ARTICLE IN PRESS

INTELL-00420; No of Pages 17

Available online at www.sciencedirect.com

ScienceDirect

Intelligence xx (2007) xxx–xxx

INTELLIGENCE

Delay discounting and intelligence: A meta-analysis

Noah A. Shamosh *, Jeremy R. Gray

Department of Psychology, Yale University, United States

Received 21 March 2007; received in revised form 31 August 2007; accepted 22 September 2007

Abstract

Delay discounting (DD), the tendency to prefer smaller, sooner rewards to larger, later ones, is an important indicator of self-control. Assessments of DD superficially require individuals to make choices based on motivational processes. However, several lines of evidence suggest that DD may be systematically related to cognitive ability. We sought to provide a definitive assessment of the relation between DD and intelligence via quantitative research synthesis. A comprehensive literature search in two electronic databases yielded 24 eligible studies with 26 effect sizes in total. Meta-analysis revealed that, across studies, higher intelligence was associated with lower DD (random effects model weighted mean $r = -0.23$). Studies using reward schemes in which payoffs were subject to chance (i.e., involving either a chance of receiving one choice or random selection of one choice) showed weaker associations between DD and intelligence than did studies in which payoffs were all hypothetical or all real. Other moderator analyses revealed no influence of DD measure, DD choice paradigm, or intelligence type. There was no evidence of publication bias. Given clear evidence for a negative relation between DD and intelligence, investigating the processes that support or moderate this relation would be worthwhile.

Keywords: Self-control; Delay discounting; Temporal discounting; Inter-temporal choice; Impulsivity; Impulsiveness; Intelligence; Cognitive ability

Contents

* Corresponding author. Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205, United States. Tel.: +1 203 535 7018.

E-mail address: noah.shamosh@yale.edu (N.A. Shamosh).

0160-2896/$ - see front matter
doi:10.1016/j.intell.2007.09.004

Please cite this article as: Shamosh, N. A., & Gray, J. R., Delay discounting and intelligence: A meta-analysis, *Intelligence* (2007), doi:10.1016/j.intell.2007.09.004

1. Introduction

1.1. Background

Self-control plays a critical role in every sphere of our lives. We balance our diets to maintain our health, sacrifice for others to strengthen our relationships, and juggle work and play to succeed in our jobs. Self-control can be characterized as a person's tendency or ability to behave in a way that favors distal, more important goals over proximal, less important ones.

Several lines of research highlight the importance of having high self-control: It predicts positive outcomes across the board, from better academic performance and social adjustment to lower psychopathology and criminal behavior (Kirby, Winston & Santiesteban, 2005; Pratt & Cullen, 2000; Shoda, Mishel & Peake, 1990; Tangney, Baumeister & Boone, 2004). Intelligence is also a well-established predictor of various life outcomes (e.g., Jensen, 1998). Although recent investigations have focused on disentangling influences of self-control and intelligence on success (e.g., Duckworth & Seligman, 2005, 2006), it is also possible that the two constructs are related, having joint effects.

An important indicator of self-control is delay discounting (DD), which refers to an individual's tendency to prefer smaller, more immediate rewards to larger, more delayed ones.

DD tasks typically require an individual to make choices between two rewards that vary in both the amount and delay until receipt (e.g., Myerson & Green, 1995). For example, participants might respond to several questions of the following form: "Would you rather receive \$170 now or \$200 in six months?" with reward size or delay interval varying by question. In this example, individuals who tend to prefer the smaller of the two rewards are considered to demonstrate greater DD. Thus, higher DD reflects lower self-control or a preference for the immediate—the extent to which a more distant reward is valued less right now because of the delay associated with receiving it.

Assessments of DD are reliable and stable over time (Ohmura, Takahashi, Kitamura & Wehr, 2006), making them useful for studying self-control in the context of various individual differences. For instance, performance on DD tasks has been associated with age (Green, Myerson, Lichtman, Rosen & Fry, 1996), personality (Ostaszewski, 1996), substance abuse and addiction (Jaroni, Wright, Lerman & Epstein, 2004), and clinical disorders such as ADHD (Barkley, Edwards, Laneri, Fletcher & Metevia, 2001). The DD framework also lends theoretical insight into numerous behaviors relevant to mental and physical health (for reviews, see Bickel & Marsch, 2001; Critchfield & Kollins, 2001) as well as personal financial wellbeing (Angeletos, Laibson, Repetto, Tobacman & Weinberg, 2001).

Although DD relates to important behavioral phenomena, its association with intelligence, a critical dimension of human individual differences, remains unclear. Aside from a handful of studies, data that bear on this relation typically are buried within reports' secondary analyses. Moreover, as we will show later, studies that report these data have produced heterogeneous results. For example, Monterosso, Ehrman, Napier, O'Brien and Childress (2001) found virtually no association between DD and intelligence in cocaine-dependent individuals, but de Wit, Flory, Acheson, McCloskey and Manuck (2007) found a negative, moderately sized relationship between

Please cite this article as: Shamosh, N. A., & Gray, J. R., Delay discounting and intelligence: A meta-analysis, *Intelligence* (2007), doi:10.1016/j.intell.2007.09.004

preference for immediate rewards and IQ in middle-aged adults. Indeed, the relation between DD and intelligence could be very small or nonexistent. Perhaps performance on DD tasks primarily reflects motivational individual differences, such as conscientiousness or reward sensitivity. This possibility seems reasonable given that DD tasks measure preferences: There are no objectively correct responses in the same way as there are on intelligence tests.

