Impatience and grades: Delay-discount rates correlate negatively with college GPA

Kris N. Kirby\textsuperscript{a,}\textdagger, Gordon C. Winston\textsuperscript{b}, Mariana Santiesteban\textsuperscript{a}

\textsuperscript{a}Department of Psychology, United States
\textsuperscript{b}Department of Economics, United States

Received 15 July 2004; received in revised form 31 January 2005; accepted 31 January 2005

Abstract

Because the rewards of academic performance in college are often delayed, the delay-discounting model of impulsiveness (Ainslie (1975) \textit{Ainslie, G. Specious reward: A behavioral theory of impulsiveness and impulse control. Psychological Bulletin, 82 (4), 463–496}) predicts that academic performance should tend to decrease as people place less weight on future outcomes. To test this hypothesis, we estimated (hyperbolic) discount rates for real delayed monetary rewards (US$10 to US$20) using second-price auction procedures with 247 undergraduates at two liberal arts colleges. College GPA was reliably correlated with discount rates, $r=-.19$ ($p=.003$), and remained reliable after partialling out SAT scores. The results add to the external validity of the discounting model of impulsiveness, and point to a possible contributor to academic performance of interest in the study of higher education.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Academic achievement; Reinforcement delay; Delay-discounting; Impulsiveness

Although we teachers may not care to admit it, for many college students most of the rewards of academic success are rather remote in time. The rewards of studying for an exam may be days or weeks away. The career-related rewards of earning a good grade point average (GPA) may be years away. It is perhaps common sense that students who are more future-oriented should perform better in college than those who are less so. In this paper we test a prediction of this common sense notion by examining the
relationship between the rate at which a person discounts future monetary rewards and college academic performance as measured by GPA.

Delay-discounting refers to the reduction in the present value of a future reward as the delay to that reward is increased. The discount rate determines the steepness of this reduction in value with delay, such that the larger the rate, the more the person discounts the future, and the less value future rewards (or costs) should have in current decisions. A person with a discount rate of zero would not be sensitive to delay, and would value future and immediate rewards equally. A person with a high discount rate would place little value on future rewards. According to the delay-discounting model of impulsiveness (Ainslie, 1975, 1992; Rachlin & Green, 1972), as the discount rate increases, people should also become more susceptible to making impulsive decisions (cf., Strotz, 1956; Winston & Woodbury, 1991). Thus, this model predicts that college students with higher discount rates should tend to place less weight on the future in current decisions, and should be more impulsive than those with lower rates. Consequently, we predict that discount rates and college GPA should be negatively correlated.

To examine the relationship between GPA and discount rates, we obtained college GPA and SAT scores for students at two liberal arts colleges for whom discount rates had been estimated in previous studies (Kirby, 1997, 1998; Kirby & Santiesteban, 2003).

1. Method

1.1. Participants

The participants were 154 Williams College undergraduates (73 men and 81 women) and 93 Massachusetts College of Liberal Arts (MCLA) undergraduates (28 men and 65 women). In 2002 MCLA had a mean total SAT (verbal plus math) of 1070, with an interquartile range of 933 to 1160. Williams College had a mean total SAT of 1400, with an interquartile range of 1310 to 1510. Participants had received either monetary payment, course credit in introductory psychology, or both, as compensation for participation.

1.2. Procedure

1.2.1. Estimating discount rates

Previous research has demonstrated that the discount rate itself may be a decreasing function of the delay to the future reward (see Kirby, 1997, for a review). This phenomenon can be modeled using hyperbolic discount functions, which typically provide a better fit to empirical data than do exponential functions in which the discount rate itself is insensitive to delay. One hyperbolic function, which has been shown to fit animal and human discounting data quite well, is the following (Mazur, 1987):

\[ V = \frac{A}{1 + kD}, \]

where \( V \) is the present value of a reward of amount \( A \) that is available at delay \( D \), and \( k \) is the discount rate parameter. In this paper all delays are measured in days, and \( k \) is scaled accordingly. Estimates of the discount rate parameter \( k \) were obtained for each participant based on their valuations of delayed rewards presented at a series of delays. The discount rate data was previously reported in Kirby (1997, 1998), and
Kirby and Santiesteban (2003). Details of the valuation procedure and the differences between the procedures that were used in the various experiments are provided in those papers. Here we give a brief overview.

