Published in final edited form as:
Exp Clin Psychopharmacol. 2008 June ; 16(3): 264–274. doi:10.1037/1064-1297.16.3.264.

An Algorithm for Identifying Nonsystematic Delay-Discounting Data

Matthew W. Johnson and
Behavioral Pharmacology Research Unit, Johns Hopkins University School of Medicine

Warren K. Bickel
Center for Addiction Research, Psychiatric Research Institute, University of Arkansas for Medical Sciences

Abstract

Several previous discounting studies have used the R^2 measure to identify data sets with poor fits to a mathematical discounting model as nonsystematic data to be eliminated before further analyses are conducted. Data from three previous delay-discounting studies (6 separate groups, with a total of 161 individuals) were used to demonstrate why using R^2 to assess the fits of discounting data is problematic. A significant, positive correlation between discounting rate parameter and R^2 was found in most groups, showing that R^2 is more stringent as a measure of fit for low discounting rates than for high discounting rates. Furthermore, it is suggested that identifying nonsystematic data based on any measure of fit to a mathematical discounting model may be problematic because it confounds discounting rate comparison with the issue of discounting model assessment. Therefore, a model-free method to identify nonsystematic data is needed. An algorithm for identifying nonsystematic data is presented that is based on the expectation of a monotonically decreasing discounting function. This algorithm identified 13 cases out of the 161 reanalyzed data sets as nonsystematic. These nonsystematic data are presented, along with examples of data not identified as nonsystematic. This algorithm, or modifications of it, may be useful in a variety of human and nonhuman animal discounting studies (e.g., delay discounting, probability discounting) as an alternative to the R^2 measure for identifying nonsystematic data. The algorithm may be used in empirical investigations to improve discounting methodology, and may be used to identify outliers to be removed from analyses.

Keywords

delay discounting; probability discounting; hyperbolic discounting; R^2; outliers

Delay discounting (also referred to as “time” or “temporal” discounting) is the concept that delaying a consequence decreases its effect on behavior (Critchfiled & Kollins, 2001). Such devaluation of delayed consequences has been demonstrated in both human and nonhuman animals, typically with choice procedures revealing that, all else being equal, sooner reinforcers are preferred over delayed ones, and delayed losses or punishments are preferred over sooner ones (e.g., Ainslie, 1975; Baker, Johnson, & Bickel, 2003; Benzion, Rapoport, & Yagil, 1989; Chung & Herrnstein, 1967; Deluty, 1978; Green, Fisher, Perlow, & Sherman, 1981; Green, Fry & Myerson, 1994; Loewenstein, 1988; Logue, 1988; Mischel, Shoda & Rodriguez, 1989; Rachlin, Raineri, & Cross, 1991; Rachlin & Green, 1972; Shelly, 1993; Thaler, 1981).

Correspondence: Matthew W. Johnson, Behavioral Pharmacology Research Unit, Johns Hopkins University School of Medicine, Baltimore, MD 21224-6823, Phone: (410) 550-0056, mwj@jhu.edu.

Delay discounting is considered a behavioral model of impulsivity, and the use of laboratory choice procedures may represent the best way to empirically quantify delay discounting (Ainslie, 1975).

Application of delay-discounting methods and analyses has rapidly expanded to address a variety of issues of clinical and theoretical importance. For example, several drug-dependent populations discount future consequences at high rates, indicating that delay discounting may serve a fundamental role in drug dependence (for reviews see Bickel & Johnson, 2003; Reynolds, 2006). Studies in rats suggest that high delay-discounting rates are associated with vulnerability to acquisition of cocaine self-administration (Perry, Larson, German, Madden & Carroll, 2005), greater nicotine seeking during nicotine extinction, and greater reinstatement of nicotine seeking by re-exposure to nicotine-associated cues (Diergaarde et al., 2008). Human studies suggest that high delay-discounting rates may predict lack of success in drug-dependence treatment (Dallery & Raiff, 2007; Krishnan-Sarin et al., 2007; Yoon et al., 2007). Similarly, individuals with other psychiatric problems have also been shown to discount at high rates (Alessi & Petry, 2003; Barkley, Edwards, Laneri, Fletcher, & Metevia, 2001; Crean, de Wit, & Richards, 2000; Rounds, Beck, & Grant, 2007). In addition to behavioral pathology, delay discounting may further the analysis of virtually any human behavior involving delayed consequences, from academic study patterns to environmental conservation (Critchfiled & Kollins, 2001). Human studies have demonstrated that delay-discounting rate differs across lifespan development (Green, Myerson, Lichtman, Rosen, & Fry, 1996; Green, Myerson, & Ostaszewski, 1999), levels of income (Green et al., 1996), and cultures (Du, Green, & Myerson, 2002). Furthermore, delay discounting has diverse theoretical implications. For example, the mathematical form of delay discounting may inform theories of animal foraging (Critchfield & Atteberry, 2003; Green & Myerson, 1996), and may account for compulsions, the perception of pain, and willpower (Ainslie, 2001). Delay discounting has proven to be methodologically applicable and theoretically relevant to a variety of scientific interests, and has resulted in an increasingly expanding scientific literature.

Given the myriad current and potential applications of delay discounting, a critical goal for investigators is to establish valid and sensitive procedural and analytic methods for future research. In fact, several reports have focused primarily on procedural issues, such as the development or comparison of delay-discounting procedures (e.g., Epstein, Richards, Saad, Paluch, Roemmich, & Lerman, 2003; Kowal, Yi, Erisman, & Bickel, 2007; Lane, Cherek, Pietras, & Tcheremissine, 2003; Reynolds & Schiffbauer, 2004), the reliability of delay-discounting measures (Ohmura, Takahashi, Kitamura, & Wehr, 2006; Simpson & Vuchinich, 2000), and the use of hypothetical instead of real rewards in delay-discounting procedures (Johnson & Bickel, 2002; Lagorio & Madden, 2005; Madden, Begotka, Raiff, & Kastern, 2003; Madden et al., 2004). However, a relatively unexplored issue is the identification of nonsystematic data in delay-discounting research. In discounting studies, some subjects provide more systematic data than others. For example, some subjects may generate data showing the value of delayed rewards to increase and decrease across delays in a haphazard fashion. An objective method for identifying such cases may assist in improving research methods and conducting statistical analyses.

Use of R^2 for Testing Model Fit and Data Elimination

Often the R^2 measure is used to assess the fit of a set of indifference points (magnitude of immediate outcome subjectively equivalent to the delayed larger outcome) to a mathematical discounting model. Moreover, researchers have commonly identified those data sets with an R^2 less than some threshold (e.g., $R^2 < 0$, .3, .4, or .5) and eliminated these data sets from comparative analyses (e.g., Acheson, Richards, de Wit, 2007; Epstein et al., 2003; Green & Myerson, 1995; Hamidovic, Kang, & de Wit, 2008; Myerson & Green, 1995; Ohmura et al.,

2006; McDonald, Schleifer, Richards, & de Wit, 2003; Reynolds, Richards, & de Wit; Richards, Zhang, Mitchell, & de Wit, 1999; Reynolds & Schiffbauer, 2004; Simpson & Vuchinich, 2000). An R^2 value is obtained by using nonlinear regression to fit the data to a mathematical discounting model. The hyperbolic decay model, developed by Mazur (1987) and expressed in Equation 1, is a widely utilized single parameter model for discounting data, and will be used to obtain R^2 in subsequent analyses. In this equation, *I* is the indifference point expressed as a proportion of the later outcome amount. *D* is the delay from the choice until the occurrence of the outcome, and *k* is a free parameter that serves as an index for discounting rate (i.e., impulsivity). We use the term "rate" for convenience. However, in hyperbolic discounting, *k* does not represent a global discounting rate across the discounting function (i.e., an equal proportional decrement in value with every equal unit of delay), as is the case with exponential discounting. Instead, hyperbolic discounting implies that discounting rate is greatest at the shortest delays and declines as a function of increasing delay (Frederick, Loewenstein, & O'Donoghue, 2002).