Other work, however, provides numerous reasons to suspect that higher DD is associated with lower intelligence. For instance, some research has demonstrated a link between DD and working memory function. In three studies with both hypothetical and real rewards, Hinson, Jameson, and Whitney (2003) found, that imposing working memory load during a DD task increased participants' preference for smaller, sooner rewards over larger, later ones. This finding is consistent with a relation between DD and intelligence because working memory function and intelligence are highly and robustly correlated (albeit not isomorphic; for a meta-analysis and related discussion, see Ackerman, Beier & Boyle, 2005; Beier & Ackerman, 2005; Kane, Hambrock & Conway, 2005 and Oberauer, Schulze, Wilhelm & Süß, 2005; for a qualitative review, see Conway, Kane & Engle, 2003). Therefore, DD paradigms appear to require some of the specific abilities related to both working memory and intelligence, namely the active maintenance of goal-relevant information in the face of potentially distracting information (Engle, Tuholski, Laughlin & Conway, 1999; Kane & Engle, 2003), as well as the integration of complex or abstract information (e.g., Carpenter, Just, & Shell, 1990; Gottfredson, 1997).

Specifically, during a DD task, a participant must actively maintain representations of value for the rewards under consideration. These representations are subject to disruption by many possible concurrent processes that are relevant to deciding between the rewards, including affect management (e.g., controlling one's excitement over the prospect of receiving an immediate reward), strategy deployment (e.g., calculating the opportunity cost of not investing the immediately available reward), and both recall of previous choices and speculation about future choices (e.g., to achieve consistency or optimize a sequence of outcomes). Thus, participants in a DD paradigm have to manage and integrate considerable amounts of complex information, all while maintaining and updating representations of reward value in working memory. Insofar as more intelligent individuals are better at these types of processes, their behavior on DD tasks might be different than that of individuals with lower intelligence. Of course, it is also possible that people with better working memory (and higher intelligence) are just more consistent in their DD behavior, but do not discount more or less, because their representations of reward value are less susceptible to interference (cf. Franco-Watkins, Pashler & Rickard, 2006; but see Hinson & Whitney, 2006).

In a similar vein, DD may require intellectual abilities that assist in the interplay between more cognitive and more affective types of executive control. For example, Metcalfe and Mischel (1999) have characterized DD as involving a dynamic relationship between "hot" and "cool" systems, such that a balance in favor of the cool system over the hot will increase the tendency to delay rewards. There are numerous ways in which differences in intelligence can potentially tip this balance. For instance, children with higher intelligence tend to be better at shifting attention away from the affective properties of the rewards, which decreases hot system activation and, consequently, DD (Rodriguez, Mischel & Shoda, 1989). Another possibility is that individuals with higher intelligence are more adept at transforming reward representations to make them more abstract, which increases cool system activation and decreases DD in children (Mischel & Baker, 1975). Other work suggests that DD involves "hot executive function," a system chiefly involved in the regulation of affect and motivation (Hongwanishkul, Happaney, Lee & Zelazo, 2005). Hot executive function appears to be subserved by regions of ventromedial prefrontal cortex (e.g., Bechara, Damasio, Damasion & Anderson, 1994), and fluid intelligence (reasoning and novel problem-solving ability; "cool executive function") is typically associated with dorsolateral prefrontal cortex (Kane & Engle, 2002), which suggests that performing DD tasks involves mechanisms subserved by both brain regions. A recent neuroimaging study of DD in humans revealed patterns of brain activity consistent with this possibility (McClure, Laibson, Loewenstein & Cohen, 2004). Thus, a relation between intelligence and DD seems likely to the extent that DD involves executive functions, which are linked to intelligence through the functions of the prefrontal cortex (e.g., Kane & Engle, 2002; but see Friedman et al., 2006).

Additionally, DD paradigms often highlight inconsistencies in individuals' preferences that may relate to intelligence. For many types of DD tasks, participants' present subjective value for a future reward decreases with longer delays to that reward. The function that best describes this relationship is not, however, exponential (like that of compounding interest), but hyperbolic (e.g., Kirby & Marakovic, 1995; Myerson & Green, 1995). This hyperbolic function departs from normative economic theory, because it entails that individuals' preferences depend not only on the delay between two rewards, but also on the delay to the sooner reward. Specifically,

individuals who prefer a larger, later reward to a smaller, sooner one typically reverse their preference when the delay to both rewards decreases by the same amount (Kirby & Herrnstein, 1995). Individuals can discount more or less steeply (Myerson & Green, 1995), which, given hyperbolic discounting, can entail more or less consistency in preferences over time. Some evidence links higher intelligence to greater consistency in several types of decision making (Parker & Fischhoff, 2005). Therefore, to the extent that lower DD entails greater consistency in preferences, it may also be associated with higher intelligence.

