effect on satisfaction for those in the maximal condition, $F(1, 500) = 47.18$, $p < .001$, $\eta_p^2 = .09$, CI [0.05, 0.13], than for those in the minimal condition, $F(1, 500) = 6.81$, $p = .009$, $\eta_p^2 = .01$, CI [0.002, 0.03]. This pattern thus supports the failure condition element of our Hypothesis 2 (see Fig. 4).

8.3. Discussion

Study 3 again provides support for Hypothesis 2 (in the failure condition). While the null hypothesis for testing Hypothesis 1 could not be rejected, the pattern of results – albeit showing only a small differences – is in line with our prediction and with the findings of the other studies. One reason why we may not have found as strong a main effect could be the exclusion of the success condition. It might be more likely, however, that it is due natural variation in effect sizes (Kenny & Judd, 2019). Importantly, in this study actual performance scores were given to participants and a goal was included that was the same for all participants in terms of its objective difficulty. Thus, the way goal standards influence the relationship between goal–performance discrepancy and satisfaction cannot be explained in terms of the objective difficulty of the goal.

9. Study 4

In our final study we aim to address two limitations of the first three studies. First, the earlier studies were conducted largely with undergraduate students (Studies 1 and 3) and used tasks that were not directly related to organizational behavior. As a result, one might question whether we can generalize our findings to other organizational contexts. Second, we have focused on the level of satisfaction as our main outcome but have not yet outlined how this might lead to further behavioral outcomes. Our last study aims to address both of these shortcomings.

In this study, we asked participants to take part in a negotiation game in which they appeared to be negotiating with another player. They would take on the role of either a buyer or a seller. In reality, all participants were placed in the seller role. In this game, we manipulated the performance, goal standard, and achievement level. Afterwards, we asked our participants about their satisfaction levels. This provided another test of Hypotheses 1 and 2. In addition, we asked our participants whether they would accept the offer from the buyer. Here, we

---

[6] The difference between the planned sample size and the actual sample size (delta n = 4) is due to the procedure of advertising for participants and conducting research in the laboratory. We had to open a certain number of slots, take care of any no-shows, and conduct the study for all those who had signed up.
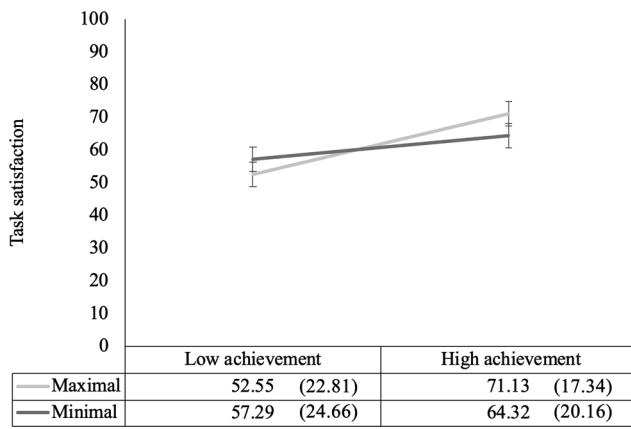
Fig. 4. Mean satisfaction scores in Study 3. Error bars denote 95% confidence intervals around the mean. The table shows means and standard deviations (in brackets).

	Low achievement	High achievement
Maximal	52.55 (22.81)	71.13 (17.34)
Minimal	57.29 (24.66)	64.32 (20.16)

in a desired sample size of 351 participants. Given that we expected to have to exclude some individuals, we aimed for a total sample size of 600.

9.1. Methods

9.1.1. Participants and design

In total 600 students took part in this study. Participants were on average 34.36 years old (SD = 12.19); 354 were female and 1 was transgender. We used a 2 (*goal standard*: maximal vs. minimal) by 2 (*performance*: success vs. failure) by 2 (*achievement level*: low vs. high) between-subjects design.

9.1.2. Procedure and measures

The data were collected using the online panel Prolific Academics (https://app.prolific.ac). We used pre-screening, choosing participants whose first language was English, who had not taken part in one of our previous studies on this topic during the last two years, and who had an approval rate of at least 90 percent. The study was advertised as a
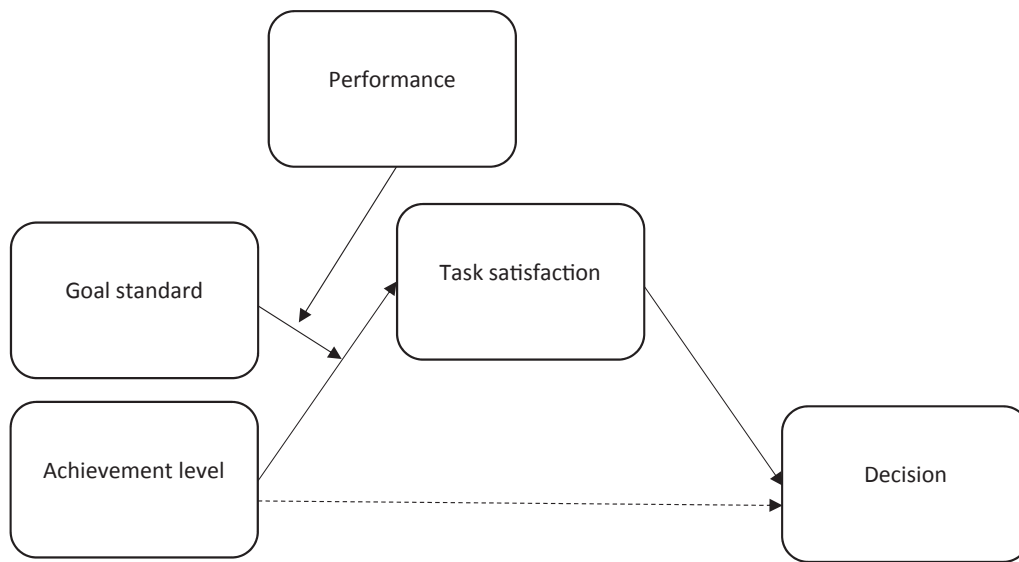

