Honest on Mondays: Honesty and the temporal separation between decisions and payoffs

Bradley J. Ruffle a,b, Yossef Tobol c,d

a Department of Economics, Wilfrid Laurier University, Waterloo, ON, Canada N2L 3C5
b Department of Economics, Ben-Gurion University, Beer Sheva 84105, Israel
c School of Management, Jerusalem College of Technology, Jerusalem 91160, Israel
d IZA, Bonn, Germany

Abstract

We show that temporally distancing the decision task from the payment of the reward increases honest behavior. Each of 427 Israeli soldiers fulfilling their mandatory military service rolled a six-sided die in private and reported the outcome to the unit’s cadet coordinator. For every point reported, the soldier received an additional half-hour early release from the army base on Thursday afternoon. Soldiers who participated on Sunday (the first work day of the week) are significantly more honest than those who participated later in the week. We derive practical implications for eliciting honesty.

1. Introduction

In this paper, we show that individuals behave more honestly when the decision task is temporally distanced from the reward. We report a field experiment conducted on 427 Israeli soldiers. Following Fischbacher and Föllmi-Heusi’s (2013) innovative paradigm to measure (dis)honesty, each soldier rolls a six-sided die in private and reports the outcome to his unit’s cadet coordinator. For every point reported, the soldier received an additional half-hour early release from the army base on Thursday. We conducted these experiments on different days of the week. Thus, while the time of payment is held fixed (Thursday afternoon), the time of participation in the experiment varied from Sunday (four days until early release) through Thursday (the same day).

Soldiers who participate on Sunday are honest on average. Namely, their distribution of reported die outcomes cannot be rejected as coming from a uniform distribution. Moreover, they report significantly lower outcomes on average than those who participate on later days of the week whose outcomes do not resemble a uniform distribution. This first-work-day-of-the-week effect is highly robust to the inclusion of individual characteristics and peer effects.

While the traditional economic view claims that individuals merely weigh their personal monetary benefit versus their expected cost of lying, numerous experiments offer a more expansive view of the considerations that individuals bring to bear in deciding whether to tell the truth. For example, a heightened self-awareness can increase honesty, whereas lies that benefit others (Erat and Gneezy, 2012), the perception of having been treated unfairly (Houser et al., 2012) and the ability to reframe a situation to justify one’s lie to oneself (Shalvi et al., 2011, 2012; Gino and Ariely, 2012) all induce dishonest behavior (see also...
Mazar and Ariely, 2006 for a survey). Our paper’s main contribution is to demonstrate for the first time the sensitivity of honesty to the temporal separation between the decision task and the receipt of payment from the task. We conclude the paper with some guidance designed to elicit more honest behavior from employees, clients, students and children.

Our experiment is distinct from other studies on honesty in three important respects. First and foremost, many studies on (dis)honesty vary either the material benefit from dishonesty or the cost of dishonesty (i.e., the probability of detection or the punishment from getting caught). Both of these considerations are held constant in our study. With the die rolled in private, the probability of detection is zero, regardless of the day of the week. Moreover, the potential material benefit to dishonesty is the same on all days of the week. Instead, the subtle distinction between days of the week lies in the perceived benefit of dishonesty. According to our results, soldiers value early release from their army base more when the release date is near. To the best of our knowledge, the finding that honest behavior is more likely when the decision task is removed from the payoff is new to the literature.

A second distinctive feature of our experiment is the novelty of the subject pool: soldiers completing their mandatory military service. Unlike student subject pools or even most field experiments targeted at a particular population, soldiers completing their mandatory military service constitute a representative cross-section of 19-year-olds in Israeli society as a whole. Third, subjects in our experiment cheat not an anonymous firm (e.g., Levitt, 2006; Pruckner and Sausgruber, 2013), unfamiliar wait staff at a restaurant (Azar et al., 2013), anonymous subjects (e.g., Gneezy, 2005) nor the experimenter (e.g., Fischbacher and Föllmi-Heusi, 2013). Rather, our subjects cheat first and foremost their boss (i.e., commanding officer) with whom they interact on a daily basis. To a lesser extent they also cheat their fellow soldiers: a soldier who leaves the army base early necessitates the distribution of his uncompleted duties among those who remain behind.

In the next section, we describe our experimental design, procedures and sample. In Section 3, we present the results of our experiment. Section 4 concludes with some policy implications.

2. The experiment

2.1. Experimental design and procedures

Between December 28, 2010 and June 19, 2011, 427 soldiers from 27 different permanent and provisional military bases throughout Israel and 15 distinct army units participated in our experiment. The participating soldiers knew nothing of the experiment in advance. Our arrival was unannounced. Moreover, to avoid possible subject-pool contamination, we visited each army base only once. Within each participating army company, all soldiers took part. All soldiers were in training, meaning they were serving their first of three years (first of two years for female soldiers) of required military service.

