Aspirational Rules*

Adi Leibovitch†
Hebrew University of Jerusalem

Alexander Stremitzer‡
ETH Zurich

Mila Versteeg§
University of Virginia

January 30, 2019

Abstract

An agent promises to do his duty in exchange for a fixed wage. The content of his duty is subsequently defined by a set of rules that specify the extent to which a multidimensional task needs to be performed. The agent then chooses his effort level on each dimension. His choice on each dimension affects the welfare of others. There is no third party enforcement of the agent's promise. In an experimental study, we find that requiring extreme as opposed to reasonable effort on one dimension not only reduces mean effort provision on that dimension, but may also have a negative spillover effect on rule compliance on other dimensions. This cautions against the setting of aspirational rules.

Keywords: aspirational rules, overpromising, legal design.

JEL-Classification: C90, D02, K23, K33, K37.

*The authors are grateful to ### We are also grateful to Seminar Audiences at the Max-Planck-Institute in Bonn and ETH Zurich. We thank Ashly Eckerling, David Gerber, Stefan Wehrli and Henry Kim for excellent research assistance.

†Associate Professor of Law, Hebrew University of Jerusalem, aleibovitch@mail.huji.ac.il.

‡Professor of Law, Economics, and Business, ETH Zurich, Center for Law & Economics, IFW E49, Haldeneggsteig 4, 8092 Zurich, Switzerland, astremitzer@ethz.ch.

§Professor of Law, University of Virginia School of Law, versteeg@law.virginia.edu.

1 Introduction

Employees promise to perform what is required of them. Their actions affect the stakeholders of the companies they work for. Civil servants, judges, ministers, and presidents take solemn oaths of office to faithfully execute the laws of their country. What they do affects public welfare. In many of these cases, third party enforcement of rules is not available. This is true for a country that has agreed to an international treaty but does not want to abide by its commitment. It is equally true for a president who is bound by the constitution but may choose to ignore the pronouncement of the constitutional court. And it is even true for employees, civil servants, and judges who can get away with breaking the law because of institutional protections or because their actions are not verifiable in court. In these setting, should rules be aspirational or realistic?

This is of interest to managers setting goals for their workers, to government agencies issuing regulation, to legislatures passing laws or ratifying international treaties, to supreme courts issuing rulings, or to the framers of a constitution. Setting aspirational goals could be inspiring for agents and make them work harder. On the other hand, being confronted with overly ambitious requirements might be discouraging. Similarly, if aspirational goals are part of a bundle of goals they could motivate agents to work harder on other tasks (positive spillovers), have negative effects on other tasks (negative spillovers), or have no impact on other tasks at all. We find that aspirational goals not only depress mean effort on the dimension on which the aspirational goal was set, but that it may also have negative spillover effects on other dimensions.

We conducted an online experiment asking participants to perform three tasks, and varying the level of performance they were requested to promise in one of the tasks. The tasks presented to participants included two real effort tasks and one no-effort task. The first real effort task asked participants to count how often the digit "1" occurred in a table with 200 digits between 1 and 9. The second real effort task asked participants to transcribe words from the Ancient Greek alphabet to the Latin alphabet, using a transcription table.

The third task consisted of a dictator game where participants were awarded 100 coins and had to decide how many coins to give to another participant.

We run three treatments. The baseline condition—"No Promise"—did not prescribe any targeted quantity, but simply asked participants if they want to participate and perform the tasks. In the "Aspirational Promise" condition, participants promised to perform tasks to the extent they would be required to, and were subsequently asked to count 50 tables, transcribe 20 words, and transfer between 60–80 coins.[1] In the "Realistic Promise" condition, participants could make the same promise and where subsequently asked to count 15 tables, transcribe 20 words, and transfer between 60–80 coins. Under each of the three conditions, participants payoff remained identical and included only a flat fee of $3, regardless of performance levels. To remove any concern that participants might be trading-off performance levels across the different tasks, participants were informed that the proceeds from their performance will be paid to three different third parties—a charity, and two other players with whom they were randomly matched.

The basic results from our experiment indicate that setting the aspirational goal for the tables task, not only depressed mean effort on the tables task, but also had a negative spillover effect on performance in the words task. The results demonstrate how setting an aspirational goal can backfire in two separate ways: First, the aspirational goal lead to decreased, rather than increased, performance of the relevant task. In addition, the aspirational goal created a spillover effect—decreasing performance in the other real-effort task, despite the fact that a completely realistic goal was set for that task. We conclude, that contrary to common wisdom, setting an aspirational goal might move us further from, not closer to, reaching that goal, and in the process it might also move us further from reaching other independent goals.

