

BRIEF REPORTS

One-year temporal stability of delay-discount rates

KRIS N. KIRBY

Williams College, Williamstown, Massachusetts

The temporal stability of delay-discount rates for monetary rewards was assessed using a monetary choice questionnaire (Kirby & Marakovic, 1996). Of 100 undergraduate participants who completed the questionnaire at the initial session, 81 returned 5 weeks later and 46 returned 57 weeks later for subsequent sessions. The 5-week test-retest stability of discount rates was .77 (95% confidence interval = .67–.85), the 1-year stability was .71 (.50–.84), and the 57-week stability was .63 (.41–.77). Thus, at least when similar testing situations are reinstated, discount rates as individual differences have 1-year stabilities in the range that is typically obtained for personality traits. Discount rates index an attribute of the person that is relatively stable over time but that is moderated by aspects of the situation, such as reward type and deprivational state.

People differ in how they value the future consequences of their actions. Some people value future consequences very highly, which, according to the discounting model of impulsive choice (Ainslie, 1975; Rachlin, 2000), tends to make them more self-controlled. Others appear to place little value on future consequences, which tends to make them more impulsive. The primary question addressed in this article is whether such individual differences in the way we value future consequences are stable over time.

Delay discounting refers to the decrease in the present value of a future outcome as the delay to that outcome increases. That is, future outcomes are discounted, in part, because of their delays. (Because the research below used only positive outcomes, I will henceforth refer to outcomes as *rewards*.) Figure 1 illustrates discount curves for two hypothetical individuals. Time moves left to right, and, thus, the delays to future rewards increase from right to left. The height of the vertical line at delay = 0 indicates the present value of a reward of amount A when it is available immediately. The curves show how the present value of the reward declines as the delay to the reward increases. The person represented by the solid curve has a lower discount rate than does the person represented by the dashed curve and, thus, tends to value future rewards more highly.

Previous research (for a review, see Green & Myerson, 2004) has shown that discount curves are well fit by the hyperbolic function (Mazur, 1987)

$$V = \frac{A}{1 + kD}, \quad (1)$$

where V is the present value of reward A available at delay D , and k is the discount rate parameter. Although k determines the rate of discounting, it does not correspond to a fixed proportional rate of decrease. Thus, to get an intuitive sense of the value of k , it can be helpful to convert it to a *half-life*—that is, the delay at which the delayed reward has a present value equal to half its full amount (Kirby & Marakovic, 1996). For Equation 1, the half-life is equal to the inverse of k . For example, the hypothetical person represented by the dashed curve in Figure 1 has $k = 0.02$, and, thus, the reward A has half of its value

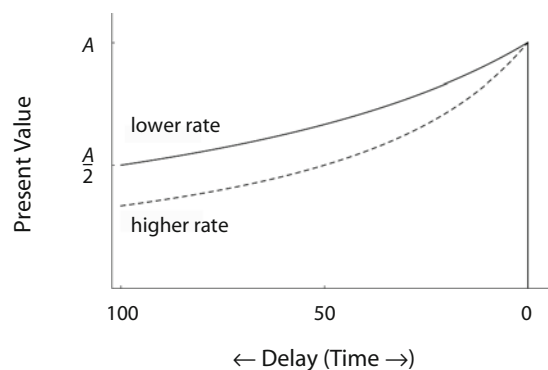


Figure 1. Discount curves for a reward of amount A available at a future time denoted by delay = 0 for two different people: one with a higher discount rate (dashed curve) and one with a lower discount rate (solid curve). For the person with the higher discount rate, the reward has only half of its full value ($A/2$) at a delay of 50 days. For the person with the lower discount rate, the reward still has half of its value at twice that delay (100 days).

K. N. Kirby, kkirby@williams.edu

when delayed by 50 days. In contrast, the hypothetical person represented by the solid curve in Figure 1 has $k = 0.01$, and so reward A has half of its value when delayed by 100 days.