To coordinate the experiments, we contacted the commanding officer of each participating army unit and requested a block of time prior to the soldiers’ breakfast hour for the purpose of conducting the experiment. To avoid diffusing our observations across all days of the week, we requested Sunday, Wednesday or Thursday when possible.

All of the experiments were conducted just prior to the soldiers’ breakfast hour in the dining hall. The cadet coordinator (CC) of the participating army unit called each soldier by name one-at-a-time to a room or large tent with two entrances/exits located on the army base that was used for the purpose of the experiment. Each participating soldier entered through one designated entrance. The CC then read the rules of the experiment to the soldier from a script as follows. The soldier was first told that he would be asked to roll a six-sided die in private and then to report the outcome to the CC. For each point on the die, the soldier would be released on Thursday half an hour ahead of the scheduled time. To avoid any possible confusion, the exact payment in the form of hours of early release for each of the six possible outcomes was enumerated. The soldier was explained that after all soldiers in the unit had completed the experiment, the CC would submit the list of early release times to the unit commander who had previously approved the experiment and the terms of early release.

After the soldier was handed a six-sided die, he proceeded to a table at the other side of the room or tent where, out of sight of the CC, he rolled the die in private. After rolling the die, the soldier returned to the CC to report the outcome.

---

1 Goette et al. (2012) compare the in-group cooperativeness and willingness to punish of extant groups of Swiss soldiers with those of randomly formed minimal groups of soldiers. The distinction enables the researchers to demonstrate the importance of social ties for behavior, beyond the arbitrary labeling of a group upon which the minimal groups paradigm focuses. Lahav et al. (2011) distribute questionnaires on trains traveling between major Israeli cities to soldiers, teenagers and university students and show that soldiers have higher subjective discount rates than non-soldiers. Warner and Pfeeter (2001) exploit a natural experiment conducted by the U.S. Department of Defense to reduce military personnel in which mid-career personnel were offered the choice between a lump-sum separation payment and an annuity valued at considerably more in present terms. The majority’s preference for the lump-sum payment implies personal discount rates exceeding 18 percent.

2 We discuss further the representativeness of our sample in Sections 2.2 and 3.2.

3 A military unit may be spread over more than one military base. Think of a combat unit located at numerous military bases along the border(s). Nonetheless, the geographic separation between the military base and the military culture are such that bases operate independently from one another with minimal or no contact between them. Throughout the entire time period of our experiments, relative quiet prevailed in Israel. There were no wars or military confrontations with Hizbullah or Hamas, nor were any flotillas sent to the Gaza Strip from Turkey or elsewhere.

4 We had the cadet coordinator conduct the experiment to put subjects at ease because the CC is one of them. Namely, the CC is a soldier chosen (on a rotating basis) from the same military unit for which he serves. The CC’s job is to serve as a liaison between the soldiers of the unit and the commanding officer.

5 While we use the masculine pronoun throughout the text, it is meant to apply equally to women and men for the soldiers, cadet coordinators and commanding officers. Indeed, Israeli women are also required to complete mandatory military service and 42% of our sample of soldiers is female (see Table 1).
Finally, the soldier completed a brief post-experiment questionnaire (included in the Appendix), submitted it to the CC and was directed to proceed to the dining hall through the door or tent opening designated as the exit and through which he had not entered. The distinction between the two doors or tent openings as entrance and exit was maintained to prevent soldiers from having contact with those who had not yet participated in the experiment. The CC called in the next soldier who had not entered. The distinction between the two doors or tent openings as entrance and exit was maintained to prevent soldiers from having contact with those who had not yet participated in the experiment. The CC called in the next soldier according to the list and so on until all soldiers in the unit had completed the experiment.

The entire experiment including the questionnaire took about 7 min for each soldier. In view of the value soldiers attribute to an early release of half an hour (median $= 30$ NIS, mean $= 43$ NIS, see Table 1) and 3 h (median $= 100$ NIS, mean $= 194$ NIS), the experimental payment can be viewed as salient.6

Information collected on each soldier from his commanding officer, including his military entrance exam score, and from the experiment itself was directed to the CC for processing in the form of a questionnaire that the soldier completed. The entire experiment including the questionnaire took about 7 min for each soldier. In view of the value soldiers attribute to an early release of half an hour (median $= 30$ NIS, mean $= 43$ NIS, see Table 1) and 3 h (median $= 100$ NIS, mean $= 194$ NIS), the experimental payment can be viewed as salient.6

Information collected on each soldier from his commanding officer, including his military entrance exam score, and from the post-experiment survey will be used in Section 3 to explain the variance in the reported die outcomes.