Intelligence may also proxy for other individual differences that may influence DD. In particular, individuals with higher intelligence are more likely to adopt a deliberative, controlled mode of processing on certain tasks (Frederick, 2005). Some evidence suggests that such an approach may be associated with lower DD. For example, one forthcoming study found that DD decreased when participants were instructed to provide reasons for their choices (Benjamin, Brown, & Shapiro, 2006).

Finally, several studies suggest that changes in children's DD behavior track with the maturation of intellectual functioning. In particular, as children become older, their preference for larger, delayed rewards increases relative to smaller, immediate rewards (e.g., Green, Fry, Myerson, 1994; Krietler & Zigler, 1990). This evidence suggests, albeit indirectly, that higher intelligence is associated with lower DD.

To summarize, several lines of evidence suggest that a relation between DD and intelligence is plausible, but the DD literature itself has not produced clear evidence of a relation. The purpose of the present study is to use meta-analytic techniques to tease out this relation, if it exists: We aimed to ascertain the direction and magnitude of the hypothesized association between DD and intelligence and explore potential moderators of this relation. Doing so could help to constrain hypotheses about why humans differ in the extent to which they discount delayed rewards. One possibility is that the mean association between DD and intelligence is effectively zero. This result would imply that performance on DD tasks depends exclusively on motivational processes—that people behave on such tasks according to wants and needs that are impervious to differences in cognitive ability (i.e., "there's no accounting for taste"). In contrast, the relationship between DD and intelligence could be substantial. This case would suggest that, on DD tasks, differences in intelligence might influence how individuals appraise sooner and later rewards, or how they integrate appraisals to decide between the rewards.

1.2. Potential moderators

Assuming that DD and intelligence do demonstrate a reliable and nontrivial relation, investigating which variables might influence the magnitude of this association can potentially help reveal its underlying bases. Although numerous variables could moderate an association between DD and intelligence, we drew upon *a priori* hypotheses, available information, and statistical power considerations to focus on three. One variable that might influence the relation between DD and intelligence is the type of DD measure a study used. Measures of DD fall into two general categories. The first includes measures of time preference, such as the percentage of time a participant chooses an immediate reward. These measures capture an individual's overall tendency to prefer smaller, sooner rewards to larger, later ones. The second comprises measures of discount rates, which reflect how rapidly the present subjective value of a reward decreases as a function of delay. Estimation of discount rates typically assumes hyperbolic discounting, so in addition to reflecting greater sensitivity to delay, higher discount rates correspond to a greater likelihood of exhibiting preference reversals (e.g., Kirby & Herrnstein, 1995). Therefore, although both time preference and discount rate assess DD, they are not identical (see Myerson, Green & Warusawitharana, 2001). Specifically, a given level of time preference can correspond to a higher or lower discount rate depending on how the percentage of immediate rewards chosen varies with delay. Moreover, discount rates reflect not only preferences, but also the consistency of those preferences. The distinction between the two types of DD measures may potentially lead to divergent relations to intelligence. For instance, to the extent that more intelligent individuals are more consistent in decision making generally (Parker & Fischhoff, 2005), discount rates may be more highly correlated with intelligence than are measures of time preference.

A second potential moderator is the type of choice paradigm used in the DD task. As discussed earlier, DD tasks typically require individuals to make one or more forced choices between immediate and delayed rewards of lesser and greater value, respectively. Such tasks can be considered "commitment-choice" procedures, because each choice requires individuals to commit to one reward or the other (Reynolds & Schiffbauer, 2005). There is, however, another type of task that also can be considered a measure of DD. This type of task, often described as measuring delay of gratification, requires individuals to wait for a reward during a delay period while a smaller reward is constantly available. How long

the individual waits before consuming the smaller, immediate reward (or whether he or she persists to the end) is the dependent measure. These tasks can be considered "sustained-choice" procedures because the choice between the smaller, sooner and larger, later reward is available for the duration of the delay (Reynolds & Schiffbauer, 2005). Although some have argued that the additional requirement of coping with delay makes sustained-choice procedures distinct from commitment-choice tasks (e.g., Reynolds & Schiffbauer, 2005), both paradigms measure individuals' tendency to prefer smaller, sooner to larger, later rewards, and may represent the same construct (Rachlin, 2000). Thus, we included both types of studies, but coded commitment-choice versus sustained-choice as a dichotomous variable and investigated it as a potential moderator. Either choice procedure could plausibly yield a stronger relation between DD and intelligence than the other. Sustained-choice procedures, for example, might tap intelligence more because participants need to deploy coping strategies during the delay interval, which require operation of "cool" executive systems potentially related to intelligence (Hongwanishkul et al., 2005; Metcalfe & Mischel, 1999; Reynolds & Schiffbauer, 2005). Commitment-choice procedures, however, also have potentially unique requirements for intelligence—for instance, reasoning about opportunity costs of choosing the delayed reward. Of course, unique intelligence demands may balance out between the two choice paradigms, or both paradigms may require intelligence only for the processes that they share.