$$I=\frac{1}{1+kD} \tag{1}$$

R^2 is defined in Equation 2, in which SSE represents the sum of squares error. In this equation, the mean of the data represents the alternative or null hypothesis of the model, i.e., that the independent variable (delay) has no relation to the dependent variable (indifference points).

$$R^2=1-\frac{SSE_{model}}{SSE_{mean}} \tag{2}$$

R^2 is typically interpreted as the proportion of variance explained by the model. This description is not appropriate as it is with linear regression (Ratkowsky, 1990). In linear regression the ratio in Equation 2 can never be greater than 1, and R^2 can never be less than zero, because the SSE for the mean represents the maximum SSE possible for the model. That is, a horizontal line through the mean of the data will always result in equivalent or greater error than a line with its slope and intercept free to vary (i.e., the linear model). In contrast, with nonlinear regression the SSE for the mean does not represent the maximum SSE possible for the model. The mean may provide less error than the nonlinear curve, resulting in a negative R^2 value (Motulsky & Christopoulos, 2003). Therefore, the SSE for the mean in nonlinear regression represents a rather arbitrary comparator for the fit of the nonlinear model, which calls into question the interpretation of R^2 as the proportion of explained variance. Indeed, it has been argued that R^2 does not provide any clear meaning when applied to nonlinear regression (Ratkowsky, 1990).

An Empirical Demonstration of the Problematic Nature of R^2 in Discounting Analyses

This and subsequent analyses used indifference points for a hypothetical $1000 reward collected from 161 research participants in three previously published studies: 18 opioid-dependent participants and 38 control participants from Madden, Petry, Badger & Bickel, 1997; 23 heavy tobacco smokers and 22 control participants from Bickel, Odum & Madden, 1999; and 30 heavy tobacco smokers and 30 control participants from Baker et al., 2003. The procedures used to assess delay discounting in these participants were typical for delay-discounting studies with humans. All studies used methods generally developed by Rachlin et

al. (1991) in which a series of hypothetical choices between an immediate amount of money and a delayed $1000 reward was presented at each of seven delays. At each delay, the magnitude of the immediate reward was titrated across trials until the participant's choices indicated indifference between the immediate reward amount and the delayed $1000. As shown in Table 2, the studies differed regarding choice presentation method (computer or index cards), method for titrating immediate reward value, and delays assessed. These methodological variations, along with the examination of differing populations (opioid-dependent, tobacco-dependent, and non-dependent control participants) helped to evaluate whether analysis results were robust across methodologies and populations.

Table 2 shows Spearman correlations between the discounting parameter *k* (Equation 1) and R^2 for each of the six groups. In four of the six groups, *k* was significantly (i.e., $p<.05$) and positively correlated with R^2. The smoking group in the Baker et al. (2003) study showed a nonsignificant trend toward a positive correlation. The opioid-dependent group in the Madden et al. (1997) study showed no substantial trend, possibly because this group had the lowest number of participants (less power than other correlations). These results show that as a measure of discounting function fit, R^2 is biased toward reporting a better fit for high-rate (i.e., more impulsive) discounting data. Because R^2 assesses the fit of the model *relative* to the fit of the mean of the data, R^2 is more stringent on discounting functions as they approach zero discounting. The error associated with the mean will tend to be smaller with low-rate discounting because all indifference points, even those at long delays, will be relatively close to 1 and resemble a horizontal line. With higher-rate discounting, indifference points may range from values close to 1 at short delays, to values close to 0 at long delays, resulting in a relatively large SSE associated with their mean, a large denominator in the ratio in Equation 2, and a relatively large R^2.

To illustrate the drawback of R^2 for delay-discounting data, Figure 1 displays indifference points and best-fitting hyperbolic functions for two individual participants from the Bickel et al. (1999) control group. The upper curve and squares are the data of participant EA, and the lower curve and circles are the data of participant WM. The SSE (which is based upon the residuals for the model tested) is smaller for participant EA (SSE=.02) than for participant WM (SSE =.16), indicating a better fit for participant EA. However, the R^2 measure indicates a superior fit for participant WM (R^2=.79) relative to participant EA (R^2=.36). The discrepancy results from the fact that the error associated with the mean of the data (SSE_{mean} in Equation 2) is greater for the higher-rate (i.e., more impulsive) discounting curve, simply because the indifference points span a greater range of values across delays.

Unlike R^2, which confounds model fit with discounting rate, goodness-of-fit measures such as SSE and the root mean squared error (RMSE) are not influenced by a comparison to the mean of the data (i.e., the ratio in Equation 2). Such measures have occasionally been used to quantify model fit in the delay-discounting literature (two instances are Kirby & Santiesteban, 2003, using RMSE, and Madden et al., 1997, using SSE). In addition, a variety of other approaches have been recommended as methods for evaluating nonlinear regression models based on model generalizability rather than goodness-of-fit, such as the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) (Burnham & Anderson, 2002; Forster, 2000; Pitt & Myung, 2002; Zucchini, 2000). Perhaps the best method for evaluating models is to empirically test divergent predictions of competing models (Roberts & Pashler, 2000; Shull, 1991), a strategy utilized in a few analyses of competing discounting models (Ainslie, 1975, Mazur, 1987, Green, Fristoe & Myerson 1994). While there are advantages to these methods as well as goodness-of-fit methods (e.g., Forster, 2000; Pitt & Myung, 2002; Roberts & Pashler, 2000; Zucchini, 2000), they are all intended to assess the appropriateness of a given mathematical model (or compare competing models) for describing empirical data. The use of such methods for identifying data as nonsystematic (as has been done with R^2 in several

discounting studies) turns this situation on its head. Instead of using the data to determine the appropriate model, this approach assumes the given model (e.g., Equation 1) is a perfect standard by which the validity of empirical data is to be judged. Because support exists for multiple models of discounting (e.g., Green, Fry & Myerson, 1994; Loewenstein & Prelec, 1992; Mazur, 1987; Myerson & Green, 1995; Rachlin, 1989), using the fit or generalizability index of any single model to assess the validity of data may be problematic. While the use of measures such as SSE or RMSE for identifying nonsystematic data would be superior to R^2 because they are not typically correlated with discounting rate, their use nonetheless unnecessarily confounds the issue of discounting rate comparison with the issue of model assessment. A preferable method for identifying nonsystematic discounting data would make no assumption about the optimum mathematical model to be employed, and would be based instead on basic expectations of the data that may relate to their validity.

Algorithm for Identifying Nonsystematic Data

An algorithm was developed to identify nonsystematic discounting data. This algorithm was intended to identify instances in which indifference points were not monotonically decreasing with delay. Specifically, data were identified as nonsystematic if either or both of the two following criterion were met: 1) if any indifference point (starting with the second delay) was greater than the preceding indifference point by a magnitude greater than 20% of the larger later reward (i.e., $200); 2) if the last (i.e., 25-year) indifference point was not less than the first (1 day or 1 week, depending on the study) indifference point by at least a magnitude equal to 10% of the larger later reward (i.e., $100). Instances of either criterion call into question the validity of results. Violation of criterion 1 suggests that further delay causes reward value to *increase* rather than decrease. We allowed increases up to 20% to avoid being overly stringent, and to allow some variability in the data. Violation of criterion 2 in the data reanalyzed here indicates that an individual does not discount by at least 10% in 25 years (e.g., would prefer to wait 25 years to receive $1000 rather than receiving $900 now). Such data suggest that delay had no effect on reward value. The specific parameters utilized (i.e., 20% for criterion 1 and 10% for criterion 2) were, and selected after applying several combinations of parameters and judging the resulting excluded and included cases by visual inspection.