### 2.2. Sample

We view our sample of soldiers as representative of the population of Israeli combat and non-combat soldiers with one exception: we purposefully recruited a disproportionate number of religious soldiers. With honesty as a central tenet in Judaism and established religions more generally, one plausible hypothesis is that religious individuals are more honest than secular ones.7 Also, in examining whether the degree of honesty varies across different day of the week, we might expect the approaching of the Sabbath (beginning on Friday at sunset) to play a different role for religious and non-religious soldiers. Thus, to obtain a sufficient number of religious soldiers to perform powerful statistical tests, we intentionally oversampled religious military companies.8

### 3. Results

#### 3.1. Overall distribution

Fig. 1 displays the distribution of reported die outcomes for our entire sample of soldiers ($N = 427$). If all soldiers reported the truth, we would expect a uniform distribution. The Pearson chi-square test rejects the hypothesis that the sample distribution is drawn from a uniform population distribution, $\chi^2(5) = 16.2, p = .001$. Soldiers clearly inflate their reported

---

**Table 1**

Descriptive statistics by day of the week the experiment was conducted.

<table>
<thead>
<tr>
<th>Variable (possible values)</th>
<th>Overall</th>
<th>Sunday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reported die outcome (1, 2, 3, 4, 5, 6)</td>
<td>3.87 (1.61)</td>
<td>3.41 (1.75)</td>
<td>3.98 (1.34)</td>
<td>4.05 (1.35)</td>
<td>4.06 (1.36)</td>
</tr>
<tr>
<td>Military test score (45, ..., 56)</td>
<td>51.1 (2.5)</td>
<td>51.7 (2.0)</td>
<td>51.9 (2.3)</td>
<td>51.2 (2.7)</td>
<td>50.3 (2.5)</td>
</tr>
<tr>
<td>City residents (0, 1)</td>
<td>.72 (.46)</td>
<td>.74 (.44)</td>
<td>.66 (.48)</td>
<td>.65 (.46)</td>
<td>.81 (.40)</td>
</tr>
<tr>
<td>Female (0, 1)</td>
<td>.42 (.50)</td>
<td>.43 (.50)</td>
<td>.38 (.49)</td>
<td>.47 (.50)</td>
<td>.39 (.49)</td>
</tr>
<tr>
<td>Religious (0, 1)</td>
<td>.34 (.47)</td>
<td>.33 (.47)</td>
<td>.21 (.41)</td>
<td>.28 (.45)</td>
<td>.46 (.50)</td>
</tr>
<tr>
<td>WTP for half hour early (in Israeli NIS)</td>
<td>42.7, 30 (67.2)</td>
<td>52.7, 30 (98.9)</td>
<td>40.2, 50</td>
<td>30.8, 20 (33.7)</td>
<td>46.6, 30 (65.4)</td>
</tr>
<tr>
<td>Self-reported honesty (1 = always tell truth … 4 = truth when convenient)</td>
<td>2.12 (.99)</td>
<td>2.30 (.92)</td>
<td>1.89 (.95)</td>
<td>2.06 (1.10)</td>
<td>2.09 (.90)</td>
</tr>
</tbody>
</table>

**Notes:** For each day of the week the experiment was conducted, the mean (s.d.) soldiers’ reported die outcome; mean military entrance test score; fraction of city residents, females and religious soldiers; mean willingness to pay (WTP) for half-hour early release from the army base, mean self-reported response to “how often do you tell the truth?”; the WTP cells report the mean value followed by the sample median.

For the subset of soldiers asked whether they knew their military entrance test score ($n = 217$), the last three rows report the fraction that claimed to know it, the fraction that actually does know it (conditional on claiming to know) and the mean difference between the actual and reported score.

---

6 To appreciate the size of the stakes, consider the following back-of-the-envelope calculation. The average subject reported a die outcome of 3.87 (see row 1 of Table 1), equivalent to 1.97 h early release. Assume, for simplicity, that a soldier’s willingness to pay increases linearly with each additional half hour of early release. It follows that the average soldier received payment worth 130 NIS for 7 min of work. Contrast this with a combat soldier’s monthly wage of 700 NIS a month and non-combat soldier’s monthly salary of between 300 and 500 NIS, depending on their job.

7 Relatedly, Mazur et al. (2008) find that a religious prime (namely, writing down the Ten Commandments from memory) prior to performing a mental task and self-reporting one’s success reduces cheating compared to a neutral prime that asks subjects to list 10 books they read in high school. In an Israeli context, Shalvi and Leiser (2013) show that religious female students hold others’ behavior to a stricter standard about what constitutes a lie than do secular female students; however, religious females were not significantly more honest in their behavior than secular females.