According to the discounting model of impulsive choice (Ainslie, 1975; Rachlin, 2000), as discount rates increase, people are increasingly inclined to choose more immediate over more delayed rewards. Thus, k serves as an index of impulsivity and has acquired respectable external and criterion validity as such an index. Estimates of k correlate positively with self-reported impulsivity (Kirby & Finch, 2008; Kirby & Petry, 2004) and behaviors such as impulse buying (Dittmar, 2001). They correlate negatively with college grade point average (Kirby, Winston, & Santisteban, 2005) and adolescent academic performance, which they predict better than does IQ (Duckworth & Seligman, 2005). They are higher, on average, for substance abusers than for non-drug users (Kirby & Petry, 2004; Kirby, Petry, & Bickel, 1999) and for pathological gamblers than for controls (Petry, 2001).

Such applications have led to a rapid increase in research using estimates of k . It is clear that k is associated with meaningful individual differences. What is not clear is the degree to which k is stable over long time periods. Simpson and Vuchinich (2000) assessed discount rates for hypothetical monetary gains for 15 participants in two sessions separated by 1 week. For Equation 1, the correlation between k at Week 1 and Week 2 was .91. Baker, Johnson, and Bickel (2003) also examined the 1-week stability of k for 30 current smokers and 30 never-before smokers. For monetary gains, the test-retest correlations ranged from .71 to .90 across hypothetical/real, smoker/nonsmoker, and delayed reward magnitude conditions. Johnson, Bickel, and Baker (2007) replicated this study with a group of light smokers, and reported 1-week correlations from .55 to .90. Each of these studies suggests that discount rates are quite stable. However, a 1-week interval is short enough that memory effects—simply recalling what one chose last time—plausibly could have inflated the correlations (Remmers, 1963) and that any situational factors that might affect participants' discount rates over longer periods were not likely to have changed much during the retest interval.

Ohmura, Takahashi, Kitamura, and Wehr (2006) assessed the 3-month stability of k for hypothetical monetary rewards. For their 22 participants, the correlation between sessions was .61. Although roughly consistent with the results of previous research, it appears that the test-retest correlations in their study may have been influenced by a small number of points with very high leverage (see Ohmura et al., 2006, Figure 1), and, excluding those points, it is difficult to assess whether discount rates would have been stable.

Two of these previous studies reported mean discount rates at each time period, and each found that discount rates increased across sessions. Simpson and Vuchinich (2000) found a (geometric) mean increase from $k = 0.0058$ to $k = 0.0061$ across a 1-week interval, correspond-

ing to an increase in discount rates with slope $0.208 \ln(k)/\text{month}$. Ohmura et al. (2006) found a mean increase across a 3-month interval from $k = 0.017$ to $k = 0.031$, corresponding to a slope of $0.202 \ln(k)/\text{month}$. (For both of these studies, I converted the k values to correspond to delays scaled in days.) The consistency of positive slopes in these two studies suggests that this effect merits further consideration.

In the present study, I avoided the particular limitations of the studies reviewed above by examining the temporal stability of discount rates for probabilistic real rewards with a larger number of participants over longer time periods. Both *relative* stability (traditionally called *test-retest reliability*) and *absolute* stability (mean change over time) were addressed.

METHOD

Participants

The participants in the first session were 100 undergraduate students (48 men and 52 women) at Williams College. Mean age was 18.7 years. Eighty-one of these returned 5 weeks later for a second session, and 46 returned 1 year after that for a third session. Thirty-seven of the participants who returned for Session 2 were among those who returned for Session 3. The numbers of participants who completed each combination of sessions are shown in Figure 3. All were enrolled in an introductory psychology course during the first two sessions, for which they received course credit.

Procedure

Session 1 data collection took place in early March, about 1 month into a semester that ended in late May. Data collection for Session 2 took place in mid-April, as did that for Session 3 1 year later. For Sessions 1 and 2, participants were tested in groups of 30–60 in a large room. They completed the monetary choice questionnaire described below, along with several other questionnaires. At the end of each group session, one participant was randomly selected by lottery to receive a monetary reward, which was based on his or her choice on a randomly selected item from the questionnaire. For Session 3, all of the initial 100 participants were invited to a campus lab to complete the questionnaire again. The participants who accepted were scheduled and tested individually and were each given a 1-in-10 chance of winning one of their choices on the questionnaire. Thus, the probability of receiving a reward was 3–6 times higher in Session 3 than in the earlier sessions.