Fig. 5. Moderated moderated mediation model of Study 4.

reasoned that a greater degree of satisfaction with the offer would make someone more likely to accept the buyer's first offer. This is because previous research has shown that dissatisfaction will increase an individual's motivation to continue engaging in a behavior (Bandura & Cervone, 1983; Cervone et al., 1991; Locke & Latham, 2002) – in this case, to continue negotiating. Satisfaction, on the other hand, should make it more likely that the negotiation will come to an end. Thus, we additionally hypothesize:

Hypothesis 3: Task satisfaction will be positively related to a decision to agree to the offer that has been negotiated.

In order to test the full model (see Fig. 5) including the subsequent behavioral effect, we also test the moderated mediation effect.

Hypothesis 4: The three-way interaction predicted in Hypothesis 2 will affect the decision to agree to the offer that has been negotiated via task satisfaction.

The study was pre-registered on https://osf.io/at23u. All materials, data, and syntax can be found on https://osf.io/9sqn6/. Based on a pilot study, we did calculations using G-Power for the three-way interaction effect (Faul et al., 2009) (effect size: 0.15; ∝ error probability: 0.05; power: 0.80; number of groups: 8[7], numerator df: 1). This resulted

negotiation game.

At the beginning of the study, participants gave their informed consent. We asked them to take part in a negotiation simulation about the sale of a carpet, and told them they would be given the role of either seller or buyer. After asking questions about demographics (age and gender), we ostensibly 'matched them up' with another participant. In order to make the apparently interactive nature of the negotiation game more believable, we also said that if they could not be matched up, they would be given an alternative simple scenario. In reality, the 'matching' of all participants took only a few seconds.

All participants were in fact placed in the role of seller. We then manipulated the goal standard. We told them that their task was to sell a carpet for a client. This was an antique carpet (we provided a picture) that had seemingly already elicited some interest from a potential buyer. The participant was told to sell the carpet for £500, but in half of the cases this was expressed as a minimal goal (at least £500) and in the other half it was expressed as a maximal goal (ideally £500). Afterwards, we conducted a fair attention check (Prolific Team, 2018).

(*footnote continued*)

The calculation of 351 participants is based on eight groups and we have been using eight groups in the study. Thus, this is a typographical mistake in the pre-registration.

---

[7] Please note that we wrongly reported four groups in the pre-registration.

To be more precise, we asked on the same web page: "Based on the text above, what is the minimal/ideal selling price you should achieve (type in the exact number as stated above in pounds)?" The alleged buyer then made her/his first offer. Here, we manipulated performance (failure vs. success) and achievement level (low vs. high): £420 (failure/low achievement), £490 (failure/high achievement), £510 (success/low achievement), £580 (success/high achievement). We then asked participants to indicate on a nine-point visual scale, which uses a series of smileys (cf. Davies & Brember, 1994), how satisfied they were. Higher values represented more satisfaction. On the next page, we asked whether they would accept the offer (yes = 1, no = 0).

At the end of the survey, we included two further attention checks. The first was a goal standard attention check and similar to the one used in Study 1. Participants had to choose between four statements describing the goal they had been set: The client asked to sell the carpet for at least £500; The client asked to sell the carpet for ideally £500; The client asked to sell the carpet for at least £400; The client asked to sell the carpet for ideally £400 (the last two answers had the wrong target goal; the correct answer for the first two depended on the goal standard condition). The performance attention check was about the buyer's offer, and also gave four choices: £420, £490, £510, and £580. The correct answer depended on the performance and achievement level condition a participant has been placed in. Finally, all participants were thoroughly debriefed and thanked.

9.2. Results

9.2.1. Exclusion and attention checks

The fair attention check (Prolific Team, 2018)[8] was answered correctly by 85.3% of the participants (incorrect: n = 88). We excluded these 88 participants and had a final sample of 512 participants. Of these, 93.6% gave a correct answer in the goal standard attention check (incorrect n = 33), and there were more mistakes made by those in the maximal goal condition, χ^2 (1) = 12.51, $p < .001$ (maximal: n = 25 and minimal n = 8). For the performance attention check, 99.2% of participants gave correct answers (n = 4 incorrect), and there was no significant difference between the conditions, χ^2 (1) = 1.04, $p = .31$ (success: n = 3, failure n = 1).

9.2.2. Task satisfaction

A 2 (goal standard: maximal vs. minimal) by 2 (performance: success vs. failure) by 2 (achievement level: low vs. high) between-subjects ANOVA yielded a significant main effect for performance, $F(1, 504) = 264.50, p < .001, \eta_p^2 = .34$, CI [0.29, 0.39]. Participants in the success condition experienced higher levels of satisfaction ($M = 7.59$, $SD = 1.75$ vs. $M = 5.00$, $SD = 2.39$). A significant main effect of achievement level, $F(1, 504) = 65.37, p < .001, \eta_p^2 = .12$, CI [0.07, 0.16], indicates that participants were more satisfied with higher levels of achievement ($M = 6.95$, $SD = 2.32$) than with lower levels ($M = 5.65$, $SD = 2.43$). Our Hypothesis 1 predicted that greater task satisfaction should be experienced when individuals had maximal rather than minimal goals. The significant main effect for goal standard, $F(1, 504) = 108.61, p < .001, \eta_p^2 = .18$, CI [0.13, 0.23], supports this, because participants in the maximal goal condition also showed more satisfaction ($M = 7.13$, $SD = 2.07$) than those in the minimal condition ($M = 5.59$, $SD = 2.56$)