8 In Israel, religious soldiers form their own military companies to accommodate their religious practices, including prayer and kosher food.
outcomes, but do not profit maximize. Up to and including the outcome of 5, the reported frequency increases monotonically. The frequencies of 11.5% and 11.9% associated with 1 and 2, respectively, are each significantly less than the true percentage of 16.67% of a uniform distribution (\( p < .001 \) from one-sided Binomial tests in both cases); whereas, the frequencies of 19.9% and 24.8% associated with 4 and 5, respectively, are significantly greater than 16.67% (\( p = .04 \) and \( p < .001 \), respectively). We cannot quite reject that the frequency of 14.5% associated with 3 is significantly less than 16.67% (\( p = .13 \)).

The most surprising and unanticipated aspect of the outcome distribution is the decline in frequency from 5 to 6. In fact, the reported frequency of 6 of 17.3% is not significantly more than 16.67% (\( p = .38 \)). Fischbacher and Föllmi-Heusi (2013) also observe incomplete cheating in that significantly more than 1/6 of their subjects report one number below the income-maximizing outcome. Nonetheless, they find a still higher percentage of subjects who report the highest outcome, while we witness a sharp decline.9 One explanation for our observed decline in reported 6 s is that payments are publicly observable. A soldier seen leaving the base 3 h early on Thursday may be concerned that his peers or his commanding officer will view him as dishonest.10

The distributions of reported die outcomes for combat soldiers (mean = 3.83, s.d. = 1.74, \( n = 129 \)) and non-combat soldiers (mean = 3.88, s.d. = 1.55, \( n = 298 \)) are nearly identical (Wilcoxon-Mann-Whitney \( z = -.013, p = .99 \)). Thus, we combine these two groups of soldiers for the remaining analyses.

3.2. Honesty and the day of the week

Although all soldiers received their “payment” on the same day (namely, Thursday afternoon of the week during which the experiment was conducted), they participated in the experiment on different days.11 Thus, the distance in time between the experiment and the date of payment varies across participants. In more usual time-preference experiments (see Frederick et al., 2002 for a review), the source of variation between the point in time at which the decision is taken and the payment received derives from holding the former constant while varying future payment dates.12 The time to future payment matters because of foregone opportunities that a delayed payment represents. In our setup, with no opportunity to receive advance payment or collect interest on delayed payments, these standard economic explanations for discounting the future are not applicable. And yet soldiers’ choices are consistent with temporal discounting.13

The first row of Table 1 reveals that the earlier in the week the experiment was conducted, the lower is the average reported die outcome. On Sunday (the first day of the week in Israel and equivalent to Monday in Western countries), the mean die outcome is 3.41, over a half-point less than the mean outcomes of 3.98 on Tuesday, 4.05 on Wednesday and 4.06.
on Thursday. Fig. 2 presents a histogram of the reported die outcomes for each of the four days of the week on which the experiment was run. The distribution for Sundays strikingly resembles a uniform distribution. In fact, for 4/6 outcomes the reported frequencies are not significantly different from 16.67%. Only the frequency of 4 is significantly below 16.67%, while the frequency of 5 is significantly above this percentage. Overall, we cannot reject the hypothesis that this distribution of outcomes was drawn from a uniform distribution on 1–6 \( (\chi^2(5)=4.2, p=.52) \). In other words, subjects on average are honest on Sundays. For all other days of the week, the sample distribution of outcomes differs significantly from the uniform distribution \( (p\text{-values range from .02 to .10}) \).

Regression (1) in Table 2 confirms that soldiers on Sunday are significantly more honest than soldiers on each of the remaining days of the week on which the experiment was conducted. There are no significant differences between Tuesday, Wednesday and Thursday \( (p\text{-values from t-tests of coefficients range from .72 to .93}) \). The significance levels of the day-of-the-week indicators and lack of significant differences between them are all preserved with the inclusion of a time trend variable, \( p \) which itself does not differ significantly from zero \( (p=.45) \).

One possible explanation for our “honest-on-Mondays” result is that the composition of soldiers differs across days of the week. That is, perhaps there are individual characteristics that are correlated with honesty and these characteristics are unevenly distributed between different days of the week. To examine this possibility, we collected socio-demographic data on each soldier through a post-experiment questionnaire. Furthermore, the commanding officer of each unit provided us with each soldier’s military entrance exam score. Table 1 summarizes this data for each day of the week as well as for the overall sample. The first striking observation about this table is that the mean value for each variable for Sunday is remarkably similar to the overall sample mean. While some of the variables do vary in their mean values across days of the week, no discernible pattern or distinguishing feature about Sunday is evident. More to the point, we include each of these socio-demographic variables as controls in regression (2) of Table 2. The regression shows that soldiers from cities report .44 points higher than soldiers from rural areas \( (p=.01) \). In addition, secular soldiers on the whole are not significantly less (and not more) honest than their religious counterparts. In fact, among the gender-religiosity subgroups, the only even weakly significant difference in die outcomes is secular females who report on average .34 points more \( (i.e., \text{are less honest}) \) than secular males \( (p=.06) \). Most importantly, the magnitudes of the day-of-the-week indicators and their significance levels remain unchanged with the inclusion of these socio-demographic controls.