Materials

The monetary choice questionnaire (Kirby et al., 1999) contained 27 items, and, for each item, the participant made a choice between a smaller, immediate amount (S) and a larger, delayed amount (L). Table 1 shows the values used for each item. For example, for the item in the top row of Table 1, participants were asked, "Would you prefer (a) \$34 today or (b) \$35 in 186 days?" The 27 items were grouped into three categories on the basis of the approximate magnitudes of the delayed rewards: small (\$25–\$35), medium (\$50–\$60), and large (\$75–\$85). Estimating discount rates separately for each size allows us to assess the magnitude effect on discount rates—that is, the tendency for discount rates to decrease as the magnitudes of the delayed rewards increase (Green, Fisher, Perlow, & Sherman, 1981).

Data Analysis

Discount rates were estimated from the pattern of choices that participants made across the nine questions in each size category.

Table 1
Questionnaire Item Values, and Their
Associated Discount Rates (k) at Indifference

	Reward Values			
Order	S (\$)	L (\$)	Delay (days)	Indifference k
Small Delayed Rewards (\$25–\$35)				
13	34	35	186	0.00016
20	28	30	179	0.00040
26	22	25	136	0.0010
22	25	30	80	0.0025
3	19	25	53	0.0060
18	24	35	29	0.016
5	14	25	19	0.041
7	15	35	13	0.10
11	11	30	7	0.25
Medium Delayed Rewards (\$50–\$60)				
1	54	55	117	0.00016
6	47	50	160	0.00040
24	54	60	111	0.0010
16	49	60	89	0.0025
10	40	55	62	0.0060
21	34	50	30	0.016
14	27	50	21	0.041
8	25	60	14	0.10
27	20	55	7	0.25
Large Delayed Rewards (\$75–\$85)				
9	78	80	162	0.00016
17	80	85	157	0.00040
12	67	75	119	0.0010
15	69	85	91	0.0025
2	55	75	61	0.0060
25	54	80	30	0.016
23	41	75	20	0.041
19	33	80	14	0.10
4	31	85	7	0.25

Note—Indifference k indicates the discount rate parameter for which the smaller, immediate reward (S) and the larger, delayed reward (L) are of equal value according to Equation 1. Order, the order of items as presented on the questionnaire.

Using Equation 1, we can solve for the value of k that would yield indifference between S and L for each item in Table 1. These values are shown in the last column of Table 1. A participant who chooses S on a given item reveals a discount rate for L that is greater than the indifference k , whereas a participant who chooses L reveals a discount rate that is less than the indifference k . By making use of these inequalities, a participant's pattern of choices across the nine items allows us to estimate the participant's discount rate. For example, choosing S on the top four items in Table 1 and L on the next five

would imply a discount rate greater than 0.0025 but less than 0.006. The geometric mean of this interval (0.0039) would then be used as the estimate of k . The nine items within each reward magnitude define eight such bounded intervals, and the two endpoints (0.00016 and 0.25) were assigned to participants who chose all nine L or S rewards, respectively. The questionnaire therefore defines 10 possible k assignments. The particular k values shown in Table 1 were chosen because they yield equal intervals on a log scale.

Unlike the example above, however, participants' choices are not always perfectly consistent with a single value of k . Therefore, the proportion of each participant's choices that was consistent with assignment to each of the 10 values of k was computed, and the participant was assigned the k that yielded the highest consistency among his or her choices (see Kirby & Marakovic, 1996). If two or more k values were tied, their geometric mean was used. This procedure minimizes the number of choices that would be errors with respect to the assigned k value.

The distributions of k were approximately normalized using the natural log transformation. All means reported below are geometric means—which correspond to means of the log-transformed values—and thus provide better measures of central tendency for positively skewed metrics, such as discount rates, than do arithmetic means. All confidence intervals (CIs) below are 95% intervals, and, thus, any effects whose CIs do not contain zero can be considered statistically significant in the traditional sense. The correlations reported below are all Pearson product-moment correlations, and exact 95% CIs for the correlations were computed using the program R2 (Steiger & Fouladi, 1992).

RESULTS

The estimation procedure assigned k values such that over 99% of choices were consistent with the assigned k s. Two participants made choices for a single reward magnitude in a single session that yielded less than 75% consistency with any value of k , and those two values were dropped. For all other reward magnitudes in all sessions, at least 88% of the participants had 100% consistency, and for no estimate were choices less than 89% consistent with the assigned k .