These effects were qualified by two-way interactions between goal standard and performance, $F(1, 504) = 6.94, p = .009, \eta_p^2 = .01$, CI [0.002, 0.23], as well as between achievement level and performance, $F(1, 504) = 20.04, p < .001, \eta_p^2 = .04$, CI [0.02, 0.07]. More importantly, however, we found a significant three-way interaction, as we had predicted in Hypothesis 2, $F(1, 504) = 12.82, p < .001, \eta_p^2 = .02$,
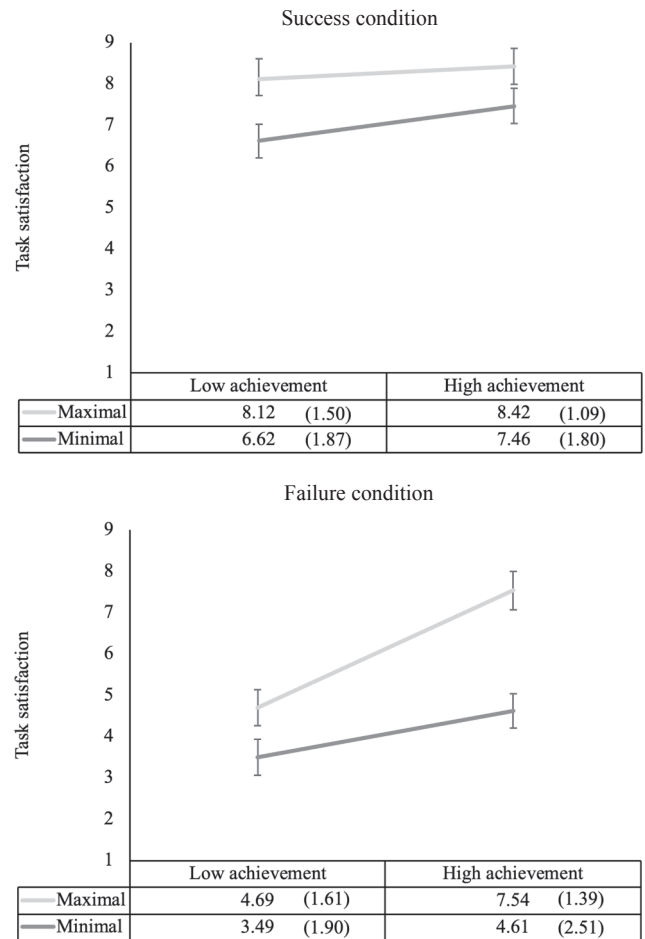

Fig. 6. Mean satisfaction scores in Study 4. Error bars denote 95% confidence intervals around the mean. The table shows means and standard deviations (in brackets).

Success condition

	Low achievement		High achievement	
Maximal	8.12	(1.50)	8.42	(1.09)
Minimal	6.62	(1.87)	7.46	(1.80)

Failure condition

	Low achievement		High achievement	
Maximal	4.69	(1.61)	7.54	(1.39)
Minimal	3.49	(1.90)	4.61	(2.51)

CI [0.007, 0.05] (see Fig. 6 for means, SD and CIs). To better understand this three-way interaction and its pattern, we ran simple interaction and simple main effects analyses.

Simple interaction effect analyses in the performance conditions yielded a stronger effect between goal standards and achievement level in the failure condition, $F(1, 250) = 12.56, p < .001, \eta_p^2 = .05$, CI [0.01, 0.09], than compared to the success condition, $F(1, 254) = 1.78$, $p = .18, \eta_p^2 = .01$, CI [0.00, 0.28]. More importantly, however, the simple main effect analyses provided support for the pattern we had predicted in Hypothesis 2. In the failure condition, achievement levels had a stronger effect on satisfaction when the goal standards were maximal $F(1, 250) = 65.37, p < .001, \eta_p^2 = .21$, CI [0.14, 0.28], rather than minimal goal standards, $F(1, 250) = 11.22, p = .001, \eta_p^2 = .04$, CI [0.01, 0.09]. However, in the success condition, this pattern is reversed, as here achievement levels have a stronger effect on satisfaction in the minimal goal condition, $F(1, 254) = 9.56, p = .002, \eta_p^2 = .04$, CI [0.008, 0.08], than in the maximal goal condition, $F(1, 254) = 1.00$, $p = .32, \eta_p^2 = .004$, CI [0.00, 0.03].

9.2.3. Satisfaction and decision

In hypothesis 3, we predicted that task satisfaction is positively related to a decision to agree to the offer that has been negotiated. To test this, we ran a logistic regression predicting the decision from satisfaction. The analysis supports Hypothesis 3, showing that satisfaction was positively related to the decision to accept the offer, b = 1.19, SE = 0.10, chi^2 (1) = 370.82, p < .001.

Table 1
Test of hypothesis 4 with conditional process analysis (Hayes, 2018).

Mediator model (task satisfaction)						
	b	SE	t	p	LLCI	ULCI
Constant	6.37	0.08	80.67	< 0.001	6.21	6.52
Goal standard (A)	0.82	0.08	10.42	< 0.001	0.67	0.98
Achievement level (B)	0.64	0.08	8.08	< 0.001	0.48	0.79
Performance (C)	1.28	0.08	16.26	< 0.001	1.13	1.44
A × B	0.15	0.08	1.87	0.06	−0.007	0.30
A × C	−0.21	0.08	−2.63	0.009	−0.36	−0.05
B × C	−0.35	0.08	−4.48	< 0.001	−0.51	−0.20
A × B × C	−0.28	0.08	−3.58	< 0.001	−0.44	−0.13
Dependent variable model – logistic regression (decision)						
Constant	−7.73	0.72	−10.65	< 0.001	−9.16	−6.31
Achievement level	−0.05	0.14	−0.33	0.74	−0.33	0.24
Task satisfaction	1.20	0.10	11.41	< 0.001	0.99	1.40