One might hypothesize that the honest on Mondays effect will not hold for religious soldiers. As the week advances, religious soldiers become more conscious of the approaching Sabbath and reminded of their moral duty to be honest in their dealings. To test for differences in the first-work-day-of-the-week effect between secular and religious soldiers, we interact each day of the week with an indicator for whether the soldier is religious. Regression (3) reveals that all three \( (\text{non-interacted}) \) day-of-the-week indicators \( \text{now interpreted as the mean difference in reported die outcomes on the indicated day compared to Sunday for secular soldiers only}) \) remain unchanged and highly significant. Furthermore, none of the four day-of-the-week indicators interacted with religious \( \text{differs from zero at conventional levels of significance, implying that secular and religious soldiers behave similarly on average for each day of the week. The linear combinations of estimates below the constant in (3) test whether religious soldiers also behave significantly more honestly on Sunday than on other days of the week. As evidence that they do, all three computed estimates are negative. Yet with only 11 religious soldiers who participated on a Tuesday (see also Table 1), the difference between Sundays and Tuesdays is not significant, while differences between Sundays and Wednesdays or Thursdays are highly significant. In summary, the honest on Mondays finding holds for secular and religious soldiers alike.

Next we explore a possible relationship between soldiers’ honesty and their military entrance exam scores. Months prior to being drafted, every candidate soldier is evaluated on the basis of his educational background and a series of computerized psycho-technical exams. Numerous questions on these exams are designed to evaluate the soldier’s honesty through, for instance, framing the same question in different ways to test for consistent responses. In addition, males undergo a lengthy personal interview in which female interviewers aim to “assess body language, to identify lies and individuals who are unreliable” (Hebrew Wikipedia under the title, “recruitment to the Israeli military”). For male soldiers, the final entrance test score \( \text{(known as kaba in Hebrew and to be subsequently referred to as such for brevity) is made up of the interview (33%), the psycho-technical exams (50%) and the candidate’s educational background (17%)}. \text{Women are not subjected to the interview. Instead, their kaba is based on the psycho-technical exams (60%) and their educational background (40%). An individual’s kaba determines the unit and job to which he is assigned for his military service.}

\[ \text{We define the time trend variable as the number of days elapsed between the date the experiment was conducted (i.e., the date the base was visited) and the date of the first experiment (i.e., December 28, 2010). The results are identical if instead we assign a value of 1–27 corresponding to the order in which the bases were visited.} \]

\[ \text{We classified each soldier according to whether he lives in a city (more than 20,000 residents according to the Israeli Central Bureau of Statistics), Non-cities consist of moshavim and kibbutzim (Israeli cooperative settlements), towns and villages. Consistent with our finding, Bowles and Gintis (2002) discuss the ability of smaller, tight-knit communities to uphold social norms. Gächter and Herrmann (2011) observe higher levels of cooperation among rural than urban Russians.} \]

\[ \text{Among the many moral imperatives not to lie found in the Five Books of Moses, the best known are the commandment to “not bear false witness against your neighbor” (Exodus, 20:16 and Deuteronomy 5:20) and “thou shalt not steal, neither deal falsely, neither lie one to another” (Leviticus 19:11).} \]

\[ \text{These identical findings continue to hold in all subsequent regressions. For ease of reporting, we omit these day-of-the-week-religious interaction indicators and return to reporting the sex and religiosity indicators and interaction between them as in (2).} \]
All of the component tests and interviews are taken once and, unlike many other outcomes in Israel, the results are not subject to appeal. The kaba scores range from 41 to 56. Scores below 44 exempt one from regular military service, while scores of 52 or more qualify one for privileged jobs and an officer’s course. In our sample, 48.2% of soldiers qualify to be officers (i.e., a kaba of 52 or higher). As evidence of the representativeness of our sample, Lerer (2009) reports an identical figure (48%) for the fraction of soldiers with a kaba of 52 or higher in 1995 (the most recent year for which he obtained data).19

Through each unit’s commanding officer and unbeknown to the soldier, we obtained every participating soldier’s kaba. Regression (4) includes each soldier’s kaba (expressed as the difference between the kaba and 45 (the lowest entrance score in our sample)). The highly significant coefficient of \(\beta_0\) on kaba suggests that the entrance score does indeed reflect the soldier’s degree of honesty as intended by the test designers and interviewers. Curiously, this highly significant relationship contrasts with the prevalent view in the personnel psychology literature that personality tests aimed at evaluating personality traits, including honesty, consisting primarily of self-reported measures and used in hiring decisions have low predictability of job performance (see, e.g., Morgeson et al., 2007 for an exhaustive survey). Again, the first-work-day-of-the-week effect remains robust to the inclusion of the soldier’s test score.