The temporal stabilities of $\ln(k)$ between sessions are presented in Table 2, along with their 95% confidence intervals. Univariate histograms and scatterplots of $\ln(k)$ for each pair of sessions are shown in Figure 2 (see next page). The univariate distributions were roughly normal, with small negative skew. There were a few points with high leverage in these analyses (e.g., the lower left data point in Figure 2A), but none had meaningful influence on the correlations, with all Cook's D values less than 0.5.

Table 2
Test–Retest Correlations Between $\ln(k)$ Estimates

	Test–Retest Interval					
	5 Weeks		1 Year		57 Weeks	
	Sessions 1–2	95% CI	Sessions 2–3	95% CI	Sessions 1–3	95% CI
Number of participants	81		37		46	
Mean $\ln(k)$.77	.67–.85	.71	.50–.84	.63	.41–.77
$\ln(k)$ for small rewards	.71	.58–.80	.71	.50–.84	.59	.36–.75
$\ln(k)$ for medium rewards	.75	.63–.83	.71	.49–.83	.61	.38–.76
$\ln(k)$ for large rewards	.66	.51–.76	.59	.33–.76	.57	.33–.73

Note—CI, confidence interval.

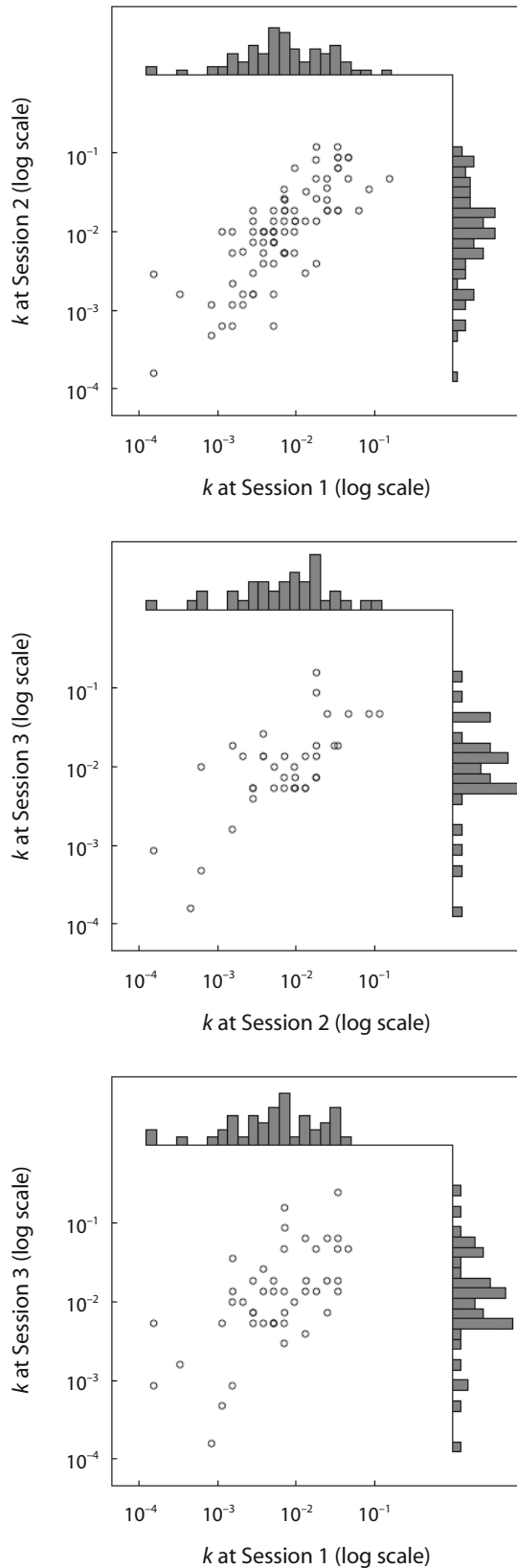


Figure 2. Scatterplots of discount rates at a later testing session as a function of rates at an earlier session and univariate histograms for each variable.

Thus, the high test–retest correlations in Table 2 were not attributable to a few points with high influence.