Our paper proceeds as follows: Section 2 presents our experimental design. Our results are presented in Section 3. Section 4 discusses results and explores policy implications.

[1] In different runs of the experiment we tested different thresholds for the coins tasks—60, 70, and 80 coins. In each run, the same amount of coins was requested in both the realistic promise and the aspirational promise treatments. Results for all three tasks were similar regardless of the amount of coins promised, and we therefore present them together.

2 Design and Procedure

Participant were given an opportunity to perform three tasks. The first task—the tables task—consisted of counting how often the digit "1" occurred in a table containing 200 digits ranging from 1 to 9. Given that the task did not require any special skills, we assume that performance was a function only of effort (see Abeler et al, 2011). After scanning the table and counting the occurrences of the specified digit, participants entered the count into a dialog box on the screen. An answer was considered correct if it fell within a range of +/-2 of the true value. For example, if the correct number of $1s$ was 42 while the participant counted 40, the result was treated as correct. If participants gave an answer outside of this tolerated margin, they could retry counting the table as often as they wished but had to wait 15 seconds after each failed trial before they could make a new input.[2] Participants were required to enter some input at least every 90 seconds; failure to do so terminated the experiment. After each completed screen (whether successfully or unsuccessfully), participants were asked whether they wanted to continue to the next screen. If they decided to go on, a new table was displayed; if subjects chose to stop the task, they were directed to the next task.

The second task—the words task—consisted of transcribing words from the Ancient Greek alphabet to the Latin alphabet. Participants were presented with a transcription table displaying all the letters is the Ancient Greek alphabet and their equivalents in the Latin alphabet, and we assume here as well that performance was a function only of effort. On each screen one word was presented, and participants entered the transcription of the full word into a dialog box on the screen. An answer was considered correct only if the entire word was accurate. If participants gave a wrong answer, they could retry transcribing the word as often as they wished. After each completed screen (whether successfully or unsuccessfully), participants were asked whether they wanted to continue to the next screen. If they decided to go on, a new word was displayed; if subjects chose to stop the task, they

[2]We introduced this waiting time after each failed trial in order to dissuade participants from guessing repeatedly without counting any digits. Counting one table takes between 45 seconds and 1 minute, so a wait time of 15 seconds seems sufficient to achieve this goal.

were directed to the next task.

The third task—the Coins task—was a dictator game. Participants were given 100 extra coins in addition to their pay and were asked to decide how many of the coins to transfer to another player.

As previously described, we implemented two treatments that varied the requested quantity to be performed in the tables task (the "aspirational goal" and the "realistic goal" conditions), and one baseline condition (the "no goal" condition) that did not express any target quantity. In both goals treatments, only the performance goal for the tables task is manipulated—15 tables in the realistic goal condition, and 50 tables in the aspirational goal condition. Performance goals for the other two tasks—20 words and between 60-80 coins—remain constant under both conditions. That enables us to test both for the direct effect of setting an aspirational goal on the performance in the tables task, and for the spillover effect of setting an aspirational goal in the tables task on performance in the words and coins tasks.

The payment scheme was set to ensure that participants have no direct monetary motivation to perform the tasks. Rather participants received a flat fee for participating in the experiment and beneficiaries for all three tasks were different third parties: a charity in the tables task, and two other participants with whom they are randomly matched in the words and coins tasks. By using different recipients for each task, we overcome the concern that participants might be trading-off performance across the tasks targeting some overall pay to transfer. Rather, all participants were informed in advance that they were already matched with the recipients, and that the recipient for each task is different.

Both active players and recipients were given an opportunity to practice all three task before the experiment began. But they were informed in advance of their roles, and of the fact that only players will actually perform the tasks. After practicing the tasks, players were asked to decide whether they want to participate in the experiment (in the no goal condition) or to make a promise to fulfill the set goal (in the goal conditions). Participants who chose to make a promise were then randomized between the realistic and aspirational

goal conditions.