A third potential moderator that we examined was the type of intelligence that a study measured. In particular, studies measuring verbal intelligence may demonstrate stronger associations between DD and intelligence because verbal intelligence may influence the effectiveness of strategies that individuals use on some DD paradigms (e.g., Shoda et al., 1990). For example, individuals with higher verbal intelligence may be particularly adept at regulating stimulus-induced affect during a delay period through the use of cognitive reappraisal, an effective emotion regulation strategy that sometimes can be verbally intensive (Ochsner, Bunge, Gross & Gabrieli, 2002). It is also plausible that DD depends on more general cognitive abilities, which tend to be intercorrelated (Jensen, 1998), or other specific (nonverbal) abilities that comprehensive tests assess, such as quantitative reasoning. In these cases, studies using intelligence measures that capture both verbal and nonverbal abilities may show stronger effect sizes than those using measures of verbal intelligence exclusively.

2. Method

2.1. Inclusion/exclusion criteria

Only studies published in peer-reviewed journals were eligible for inclusion. Because the first major study of DD and intelligence appeared in 1962 (Mischel & Metzner, 1962), we limited our scope to studies published in 1962 and afterwards. We included a study in the meta-analysis if it reported an effect size (or enough data to compute one) representing the association between performance on a DD task (see below) and scores on an objective intelligence test (see below).

To maximize the validity and reliability of the intelligence construct, we included a study only if it assessed intelligence using a standard intelligence test or a test that can be converted reliably into an IQ (e.g., the SAT; see Frey & Detterman, 2004). Thus, studies reporting correlations between DD and a known correlate of intelligence were ineligible. For example, Sonuga-Barke, Dalen and Remington (2003) reported the correlation between DD and working memory performance. Although measures like working memory do correlate with intelligence, they likely do not measure intelligence *per se* (e.g., Conway et al., 2003). Therefore, we used a strict operationalization of intelligence to avoid potential confusion.

We also included a study in the meta-analysis only if it measured DD using a task that required an overt choice and presented an unambiguous tradeoff between delay and reward size. For example, Bishai (2004) estimated individuals' DD rates based on the tradeoff between wages and occupational fatality risk across the lifespan. This type of assessment represents a potentially useful alternative to measuring discount rates using behavioral measures, but may or may not reflect overt choice behavior. Funder, Block and Block (1983) measured whether children abstained from playing with a toy when told that they should wait until later. Although in this case delaying confers the additional benefit of adhering to experimenter demands, whether doing so is ultimately more rewarding is unclear. Kendall and Wilcox (1979) measured children's preference for a pencil now versus a pen tomorrow, but acknowledged that no pre-test was used to determine which would be preferred overall. As these three cases illustrate, constraining the eligibility requirement for DD tasks to those that measure clear tradeoffs of time and reward magnitude clarifies the construct under investigation.

There were several conditions under which we excluded studies that otherwise met the above inclusion criteria. First, we excluded studies that did not present an association between IQ and DD within the same sample (or provide enough data to compute one). Several studies (e.g., Bickel, Odum & Madden, 1999; Coffey, Gudleski, Saladin & Brady, 2003; Madden, Petry, Badger & Bickel, 1997) presented means and standard deviations of both variables for two separate groups, but the separate groups were sampled from different populations, making it difficult (or, in cases where one of the groups consisted of matched controls, impossible) to disentangle group membership or condition from the primary variables. For instance, Madden, Petry, Badger and Bickel (1997) showed that patients with opioid dependency had

higher discount rates and lower IQ than controls, but given the information in the paper, it could not be determined whether the higher discount rate for the patients was due to their lower intelligence or their drug dependency. Thus, we retained studies reporting means and standard deviations for separate groups only if the two groups were from the same sample (e.g., if group membership was assigned based on DD scores).

Second, we excluded studies that did not report zero-order effect sizes or enough data to calculate them. Controlling for other variables in a statistical analysis can significantly influence effect sizes (either by eliminating unwanted variance or exerting a suppressor effect), and so including such effect sizes in a meta-analysis can bias the distribution of effect sizes across studies unpredictably. This issue was most problematic for studies that reported effect sizes as coefficients in a regression with multiple variables entered simultaneously. For example, Kahana and Kahana (1975) report the coefficient for DD and eleven other variables in a regression used to predict IQ. Therefore, in order to converge on the zero-order relation of interest, we excluded studies like these, as well as those reporting partial correlations or other residualized associations. If there were consistent ways in which studies controlled for third variables (e.g., if half of all eligible studies controlled only for socioeconomic status) then it would have been possible to retain such studies and code them accordingly. However, the introduction of third variables was infrequent and unsystematic, and so we dropped these studies. Similarly, we excluded studies that presented either of the constructs of interest solely within a factor score or composite measure. For instance, Duckworth and Seligman (2005) report an association between IQ and a composite score of self-control that included a DD measure. Because the composite score reflects self-control more broadly rather than DD alone, the effect size does not validly reflect the relation of interest.