Next, we check whether soldiers’ more modest claims on Sunday show up in the form of a lower willingness to pay (WTP) for early release. The mean and median amounts that soldiers are willing to pay for a half-hour early release reveal no consistent pattern across days of the week (middle of Table 1).20 Regression (5) includes soldier i’s WTP for a half-hour early release. The differences between mean reported die outcomes on Sundays and Tuesdays increase slightly, while the estimates on the other days of the week remain within the previously observed range and are highly significant. The coefficient of .001 on the WTP is neither statistically significant nor economically meaningful: for each additional 66 NIS a soldier values a half-hour early release (two to three times the mean valuation), he inflates his reported outcome by a mere .1 points.

The standard deviation of 67.2 exceeds the mean WTP by more than 50%, thus attesting to the presence of outliers. If we exclude observations that deviate from the mean by more than two standard deviations, the median remains unchanged, while the mean and standard deviation drop to 33.9 NIS and 28.1 NIS respectively (N=412). Regression (6) excludes these 15 outlying observations, but is otherwise identical to (5). The coefficient on the WTP variable is now highly significant and five

19 The Israel Defense Forces do not make publicly available the distribution of military scores.
20 The relatively high mean WTP of 52.7 on Sunday is driven by an outlier (a valuation of 1000 NIS). Thus, we also include the median for comparison. If we exclude outliers (defined as two or more standard deviations from the mean), the absence of a relationship between WTP and the day of the week persists. Moreover, midway through the experiments (with 217 participants remaining), we introduced an additional survey question eliciting the soldier’s WTP for three hours early release from the base. Again, no discernible relationship – with or without outliers – exists between this measure and the day of the week.
times the magnitude of that in regression (5).21 It now takes only a 13 NIS increase in a soldier’s WTP on average to inflate his reported die outcome by .1 units. Once more, the highly significant honest on Mondays effect persists.

To the extent that military units or bases differ in their ethos or morale, we would anticipate soldiers’ reported die outcomes to be correlated with those of fellow members of their unit or base. To test for military unit peer effects, we include in regression (7) an explanatory variable for soldier i in unit j the unit’s mean reported die outcome based on every other soldier’s (i.e., not including soldier i) reported die outcome in the same unit (unit peer effects). Explicitly, the unit peer-effect measure is \( \sum_{k \neq i} \text{outcome}_{kj} / n - 1 \). Similarly, we define an analogous measure for soldier i in base q (base peer effects). The lack of significance of the unit and base peer-effects measures (\( p = .24 \) and .88, respectively) in (7) as well as in all of the other regressions in Table 2 (not reported) when these measures are added attests to the independence of observations within a military unit and base.

### 3.3. Discounting or unobservables? A second test of honesty

To test whether our observed honest on Mondays effect could be due to other unobserved differences between the days of the week, we devised an additional test of soldiers’ honesty for which the incentive to lie (or tell the truth) is identical on all days of the week. Consequently, if soldiers are again found to be more honest on Sundays, then unobservable variables would seem a highly plausible explanation for the increased honesty observed on Sundays in our die-rolling experiment. If, however, no differences in honesty between days of the week are observed in this second test, then the unobservables explanation seems unlikely.

---

21 A squared term for the WTP measure is not significant and its inclusion does not affect the significance of any of the days-of-the-week variables.
Midway through the implementation of the experiments, we introduced two questions to the post-experiment survey. The questions ask the soldier whether he knows his army test score and, if so, to write it down. These questions enable us to determine whether soldiers truthfully report their test score and, if not, by how much their reported score deviates from their true score.

The last three rows of Table 1 report the summary statistics for each day of the week. For a soldier who claims not to know his score (14% of our sample with no significant difference in the percentage reported by day of the week ($\chi^2(2)=.98$, $p=.61$)), we cannot know whether the soldier truly does not know his score or simply chooses not to reveal it to us. For this reason, our analysis focuses on those who claim to know their score. Among the 86.2% of soldiers in our sample who claim to know their true test score, only about half (47.6%) accurately report it with no apparent differences across days of the week according to the second-to-last row of Table 1. Regressions (8) and (9) in Table 3 report the estimates from a linear probability model on whether the soldier truthfully reports his score as a function of the day-of-the-week indicators only in (8) and along with the same set of socio-demographic controls included in regression (2) from Table 2 in (9). In both regressions the estimates on “Wednesday” and “Thursday” are close to and not significantly different from 0, thereby confirming that the tendency to report honestly one’s score does not differ significantly across days of the week.