Index of moderated moderated mediation				
	Index	Boot SE	Boot LLCI	Boot ULCI
	−0.34	0.10	−0.56	−0.15
Conditional moderated mediation in performance conditions				
Failure	0.51	0.15	0.24	0.83
Success	−0.16	0.12	−0.42	0.09
Conditional indirect effect in performance/goal standard conditions				
Failure, maximal goal	1.70	0.26	1.28	2.28
Failure, minimal goal	0.67	0.25	0.22	1.21
Success, maximal goal	0.18	0.15	−0.10	0.50
Success, minimal goal	0.50	0.21	0.14	0.95

9.2.4. Conditional indirect effects modeling

In Hypothesis 4, we predicted that three-way interactive effect between goal standard, performance and achievement level on task satisfaction will affect the decision to agree to the offer that has been negotiated. To test this, we ran a mediated moderation analysis using the process macro in SPSS (Hayes, 2018). This step-wise procedure (see Table 1) tests several regression models and calculates indices for moderated mediation and mediation effects for each condition to estimate (a) whether there is an overall conditional indirect effect (i.e., our Hypothesis 4) and (b) how effects are mediated in each condition. The model of the process macro corresponding to our Hypothesis 4 is model 11 (see Fig. 5). The first regression we tested was the regression on the mediator, thus on task satisfaction. The results resemble those we reported for Hypotheses 1 and 2, but we report it as a regression in Table 1. To provide support for Hypothesis 4, the three-way interaction should be significant and the pattern should follow Hypothesis 2. We have already shown this to be the case (see analysis above). The second logistic regression was conducted to test how satisfaction and achievement level jointly predicted the participant's decision. This analysis showed that task satisfaction has a significant effect on the decision. Finally, supporting Hypothesis 4, the index of moderated moderated mediation is significant.

To explore this complex model in more detail, we report the conditional moderated mediation in each of the performance conditions. Here, the effect is only significant in the failure condition – which is not a surprise, given that the simple interaction effect on task satisfaction was non-significant in the success condition. Looking at the conditional indirect, we expect a stronger mediation effect for maximal (compared to minimal) goal standards in the failure condition and the reverse (i.e., stronger effect of minimal compared to maximal goal standard) in the success condition. This is because we expect that deviations from the goal matter more for maximal (vs. minimal) goal standards within failure and for minimal (versus maximal) goal standards within success. The effects in the performance/goal standard condition indicates that all of the mediation effects are significant, except for the one in the success/maximal goal condition. That the mediation effect is non-significant in this condition is due to the fact that the simple main effect of achievement level on satisfaction is non-significant. More importantly, however, the conditional indirect effect of the maximal goal is stronger than that of the minimal goal in the failure condition. In the success

condition this pattern is reversed. This provides further support for our hypotheses.

9.3. Discussion

Study 4 again provides support for Hypotheses 1 and 2. Extending the previous studies, we show subsequent behavioral effects in a negotiation context. In line with Hypothesis 3, we demonstrate that task satisfaction is positively related to the decision to agree to the offer that has been negotiated. In line with Hypothesis 4, we provide evidence that the interactive effect of achievement level, goal standard, and performance affects the decision to accept the buyer's offer via satisfaction.

10. General discussion

Goals are important for self-regulation (Brendl & Higgins, 1996; Locke & Latham, 2002). This is because their function is to provide information that helps individuals to make sense of their performance by comparing it to their goals (goal–performance discrepancies). Early goal-setting theory (Locke, 1969) proposed that there is a linear relationship between goal–performance discrepancies and satisfaction – in other words, the smaller the discrepancy in the case of failure, and the larger the discrepancy in the case of success, the more satisfied one will feel with one's performance on a task. Later research suggested that the relationship may be better described using the value function as detailed in prospect theory (Heath et al., 1999; Kahneman & Tversky, 1979). The current research extends this view by exploring the role of goal standards in the relationship between goal–performance discrepancies and satisfaction. We reasoned that this relationship differs depending on whether the goal standard is minimal or maximal and also depending on the outcome (i.e., success or failure). Throughout our studies, we found substantial support for Hypothesis 1, namely that maximal (versus minimal) goals generally result in higher satisfaction levels. Only in Study 3 this effect was non-significant in the failure condition, although even in this study the results were in the predicted direction. More importantly, all of our studies provided support for the interaction effect predicted in Hypothesis 2. The overall pattern was relatively consistent and supported our theoretical model (see Fig. 1). Finally, our last study showed that satisfaction levels can have consequences for subsequent decision-making (Hypothesis 3), demonstrating there are thus behavioral consequences of the goal–performance discrepancies depending on the goal standard (Hypothesis 4).

The variation in significance might be explained due to natural variation in effect sizes. Recent large-scale replication projects like Many Labs and Many Labs 2 (Klein et al., 2014, 2018) show that the use of the same methodological procedure still yields heterogeneity in effect sizes. Kenny and Judd (2019) argue that it might be advisable to conduct more than one study and to accept heterogeneity in results even if the same or similar methodology is used. While we aimed for large samples and made a-priori power sample size calculations for Studies 2 to 4, this does not prevent potential heterogeneity in effect sizes. Thus, in line with these considerations, one should be cautious about interpreting the results of a single study and should rather interpret the overall findings of our research. From this perspective, we believe that our studies provide very strong overall support for our hypotheses.

10.1. Theoretical and practical implications

Our results show that maximal and minimal goal standards change the satisfaction function after performance feedback. Individuals who fail to meet minimal (versus maximal) goals feel less satisfied and are less sensitive to the actual goal–performance discrepancy. Thus, using the language of prospect theory (Kahneman & Tversky, 1979), when failure is experienced, diminished sensitivity is more marked for

minimal goals than for maximal goals. The reverse seems to be true for success. While participants are more satisfied when they have maximal rather than minimal goal standards, goal–performance discrepancies have less impact on satisfaction when the goals are maximal.