All participants were also informed about the overall payment scheme for everyone. All participants (both active players and recipients) received 50 cents for reading the instructions and practicing one example of each task, and an additional 50 cents for completing an exit questionnaire. Players were paid an additional $2 if they chose to promise to fulfil the goal (or to participate, in the no goal condition). To create a sense of higher stakes, we used "coins" for the payment transferred to recipients by the players, with each coin worth 1 cent. Participants were told in advance that for every table the players count ten coins will be donated to the Save the Children Foundation; for every word they transliterate ten coins will be given to one recipient with whom they were already randomly matched, and that the coins they choose to transfer will be paid to a second random recipient. Players and recipients were also asked about their first and second order believes regarding the players' performance in the exit questionnaire, and could gain additional bonuses if their estimates were accurate.

We evaluated the effort that participants exerted under the different treatments along three dimensions: quantity, quality, and time spent. Quantity is measured by the number of completed tables, and the percentage of players who kept their promise to count at least 20 words or transfer at least 60-80 cents. Quality of performance is determined by the accuracy with which participants completed the task. We recorded each entry by participants in each task, and calculated the ratio between unsuccessful and successful entries in the tables and the words tasks. Time spent is the overall time participants invested in the experiment.

The experiment was conducted online using the Amazon Mechanical Turk platform. We decided to conduct the experiment online because we wanted participants to have real opportunity costs when deciding whether to continue with the task. We measured effort by eliciting the point at which participants prefer some other activity over continuing with the experiment. In a laboratory setting, participants have low opportunity costs because they cannot leave until the session is over. By contrast, at home, participants can easily stop and switch to a preferred activity.

Overall 369 participants were recruited: 48 in the no goal condition, 164 in the realistic goal condition, and 157 in the aspirational goal condition. Participants had to complete the stages of the experiment within strict time limits and were kept informed of this fact with recurring screen messages. If participants logged out or did not finish stages within these time limits, the experiment was automatically aborted, and participants were notified that they were excluded from the experiment. We set those time limits to force participants to focus on the task and block Internet distractions that can easily distort the results of online studies. Participants were informed up front about the amount of time they would need to complete the whole study, which thereby reduced the likelihood that they would have to break off the experiment because they ran out of time. The online instructions given to participants and screenshots of the experiment are presented in the appendix.

3 Results

3.1 Treatment Effect on Quantity and Promise Keeping

We first present how participants' mean performance differed across treatments (Figure 1). For the manipulated task (the tables task), the realistic goal leads to a mean effort of 8.5 screens (Median = 6), while the aspirational goal leads to a mean effort of only 5.3 screens (Median = 2). The treatment effect is statistically significant at the 1% level (Student's t-test $p < .001$, Wilcoxon rank sum $p < .001$; Table A1 reports all p-values).[3]

Compared with the no goal condition, where the mean effort was 2.8 screens (Median= 2), while setting a realistic goal significantly improved performance in the tables task (Student's t-test $p < .001$, Wilcoxon rank sum $p < .001$), setting an aspirational goal had a much more modest effect, driven mostly by performance of players in the top quartile (Student's t-test $p = .04$, Wilcoxon rank sum $p = .98$).

In addition, the goal set for the tables task has also affected the likelihood that participants reach the goal (held constant) in the words task. Despite the fact that under

[3]Note that Student's t-test is not appropriate for evaluating the significance of the differences in means involving the realistic goal condition, because, as can be seen from Figure 2, that distribution is bimodal. Wilcoxon rank sum test indicates that the difference between the two distributions is significant.

both conditions participants were requested to transcribe the same number of words, the frequency of fulfilling the goal was 48% under the realistic condition, but only 33% under the aspirational condition. This difference is statistically significant at the 1% level (Fisher's exact test $p < .001$).

Importantly, this spillover effect of the treatment is evident both for the likelihood of reaching the goal and for the mean quantity. It is not only that more people keep their promise and transcribe the requested number of words under the realistic condition. There was also a difference in mean performance levels: the realistic goal condition leads to a mean effort of 11.5 words (Median = 10.5), while the aspirational condition leads to a mean effort of 9.4 words (Median = 5). To evaluate the statistical significance of the difference, Student's t-test is significant ($p = .03$), but is again inappropriate because of the bimodal shape of the distributions. The results of Wilcoxon rank sum test are not statistically significant ($p = .41$), but this is likely because the shape of the distributions remain similar under the two conditions (see Figure 2). However, there are substantial differences in the likelihood of keeping the promise, which is what our above analysis using the Fisher's exact test enables to capture.