The last row of Table 1 reports the mean difference between soldiers’ true test scores and their reported test scores. Soldiers inflate their scores by about one point on average, both overall and on different days of the week (two points if we exclude those who accurately reported their scores). In fact, of the 97 soldiers who falsely report their scores, 92 inflate them, adding between one and five points. Regressions (10) and (11) in Table 3 attest to the lack of any significant differences across days of the week in the magnitude by which soldiers misreport their entrance scores.

Critically, if we restrict our investigation of the honest on Mondays effect to this same subset of 185 soldiers who claim to know their entrance score, the effect remains largely intact. Regressions (12) and (13) show that soldiers report .84 and .94 points higher on Thursdays than on Sundays, respectively, and .47 and .67 points higher on Wednesdays than on Sundays. The estimate of .47 is not quite marginally significant ($p=.11$), while the other three differences are all highly significant ($p < .01$).

In summary, the finding that soldiers’ reports of their entrance scores are no more (and no less) honest on Sundays than on other days of the week contradicts the unobserved variables story, thereby lending credence to the interpretation that soldiers are honest on Sundays in reporting their die outcomes because of the lower perceived benefit from early release when the weekend is distant and thus the temptation to cheat reduced.

One curiosity remains in our results: why are soldiers not significantly more honest on Tuesdays than on Wednesdays and on Wednesdays than on Thursdays? In other words, why is the effect of temporally distancing decision from outcome mostly limited to Sundays? We can only speculate about the explanation. The inflated reported die outcomes on Tuesdays, Wednesdays and Thursdays attest to the temptation to cheat in contrast to Sundays on which honest reporting occurs. Plausibly, by Tuesday or Wednesday (i.e., a day or two before the weekend in Israel), soldiers are exhausted from the week’s grueling military routine and may already have plans for the weekend whereas on Sunday they are fresh from the weekend and have not given the next weekend much thought. Thus, in the compressed timespan covered by our experiment (i.e., days within the same week), little distinguishes between the later days of the week and, as a result, significant differences between them in observed honesty do not arise.

4. Conclusions

In a field experiment in which a soldier’s self-reported outcome of a die roll determines how early he is released from the army base for the weekend on Thursday afternoon, soldiers participating on Sundays are honest on average and significantly more honest than soldiers on later days of the week. The simplest, most plausible explanation for honesty on Sundays is that soldiers still feel refreshed after a weekend off and have scarcely thought of the seemingly faraway weekend ahead. Thus, the perceived value of early release is discounted. Becker and Mulligan (1997) note that “according to Böhm-Bawerk and most others who have written on this subject [of discounted future consumption] ... discount factors are less than unity because of an imperfect ability to imagine the future” (p. 734). On Sunday, the inability to visualize the upcoming weekend constrains soldiers’ behavior, whereas on later weekdays the weekend’s temporal proximity enhances its salience and the temptation to cheat.

This explanation begs the question. Why are soldiers not significantly more honest on Tuesdays than on Wednesdays and on Wednesdays than on Thursdays? That is, why is the effect of temporally distancing decision from outcome mostly limited to Sundays? We can only speculate about the answer. By Tuesday or Wednesday (i.e., a day or two before the weekend in Israel), soldiers likely have plans for the weekend. With the “taste” of the weekend and its associated freedom so palpable, the temptation to inflate one’s report is irresistible. Thus, in the compressed timespan covered by our experiment (i.e., days within the same week), little distinguishes between the later days of the week and, as a result, significant differences between them in observed honesty do not arise.

22 There are no observations on Tuesdays because, midway through the design, we elected to focus our data collection on the extreme days of the week.
23 In other contexts in which the temporal distance between action and reward is measured in weeks or months, we might expect strictly increasing honesty as the temporal distance between decision and reward increases.
24 In other contexts in which the temporal distance between action and reward is measured in weeks or months, we might expect strictly increasing honesty as the temporal separation between decision and reward increases.
Whatever the source, our finding that soldiers are more honest on Sundays suggests the desirability of temporally distancing decisions from outcomes to obtain honest behavior. In principal-agent problems in which the agent’s effort as well as his output are observable only to the agent himself, temporal separation may effectively reduce misreporting. For example, insurance fraud whereby the customer overstates the value of claims or falsely reports missing or damaged items might be diminished by delaying reimbursement. Also, instead of immediately compensating company managers and employees based on their self-reported tasks, remuneration should be delayed to some (possibly unannounced) future date to promote honesty. To reduce cheating on tests and copying on assignments, do not publicize the student’s grade immediately afterwards. If feedback to the student on his performance is beneficial, then provide only the solution to the test or assignment. When posting the grade soon after the test is unavoidable (e.g., final exams), our findings advise increased monitoring to deter cheating. On a different level, parents often condition rewards to their children on good behavior or the completion of their chores or homework. The optimal time to ask your eight-year-old son whether he behaved well at school is not as you tear off the wrapper from his promised candy, but well beforehand.