Overall, our results show that goal standards appear to affect the value function of the relationship between goal achievement levels and satisfaction. Our findings therefore qualify some of the previous definitions of minimal and maximal goal standards. Giessner and van Knippenberg (2008) and Kessler et al. (2010) defined maximal goal standards as being evaluated on a continuous scale, whereas evaluations should be categorical for minimal goal standards. A problem with these previous definitions is that they imply the expected effects of the goal standard (i.e., in terms of the value function the goal will produce). In contrast, we defined minimal standards as "oughts that a person must attain or standards that must be met" (Idson et al., 2000, p.254), and maximal standards as ideals that one hopes to achieve. In line with our reasoning, our results indicate that the satisfaction function of these goal standards is more complex. It seems to be continuous for maximal (versus minimal) standards when the goals are not met (i.e., failure). This continuous relationship is also present for minimal (versus maximal) standards when goals are achieved or overachieved (i.e., success). In contrast, the evaluation seems to be more categorical – in the sense that the actual level of achievement matters less – where there is either failure to achieve minimal goal standards or success in achieving maximal goal standards. Hence, our findings support the argument for revising previous theory to include maximal and minimal goal standards as reference points.

The findings show that goal-setting theory can be extended with a focus on goal standards. As an inductive theory, goal-setting theory is well equipped to incorporate such extensions (cf. Locke & Latham, 2020). Understanding the effects of goal standards helps to clarify how the level of goal achievement affects satisfaction – one of the key variables in Locke and Lathams's goal-setting theory (1990, 2002) and in Bandura's theorizing on the effect of self-evaluations on performance (e.g., Bandura & Cervone, 1983). Both of these frameworks suggest that satisfaction is one of the ingredients for subsequent goal adjustment and motivation. Indeed, our fourth study supports these findings.

One should note, however, that the relationship between more general job satisfaction and performance (Judge, Thoresen, Bono, & Patton, 2001) is more complex, as is the relationship between task satisfaction and performance (Locke & Latham, 1990). For instance, previous research indicates that other variables such as self-efficacy (Bandura & Cervone, 1983), goal commitment, and goal changes (Locke & Latham, 1990) also affect task performance. Consequently, future research might further explore how goal standards influence the relationship between goal–performance discrepancies and other predictors of performance such as self-efficacy and self-goal-setting.

Our research further indicates that integrating the literature on self-regulation (Brendl & Higgins, 1996) and goal-setting theory (Locke & Latham, 2002) would be fruitful for developing new theoretical frameworks and models that explain individual experiences of goal achievement. While these literatures seem to have developed along rather separate lines, recent approaches (e.g., Latham & Brown, 2006) and the current findings suggest that these two theoretical fields can learn from one another.

Making the distinction between maximal and minimal standards might also help to explore further the concept of a *stretch go*al – defined as "an organizational goal with an objective probability of attainment that may be unknown but is seemingly impossible given current capabilities" (Sitkin, See, Miller, Lawless, & Carton, 2011, p. 547). While we did not study organizational goals in this research, it might be interesting to examine whether goal standards might be part of how stretch goals work. Sitkin and colleagues argue that stretch goals can fulfill a motivational function when employees have recently experienced success and when resources are available. On the one hand, stretch goals are difficult goals as they are seemingly impossible to

reach. As maximal goal standards represent ideals, these should be perceived as more difficult goals in general. On the other hand, minimal goals can also be very ambitious and seemingly impossible. As we indicated earlier, getting a tenure at a university can be regarded as a very ambitious goal, even though a minimal level of achievement is specified (e.g., getting at least a certain number of papers published in top-tier journals). Thus, it would be interesting to explore how minimal and maximal goal standards can be integrated with theory on stretch goals. We speculate that only maximal standards allow seemingly impossible goals to have a motivational effect, because not reaching such a goal can still create a positive experience in terms of satisfaction and can therefore lead to continued striving towards it.

An important insight from our findings is that the relationship between goal–performance discrepancies and satisfaction is non-linear and depends on the goal standard. This has methodological consequences for field studies because a simple correlational function might not represent this relationship well and may therefore may not capture that relationship accurately. While previous research on goal-setting showed that, when examining the relationship between performance and satisfaction, the actual goal set needs to be considered (Kernan & Lord, 1991; Locke & Latham, 1990), we extend this view and show that the goal standard is another factor that influences this relationship. Thus, when studying the effects of performance on satisfaction, we recommend measuring the goal level and the goal standard, and using non-linear modeling to test this relationship.

The current findings also provide insights for managers that can help them to better understand the satisfaction of their employees after performing on their daily tasks. We show that maximal standards generally lead to more satisfaction than minimal ones. In other words, managers might be better off framing goals as ideals. This is in line with the literature on visionary leadership (Conger & Kanungo, 1987; Stam, Lord, van Knippenberg, & Wisse, 2014), which argues that leaders should use ideals when crafting visions. A side effect of these maximally framed visions might be that employees who aim to reach such maximal goal standards are generally more satisfied with their achievements, even if they fail to reach those goals. Furthermore, given that task satisfaction should feed into overall job satisfaction, and managing the job satisfaction of employees is in itself important goal for managers (Spector, 1997), managers and leaders will be more able to influence satisfaction levels either by setting a maximal standard or intervening when a minimal goal has not been met.