In all cases, participants were asked to bid real money in auctions for real delayed rewards. There were two delayed reward amounts, large (US$20 or US$19.90) and small (US$10 or US$9.90). On each trial, one of the two rewards was presented at some delay ranging from 1 to 43 days. For example, the participant might be presented a trial in which a reward of “US$20 in 29 days” is offered. Participants were asked to specify their present values of those rewards. In some experiments they reported an amount that they would be willing to pay at the time of the session in order to receive the delayed reward. In other experiments they reported an amount that they would be willing to receive at the time of the session in lieu of the delayed reward. Although in theory these procedural variations could lead to different responses, Kirby and Santiesteban (2003) found no differences in the discount rates that they generated.

On each trial, the participant’s reported present value of the delayed reward was entered into a sealed-bid, second-price auction. In some experiments participants bid against other participants (Kirby, 1997, 1998), and in other experiments they bid against the computer (Kirby, 1998; Kirby & Santiesteban, 2003). Kirby (1998) found no difference in discount rates between these two conditions.

In the willing-to-pay auctions, if a participant had the highest bid (either compared with other participants or with a bid generated randomly by the computer), then the participant would win the auction. He or she would then pay the smaller amount given by the second-highest bid, and receive the delayed reward in return in the number of days specified. If the participant did not win the bid then no money changed hands. In the willing-to-accept auctions, if a participant had the lowest bid, then the participant would win the auction. He or she would then receive immediately the larger amount given by the second-lowest bid, and would not receive the delayed reward. If the participant did not win the bid then he or she would simply receive the delayed reward in the number of days specified. In both of these auction procedures, the second-price structure makes bidding one’s true present value the response with the highest expected value.

There were 15 trials for each of the two reward sizes. One trial was chosen at random at the end of each session and participants received rewards or not on the basis of the bidding on that particular trial. Thus, on every trial real money was potentially at stake, and participants were encouraged to treat every trial as though it might be the randomly chosen one.

Thus, for each reward amount for each participant we obtained present value estimates at 15 different delays. We fit Eq. (1) to each participant’s present value estimates for each reward magnitude using iterative nonlinear regression. This yielded an estimate of $k$ and a goodness of fit measure for each reward size for each participant.

1.2.2. GPA and SAT data

GPA and Scholastic Aptitude Test verbal (SAT-V) and math (SAT-M) scores were obtained from the registrars at the two colleges. Registrars were provided a list of student names paired with an arbitrary identification number. To preserve anonymity, after the GPA and SAT data were obtained, the students’ names were stripped from the file and the GPA and SAT data were matched with the discount rate data using the identification numbers. At no point did the experimenters have access to the identities corresponding to the GPA and SAT data.
2. Results

2.1. Summary statistics

Means and standard deviations for the GPAs and SAT scores are shown in Table 1. Williams was significantly higher than MCLA on all measures, all ps<.0001. For both schools, GPAs were on the usual 0 to 4.33 scale (where 4.33 is the numerical equivalent of a perfect A+), but the maximum observed GPA was 4.03 and the minimum was 1.40. To improve comparability of the GPA values across schools we standardized GPAs within schools. These values are denoted below as sGPA.

Plots of the raw data and normal probability plots revealed that the sGPA distributions were negatively skewed, but not markedly so. The SAT scores were also negatively skewed, especially in the Williams sample. Transformations to reduce skewness were only partly successful, and in some cases a reduction in skewness in one school caused an increase in skewness for the other. Because the transformations of sGPA and SAT scores made little qualitative difference in the results, for ease of interpretation we report results using the untransformed sGPA and SAT scores in the analyses below.

The median values of \( k \) for each school are shown in Table 2. The distribution of \( k \) s was approximately normalized using the natural log transformation. As is usually found, the mean value of \( \ln(k) \) was smaller for the US$20 than for the US$10 reward, \( t(256)=12.9, p<.0001 \). The fits of Eq. (1) were generally very good, yielding an average root mean squared error of 12% for the US$10 reward and 10% for the US$20 reward. As is typically found, the correlation between the \( \ln(k) \) values for the US$10 and US$20 rewards was very high, \( r=.93 \). Thus, we combined the two rate estimates to obtain a single discount rate estimate per participant by averaging over the \( \ln(k) \) values for the two reward magnitudes. Below we denote this mean \( \ln(k) \) value simply as \( \bar{k} \).

As described in previous papers (Kirby, 1997, 1998; Kirby & Santiesteban, 2003) the hyperbolic discount function in Eq. (1) fit better than did an exponential function for the large majority of participants. Note, however, that the fitted hyperbolic and exponential discount rates are very highly correlated, \( r=.99 \). Thus, the regression results reported below were nearly identical whether hyperbolic or exponential discount rates were used.