Four data sets out of the 161 total examined were identified as nonsystematic due to criterion 1. That is, at least one indifference point (starting with the second delay) was greater than the preceding indifference point by a magnitude greater than 20% of the larger later reward (i.e., $200). These four data sets are shown in Figure 2. Data sets are labeled by study and participant identifier. For each data set in Figure 2 and in subsequent figures, the data are shown in two ways. The left panel under each participant identifier shows the seven indifference points equidistantly along the x-axis, so that each indifference point may be clearly seen. The right panel under each participant identifier shows the seven indifference points proportionally according to their delay, along with the best-fitting hyperbolic function (Equation 1) through the points. In each of these data sets, at least one indifference point suggests a departure from a monotonically decreasing function.

Nine data sets out of the 161 total examined were identified as nonsystematic due to criterion 2. That is, the last (i.e., 25-year) indifference point was not less than the first (1-day or 1-week) indifference point by at least a magnitude equal to 10% of the larger later reward (i.e., $100). These nine data sets are shown in Figure 3. Inspection of the data sets in Figure 3 suggests that these participants were not sensitive to delay. That is, even at delays ranging from 1 day (or 1 week) to 25 years, the independent variable (delay) had no demonstrated influence on the dependent variable (indifference point). One participant (Baker Smoker 78) discounted the hypothetical $1000 reward almost completely at all delays studied, while the other eight participants showed no discounting of the hypothetical $1000 reward even at 25 years.

In order to highlight the difference between using R^2 and the proposed algorithm as a method for identifying nonsystematic discounting data, Figure 4 shows the six data sets with the lowest R^2 values (ranging from −3.59 to 0.26) out of the 148 data sets not identified as suspect by the algorithm. If a study were to use R^2 as a basis for identifying nonsystematic data, some or all of these data sets would have been eliminated from analyses (depending on the threshold used to identify a low R^2 value). However, none of these data sets were identified as nonsystematic by the algorithm. Because all six cases show evidence of a monotonically decreasing discounting function, regardless of their fit to any mathematic model, we believe eliminating these data would be inappropriate.

In order to show other instances of data (other than those shown in Figure 4) that were not identified as nonsystematic, four data sets from each of the six groups were randomly sampled from the 142 data sets not represented in Figures 2 – 4. These 12 data sets are shown in Figure 5. Comparing the data sets presented in Figures 2 and 3 to the data sets in Figure 5 provides a basis for visually comparing those data that are and are not identified by the algorithm as nonsystematic. This visual comparison supports the notion that data identified as nonsystematic (Figures 2 and 3) are more questionable in their validity than the data not identified as nonsystematic (Figure 5).

Discussion

Delay-discounting data from three previously published studies, which included six distinct groups of both drug-dependent and control individuals, were used to demonstrate that the R^2 measure of fit is positively correlated with discounting rate, with R^2 being more stringent on low-rate discounting data and more lenient on high-rate discounting data. It was shown that data sets with low R^2 values often appear valid by visual inspection (Fig. 4). An algorithm for identifying nonsystematic discounting data was introduced that is based on the expectation of a monotonically decreasing form for discounting data, and that is unrelated to the fit of the data to any particular mathematical discounting model. This algorithm can easily be implemented as a formula in a spreadsheet containing indifference point data.

Data from the three previously published studies were used to test the utility of the algorithm. We believe this algorithm categorized data sets appropriately because visual inspection of the data sets identified as nonsystematic by the algorithm demonstrated deviations from a monotonically decreasing discounting function. Data sets that were identified by criterion 1 of the algorithm (Fig. 2) showed at least one data point that deviated from a monotonic function, suggesting that for at least one point, further delay caused reward value to *increase* rather than decrease. Such data may suggest that the participant did not understand the task, that experimental conditions did not encourage careful responding (e.g., resulted in inattentiveness), or that data were based on idiosyncratic rule-governed behavior (i.e., the assignment of differential importance to one delay, for example, when rent is due) rather than reflecting a more fundamental discounting rate.

Data sets that were identified by criterion 2 of the algorithm (Fig. 3) indicated that delay had no effect on reward value, calling into question the face validity of results. It is difficult to believe that participant Baker Smoker 78 would truly prefer $100 now rather than waiting even a single day for $1000, and that the other eight individuals would truly prefer to wait 25 years for $1000 rather than receiving $900 now. It is possible that these individuals are extremely high (Baker Smoker 78) and extremely low (the other eight participants) discounters. However, it is the *inability* to distinguish nonsystematic data from high or low discounting that makes these data suspect. Such data may suggest that the participant did not understand the task, or that experimental conditions did not encourage careful responding (e.g., exclusive responding for one reward in order to finish the task easily). Furthermore, the observation that 8 out of 9

participants identified as nonsystematic by criterion 2 showed exclusive preference for delayed rewards may suggest that these participants were influenced by demand characteristics of the experiment. That is, selecting delayed rewards may be viewed as socially desirable, perhaps implying greater intelligence or responsibility. In this respect, criterion 2 may serve a function similar to scales within personality assessment instruments designed to detect deception or socially desirable responding (e.g., Nichols & Greene, 1997;Peebles & Moore, 1998).

As noted above, the specific parameters utilized in the algorithm (i.e., 20% for criterion 1 and 10% for criterion 2) were selected after applying several combinations of parameters to the 161 delay-discounting data sets analyzed here. Other parameters, or the utilization of only one of the criteria, may be appropriate for other experiments. For example, criterion 1 may be adjusted to allow less or more variability in the data. A threshold of 10% would allow for less variability and identify a greater number of data sets as nonsystematic. A threshold of 30% would allow for more variability and identify fewer data sets as nonsystematic. Another way to allow for more variability in the data would be to allow a single individual indifference point to deviate from a monotonically decreasing function. Indeed, inspection of Figure 2 suggests that in each of the four data sets identified as nonsystematic by criterion 1, only a single indifference point was deviant. If one wished to accept such data sets as systematic, while identifying as nonsystematic instances in which multiple indifference points were deviant, a two-step process may be employed. First, apply criterion 1 as described thus far. Then, identify and eliminate the first single deviant indifference point, and reapply criterion 1 to the remaining indifference points. Only those data sets that are still flagged by the criterion would be identified as nonsystematic. In data sets with only a single deviant indifference point, investigators may consider removing this point before analysis with nonlinear regression.

Criterion 2 may also be adjusted. While the failure to demonstrate at least minimal discounting of money may seem incredulous if the longest delay in the experiment is 25 years, minimal discounting may seem believable if discounting is only assessed at much shorter delays. As an example, in a human discounting procedure using delays up to 90 seconds, Lane et al. (2003) found several instances of individuals who showed no discounting (exclusive preference for the larger delayed reward). These data were likely valid because of the short delays, and because trial-by-trial consequences were used. Therefore, depending on the delays utilized and the nature of the study, it may be considered valid for an individual to discount by less than 10% from the first to last time point, and criterion 2 may be eliminated. If criterion 1 is eliminated and low rates of discounting are accepted as valid, one may encounter difficulty in computing a discounting parameter estimate (k value) because nonlinear regression will not be able to converge when indifference points are equivalent across delays. If all indifference points are 100% of the delayed reward value (i.e., exclusive preference for the larger later option), an algebraic solution may be used, which will result in $k = 0$. If indifference points are equivalent across delays but are less than 100% of the delayed reward value, the final indifference point value may be reduced by a trivial amount (e.g., by one ten-thousandth), which will allow the nonlinear regression to converge on a k value.