Table 3 (facing page) shows the pairwise within-session correlations between $\ln(k)$ estimates for each of the three reward magnitudes, along with their 95% confidence intervals. These correlations were uniformly high and indicate good internal consistency among discount rates for different reward magnitudes.

Figure 3 (see below) displays mean k across sessions for each subgroup of participants who completed different combinations of sessions. The mean within-subjects change from Session 1 to Session 2 in log units was 0.34 (0.14–0.55), which corresponds to 0.27 $\ln(k)$ /month. This is similar to the changes reported in the studies reviewed above. The mean within-subjects change from Session 2 to Session 3 in log units was 0.27 (–0.08–0.62), which corresponds to 0.02 $\ln(k)$ /month. However, as was noted above, changes from Session 2 to Session 3 were confounded with an increase in the probability of receiving a real reward. If people discount expected values, and if there is a magnitude effect for expected values (i.e., if higher expected values are discounted at lower rates), then the higher expected values in Session 3 might be predicted to yield lower discount rates than those in the previous sessions. That discount rates were highest in Session 3 suggests either that the magnitude effect is prior to the effect of expectation (if any) on discount rates or that this effect is more than offset by an increase in discount rates over sessions or time.

The magnitude effect on discount rates was observed in all three sessions. The subset of participants who completed all three sessions is shown in Figure 4 (facing page). (The graph including all participants looked nearly identical to the one in Figure 4.) Geometric mean k declined as a function of mean reward size within each of the three sessions, and the decreasing linear trends each had CIs that excluded zero.

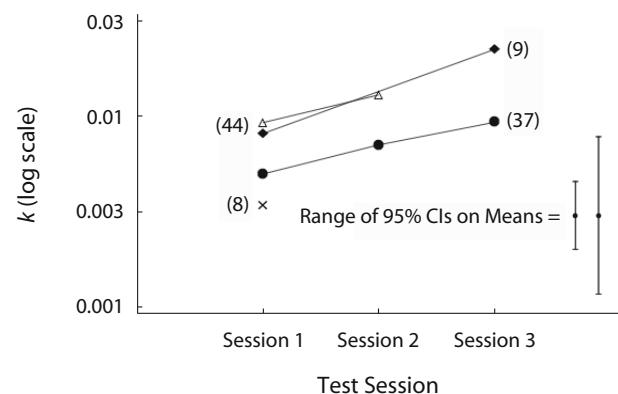


Figure 3. Geometric mean k (in log coordinates) as a function of session for participants who completed only Session 1 (\times), participants who completed only Sessions 1 and 2 (Δ), participants who completed only Sessions 1 and 3 (\blacklozenge), and participants who completed all three sessions (\bullet). The numbers of participants in each subgroup are shown in parentheses. The detached error bars indicate the 95% confidence intervals (CIs) on the single means with the largest and smallest n s, on the basis of a pooled error term (Estes, 1997).

Table 3
Within-Session Correlations Between $\ln(k)$ Estimates for Different Reward Magnitudes

	Reward Magnitude Pairs					
	Small–Medium		Small–Large		Medium–Large	
	Correlation	95% CI	Correlation	95% CI	Correlation	95% CI
Session 1 ($n = 99$)	.81	.73–.87	.79	.70–.85	.83	.75–.88
Session 2 ($n = 81$)	.88	.82–.92	.84	.75–.89	.92	.88–.95
Session 3 ($n = 46$)	.88	.78–.93	.92	.85–.95	.95	.90–.97

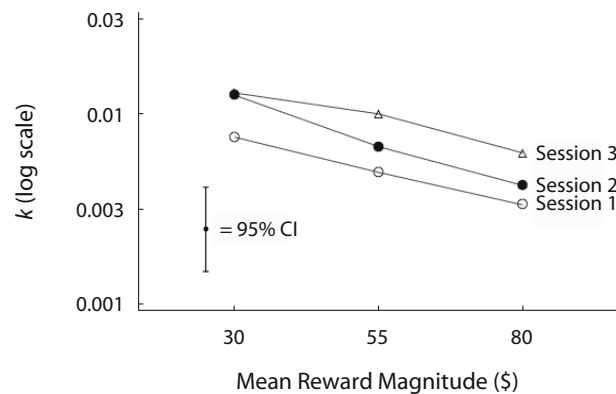


Figure 4. Geometric mean k (in log coordinates) as a function of mean reward magnitude for those participants who completed all three sessions. The detached error bars indicate the 95% confidence interval (CI) on a single mean, on the basis of the largest within-cell variance from any cell.