2.2. Simple correlation between \( \bar{k} \) and sGPA

A scatterplot of sGPA versus \( \bar{k} \) is shown in Fig. 1. The flatter of the two lines shows the best fitting linear relationship when all of the data is included. This linear relationship was reliable, \( r=-.19, t(247)=-3.03, p=.003 \). The assumptions underlying the regression appeared to be approximately met. There was no apparent departure from normality in a normal probability plot of the residuals, the

<table>
<thead>
<tr>
<th>School</th>
<th>n</th>
<th>GPA</th>
<th>SAT scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Verbal</td>
</tr>
<tr>
<td>Williams</td>
<td>154</td>
<td>3.30 (0.41)</td>
<td>649 (78.4)</td>
</tr>
<tr>
<td>MCLA</td>
<td>93</td>
<td>2.87 (0.59)</td>
<td>484 (81.9)</td>
</tr>
</tbody>
</table>

Standard deviations are shown in parentheses after each mean.
score test (Cook & Weisberg, 1982) for non-constant variance was not significant, $\chi^2(1)=0.11$, $N=247$, $p=0.73$, and there were no visible departures from linearity in the plots of the residuals versus estimated values of $\hat{k}$.

The correlation was the same to two decimal places when the case with by far the largest leverage was excluded. As another test of the sensitivity of this correlation to cases with high leverage, we reanalyzed the data using only the cases that fell within the predicted 50% bivariate ellipse, shown in Fig. 1. (In our data, 54% of the cases fell within this ellipse, which suggests that the bivariate normal assumption was roughly met.) The best-fitting line for these cases is shown by the steeper of the two lines in Fig. 1. For

![Fig. 1. Scatterplot of GPA standardized within school (sGPA) versus mean of the log-transformed discount rates across the two reward magnitudes. The flatter line shows the linear fit to all of the data. The steeper line shows the linear fit to the cases falling within the predicted 50% bivariate ellipse, which is indicated by the oval.](image-url)

Table 2
First quartile, median, and third quartile parameter values ($\hat{k}$) from the fits of Eq. (1)

<table>
<thead>
<tr>
<th>Reward size</th>
<th>US$10</th>
<th>US$20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Williams</td>
<td>Quartile 1</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Quartile 3</td>
<td>0.08</td>
</tr>
<tr>
<td>MCLA</td>
<td>Quartile 1</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Quartile 3</td>
<td>0.10</td>
</tr>
</tbody>
</table>

$k$ is the fitted parameter from Eq. (1).
these cases the correlation was slightly larger than for all cases, \( r = -0.24 \), and was reliably greater than zero, \( t(134) = -2.88, p = 0.005 \). Finally, we reanalyzed the data after excluding the 15 participants who never bid below the full amount of the delayed reward for one or both of the two reward magnitudes. By restricting the range in this way one would expect the correlation to decrease, and it did decrease but remained reliable, \( r = -0.15, t(232) = -2.27, p = 0.02 \).

### 2.3. SAT-adjusted correlations

GPA was reliably correlated with SAT-V, SAT-M, and total SAT scores within schools, all \( ps \leq 0.006 \). Across schools, sGPA also was reliably correlated with all three SAT scores. These correlations are shown in Table 3. However, SAT scores were not reliably correlated with \( \bar{k} \): for SAT-V scores \( r = -0.07, t(232) = -1.11, p = 0.27 \); for SAT-M scores, \( r = -0.08, t(232) = -1.18, p = 0.24 \); and for total scores, \( r = -0.08, t(232) = -1.21, p = 0.23 \). Although unreliable, all correlations within and across schools were negative. Thus, we wished to determine whether the correlation between \( \bar{k} \) and sGPA would remain after controlling for SAT. Controlling for total SAT, the correlation between \( \bar{k} \) and sGPA remained reliable, \( r = -0.15, t(231) = -2.40, p = 0.02 \). Including both SAT-V and SAT-M scores simultaneously resulted in substantial collinearity, but the correlation between \( \bar{k} \) and sGPA remained virtually unchanged, \( r = -0.15, t(230) = -2.39, p = 0.02 \). Thus, it does not appear that the relationship between sGPA and \( \bar{k} \) can be attributed to a relationship between \( \bar{k} \) and SAT scores.