The proposed algorithm may be useful to investigators in at least two respects. First, by identifying cases of nonsystematic data, investigators may manipulate variables such as discounting assessment procedures, environment, subject condition, and subject characteristics in order to investigate circumstances under which more systematic data are generated. For example, the algorithm may be a useful metric to include as a dependent variable in future studies similar to those that have focused on the development or comparison of delay-discounting procedures (e.g., Epstein et al., 2003; Kowal et al., 2007; Lane et al., 2003; Reynolds & Schiffbauer, 2004; Robles & Vargas, 2007). The results of such investigations may improve research methodology in the study of discounting processes.

Second, as an alternative to R^2, the algorithm, or a variation of it, may improve an experiment's power to detect discounting effects. As suggested by Fig. 4, when using R^2 to identify and eliminate nonsystematic data, several apparently orderly data sets are likely to be eliminated because their R^2 value is below an arbitrary elimination threshold (e.g., R^2 of 0, .3., .4 or .5). Some studies have eliminated large portions of their collected data by using such a threshold. To the degree that the proposed algorithm more accurately detects instances of nonsystematic data without false identifications, using the algorithm instead of R^2 to identify and eliminate nonsystematic data before group or condition comparisons may yield greater experimental power from a given number of subjects. This use of the algorithm is essentially a method of eliminating outliers from discounting data. An outlier can be viewed as an experimental result that is caused by processes other than the process being investigated, and it is generally recommended that outliers be removed from data analysis (Ratcliff, 1993). In delay discounting studies, an outlier may be conceptualized as a set of indifference points generated by processes other than delaying the consequence. As mentioned above, such processes may include inattentiveness, attempts to easily end the task, rule-governed behavior, or demand characteristics of the experiment.

One alternative to using R^2 or the proposed algorithm to identify and eliminate data is to include all data for analyses and not use any identification method. This may be appropriate for many studies in which very systematic data are generated. This was the approach used in the original published analyses for the three studies reanalyzed here (Baker et al., 2003; Bickel et al., 1999; Madden et al., 1997). It should be noted that for the purposes of this paper the drug-dependent and control groups within each of these three original studies were statistically compared (t-tests of log transformed *k* values) after eliminating data identified as nonsystematic by the algorithm. Results were consistent with original findings. That is, the drug dependent group showed statistically greater discounting than the control group in each of the three comparisons. However, other studies may use methods or subjects that may result in a greater number of instances of spurious data than the studies reanalyzed here. In these situations, application of the proposed algorithm, and elimination of the identified nonsystematic data, may allow investigators to find statistically significant effects where no significant effects would be found with the entire data set.

Although the data reanalyzed here were delay-discounting data, the proposed algorithm is applicable to other experimental forms of discounting assessment, such as probability discounting (Rachlin et al., 1991), social discounting (Jones and Rachlin, 2006), and effort discounting (Mitchell, 2004). In addition, although the data reanalyzed here were human data, the algorithm may be useful in nonhuman animal studies of discounting processes. Like human studies, nonhuman animal studies may use the algorithm as a metric to identify experimental conditions under which more orderly data may be generated, and as an alternative to R^2 for an objective method to identify outliers to exclude from analyses. Some of the hypotheses mentioned above as to why a subject's data might be nonsystematic (e.g., not understanding the task) may not apply to nonhuman animal studies. However, nonsystematic data identified by the algorithm may suggest that the experimental conditions need to be altered (e.g., additional training sessions to achieve stability). Because nonhuman animal studies run many sessions in order to achieve stability, a variation of the algorithm may also be incorporated as a metric to detect if stability has been achieved (cf., stability criterion in Winstanley, Dalley, Theobald, & Robbins, 2003).

In summary, the proposed algorithm or variations of it may be useful in a wide variety of human and nonhuman animal studies investigating any of several forms of discounting. Studies that result in a substantial number of questionable discounting data sets may use the algorithm to identify outliers to be considered for exclusion from further analyses. More importantly, the algorithm may be used as a metric in studies that investigate discounting procedures, which

may improve discounting methodology, and in turn increase our understanding of discounting process and their role in a variety of other scientific domains.

Acknowledgments

This research was supported by National Institute on Drug Abuse Grants R01 DA 11692 and T32 DA07242.

References

Acheson A, Richards JB, de Wit H. Effects of sleep deprivation on impulsive behaviors in men and women. Physiology & Behavior 2007;91:579–587. [PubMed: 17477941]

Ainslie G. Specious reward: A behavioral theory of impulsiveness and impulse control. Psychological Bulletin 1975;82:463–496. [PubMed: 1099599]

Ainslie, G. Breakdown of Will. New York: Cambridge University Press; 2001.

Alessi SM, Petry NM. Pathological gambling severity is associated with impulsivity in a delay discounting procedure. Behavioural Processes 2003;64:345–354. [PubMed: 14580703]

Baker F, Johnson MW, Bickel WK. Delay discounting in current and never-before cigarette smokers: Similarities and differences across commodity, sign, and magnitude. Journal of Abnormal Psychology 2003;112:382–392. [PubMed: 12943017]

Barkley RA, Edwards G, Laneri M, Fletcher K, Metevia L. Executive functioning, temporal discounting, and sense of time in adolescents with attention deficit hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). Journal of Abnormal Child Psychology 2001;29:541–556. [PubMed: 11761287]

Benzion U, Rapoport A, Yagil J. Discount rates inferred from decisions: An experimental study. Management Science 1989;35:270–284.

Bickel, WK.; Johnson, MW. Delay discounting: A fundamental behavioral process of drug dependence. In: Loewenstein, G.; Read, D.; Baumeister, RF., editors. Time and Decision. New York: Russell Sage Foundation; 2003.

Bickel WK, Odum AL, Madden GL. Impulsivity and cigarette smoking: Delay discounting in current, never, and ex-smokers. Psychopharmacology 1999;146:447–454. [PubMed: 10550495]

Burnham, KP.; Anderson, DR. Model Selection and Multimodel Inference. Vol. 2. New York: Springer-Verlag; 2002.

Chung SH, Herrnstain RJ. Choice and delay of reinforcement. Journal of the Experimental Analysis of Behavior 1967;10:67–74. [PubMed: 16811307]

Crean JP, de Wit H, Richards JB. Reward discounting as a measure of impulsive behavior in a psychiatric outpatient population. Experimental and Clinical Psychopharmacology 2000;8:155–162. [PubMed: 10843297]

Critchfield TS, Atteberry T. Temporal discounting predicts individual competitive success in a human analogue of group foraging. Behavioral Processes 2003;64:315–331.

Critchfield TS, Kollins SH. Temporal discounting: Basic research and the analysis of socially important behavior. Journal of Applied Behavior Analysis 2001;34:101–122. [PubMed: 11317983]

Dallery J, Raiff BR. Delay discounting predicts cigarette smoking in a laboratory model of abstinence reinforcement. Psychopharmacology 2007;190:485–496. [PubMed: 17205320]

Deluty MZ. Self-control and impulsiveness involving aversive events. Journal of Experimental Psychology: Animal Behavior Processes 1978;4:250–266.

Diergaarde L, Pattij T, Poortvliet I, Hogenboom F, de Vries W, Schoffelmeer AN, De Vries TJ. Impulsive choice and impulsive action predict vulnerability to distinct stages of nicotine seeking in rats. Biological Psychiatry 2008;63:301–308. [PubMed: 17884016]

Du W, Green L, Myerson J. Cross-cultural comparisons of discounting delayed and probabilistic rewards. Psychological Record 2002;52:479–492.