The spillover effect of setting an aspirational goal in the tables task for performance in the words task is evident not only when looking at performance on the aggregate, but also when comparing performance on the individual level. To test for consistency of individual performance we ranked all participants by their relative performance within each of the three experimental conditions. Participants' relative output across tasks is highly and positively correlated ($r = 0.74$) indicating that participants who perform less on the tables task also perform less on the words task, and vice versa.[4]

In both the realistic and the aspirational goals performance exceeded that of the no goal condition, where participants only transcribed 5.1 words on average (median = 2.5), which is expected given that the realistic goal for the words task itself set a benchmark for

[4]The correlation in output across tasks is similar when looking at each condition separately: $r = 0.8$ in the no goal condition, $r = 0.73$ in the realistic goal condition, and $r = 0.63$ in the aspirational goal condition.

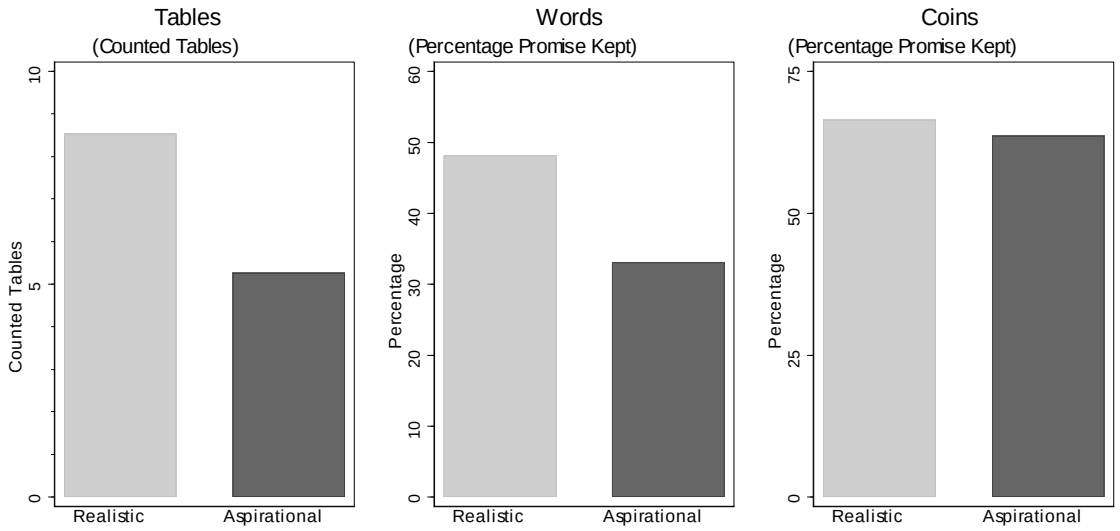

Figure 1: Mean Effort and Promises Kept Across Treatments

performance in this task.

We did not find a similarly significant difference for the coins task. 66% of participants kept their promise under the realistic condition, in comparison to 64% of participants under the aspirational condition (Fisher's exact test $p = .34$). They transferred 49.5 coins on average (Median = 60) under the realistic condition, in comparison to 48.4 coins (Median = 60) under the aspirational condition (Student's t-test $p = .74$, Wilcoxon rank sum $p = .64$). We cannot tell at this point whether these results are driven by a difference between tasks requiring real effort or payment, or whether the number of dimensions involved in the promise matters (for example, participants may wish to keep the promise in at least one task, and the coins task may be the easiest one to keep).

Again, in both the realistic and the aspirational conditions performance exceeded that of the no goal condition, where participants only transferred 23.3 coins on average (Median = 10), which is expected given that the realistic goal for the coins task itself set a benchmark for performance in this task.