Third, we excluded studies that artificially dichotomized continuous variables. Combining artificially dichotomized variables in the same analysis with product-moment correlation coefficients is questionable, because correlations between continuous variables that have been dichotomized often are attenuated, biasing the overall effect size (Lipsey & Wilson, 2001). Barkley, Edwards, Laneri, Fletcher and Metevia (2001) and Weisz (1978) dichotomized and trichotomized, respectively, the IQ distributions in their analyses. Although the effect sizes from these studies can be adjusted mathematically to resemble more accurately those captured in the product-moment coefficient (cf. Hunter & Schmidt, 2004), we decided to avoid risking effect size distortion by excluding these two studies.

Fourth, we excluded studies that instructed participants to use specific strategies during the DD task. Such manipulations can dramatically and systematically influence performance on DD tasks (for a review, see Mischel, Shoda, Rodriguez, 1989), and intelligence may potentially influence DD through individual differences in strategy use. For example, Shoda, Mischel and Peake (1990) told some participants to think about specific things during the delay to reward. We excluded the effect sizes based on participants in these conditions, but retained those based on participants in other conditions that did not involve strategy instruction.

Finally, we excluded studies that were not written in English, as accurate translation of non-English text was not possible given our resources.

2.2. Search strategy

We developed a search strategy based on recommendations from a leading meta-analysis text (Lipsey & Wilson, 2001) and a Yale Medical Library reference librarian specializing in psychiatry. We obtained eligible studies through a two-stage process. The first stage involved searching PsycINFO and SCOPUS electronic databases using the following terms: delay$ adj2 gratification; delay discounting; temporal discounting; time preference; delay aversion; time discounting, inter-temporal choice, inter-temporal choice. (In PsycINFO, the '$' character matches zero, one, or multiple non-space characters. For example, delay$ matches "delays" and "delaying." The 'adj2' command returns any instance of the two flanking strings separated by two or fewer words. For instance, 'delay adj2 gratification' returns both "delay gratification" and "delay of gratification.") Based on our inclusion criteria, we limited the search to studies printed in English that were published between 1962 and the most recent search date (January 9th, 2007).

The second stage consisted of a backward search through the references section of eligible articles. For both stages, we identified potentially eligible articles by screening titles first, and then abstracts. If an article appeared eligible, we retrieved it and used a find function within the .pdf to scan for the strings "intel," "IQ," and "ability". If the function encountered a match, we read the article text. If the find function did not find matching text, or if using the find function was not possible, we searched the article text visually instead for assessment of intelligence. Through the entire search process, we screened over 2000 hits for inclusion, resulting in the retrieval of 110 potentially eligible articles. Of these, 24 studies met the inclusion criteria (without meeting any exclusion criteria) and were retained for meta-analysis.

2.3. Definition of primary variables

2.3.1. Delay discounting

The present investigation focuses on studies that measured preference for smaller, sooner rewards versus larger, later rewards. As discussed above, DD paradigms fall into two classes: commitment-choice and sustained-choice procedures.

Commitment-choice procedures assess people's tendency to prefer one reward to another, and sustained-choice procedures assess their tendency or ability to sustain a choice over delay while an alternative is still available (Reynolds & Schiffbauer, 2005). Despite potential differences between the two paradigms (Reynolds & Schiffbauer, 2005; but see Rachlin, 2000), both measure the core construct of DD. By retaining studies that measure DD in both ways and coding the disparity, we were able to investigate this core construct along with potential differences in the paradigms' relation to intelligence. Moreover, as reflected in the search strings described above, the terminology in the DD

literature is somewhat loose, in that both classes of paradigms sometimes go by the same names (e.g., "delay discounting and "delay of gratification" can refer to either commitment- or sustained-choice procedures). Therefore, our decision to include both types of studies also protected against inadvertent exclusion of relevant studies.

2.3.2. Intelligence

As discussed above, we included a study in the meta-analysis only if it administered a test of intelligence. Many studies involved tests that are widely considered to measure intelligence (e.g., Wechsler tests, Otis–Lennon), or gave ample descriptions of the test they used. Other cases required consulting other studies that used the test, or information from its distributor, for supplementary descriptions. In most cases the consensus on what the test measured was clear. An exception is the Porteus Maze Test, which was presented in a few studies as a measure of intelligence. Because this assessment's appropriateness as an intelligence test is questionable, given its associations with impulsivity (Paulsen & Johnson, 1980) and potentially weak association with other IQ measures (Krikorian & Bartok, 1998), we excluded studies that used it as the sole measure of intelligence. We also excluded reports measuring correlates of intelligence such as working memory and digit span, as well as those using achievement tests, for reasons described above.

2.4. Study coding

Once we retrieved all eligible studies, the first author coded them for sample characteristics, study level descriptors, and effect size data using a coding form that we prepared before the coding process began and adjusted iteratively during the process to accommodate exceptions. The coding process involved both reporting raw numbers (e.g., percentage of the sample that was male) and assigning data to categories (e.g., the category that best described the DD choice paradigm). See the Appendix for a complete list of coded variables and their values for each study.