DISCUSSION

The relative temporal stability of k as assessed by this questionnaire was impressive. The 1-year and 57-week test–retest reliabilities in Table 2 compare favorably with the 1-year test–retest stabilities of a variety of personality scales (Schuerger, Tait, & Tavernelli, 1982) and suggest that the discount rate for monetary rewards is a stable individual difference, at least within the context in which these participants were tested. All were students tested on their college campus, and, thus, subsequent test sessions would have reinstated many of the situational factors in place at earlier test sessions. Transient variables that are related to discount rates, such as reward-specific deprivation (e.g., Field, Santarcangelo, Sumnall, Goudie, & Cole, 2006) and sleep deprivation (Acheson, Richards, & de Wit, 2007), presumably varied unsystematically across testing sessions and would have decreased the observed reliabilities. Other variables that are associated with discount rates, such as educational level (Kirby et al., 2002) and income (Green, Myerson, Lichtman, Rosen, & Fry, 1996), would not have changed much for most participants over the course of 1 year in this context. However, stability might decline over longer time periods as these variables change.

Mean discount rates increased across sessions, consistent with the results of previous studies. The increases observed here were nontrivial: For the participants represented by the filled circles in Figure 3, the mean reward half-life decreased from 200 to 107 days across sessions. This effect runs counter to the typical decline in discount

rates with age across the life span (Green et al., 1996; Green, Myerson, & Ostaszewski, 1999; Kirby et al., 2002). It is possible that the college years constitute an exception to the longer-term trend, yielding a mean increase in discount rates and perhaps in impulsivity. Further research will be required to assess this possibility.

Overall, the attrition rate in this study from Session 1 to Session 2 was 19% and from Session 1 to Session 3 was 54%. It seems plausible that participants with the highest discount rates, who are presumably the most impulsive, might be least likely to show up for a given test session. There was no evidence for such rate-related self-selection, however. The 8 participants who participated only in Session 1 had a lower mean discount rate than did those who participated in additional sessions; those who participated in two sessions had the highest discount rates, and those who participated in all three sessions had intermediate discount rates (see Figure 3). In addition, the subgroups of participants defined by the different combinations of sessions that they completed tended to show qualitatively similar effects: Their discount rates increased over sessions with similar slopes (Figure 3), and their magnitude effects were similar. We have no evidence, therefore, that the assessment of temporal stability was compromised by attrition.

As in personality research (Fleeson, 2007), our present state of knowledge appears to support an interactionist position on discount rates: They index an attribute of the person that is relatively stable over time but moderated by aspects of the situation, such as reward type and depri-

ational state. Associations between monetary discount rates and, for example, substance abuse support the external validity of discount rates as measures of impulsivity and provide evidence that discounting is to some extent cross-situational. The results of the present study suggest that discount rates as individual differences are relatively stable over time periods of more than 1 year, at least when similar testing situations are reinstated. This is good news for applied researchers who may wish to assess associations between discount rates and other variables over extended time periods.

AUTHOR NOTE

This research was funded by NIH Grant MH55262, and preparation of the manuscript was supported in part by NSF Grant SES-0617868. The author thanks Jeremy Wilmer for data collection. Correspondence concerning this article should be addressed to K. N. Kirby, Department of Psychology, Williams College, Williamstown, MA 01267 (e-mail: kkirby@williams.edu).