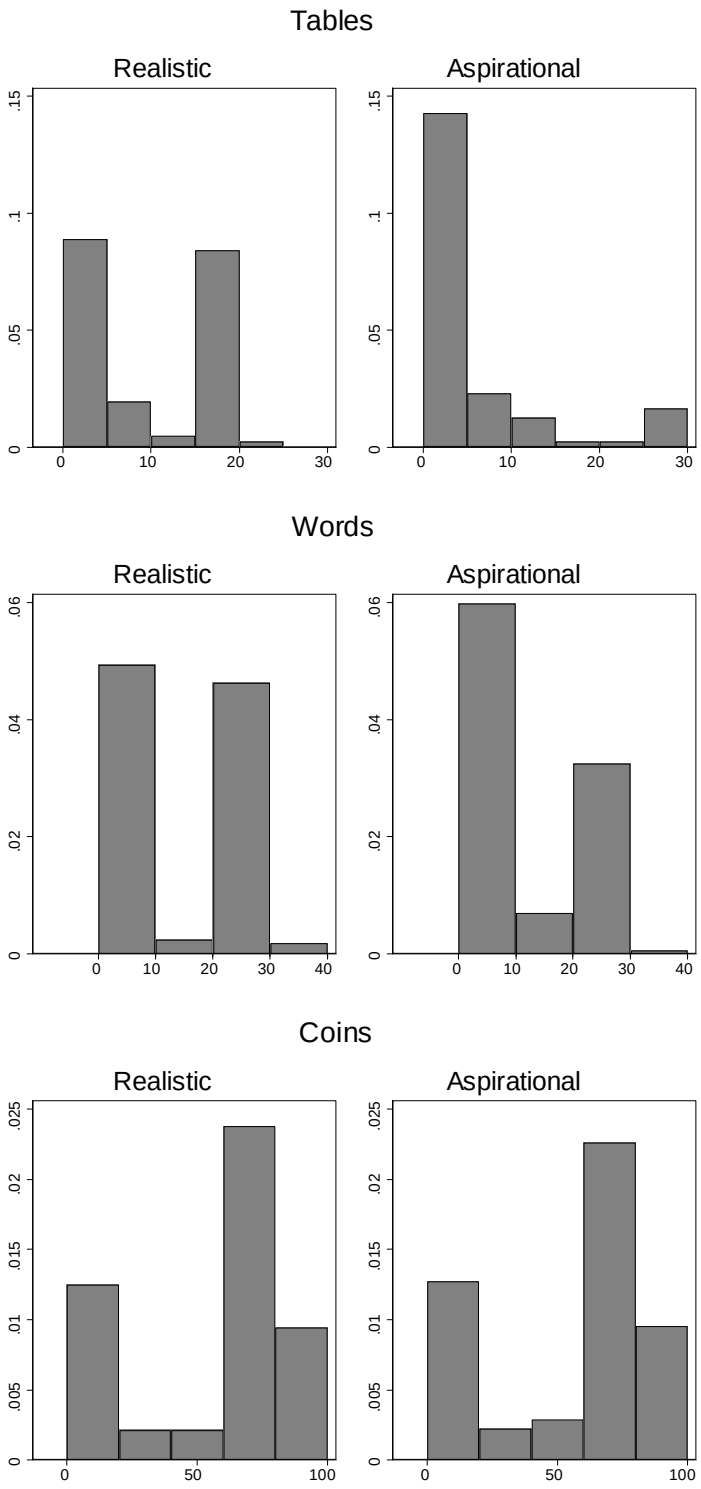

Figure 2: Distribution of Effort Levels Across Treatments

3.2 Treatment Effect on Quality

In addition to effort measured in terms of quantity, we also explore the quality of effort exerted, measured by the accuracy with which participants perform the tasks. One concern could be that quantity and quality may be in conflict. In other words, the aspirational goal may lead to lower quantity but higher quality (higher accuracy of performance). To test for accuracy of performance we form an inaccuracy score for each participant measuring the ratio between failed attempts and correct answers in each task. Figure 3 shows how the mean inaccuracy score differs across treatments.

For the tables task, we find that participants are significantly more accurate in the realistic goal condition than in the aspirational goal condition, with mean inaccuracy scores of 0.14 and 0.23 respectively (Student's t-test $p = .09$). For the words task, we find that participants exerted similar accuracy in both treatments, with a mean inaccuracy score of 0.129 in the realistic condition and of 0.131 in the aspirational condition.

The positive relationship between quality and quantity is in line with the findings of Brooks, Stremitzer, and Tontrup (2017), showing that contract frames that motivate participants make them exert more effort on both the quality and the quantity dimensions of effort. The relationship between quality and quantity also strengthen our results on the quantity dimension. For the manipulated task, setting an aspirational goal leads participants to perform less both in quality and in quantity, in comparison to when a realistic goal is set.