Finally, previous research suggests that individuals' perceptions of goals might be influenced by the self-regulatory systems of promotion and prevention focus (Higgins, 1998; Idson et al., 2000; Kluger, Stephan, Ganzach, & Hershkovitz, 2004). *Promotion focus* is concerned with advancement, growth, and accomplishment – with the individual being driven by the notion of reaching a desired end state. In contrast, *prevention focus* is concerned with security, safety, and responsibility – with the individual following a strategic inclination to be prudent and cautious, and to avoid anything that does not fit with the desired end state. When faced with difficult tasks, promotion-focused individuals are more eager to experience success than prevention-focused individuals, who are more sensitive to potential failure or mistakes (Crowe & Higgins, 1997; Higgins et al., 1997). These regulatory foci represent a more distal motivational system that influences the more proximal goal standards. Idson et al. (2000, p. 254) proposed that "inducing a promotion focus will induce representation of the goal as maximal (a standard one hopes to achieve), whereas inducing prevention focus will induce representation of it as a minimal goal (a standard one must achieve)". A recent study by Lalot, Quiamzade, and Falomir-Pichastor (2018) found support for this notion. Consequently, our findings might also have implications for the literature on regulatory focus.

10.2. Limitations and strengths

As with any research, our studies have limitations and strengths. Each of our paradigms has certain specific weaknesses. In all of the studies, participants were engaged in an actual task and evaluated their own satisfaction. However, we used simple concentration tasks in Studies 1–3, which may be difficult to generalize to other settings. Furthermore, participants might have been suspicious about our experimental manipulations – although from the comments we later received from our participants, we do not feel that this was the case. We believe, however, that the strength of our studies lies in the overall consistency of the findings (see Kenny & Judd, 2019). Another notable strength is that pre-registered studies were used.

More generally, using an experimental approach makes it difficult to judge the extent to which such goal standards have a practical impact. Nevertheless, our last study provides support for the potential impact such standards can have on subsequent decisions as a result of the satisfaction or dissatisfaction experienced when trying to achieve them. However, we cannot generalize our findings to all contexts. Furthermore, previous studies have shown that in certain contexts other variables are also important. For instance, Medvec, Madey, and Gilovich (1995) studied satisfaction among Olympic medalists. They found that silver medalists showed less satisfaction than bronze medalists – a finding that our studies cannot explain. This context differs from the one used in our research, however, in several important ways. First, in these competitions, the actual performance scale is limited in terms of placement (i.e., one cannot do better than the gold medalist). In our studies, the actual performance measurement is on a continuous scale and the goal standard sets a mere goal (i.e., meaning that no reinforcing resource is related to the goal) with a reference point to differentiate between success and failure. Second, when it comes to ranking in terms of first, second, and third place, these ranks have more connotations. It is thus not a mere goal. This is not the case in our research context. Finally, the research by Medvec et al. (1995) looks at competition between different individuals (see also Garcia, Tor, & Gonzalez, 2006). Our studies focus on one's own goal and how one judges goal–performance discrepancies in terms of one's satisfaction levels. Thus, in the current research, we have set specific parameters (i.e., no social comparison; no a-prior meaning attached to specific performance levels; task performance can be evaluated on a continuous scale; mere goals). Future research should explore how goal standards can be applied and integrated in other settings.

Furthermore, earlier negotiation research suggested that economic outcomes and subjective evaluations can be disconnected. Specifically, negotiators who focused on more ambitious anchors - more maximal in our terms - achieved better outcomes but were less satisfied compared to negotiators who focused on less ambitious anchors – more minimal goals (e.g., Galinsky, Mussweiler, & Medvec, 2002). Moreover, Galinsky et al. (2002) find that effects on satisfaction and performance become aligned when less ambitious anchors are brought to the attention of those who originally focused on more ambitious anchors. More recent research by Schaerer, Schweinsberg, Thornley, and Swaab (2020) extended this finding by showing that shifting the focus of the other negotiating party to the reservation price of this party (i.e., minimal levels) leads to more favorable economic outcomes for the offering party (lower counteroffers of recipients) and to more favorable relational outcomes (higher satisfaction of recipients). This research relates to the current work since it looks at maximal and minimal goal standards in the context of negotiations and also demonstrates influences on satisfaction and performance (that sometimes are aligned and sometimes are not). We note that one critical difference between these studies and the current work is that whereas these prior studies manipulated goal levels and goal standards (i.e., more ambitious goals are more difficult and maximal standards whereas less ambitious goals are less difficult and minimal standards), we directly manipulated goal standard with equal goal levels. Further, in the case of focusing on more ambitious

goals, participants experience failure whereas in the case of focusing on less ambitious goals, participants experience success (cf. Galinsky et al. 2002). Our research addresses more directly the effects of minimal and maximal goal standards independent of goal level and considers the under and overachievement of the goal reference point (i.e., success and failure). We find that effects on satisfaction and subsequent performance (in terms of decisions) are similar possibly because goal levels and performance levels cannot bias and anchor the feeling of satisfaction in our studies. Of course, we only focused on one of the negotiating parties and it would be interesting to see what would happen when different negotiating parties would have different (or similar) goal standards to focus on.

While we predicted an overall function for the relationship between goal–performance discrepancies and satisfaction (see Fig. 1), we did not test the overall function. Using a similar approach to other research that has applied the prospect function to goals (Heath et al., 1999), we chose specific points of interest and derived hypotheses from them, which we then tested experimentally. It would, however, be interesting to test the whole function – although this might require a large sample.

The current research focused on satisfaction as an outcome variable. In line with the reasoning of Weiss (2002), we defined satisfaction as an evaluative judgment about task achievement. Locke et al. (1970), however, defined satisfaction as an affective state, thus relating it to an emotional experience. It would certainly be interesting for future research to explore further the role of emotional experiences. While our research primarily addresses the valence dimension of an emotional experience, we did not tap into the specific emotions and examine what role they may play in the relationship between goal–performance discrepancies and satisfaction depending on which goal standard applies.