REFERENCES

- ACHESON, A., RICHARDS, J. B., & DE WIT, H. (2007). Effects of sleep deprivation on impulsive behaviors in men and women. *Physiology & Behavior*, **91**, 579-587.
- AINSLIE, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, **82**, 463-496.
- BAKER, F., JOHNSON, M. W., & BICKEL, W. K. (2003). Delay discounting in current and never-before cigarette smokers: Similarities and differences across commodity, sign, and magnitude. *Journal of Abnormal Psychology*, **112**, 382-392.
- DITTMAR, H. (2001). Impulse buying in ordinary and "compulsive" consumers. In E. U. Weber, J. Baron, & G. Loomes (Eds.), *Conflict and tradeoffs in decision making* (pp. 110-135). New York: Cambridge University Press.
- DUCKWORTH, A. L., & SELIGMAN, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, **16**, 939-944.
- ESTES, W. K. (1997). On the communication of information by displays of standard errors and confidence intervals. *Psychonomic Bulletin & Review*, **4**, 330-341.
- FIELD, M., SANTARCANGELO, M., SUMNALL, H., GOUDIE, A., & COLE, J. (2006). Delay discounting and the behavioural economics of cigarette purchases in smokers: The effects of nicotine deprivation. *Psychopharmacology*, **186**, 255-263.
- FLEESON, W. (2007). Situation-based contingencies underlying trait-content manifestation in behavior. *Journal of Personality*, **75**, 825-862.
- GREEN, L., FISHER, E. B., JR., PERLOW, S., & SHERMAN, L. (1981). Preference reversal and self-control: Choice as a function of reward amount and delay. *Behaviour Analysis Letters*, **1**, 43-51.
- GREEN, L., & MYERSON, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, **130**, 769-792.
- GREEN, L., MYERSON, J., LICHTMAN, D., ROSEN, S., & FRY, A. (1996). Temporal discounting in choice between delayed rewards: The role of age and income. *Psychology & Aging*, **11**, 79-84.
- GREEN, L., MYERSON, J., & OSTASZEWSKI, P. (1999). Discounting of delayed rewards across the life span: Age differences in individual discounting functions. *Behavioural Processes*, **46**, 89-96.
- JOHNSON, M. W., BICKEL, W. K., & BAKER, F. (2007). Moderate drug use and delay discounting: A comparison of heavy, light, and never smokers. *Experimental & Clinical Psychopharmacology*, **15**, 187-194.
- KIRBY, K. N., & FINCH, J. (2008). *The hierarchical structure of self-reported impulsivity in a college-age population*. Manuscript submitted for publication.
- KIRBY, K. N., GODOY, R., REYES-GARCÍA, V., BYRON, E., APAZA, L., LEONARD, W., ET AL. (2002). Correlates of delay-discount rates: Evidence from Tsimane' Amerindians of the Bolivian rain forest. *Journal of Economic Psychology*, **23**, 291-316.
- KIRBY, K. N., & MARAKOVIC, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic Bulletin & Review*, **3**, 100-104.
- KIRBY, K. N., & PETRY, N. M. (2004). Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction*, **99**, 461-471.
- KIRBY, K. N., PETRY, N. M., & BICKEL, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General*, **128**, 78-87.
- KIRBY, K. N., WINSTON, G. C., & SANTIESTEBAN, M. (2005). Impatience and grades: Delay-discount rates correlate negatively with college GPA. *Learning & Individual Differences*, **15**, 213-222.
- MAZUR, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: The effect of delay and of intervening events on reinforcement value* (Vol. 5, pp. 55-73). Hillsdale, NJ: Erlbaum.
- OHMURA, Y., TAKAHASHI, T., KITAMURA, N., & WEHR, P. (2006). Three-month stability of delay and probability discounting measures. *Experimental & Clinical Psychopharmacology*, **14**, 318-328.
- PETRY, N. M. (2001). Pathological gamblers, with and without substance abuse disorders, discount delayed rewards at high rates. *Journal of Abnormal Psychology*, **110**, 482-487.
- RACHLIN, H. (2000). *The science of self-control*. Cambridge, MA: Harvard University Press.
- REMMERS, H. H. (1963). Rating methods in research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally.
- SCHUERGER, J. M., TAIT, E., & TAVERNELLI, M. (1982). Temporal stability of personality by questionnaire. *Journal of Personality & Social Psychology*, **43**, 176-182.
- SIMPSON, C. A., & VUCHINICH, R. E. (2000). Reliability of a measure of temporal discounting. *Psychological Record*, **50**, 3-16.
- STEIGER, J. H., & FOULADI, R. T. (1992). R2: A computer program for interval estimation, power calculation, and hypothesis testing for the squared multiple correlation. *Behavior Research Methods, Instruments, & Computers*, **4**, 581-582.

(Manuscript received September 23, 2008;
revision accepted for publication January 12, 2009.)