3.3 Treatment Effect on Time Investment

Another measure of agents' devotion to the task is how much time they invest in it. Agents might slack through how diligently they work in a given amount of time, or in how much of their time they devote to fulfilling the set goals. Testing for the duration of time that participants spent on the experiment we find participants spent 24.1 minutes on average (Median = 23.6) in the realistic condition, in comparison to 20.6 minutes on average (Median = 18.8) in the aspirational condition. This difference is statistically significant at the 1%

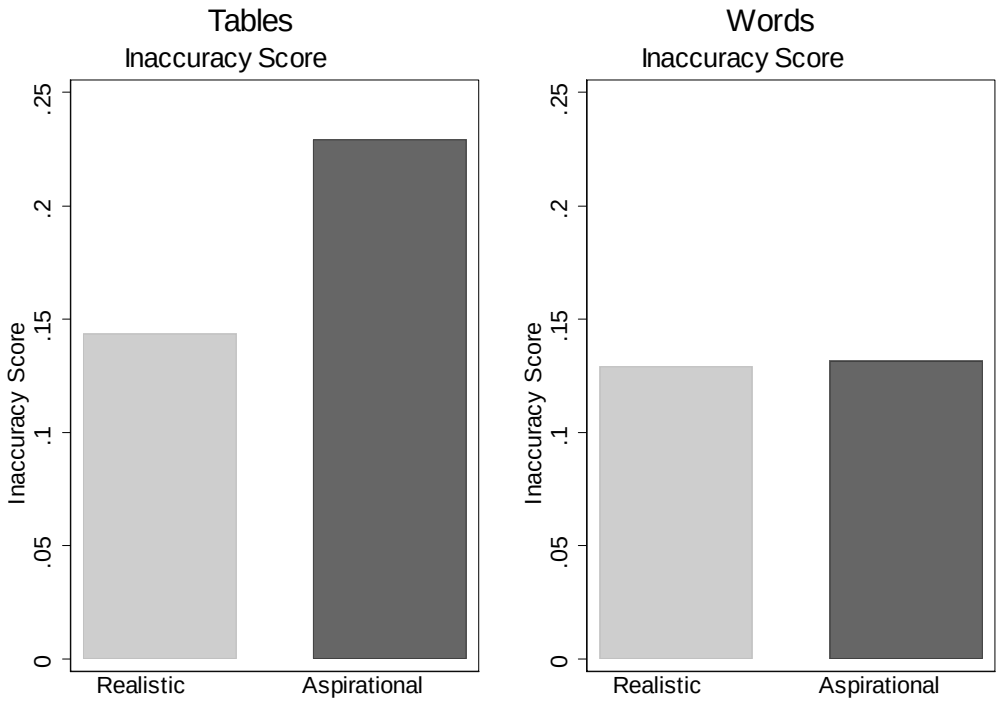

Figure 3: Inaccuracy Level Across Treatments

level (Student's t-test $p < 0.01$; Wilcoxon rank sum $p < 0.01$).

We only have a measure for the duration of time spent on the overall experiment, not on each of the tasks. Because outside the tables and the words tasks the remainder of the experiment required a relatively constant time investment (in reading the instructions, transferring coins in the dictator game, answering the exit questionnaire etc.) it is likely that our estimates here understate the difference in time investment in the tasks directly. Our results indicate that setting the aspirational goal did not motivate participants to invest more time in performing the tasks, but rather less, in addition to the decreased quantity and quality of performance we observe above.

3.4 Believes and Expectations

In addition to testing players' performance in the experiment, we also asked the recipients of the proceeds from the players' performance in the tasks to predict how they think players

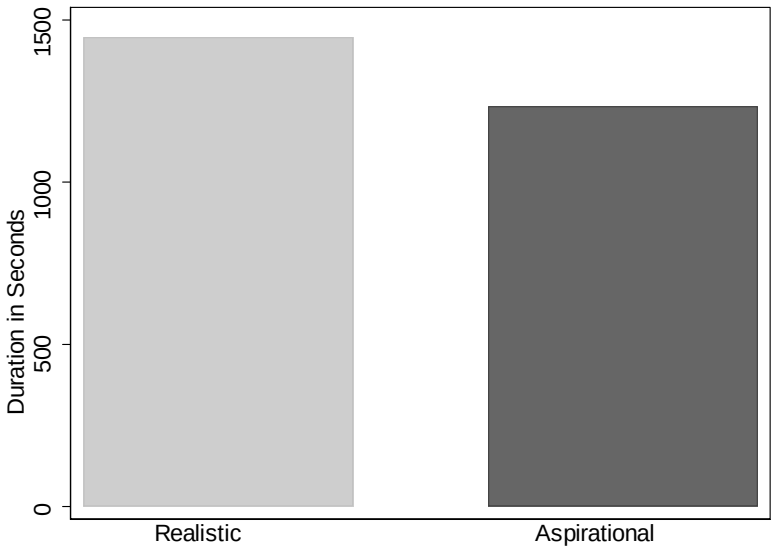

Figure 4: Mean Duration Across Treatments

behave. The recipients were notified of the goals set for players, and then asked to predict the players' performance on each of the three tasks.