### 2.4. Within schools

For the Williams College group the correlation between GPA and \( \bar{k} \) was reliable, \( r = -0.24, t(152) = -3.05, p = 0.003 \). For MCLA the correlation between GPA and discount rate was negative, but not reliable, \( r = -0.13, t(91) = -1.27, p = 0.21 \). The two correlations were not reliably different, \( Z = 0.86, N = 247, p = 0.39 \). Partialing-out the contribution of SAT-V and SAT-M scores reduced the correlation between sGPA and \( \bar{k} \) only slightly for the Williams group, \( r = -0.22, t(151) = -2.83, p = 0.005 \). However, the

### Table 3

Correlations among all variables across schools (upper panel) and within schools (lower panel)

<table>
<thead>
<tr>
<th>Variable</th>
<th>( \bar{k} )</th>
<th>sGPA</th>
<th>SAT-V</th>
<th>SAT-M</th>
<th>SAT-T</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{k} )</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sGPA</td>
<td>-0.19</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT-V</td>
<td>-0.07</td>
<td>.26</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT-M</td>
<td>-0.08</td>
<td>.24</td>
<td>.79</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>SAT-T</td>
<td>-0.08</td>
<td>.26</td>
<td>.94</td>
<td>.95</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>( \bar{k} )</th>
<th>sGPA</th>
<th>SAT-V</th>
<th>SAT-M</th>
<th>SAT-T</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{k} )</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sGPA</td>
<td>-0.24</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT-V</td>
<td>-0.26</td>
<td>.40</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT-M</td>
<td>-0.29</td>
<td>.39</td>
<td>.44</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>SAT-T</td>
<td>-0.32</td>
<td>.46</td>
<td>.87</td>
<td>.82</td>
<td>1</td>
</tr>
</tbody>
</table>

In the lower panel, the lower left triangle contains the correlations within the Williams College sample, and the upper right triangle contains the correlations within the MCLA sample.
correlation for the MCLA groups was reduced substantially, \( r = -0.05 \), \( t(79) = -0.47 \), \( p = 0.64 \). These two partial correlations were not reliably different, \( Z = 1.25 \), \( N = 234 \), \( p = 0.21 \). Given our limited power to detect differences between correlations of these magnitudes, we can conclude little from these comparisons beyond noting that the correlations are consistently negative.

Finally, for the Williams sample we also were able to obtain “Dartmouth” GPAs, in which the students’ grades are expressed as deviations from the mean grade in each class, and then averaged together to create a GPA. This way of expressing grades reduces the impact on the GPA of the student’s choice of major and self-selection into courses of varying difficulties (Sabot & Wakeman-Linn, 1991). For the Williams sample the correlation between Dartmouth GPA and \( k \) was \( r = -0.23 \), which was reliable, \( t(152) = 2.87 \), \( p = 0.005 \). It was because the results using the Dartmouth GPA and raw GPAs were so similar, and because we did not have Dartmouth GPAs for the MCLA sample, that we focused on standardized raw GPAs in the analyses above.

3. Discussion

As we hypothesized, our results are consistent with the notion that people who more highly discount the future will tend to perform less well academically. Discount rates were consistently negatively correlated with GPA, even when SAT scores were partialled-out. The correlations tended around approximately \(-0.2\), which means that discount rates were associated with approximately 4% of the variance in GPA. In spite of the fact that it is predicted by the discounting model, for several reasons we find it remarkable that the observed proportion is as high as that. First, there are large individual differences in the magnitude effect on discount rates. Kirby (1997) found that log discount rates were negatively, and approximately linearly, related to log reward magnitude, but individuals varied in the slope of this negative linear relationship. These individual differences in the magnitude effect would tend to decrease the observed correlation between individuals’ discount rates for different reward magnitudes, such as the correlation between discount rates for the relatively small monetary rewards in the present study (US$10 to US$20) and discount rates for the relatively large rewards of academic success.

Second, people may have different discount rates for different types of rewards. For example, Kirby et al. (2002) found that discount rates estimated for monetary and candy rewards using the same methodology were correlated, but only in the .4 to .6 range. (The candy rewards were smaller in value, so individual differences in the magnitude effect might have reduced that correlation also.) Kirby and Guastello (2001) found that food rewards were discounted at higher rates than monetary rewards, even though they were equated in monetary value. (The food and money participants were tested at different times during the academic year, so the comparison is confounded.) Despite their limitations, the data from both studies are consistent with the idea that different types of rewards may be discounted at different rates. To the extent that this is true it makes it all the more surprising that discount rates for money would be reliably correlated with GPA, which is affected by a large number of decisions over a large variety of reward types.