Epstein LH, Richards JB, Saad FG, Paluch RA, Roemmich JN, Lerman C. Comparison between two measures of delay discounting in smokers. Experimental and Clinical Psychopharmacology 2003;11:131–138. [PubMed: 12755457]

Forster MR. Key concepts in model selection: performance and generalizability. Journal of Mathematical Psychology 2000;44:205–231. [PubMed: 10733865]

Frederick S, Loewenstein G, O'Donoghue. Time discounting and time preference: a critical review. Journal of Economic Literature 2002;40:351–401.

Green L, Fisher EB Jr, Perlow S, Sherman L. Preference reversal and self-control: Choice as a function of reward amount and delay. Behaviour Analysis Letters 1981;1:43–51.

Green L, Fristoe N, Myerson J. Temporal discounting of preference reversals in choice between delayed outcomes. Psychonomic Bulletin & Review 1994;1:383–389.

Green L, Fry AF, Myerson J. Discounting of delayed rewards: A life-span comparison. Psychological Science 1994;5:33–36.

Green L, Myerson J. Exponential versus hyperbolic discounting of delayed outcomes: risk and waiting time. American Zoologist 1996;36:496–505.

Green L, Myerson J, Lichtman D, Rosen S, Fry A. Temporal discounting in choice between delayed rewards: The role of age and income. Psychology & Aging 1996;11:79–84. [PubMed: 8726373]

Green L, Myerson J, McFadden E. Rate of temporal discounting decreases with amount of reward. Memory & Cognition 1997;25:715–723.

Green L, Myerson J, Ostaszewski P. Discounting of delayed rewards across the life span: Age differences in individual discounting functions. Behavioural Processes 1999;46:89–96.

Hamidovic A, Kang UJ, de Wit H. Effects of low to moderate acute doses of pramipexole on impulsivity and cognition in healthy volunteers. Journal of Clinical Psychopharmacology 2008;28:45–52. [PubMed: 18204340]

Johnson MW, Bickel WK. Within-subject comparison of real and hypothetical money rewards in delay discounting. Journal of the Experimental Analysis of Behavior 2002;77:129–146. [PubMed: 11936247]

Jones B, Rachlin H. Social discounting. Psychological Science 2006;17:283–286. [PubMed: 16623683]

Kirby KN, Santiesteban M. Concave utility, transaction costs, and risk in measuring discounting of delayed rewards. Journal of Experimental Psychology: Learning, Memory, & Cognition 2003;29:66–78.

Kowal BP, Yi R, Erisman AC, Bickel WK. A comparison of two algorithms in computerized temporal discounting procedures. Behavioural Processes 2007;75:231–236. [PubMed: 17368965]

Krishnan-Sarin S, Reynolds B, Duhig AM, Smith A, Liss T, McFetridge A, Cavallo DA, Carroll KM, Potenza MN. Behavioral impulsivity predicts treatment outcome in a smoking cessation program for adolescent smokers. Drug and Alcohol Dependence 2007;88:79–82. [PubMed: 17049754]

Lagorio CH, Madden GJ. Delay discounting of real and hypothetical rewards III: steady-state assessments, forced-choice trials, and all real rewards. Behavioural Processes 2005;69:173–87. [PubMed: 15845306]

Lane SD, Cherek DR, Pietras CJ, Tcheremissine OV. Measurement of delay discounting using trial-by-trial consequences. Behavioural Processes 2003;64:287–303. [PubMed: 14580699]

Loewenstein GF. Frames of mind in intertemporal choice. Management Science 1988;34:200–214.

Loewenstein, G.; Prelec, D. Anomalies in interpersonal choice: Evidence and an interpretation. In: Loewenstein, G.; Elster, J., editors. Choice over time. New York: Russell Sage Foundation; 1992. p. 119-145.

Logue AW. Research on self-control: An integrating framework. Behavioral and Brain Sciences 1988;11:665–709.

Madden GJ, Begotka AM, Raiff BR, Kastern LL. Delay discounting of real and hypothetical rewards. Experimental and Clinical Psychopharmacology 2003;11:139–145. [PubMed: 12755458]

Madden GJ, Petry N, Badger GJ, Bickel WK. Impulsive and self-control choices in opioid-dependent subjects and non-drug using controls: Drug and monetary rewards. Experimental and Clinical Psychopharmacology 1997;5:256–262. [PubMed: 9260073]

Madden GJ, Raiff BR, Lagorio CH, Begotka AM, Mueller AM, Hehli DJ, Wegener AA. Delay discounting of potentially real and hypothetical rewards: II. Between- and within-subject comparisons. Experimental and Clinical Psychopharmacology 2004;12:251–61. [PubMed: 15571442]

Mazur, JE. An adjusting procedure for studying delayed reinforcement. In: Commons, ML.; Mazur, JE.; Nevin, JA.; Rachlin, H., editors. Quantitative Analysis of Behavior: Vol. 5. The Effect of Delay and of Intervening Events on Reinforcement Value. Hillsdale, NJ: Erlbaum; 1987. p. 55-73.

McDonald J, Schleifer L, Richards JB, de Wit H. Effects of THC on behavioral measures of impulsivity in humans. Neuropsychopharmacology 2003;28:1356–1365. [PubMed: 12784123]

Mischel W, Shoda Y, Rodriguez ML. Delay of gratification in children. Science 1989;244:933–938. [PubMed: 2658056]

Mitchell SH. Effects of short-term nicotine deprivation on decision-making: delay, uncertainty and effort discounting. Nicotine & Tobacco Research 2004;6:819–828. [PubMed: 15700917]

Motulsky, HJ.; Christopoulos, A. A Practical Guide to Curve Fitting. Graphpad Software, Inc; San Diego, CA: 2003. Fitting Models to Biological Data Using Linear and Nonlinear Regression.

Myerson J, Green L. Discounting of delayed rewards: Models of individual choice. Journal of the Experimental Analysis of Behavior 1995;64:263–276. [PubMed: 16812772]

Nichols DS, Greene RL. Dimensions of deception in personality assessment: the example of the MMPI-2. Journal of Personality Assessment 1997;68:251–266. [PubMed: 16370780]

Ohmura Y, Takahashi T, Kitamura N, Wehr P. Three-month stability of delay and probability discounting measures. Experimental and Clinical Psychopharmacology 2006;14:318–328. [PubMed: 16893275]

Peebles J, Moore RJ. Detecting socially desirable responding with the Personality Assessment Inventory: the Positive Impression Management scale and the Defensiveness Index. Journal of Clinical Psychology 1998;54:621–628. [PubMed: 9696112]

Perry JL, Larson EB, German JP, Madden GJ, Carroll ME. Impulsivity (delay discounting) as a predictor of acquisition of IV cocaine self-administration in female rats. Psychopharmacology 2005;178:193–201. [PubMed: 15338104]

Pitt MA, Myung IJ. When a good fit can be bad. Trends in Cognitive Sciences 2002;6:421–425. [PubMed: 12413575]

Rachlin, H. Judgement, Decision, and Choice: A Cognitive/Behavioral Synthesis. New York: W. H. Freeman & Co; 1989.

Rachlin H, Green L. Commitment, choice, and self-control. Journal of the Experimental Analysis of Behavior 1972;17:15–22. [PubMed: 16811561]

Rachlin H, Raineri A, Cross D. Subjective probability and delay. Journal of the Experimental Analysis of Behavior 1991;55:233–244. [PubMed: 2037827]

Ratcliff R. Methods for dealing with reaction time outliers. Psychological Bulletin 1993;114:510–532. [PubMed: 8272468]

Ratkowsky, DA. Handbook of Nonlinear Regression Models. Marcel Dekker, Inc; New York: 1990.