For the manipulated task (the tables task), recipients expected players to perform better the more ambitious the goal was. Recipients predicted that players will count 27.2 tables on average (Median = 25) under the aspirational goal, in comparison to 12.1 tables on average (Median = 13) under the realistic goal (Student's t-test $p < .001$, Wilcoxon rank sum $p < .001$; Table A2 reports all p-values), and 10.1 screens on average in the no goal condition (Median = 5). These prediction were overly optimistic on all realms, and extremely inaccurate as to the relationship across goals, as described above. Players performed less on all dimensions—counting a mean of 2.8 tables in the no goal condition, 8.5 tables in the realistic goal condition, and only 5.3 tables in the aspirational goal condition – and they performed worse, not better, on the aspirational goal in comparison to the realistic one.

As for the two unmanipulated tasks, recipients also did not predict any spillover effect, and their expectations were similar under both treatment conditions. For the words tasks, recipients predicted that players will reach the goal of counting 20 words 43% of the time, under both the realistic and the aspirational conditions. They predicted that on average

players will count 15.8 words in the realistic condition and 15 words in the aspirational one (Median = 17 for both). For the coins task, recipients predicted that players will transfer the set amount of coins 23% of the times, under both the realistic and the aspirational conditions. They predicted that on average players will transfer 42 coins in the realistic condition and 43.4 coins in the aspirational one (Median = 50 for both).

This kind of overoptimism regarding the effect of the aspirational goal on the manipulated task, and the lack of foresight of any spillover effect for the other real effort task, might explain why so many hold the view that setting aspirational goals can promote achievements. Despite the fact that recipients had the opportunity to practice each task themselves beforehand, they could not predict that the aspirational goal will have such an inverse effect on the players' performance.

4 Discussion

In many realms of private and public law setting aspirational goals is viewed as inspiring for agents to make them work harder. Our findings cast doubt on that common wisdom and practice. The findings demonstrate that setting overly ambitious goals can have the perverse effect of discouraging effort on all three dimensions: quantity of output, quality of output, and time investment. Further, the findings demonstrate that such disparaging effect is not limited only to the dimension in which the aspirational goal is set, but can also create negative spillover and discourage effort on other unrelated dimensions.

The insights from this study can be applied to multiple areas of law and public policy. One area is constitutional rights. One of the most salient trends in constitution-making worldwide is that constitutions include an ever-expanding catalogue of rights. Between World War II and today, the average number of rights in national constitutions more than doubled (Goderis and Versteeg, 2014). Today, numerous constitutions contain not only civil and political rights such as property rights and freedom of religion, but also a range of socio-economic rights— such as the right to education and the right to healthcare— environmental rights, consumer rights, and even rights for animals. Many of these rights obligations require government

officials perform certain tasks, such as increasing access to education, or guaranteeing basic healthcare.

Our findings suggest that, when it comes to constitutional rights, less might be more. That is, one possible take-away from our experimental study is that when a constitution includes a large number of rights, the document as a whole becomes aspirational. Specifically, those entasked with providing goods such as healthcare, education or ensuring animal health will view the many promises as unrealistic and scale down their efforts to fulfil them. And not only that, the aspirational nature of the bill of rights as a whole might also affect their behavior on constitutional duties that are more easily attainable: such as ensuring that the state does not torture. This possibility is consistent with observational data: studies have found a negative correlation between the number of rights in constitutions and overall rights performance (Law and Versteeg, 2013).

Another area is international human rights law. Since WWII, the number of international agreements dealing with human rights has increased dramatically. There are now nine "core" multilateral human rights treaties, as well as several optional protocols. These treaties cover a wide range of protections, including social rights, women's rights, children's rights, and rights for the disabled. Every state has committed itself to at least one of the core treaties, and the median state has ratified eight of them (Cope et al., 2019). In addition, most states have also joined one of several regional human rights mechanisms.

Like with constitutional rights, it is possible that the many different treaty commitments scale down the efforts of those who are supposed to fulfill them. Indeed, some commentators have raised this possibility. For example, Eric Posner has developed a version of this argument. According to Posner, "if there were only a few rights it would seem simple enough to determine whether states comply with them." (Posner, 2014, 92-94). Yet, when there are many rights, states are basically free to pick and choose the rights that they focus on, since it is not possible to protect them all at once. In the same vein, Mary Ann Glendon has argued large catalogues of rights will produce competing rights values, which "may well trivialize this essential core without materially advancing the proliferating causes that have

been reconceptualized as involving rights" (Glendon, 1991, xi).