Third, there is reason to believe that an individual’s discount rate may fluctuate to some extent based on current conditions. For example, it is intuitive that a hungry person might have a temporarily high preference for immediate food rewards (cf., Loewenstein, 1996). Consistent with this, Kirby et al. (2002) found that correlations between discount rates assessed repeatedly at 3 month intervals were reliable, but tended to fall only in the .2 to .4 range. In addition, discount rates for money tended to be negatively
correlated with a person’s cash income over the past 30 days. These and other factors could limit the intertemporal stability of discount rates, especially when measured for only one type of reward. This in turn would tend to reduce observed correlations between monetary discount rates observed at a single point in time and less temporally variable measures, such as GPA, which are influenced by varying discount rates over a long time interval.

Last, college GPA is clearly influenced by a large and rich variety of factors that add “noise” to our data, such as intelligence, primary and secondary school academic preparation, participation in extracurricular activities, parental pressure, selection of majors and courses, and peer influences (Goethals, 2001; Zimmerman, 1999). Although some of these factors may be related directly or indirectly to discount rates, for such a richly determined outcome as GPA it is impressive that monetary discount rates for money could account for as much as 4% of the variance.

A negative association between discount rates and academic performance is consistent with several other results in the literature. In a sample of Tsimane’ Amerindians in the Bolivian rainforest, Kirby et al. (2002) found reliable negative correlations between discount rates (k) and several educational measures, including years of education, numeracy, literacy, and the participants’ fathers’ years of education. Those results are consistent with our current hypothesis that people with higher discount rates will tend to perform less well academically. They are also consistent with Becker and Mulligan’s (1997) model, which predicts that individuals’ investments in their own education will reduce their discount rates (rates of time preferences). Of course, we cannot distinguish different causal scenarios on the basis of the present data, and it is possible that causation works in both directions.

The precise correspondence between discount rates and ability to delay gratification in the Mischel paradigm (e.g., Mischel, Shoda, & Rodriguez, 1989) is not known. However, the relationship between ability to delay gratification and academic performance is consistent with predictions of the discounting model. Preschoolers’ ability to delay gratification for time periods of seconds to minutes in the Mischel paradigm is positively associated with preschool achievement test scores (Flynn, 1985), and with scholastic performance (Mischel et al., 1989) and academic competence (Shoda, Mischel, & Peake, 1990) in adolescence. The ability to predict scholastic performance in adolescence based on delay of gratification in preschool is the best evidence to date that the ability to delay gratification is an intertemporally stable attribute of the individual.

The literature on impulsiveness also provides several examples consistent with the results of this study. Self-report impulsiveness measures appear to be negatively associated with medical school GPA (Barratt & White, 1969; Roessler et al., 1978). Academic achievement is negatively associated with a measure of impulsive (or “careless”) problem solving (Rodriguez-Fornells & Maydeu-Olivares, 2000). The data is mixed on the association between academic measures and reflection-impulsivity, as measured by the Matching Familiar Figures Task (MFFT). Some studies have found negative relationships between impulsiveness in the MFFT and academic measures, such as reading achievement (Epstein, Cullinan, & Sternberg, 1977), scholastic achievements tests (Margolis et al., 1982; Weithorn, Kagen, & Marcus, 1984), and that the relationship remains when controlling for intelligence (Miyakawa, 2001). However, other studies have not found a relationship (Harrison & Romanczyk, 1991) or have found that the relationship disappears when intelligence is partialled out (Tiedemann & Meffert, 1980). Although none of these studies address discounting per se, the weight of the evidence is consistent with predictions of the discounting model of impulsiveness (Ainslie, 1975).

Of course, we cannot conclude on the basis of these correlational data that we have shown that discount rates affect GPA. There might be some path by which GPA could affect one’s discount rate, or
by which a third variable might directly or indirectly influence both GPA and discount rates. Nevertheless, the negative relationship between GPA and discount rates is consistent with the predictions of the discounting model of impulsiveness, and increases the external validity of the model. It also suggests a possible personal attribute of interest in research on higher education. For example, a high discount rate may prove to be a risk factor for academic difficulty. Future research may address issues such as the possible influence of peers’ discount rates on a student’s own GPA, as well as possible strategies for reducing the risk factor posed by high discount rates.

Acknowledgements

This research was supported by the National Institutes of Health Grant MH55262. We thank Deborah Foss for arranging for us to collect data at MCLA, and Denise Richardello for providing the SAT and GPA data at MCLA.

References