Reynolds B. A review of delay-discounting research with humans: relations to drug use and gambling. Behavioural Pharmacology 2006;17:651–667. [PubMed: 17110792]

Reynolds B, Richards JB, de Wit H. Acute-alcohol effects on the Experiential Discounting Task (EDT) and a question-based measure of delay discounting. Pharmacology, Biochemistry and Behavior 2006;83:194–202.

Reynolds B, Schiffbauer R. Measuring state changes in human delay discounting: an experiential discounting task. Behavioural Processes 2004;67:343–356. [PubMed: 15518985]

Richards JB, Zhang L, Mitchell SH, de Wit. Delay or probability discounting in a model of impulsive behavior: effect of alcohol. Journal of the Experimental Analysis of Behavior 1999;71:121–143. [PubMed: 10220927]

Roberts S, Pashler H. How persuasive is a good fit? A comment on theory testing. Psychological Review 2000;107:358–367. [PubMed: 10789200]

Robles E, Vargas PA. Functional parameters of delay discounting assessment tasks: order of presentation. Behavioural Processes 2007;75:237–241. [PubMed: 17368757]

Rounds JS, Beck JG, Grant DM. Is the delay discounting paradigm useful in understanding social anxiety? Behaviour Research and Therapy 2007;45:729–735. [PubMed: 16890909]

Shelley MK. Outcome signs, question frames, and discount rates. Management Science 1993;39:805–815.

Shull. Mathematical descriptions of operant behavior: an introduction. In: Iverson, IH.; Lattal, KA., editors. Experimental Analysis of Behavior. Vol. 2. New York: Elsevier; 1991. p. 243-282.

Simpson CA, Vuchinich RE. Reliability of a measure of temporal discounting. The Psychological Record 2000;50:3–16.

Thaler R. Some empirical evidence on dynamic inconsistency. Economic Letters 1981;8:201–207.

Winstanley CA, Dalley JW, Theobald DE, Robbins TW. Global 5-HT depletion attenuates the ability of amphetamine to decrease impulsive choice on a delay-discounting task in rats. Psychopharmacology 2003;170:320–331. [PubMed: 12955303]

Yoon JH, Higgins ST, Heil SH, Sugarbaker RJ, Thomas CS, Badger GJ. Delay discounting predicts postpartum relapse to cigarette smoking among pregnant women. Experimental and Clinical Psychopharmacology 2007;15:176–186. [PubMed: 17469941]

Zucchini W. An introduction to model selection. Journal of Mathematical Psychology 2000;44:41–61. [PubMed: 10733857]

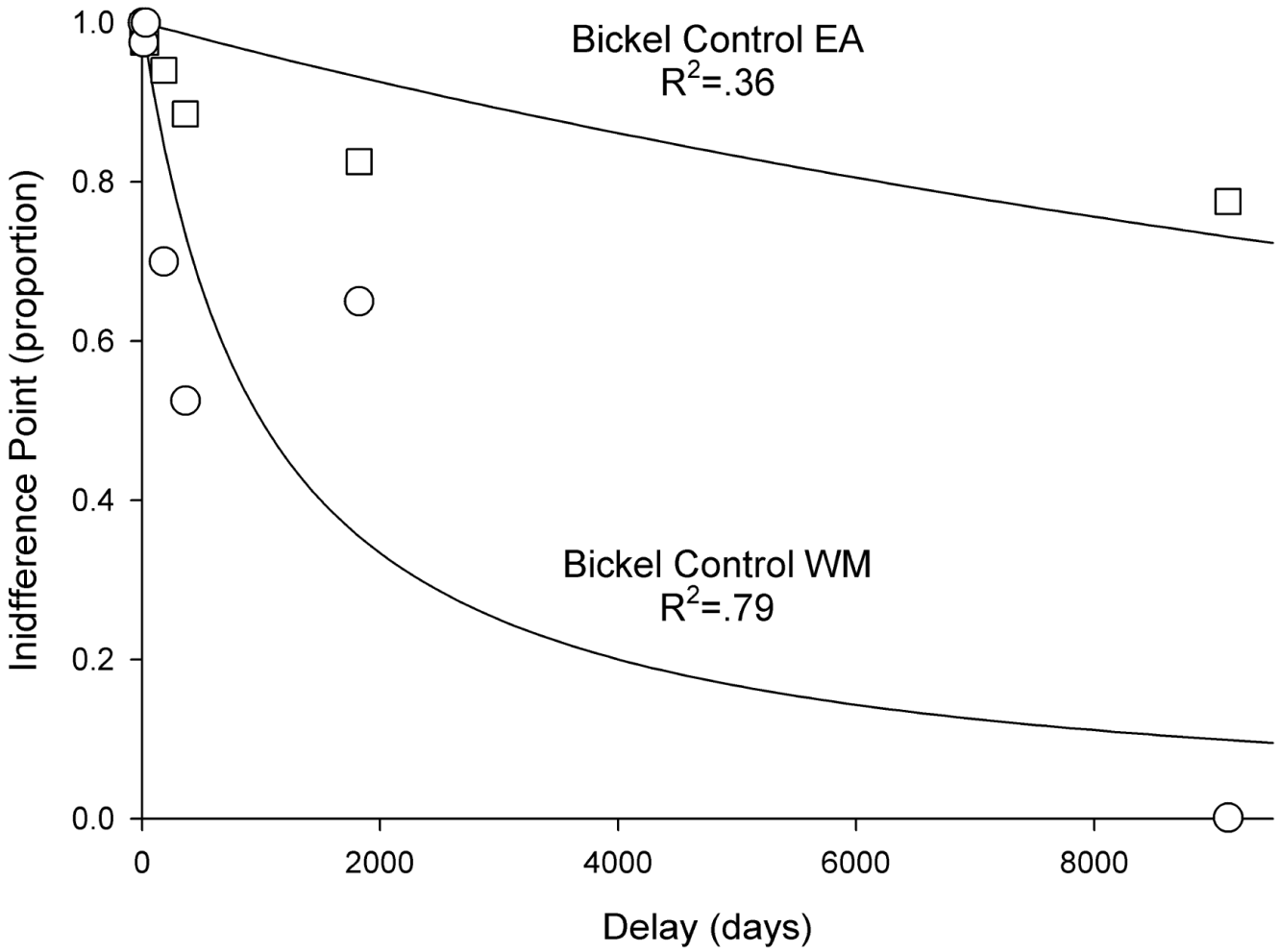

Figure 1.
Indifference points, best-fitting hyperbolic functions, and R^2 values for two individual participants from the Bickel et al. (1999) control group.