The flipside of this argument is that to elicit reliable, accurate intentions regarding a costly outcome, the question should be posed as close as possible to the outcome. Ask a person about his intention to begin exercising, dieting or saving not weeks ahead, but rather the day before the intended start date.

Acknowledgments

We thank Yuval Arbel, James Atsu Amegashie, Ofer Azar, Pat Barclay, Ronen Bar-El, Yoella Bereby-Meyer, Bram Cadsby, Naomi Feldman, Uri Gneezy, Gilles Grolleau, Todd Kaplan, Shachar Kariv, James Konow, Vova Lanin, Alex Maynard, Logan McLeod, Mattia Pavoni, Tata Pyatigorsky-Ruffle, Rupert Sausgruber, Jonathan Schulz, Shaul Shalvi, Ze’ev Shtudiner, Justin Smith, Fei Song, Michal Kolodner-Tobol, Ro’i Zultan, the editor, Jörg Oechssler, and two anonymous referees for comments. We also are grateful to Capt. Sivan Levi for research assistance, Capt. Itamar Cohen for facilitating the experiments and the Givati Infantry Brigade for their cooperation.

Appendix. Questionnaire

Note: § refers to questions introduced after 210 soldiers had already participated.

1. What is the maximum amount of money you would be willing to pay to be released to go home half an hour earlier than the scheduled time on your day of release? ___ (amount in NIS).

2. What is the maximum amount of money you would be willing to pay to be released to go home three hours earlier than the scheduled time on your day of release? ___ (amount in NIS).

3. How would you define yourself?
   1. Secular
   2. Traditional
   3. Religious
   4. Ultra-orthodox

Table 3
Regression analysis related to second test of honesty.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Soldier i truthfully reports score</th>
<th>Soldier i’s actual score – reported score</th>
<th>Soldier i’s reported die outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regressor</td>
<td>(8)</td>
<td>(9)</td>
<td>(10) (11) (12) (13)</td>
</tr>
<tr>
<td>Wednesday</td>
<td>–.12 (.08)</td>
<td>–.12 (.08)</td>
<td>–.31 (.25) –.39 (.25) .47 (.30) .67*** (.30)</td>
</tr>
<tr>
<td>Thursday</td>
<td>.04 (.10)</td>
<td>.01 (.10)</td>
<td>.08 (.25) .02 (.26) .84*** (.33) .94*** (.34)</td>
</tr>
<tr>
<td>City resident</td>
<td>–</td>
<td>–.11 (.08)</td>
<td>– .38 (.23) – .78*** (.27)</td>
</tr>
<tr>
<td>Female</td>
<td>–</td>
<td>–.03 (.09)</td>
<td>– .13 (.18) – .59* (.31)</td>
</tr>
<tr>
<td>Religious</td>
<td>–</td>
<td>–.13 (.17)</td>
<td>– .42 (.50) – .20 (.63)</td>
</tr>
<tr>
<td>Religious female</td>
<td>–</td>
<td>–.09 (.19)</td>
<td>– .21 (.55) – .45 (.70)</td>
</tr>
<tr>
<td>Constant</td>
<td>.51*** (.50)</td>
<td>.49*** (.09)</td>
<td>–.96*** (.14) –.74*** (.23) 3.40*** (.18) 2.67*** (.28)</td>
</tr>
<tr>
<td>Adj. R²</td>
<td>.00</td>
<td>.00</td>
<td>.00 .02 .09</td>
</tr>
<tr>
<td>N</td>
<td>185</td>
<td>185</td>
<td>185 185 185 185</td>
</tr>
</tbody>
</table>

Notes: Dependent variable: (8)–(9): indicator variable whether soldier i truthfully report his military entrance score; (10)–(11): the difference between soldier i’s actual score and his reported score; (12)–(13): soldier i’s reported die outcome. Regressors are a subset of those in Table 2. Heteroskedasticity-robust standard errors in parentheses.

*** Coefficient significantly different from 0 at the 1% level.

** Coefficient significantly different from 0 at the 5% level.

* Coefficient significantly different from 0 at the 10% level.
4. Which of the following sentences best describes you?
   1. I always tell the truth.
   2. I almost always tell the truth.
   3. I usually tell the truth.
   4. I tell the truth when it is convenient for me.

5. How important is it to you what others think of you?
   1 = very important, 7 = not important at all
   1 2 3 4 5 6 7

6. Where do you live (indicate the name of the town or city)? ____________________

7a. Do you know your army test score?
   0. No
   1. Yes

7b. If yes, please write it in the space provided. ____

Thank you for your participation.

References