11. Conclusion

Goal-setting is an effective way of increasing employee motivation, satisfaction, and performance. While setting specific and challenging goals is one part of the story, it seems that how managers set goals for employees is another important consideration. Failing to meet goals does not necessarily always result in a high degree of dissatisfaction, nor does succeeding at a goal always lead to strong sense of satisfaction. Rather, the degree of satisfaction depends on the goal–performance discrepancy and the goal standard. As satisfaction with task performance has implications for continued striving to meet goals, managing it is an important part of a manager's job. The current research helps to clarify the complex relationship between performance and satisfaction and thus provides a new theoretical perspective on the goal-setting literature.

CRediT authorship contribution statement

Steffen R. Giessner: Conceptualization, Writing - original draft, Writing - review & editing, Data curation. **Daan Stam:** Conceptualization, Writing - original draft, Writing - review & editing. **Rudolf Kerschreiter:** Writing - original draft, Writing - review & editing. **Danny Verboon:** Conceptualization. **Ibrahim Salama:** .

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.obhdp.2020.08.004.

References

Austin, J. T., & Vancouver, J. B. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin, 120,* 338–375.

Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology, 45,* 1017–1028.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger

than good. *Review of General Psychology, 5,* 323–370.

Brendl, C. M., & Higgins, E. T. (1996). Principles of judging valence: What makes events positive or negative? *Advances in Experimental Social Psychology, 28,* 95–160.

Brickenkamp, R. (1962). *Aufmerksamkeits-Belastungs-Test (Test d2). [The d2 Test of attention.]* (1st ed.). Göttingen: Hogrefe.

Brief, A. P. (1998). *Attitudes in and around organizations.* Thousand Oaks, CA: Sage.

Carver, C. S., & Scheier, M. F. (1982). Control theory: a useful conceptual framework for personality–social, clinical, and health psychology. *Psychological Bulletin, 92,* 111–135. https://doi.org/10.1037/0033-2909.92.1.111.

Carver, C. S., & Scheier, M. F. (2000). On the structure of behavioral self-regulation. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.). *Handbook of self-regulation* (pp. 41–84). San Diego, CA: Academic Press.

Cervone, D., Jiwani, N., & Wood, R. (1991). Goal-setting and the differ- ential influence of self-regulatory processes on complex decision- making performance. *Journal of Personality and Social Psychology, 61,* 257–266.

Chen, X., & Latham, G. P. (2014). The effect of priming learning vs. performance goals on a complex task. *Organizational Behavior and Human Decision Processes, 125*(2), 88–97.

Conger, J. A., & Kanungo, R. N. (1987). Towards a behavioral theory of charismatic leadership in organizational settings. *Academy of Management Review, 12,* 637–647.

Crowe, E., & Higgins, E. T. (1997). Regulatory focus and strategic inclinations: Promotion and prevention in decision-making. *Organizational Behavior and Human Decision Processes, 69,* 117–132.

Davies, J., & Brember, I. (1994). The reliability and validity of the 'Smiley' scale. *British Educational Research Journal, 20*(4), 447–454.

Drach-Zahavy, A., & Erez, M. (2002). Challenge versus threat effects on the goal-performance relationship. *Organizational Behavior and Human Performance, 88,* 667–682.

Elicker, J. D., Lord, R. G., Kohari, N. E., Ash, S. R., Hruska, B. J., & Medvedeff, M. E. (2010). Velocity as a predictor of performance satisfaction, mental focus, and goal revision. *Applied Psychology, 59,* 495–514. https://doi.org/10.1111/j.1464-0597.2009.00409.x.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41,* 1149–1160.

Galinsky, A. D., Mussweiler, T., & Medvec, V. H. (2002). Disconnecting outcomes and evaluations: The role of negotiator focus. *Journal of Personality and Social Psychology, 83*(5), 1131–1140.

Garcia, S. M., Tor, A., & Gonzalez, R. (2006). Ranks and rivals: A theory of competition. *Personality and Social Psychology Bulletin, 32,* 970–982.

Giessner, S. R., & van Knippenberg, D. (2008). "License to fail": Goal definition, leader group prototypicality, and perceptions of leadership effectiveness after leader failure. *Organizational Behavior and Human Decision Processes, 105*(1), 14–35.

Gould, R. (1939). An experimental analysis of "levels of aspiration". *Genetic Psychology Monographs, 21,* 3–115.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis* (2nd Ed.). New York: The Guilford Press.

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as Reference Points. *Cognitive Psychology, 38,* 79–109.

Higgins, E. T. (1990). Personality, social psychology, and person-situated relations: Standards and knowledge activation as a common language. In L. A. Pervin (Ed.). *Handbook of Personality* (pp. 301–338). New York, NY: Guilford.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52,* 1280–1300.

Higgins, E. T. (1998). Promotion and prevention as a motivational duality: Implications for evaluative processes. In S. Chaiken, & Y. Trope (Eds.). *Dual-process theories in social psychology* (pp. 503–525). New York, NY, US: Guilford Press.

Higgins, E. T., Shaw, J., & Friedman, R. S. (1997). Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology, 72,* 515–525.

Idson, L. C., Liberman, N., & Higgins, E. T. (2000). Distinguishing gains from nonlosses and losses from nongains: A regulatory focus perspective on hedonic intensity. *Journal of Experimental Social Psychology, 36,* 252–274.

Judge, T. A., Thoresen, C. J., Bono, J. E., & Patton, G. K. (2001). The job satisfaction job performance relationship: A qualitative and quantitative review. *Psychological Bulletin, 127,* 376–407.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–292.

Kenny, D. A., & Judd, C. M. (2019). The unappreciated heterogeneity of effect sizes: Implications for power, precision, planning of research, and replication. *Psychological Methods, 24,* 578–589.