Although generally we did not impute missing data, several situations called for exceptions to this rule (the Appendix specifies which variables contain imputed values). First, if the coding called for a mean but only a range was available, the mean was imputed as the mean of that range. Second, some categories required that numerical values be imputed from imprecise information. For example, if the mean age of the sample was not provided but the sample was described as 'college undergraduates', the value '20' was imputed (college students tend to be within 18–22 years of age). Although these types of imputation can potentially introduce some noise into the data, we considered them preferable to coding applicable cases as missing.

We now clarify how several variables were coded according to our scheme. We discuss only those variables for which the details of the coding scheme are essential.

2.4.1. Mean/median socioeconomic status (SES)

Information on SES was typically limited when present, and the distribution of SES data across studies exhibited positive skew. Thus, to maximize reliability and facilitate useful moderator analyses, we coded SES dichotomously and set a cut point that divided studies into subgroups of comparable size. We classified studies according to whether the mean or median income of participants fell below $8000 USD under the median income for the year in which the study was published ("lower"), or met or exceeded this value ("higher"). Median incomes were obtained from US census data (U.S. Census Bureau). We recognize that using the publication year to determine the median income is not completely accurate (data collection often precedes publication by more than a year). However, this decision likely did not influence the coding, as recalculating the categories based on publication year - 2 did not alter the category membership of any studies. Finally, in some cases, we imputed category membership based on a verbal description. We coded studies that described participants' income or class as "low" or "lower" as belonging to the lower SES category, and those with "middle" as belonging to the higher category.

2.4.2. Intelligence type

We coded intelligence tests based on the primary type of intelligence they measured. Two types of tests emerged: those that measured both verbal and nonverbal intelligence (e.g., the Wechsler Adult Intelligence Test; coded as "general") and those that measured verbal intelligence only (e.g., the National Adult Reading Test). We categorized tests based on information about the tests in the articles themselves or other relevant test information (e.g., other studies using the tests, descriptions on the proprietor's web site). See the Appendix for data on which tests were involved and how we categorized them.

2.4.3. Delay discounting choice paradigm

Based on the distinction discussed earlier, we categorized studies as involving either a commitment-choice or sustained-choice DD procedure. Specifically, if participants were unable to change a choice on a given trial after making it, we classified the study procedure as commitment-choice. If participants were able, at any time during a trial, to forego a larger reward promised at the end of a delay period for a smaller reward to be received immediately, we classified the procedure as sustained-choice.

2.4.4. Delay discounting measure

As mentioned above, DD tasks can measure either an individual's preference for immediate (smaller) versus delayed (larger) rewards, or how this preference varies as a function of delay. Accordingly, we coded studies on this dimension dichotomously. One category comprised measures of the former, time preference, and the other comprised measures of the latter, discount rate. Time preference measures typically included the number or proportion of delayed choices made, and discount rate measures usually involved parameter estimates that reflected the steepness of a hyperbolic (or hyperbola-like) discounting function.

2.4.5. Delay discounting reward material

We used four categories to code the materials that constituted the rewards on the DD tasks. Aside from money

or food, rewards sometimes comprised a mixture of items (e.g., varying across trials or between participants). We coded studies using assorted items into separate categories depending on whether money was involved in the assortment.

2.4.6. Delay discounting reward scheme

DD studies differ with regard to how they actually compensate participants. Oftentimes the compensation participants receive does not depend exclusively on how they perform on the DD task. We used three categories to represent possible reward schemes. First, rewards could be entirely hypothetical, such that a participant receives compensation that is independent of his or her choice or choices on the DD task. A second possibility is one of "chance payoff"—that is, either there is a random chance that each participant will receive the outcome of one of his or her decisions, or he or she will receive the outcome of one choice at random. Finally, a participant could receive compensation based on every choice he or she makes on the task, such that all rewards are real.

2.5. Data analysis

Before analyzing the data, we addressed the issue of multiple effect sizes per study. Several studies reported more than one effect size of interest for the same population. For example, Baker, Johnson and Bickel (2003) reported associations between IQ and two different DD measures, and Macbeth (1974) reported associations between math and verbal SAT scores and DD separately. A key assumption of meta-analysis is that the effect sizes being analyzed are independent of one another, and so reporting multiple effect sizes from the same sample violates this assumption (Lipsey & Wilson, 2001). (Multiple effect sizes from the same research report are safely assumed not have this problem, so long as they are taken from independent samples.) There are several ways to incorporate multiple, dependent effect sizes into a meta-analysis without violating the independence assumption, but each method involves tradeoffs. We decided that, where warranted, multiple effect sizes would be averaged into a single effect size, as this procedure is straightforward, and, at worst, slightly conservative (Lipsey & Wilson, 2001).

We analyzed coded data according to the recommendations of Lipsey and Wilson (2001) using Comprehensive Meta-Analysis software (CMA; Borenstein, 2005). Each entry consisted of a separate study, except in two cases that reported separate effect sizes for different segments of the sample; 26 effect sizes in total contributed to the meta-analysis. We used the software first to obtain weighted mean effect sizes and to test for effect size heterogeneity, using both fixed and random effects models (Lipsey & Wilson, 2001).