Environmental treaties may be another area where the many different multilateral environmental treaties may provide so many different obligations (according to the International Environmental Agreement Database Project, some 1300 multilateral agreements have been concluded till date) that countries scale down their overall effort. It may be for this reason that the Paris Agreement actually sets two different goals: (1) to keep the global temp well below 2°C above pre-industrial levels and (2) to pursue efforts to limit the temperature increase to 1.5 °C above pre-industrial levels. The reason for including a less ambitious goal is that keeping global temperature increase below 2°C is ambitious but plausible, while limiting global temperature increase to 1.5°C is most likely infeasible. This is supported by the International Energy Agency's (IEA) assessment of the Paris Climate Accords, which has noted that "[t]he 2 °C pathway is very tough: the road to 1.5 °C goes through uncharted territory," (World Energy Outlook, 2016, 5).

References

Abeler, J., A. Falk, L. Goette, and D. Huffman (2011): "Reference points and effort provision," *American Economic Review*, 101(2), 470–92.

Agency, I. E. (2016): *World Energy Outlook 2016*. International Energy Agency (IEA).

Brooks, R. R., A. Stremitzer, and S. Tontrup (2017): "Stretch It but DonâĂŹt Break It: The Hidden Cost of Contract Framing," *The Journal of Legal Studies*, 46(2), 399–426.

Glendon, M. A. (2008): *Rights talk: The impoverishment of political discourse*. Simon and Schuster.

Goderis, B., and M. Versteeg (2014): "The diffusion of constitutional rights," *International Review of Law and Economics*, 39, 1–19.

Law, D. S., and M. Versteeg (2013): "Sham constitutions," *Cal. L. Rev.*, 101, 863.

——— (2019): "Empirical Study of Human Rights," *Annual Review of Law and Social Science*, forthcoming.

Posner, E. A. (2014): *The twilight of human rights law*. Oxford University Press, USA.

A Appendix A: Tables

Table A1: Statistical Tests for Mean Effort Levels Across Treatments

	Fisher's Exact Test	Student's t-test	Wilcoxon Rank Sum Test
Tables			
NP vs. Real	$p < .01^{***}$	$p < .01^{***}$	$p < .01^{***}$
NP vs. Asp.	$p < .01^{***}$	$p = .04^{**}$	$p = .98$
Real vs. Asp	$p < .01^{***}$	$p < .01^{***}$	$p < .01^{***}$
Words			
NP vs. Real	$p < .01^{***}$	$p < .01^{***}$	$p < .01^{***}$
NP vs. Asp.	$p < .01^{***}$	$p < .01^{***}$	$p = .03^{**}$
Real vs. Asp	$p < .01^{***}$	$p = .03^{**}$	$p = .41$
Coins			
NP vs. Real	$p < .01^{***}$	$p < .01^{***}$	$p < .01^{***}$
NP vs. Asp.	$p < .01^{***}$	$p < .01^{***}$	$p < .01^{***}$
Real vs. Asp	$p = .34$	$p = .75$	$p = .73$

Table A2: Statistical Tests for Mean Expectations Across Treatments

	Fisher's Exact Test	Student's t-test	Wilcoxon Rank Sum Test
Tables (Tables Counted)			
NP vs. Real		$p = .21$	$p < .01^{***}$
NP vs. Asp.		$p < .01^{***}$	$p < .01^{***}$
Real vs. Asp		$p < .01^{***}$	$p < .01^{***}$
Words (Percentage Promise Kept)			
NP vs. Real	$p < .01^{***}$	$p < .01^{***}$	$p < .01^{***}$
NP vs. Asp.	$p < .01^{***}$	$p < .01^{***}$	$p < .01^{***}$
Real vs. Asp	$p = 1.00$	$p = 1.00$	$p = 1.00$
Coins (Percentage Promise Kept)			
NP vs. Real	$p < .01^{***}$	$p < .01^{***}$	$p < .01^{***}$
NP vs. Asp.	$p < .01^{***}$	$p < .01^{***}$	$p < .01^{***}$
Real vs. Asp.	$p = 1.00$	$p = .96$	$p = .96$

B Appendix B: Screenshots