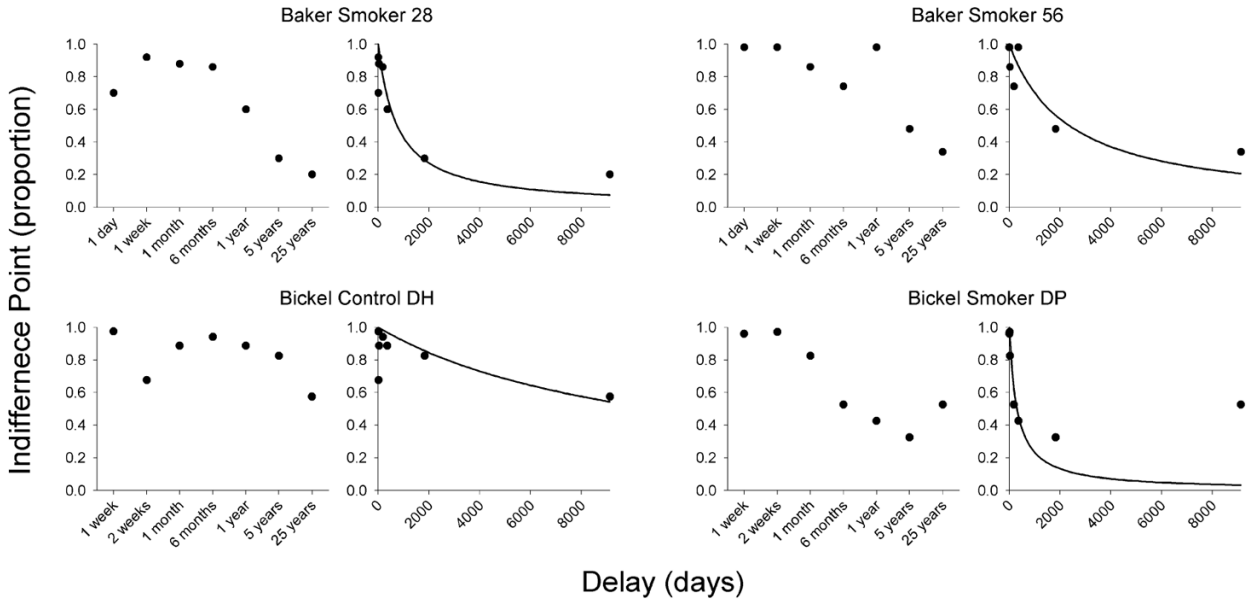

Figure 2.

Four discounting data sets out of the 161 total examined that were identified as nonsystematic due to any indifference point (starting with the first delay) being greater than the preceding indifference point by a magnitude greater than 20% of the larger later reward (i.e., $200). The left panel under each participant identifier shows the seven indifference points equidistantly along the x-axis, and delay units are shown on the x-axis. The right panel under each participant identifier shows the seven indifference points proportionally according to their delay (shown in units of days), along with the best-fitting hyperbolic function (Equation 1) though the points.

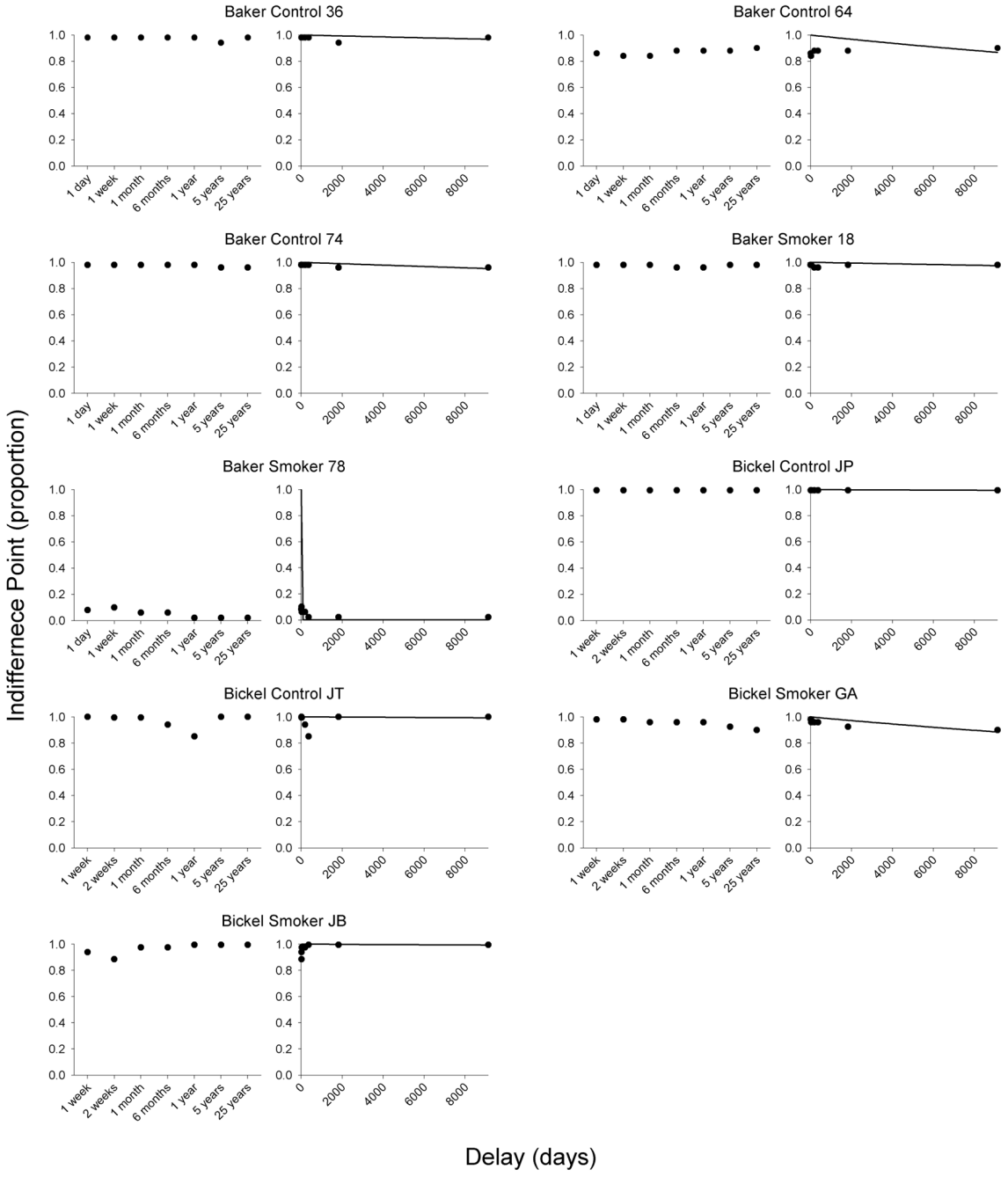

Figure 3.

Nine discounting data sets out of the 161 total examined that were identified as nonsystematic due to the last indifference point (25-year) not being less than the first indifference point (1-day) by at least a magnitude equal to 10% of the larger later reward (e.g., $100). Other details as in Fig. 2.