We then conducted moderator analyses to test for systematic differences between studies that could potentially explain the observed heterogeneity among effect sizes. Here, our analyses involved both fixed and mixed effects models (Lipsey & Wilson, 2001). We tested for categorical moderators using the analog to the ANOVA method, and continuous moderators using weighted least-squares regression, with method-of moments-estimation for all mixed effects models (Borenstein, 2005; Lipsey & Wilson, 2001). For continuous moderator analyses, we used Lipsey and Wilson's (2001) MetaReg macro for SPSS (instead of CMA).

Finally, to assess potential publication bias, we computed two types of fail-safe N and performed funnel plot analyses.

3. Results

3.1. Mean effect size and homogeneity

Under a fixed effects model, the weighted mean effect size across studies (expressed as an r value) was -0.246 ($p<0.001$; 95% CI: -0.280, -0.212). The fixed effects Q-statistic indicated that the null hypothesis of homogeneous effect sizes could be rejected ($Q(25)=63.10$, $p<0.001$), suggesting that the variance in effect sizes could not be attributed to sampling error alone. In light of this heterogeneity, and to achieve a more generalizable result (see Raudenbush, 1994), we re-computed the weighted mean effect size using a random effects model: $r=-0.234$ ($p<0.001$; 95% CI: -0.295, -0.171). The random effects model ultimately may not be the optimal specification, but we include both fixed and random effects estimates to provide less and more conservative statistics, respectively (see Overton, 1998). The point estimates from the fixed and random effects models both indicate that the association between DD and intelligence is negative in direction, as hypothesized, and small-to-moderate in magnitude, according to Cohen's (1988) criteria, with approximately 5.48% of shared variance between the constructs (based on the random effects estimate). Thus, more intelligent people demonstrate less of a preference for smaller, immediate rewards versus larger, delayed rewards.

3.2. Moderator analyses

Although the fixed effects Q-statistic was significant, indicating heterogeneity of effect sizes, a random effects model may not be the correct specification; rather, the heterogeneity may be explained by systematic factors (moderators), or both systematic and random factors, and it is not always clear *a priori* which model is preferable (Lipsey & Wilson, 2001). Moreover, although fixed effects models have been criticized as being less generalizable and more prone to Type I error (for a discussion, see Hunter & Schmidt, 2000), mixed effects analyses are more susceptible to Type II error and have theoretical shortcomings of their own (for a discussion, see Overton, 1998). Therefore, we tested moderators under fixed effects assumptions first, modeling further unexplained variance as a random effect (mixed effects model) only when the appropriate Q-statistic (Q_{Within} or Q_{Residual}) indicated that unexplained variability remained.

3.2.1. Categorical moderators

Analyses of categorical moderators focused on three variables selected *a priori*: DD choice paradigm, DD measure, and type of intelligence assessed. Neither fixed nor mixed effects analyses revealed a significant moderator effect of any of the three variables (all fixed effects $Qs_{\text{Between}}<1.02$, $ps>0.32$). That is, there was no evidence that any of the three variables accounted for heterogeneity in the overall effect size distribution.

We also explored a fourth categorical moderator *post-hoc*: DD reward scheme. This variable was chosen based on theoretical interest and completeness of the coded data. The fixed effects analysis revealed significant heterogeneity across all three groups ($Q_{\text{Between}}(2)=15.15, p=0.0005$). Because of the significant residual heterogeneity within groups ($Q_{\text{Within}}(23)=47.95, p<0.0001$) we decided to re-run this analysis as a mixed effects model (which combined studies within each subgroup using random effects and combined subgroups using fixed effects, and computed unique estimates of between-studies variance for each subgroup; Borenstein, 2005). This more conservative model was consistent with the earlier fixed effects approach, demonstrating significant heterogeneity across groups ($Q_{\text{Between}}(2)=9.91$, $p=0.007$). We followed up this result with *post-hoc* tests that compared each of the three subgroups to one another, and each subgroup to a combination of the other two (six tests in total). There was no difference in effect sizes between studies using hypothetical rewards and studies using all real rewards (or real and chance payoff combined) rewards, nor was there a difference between studies using all real rewards and hypothetical and chance payoff rewards combined (Q_{Between}s$(1)=1.69$, ps>0.19). There was, however, a difference between studies using chance payoff and those that used either all real or hypothetical rewards ($Q_{\text{Between}}(2)=8.10$, $p=0.0042$, surviving Bonferroni correction at the 0.05 level), with the former group showing weaker effect sizes than the latter (−0.12 versus −0.27). (After Bonferroni correction, chance payoff rewards studies were also significantly different from hypothetical rewards studies by themselves, but not from all real rewards studies by themselves.) Thus, we can conclude that the association between DD and intelligence is weaker when receipt of rewards on the DD task is subject to chance (i.e. when there is either a chance of a reward or one reward is received at random), but does not depend on whether such rewards are hypothetical or real.