Kernan, M. C., & Lord, R. G. (1991). An application of control theory to understanding the relationship between performance and satisfaction. *Human Performance, 4,* 173–185.

Kessler, T., Neumann, J., Mummendey, A., Berthold, A., Schubert, T., & Waldzus, S. (2010). How do we assign punishment? The impact of minimal and maximal standards on the evaluation of deviants. *Personality and Social Psychology Bulletin, 36,* 1213–1224.

Klein, R. A., et al. (2014). Investigating variation in replicability: A "Many Labs"

replication project. *Social Psychology, 45,* 142–152.

Klein, R. A., et al. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science, 1,* 443–490.

Kluger, A. N., Stephan, E., Ganzach, Y., & Hershkovitz, M. (2004). The effect of regulatory focus on the shape of probability-weighting function: Evidence from a cross-modality matching method. *Organizational Behavior and Human Decision Processes, 95,* 20–39.

Lalot, F., Quiamzade, A., & Falomir-Pichastor, J. M. (2018). Is regulatory focus related to minimal and maximal standards? Depends on how you ask!. *European Journal of Social Psychology, 48,* 174–186.

Latham, G. P., & Brown, T. C. (2006). The effect of learning vs. outcome goals on self-Efficacy, satisfaction and performance in an MBA program. *Applied Psychology, 55*(4), 606–623.

Lewin, K., Dembo, T., Festinger, L., & Sears, P. S. (1944). Level of Aspiration. *Personality and the Behavior Disorders* (pp. 349–394). Oxford, England: Ronald Press.

Locke, E. A. (1969). What is job satisfaction? *Organizational Behavior and Human Decision Processes, 4,* 309–336.

Locke, E. A., & Bryan, J. F. (1968). Grade Goals as Determinants of Academic Achievement. *The Journal of General Psychology, 79,* 217–228.

Locke, E. A., & Bryan, J. F. (1969). The directing function of goals in task performance. *Organizational Behavior and Human Performance, 4,* 35–42.

Locke, E. A., Cartledge, N., & Knerr, C. S. (1970). Studies of the Relationship between Satisfaction, Goal-Setting, and Performance. *Organizational Behavior and Human Performance, 5,* 135–158.

Locke, E. A., & Latham, G. P. (1990). *A Theory of Goal Setting & Task Performance.* Englewood Cliffs, NJ, US: Prentice-Hall Inc.

Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation. A 35-Year Odyssey. *The American Psychologist, 57,* 705–717.

Locke, E. A., & Latham, G. P. (2020). Building a theory by induction: The example of goal setting theory. *Organizational Psychology Review.* https://doi.org/10.1177/2041386620921931.

Lord, R. G., & Hanges, P. J. (1987). A control systems model of organizational motivation: Theoretical development and applied implications. *Behavioral Science, 32,* 161–178. https://doi.org/10.1002/bs.3830320302.

Medvec, V. H., Madey, S. F., & Gilovich, T. (1995). When less is more: Counterfactual thinking and satisfaction among Olympic medalists. *Journal of Personality and Social Psychology, 69,* 603–610.

Mento, A., Locke, E., & Klein, H. (1992). Relationship of goal level to valence and instrumentality. *Journal of Applied Psychology, 77,* 395–405.

Merriam-Webster. (n.d.). Ideal. In Merriam-Webster.com dictionary. Retrieved June 3, 2020, from https://www.merriam-webster.com/dictionary/ideal.

Merriam-Webster. (n.d.). Minimal. In Merriam-Webster.com dictionary. Retrieved June 3, 2020, from https://www.merriam-webster.com/dictionary/minimal.

Nagy, M. S. (2002). Using a single-item approach to measure facet job satisfaction. *Journal of Occupational and Organizational Psychology, 75,* 77–86.

Podsakoff, P. M., & Farh, J. L. (1989). Effects of feedback sign and credibility on goal setting and task performance. *Organizational Behavior and Human Decision Processes, 44,* 45–67.

Prolific Team (2018). Using attention checks as a measure of data quality. Retrieved on 30th March 2020 at https://researcher-help.prolific.co/hc/en-gb/articles/360009223553-Using-attention-checks-as-a-measure-of-data-quality.

Schaerer, M., Schweinsberg, M., Thornley, N., & Swaab, R. I. (2020). Win-win in distributive negotiations: The economic and relational benefits of strategic offer framing. *Journal of Experimental Social Psychology, 87,* Article 103943.

Seijts, G. H., Latham, G. P., Tasa, K., & Latham, B. W. (2004). Goal setting and goal orientation: An integration of two different yet related literatures. *Academy of Management Journal, 47,* 227–239.

Sitkin, S. B., See, K. E., Miller, C. C., Lawless, M. W., & Carton, A. M. (2011). The paradox of stretch goals: Organizations in pursuit of the seemingly impossible. *Academy of Management Review, 36,* 544–566.

Spector, P. E. (1997). *Job satisfaction: Application, assessment, causes, and consequences.* Thousand Oaks, CA: Sage Publications.

Stam, D. A., Lord, R. G., van Knippenberg, D., & Wisse, B. (2014). An image of who we might become: Vision communication, possible selves, and vision pursuit. *Organization Science, 25*(4), 1172–1194.

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory – Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty, 5,* 297–323.

Vallacher, R. R., & Wegner, D. M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology, 57,* 660–671.

Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: How good are single-item measures? *Journal of Applied Psychology, 82,* 247–252.

Weiss, H. M. (2002). Deconstructing job satisfaction: Separating evaluations, beliefs, and affective experiences. *Human Resource Management Review, 12,* 173–194.

Wood, R. E., & Locke, E. A. (1990). Goal-setting and strategy effects on complex tasks. *Research in Organizational Behavior, 12,* 73–109.