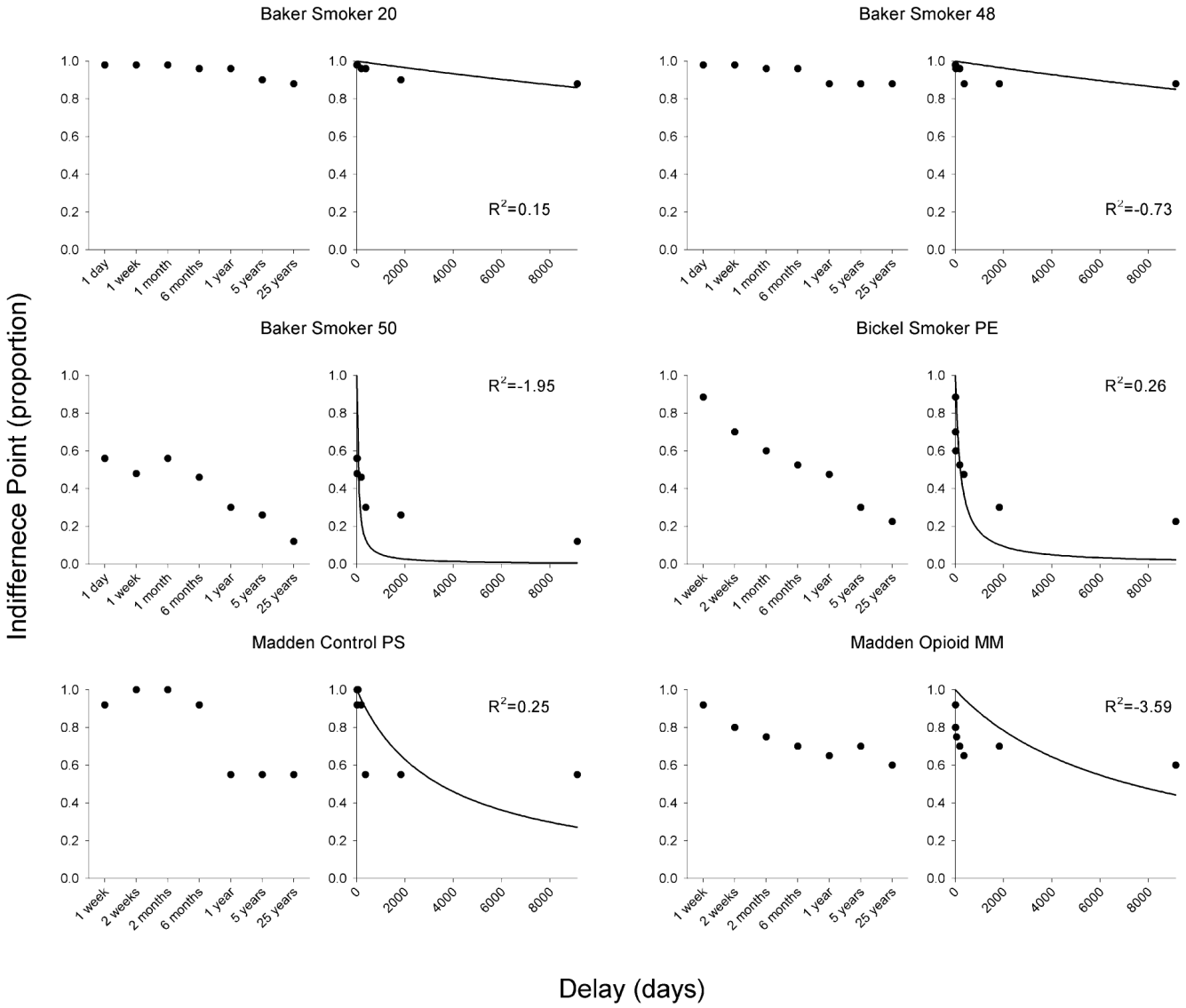

Figure 4.

The six data sets with the lowest R^2 values (ranging from −3.59 to 0.26) out of the 148 data sets not identified as nonsystematic by the algorithm. Other details as in Fig. 2.

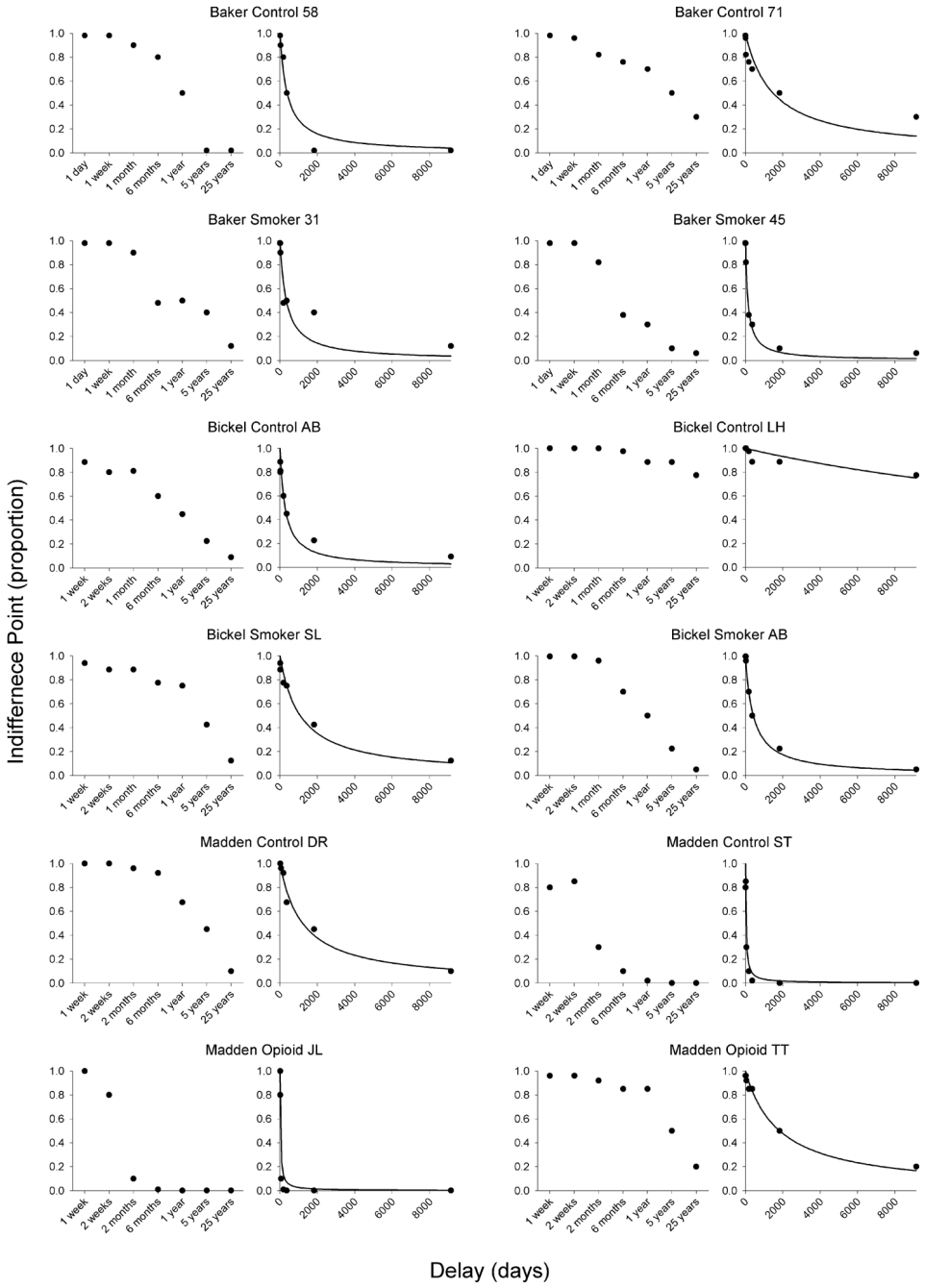

Figure 5.
Four data sets from each of the six groups randomly sampled from the 142 data sets not represented in Figures 1 – 3. (i.e., not identified as nonsystematic by the algorithm, and not among the 6 lowest R^2 data sets in Fig. 3). Other details as in Fig. 2.

Table 1

Methods used to collect delay discounting data in original studies.

Study	Choice Presentation	Method of titrating immediate reward value across trials within each delay assessment	Delays
Baker et al. (2003)	Computer	Pseudorandom selection from within limits that converged on an indifference point via algorithm	1 day, 1 week, 1 & 6 months, 1 & 5 & 25 years
Bickel et al. (1999)	Index cards	Ascending and descending titration performed separately, and the two resulting indifference points were averaged	1 & 2 weeks, 1 & 6 months, 1 & 5 & 25 years
Madden et al. (1997)	Index cards	Ascending and descending titration performed separately, and the two resulting indifference points were averaged	1 & 2 weeks, 2 & 6 months, 1, 5 & 25 years

Table 2

Spearman correlations between discounting parameter *k* and R^2 for each of the six groups.

Group	n	Spearman's ρ	p value
Baker Controls	30	0.559	.001
Baker Smokers	30	0.261	.164
Bickel Controls	22	0.570	.006
Bickel Smokers	23	0.505	.014
Madden Controls	38	0.385	.017
Madden Opioid	18	0.051	.842