A fifth moderator analyses that we also pursued *post-hoc* was the influence of socioeconomic status (SES). Studies involving participants with lower SES (as defined above) showed weaker associations between DD and intelligence than did studies involving participants with higher SES (−0.23 versus −0.36; $Q_{\text{Between}}(1)=9.44$, $p=0.0041$). However, this effect did not survive a mixed effects analysis ($Q_{\text{Between}}(1)=0.746$, $p=0.388$), so the inferential scope is limited.

Finally, we examined psychological disturbance (e.g., substance abuse, personality disorders) and DD reward material *post-hoc* as potential moderators. Consistent with a lack of moderator effects, neither of the fixed effects between group homogeneity tests reached the 0.01 level of significance, which was set *a priori* to control Type I error (cf. Hunter & Schmidt, 2000).

3.2.2. Continuous moderators

Exploratory analyses focused on four variables: publication year, mean IQ, age, and education level. All four variables of interest moderated the DD × intelligence relation under a fixed effects model, but had no influence when the model was re-estimated under mixed effects assumptions (see Table 1). The conservative nature of mixed effects analyses is apparent in these tests: all were non-significant despite significant residual variance under fixed effects (which typically indicates that a mixed effects model should be attempted; Lipsey & Wilson, 2001). Nonetheless, we can conclude that, at least in the set of studies under consideration in the present meta-analysis, more recent studies have shown a stronger association between DD and intelligence, as have studies involving more intelligent, older, and more educated participants. Unfortunately, we were unable to tease apart the unique moderating influences of age and education, because the two variables were highly collinear ($r(15)=0.91, p<0.001$), nor could we tease apart publication year from age ($r(23)=0.64$, $p=0.0011$) and education ($r(18)=0.61$, $p=0.0074$) for similar reasons.

3.3. Publication bias

We conducted two separate analyses to ascertain that the mean effect size was not unduly influenced by missing or file-drawer studies. First, we calculated two versions of fail-safe *N*. Rosenthal's (1979) fail-safe *N* indicated that 880 missing studies with null results ($p>0.05$) would have to be retrieved to render the observed mean effect size non-significant by the same criterion. Orwin's (1983) fail-safe *N* indicated that even to reduce the observed effect size from its small-to-moderate level to a small level ($r=-0.1$), 40 studies with a mean correlation of zero would have to be retrieved—more than the present number found by an exhaustive search of the literature.

Second, we constructed two funnel plots, one with the inverse of the standard error of the effect size and the other with sample size plotted against the effect size, respectively (Fig. 1).

We provide both because the two methods of funnel plot construction can potentially produce disparate results (Tang &

Table 1
Continuous moderator effects

	Fixed effects			Mixed effects		
Moderator	*B* (SE)	*β*	Q_{Residual}	*B* (SE)	*β*	$\nu_{\text{Random effects}}$
Publication year	−0.003 ** (0.0011)	−0.33	56.11 ***	−0.0037 (0.0023)	−0.30	0.014
Mean IQ	−0.0064 ** (0.0025)	−0.37	4.77 ***	−0.0023 (0.0046)	−0.12	0.013
Age	−0.004 * (0.0013)	−0.40	48.50 ***	−0.0024 (0.0024)	−0.19	0.016
Education	−0.013 * (0.0044)	−0.43	40.95 ***	−0.0096 (0.0071)	−0.30	0.013

* Significant at 0.025 level.
** Significant at 0.01 level.
*** Significant at 0.001 level.

Liu, 2000); both methods were concordant in our study. Because effect sizes from small-*n* studies are less precise estimators of the population effect size, there should be more variability among them if studies reporting small or null findings are not relegated to the file drawer. In contrast, studies with large samples should demonstrate less variability in effect sizes because their estimates of the population effect size are more precise. Therefore, a funnel-like shape, like that observed in the both plots, argues against publication bias. This shape (and our conclusion of no evident bias) was corroborated by formal significance tests, which are based on the intercept of the regression of standardized effect size (or the product of effect size and the square root of sample size) on the inverse of the standard error of the effect size (or the square root of the sample size; see Egger, Smith, Schneider & Minder, 1997; Tang & Liu, 2000). Intercepts for both regressions were not significantly different than zero ($ps > 0.277$).

4. Discussion

The present study demonstrated, using meta-analytic techniques, that individuals with higher intelligence demonstrate significantly less of a tendency to prefer smaller, sooner rewards to larger, later ones. This association was present to a small-to-moderate degree in the 24 eligible research reports (comprising 26 independent effect sizes) that we gathered via a systematic search of the literature. As the fail-safe *N*, funnel plots, and associated statistical tests indicate, publication bias did not likely influence the findings of the present report. Moderator analyses suggested that studies employing either all real or all hypothetical rewards, versus studies using rewards whose payouts were subject to chance, showed larger associations between DD and intelligence. Socioeconomic status, age, education level, publication year, and sample mean IQ also accounted for heterogeneity in effect size magnitude, albeit under a liberal statistical model with limited generalizability. DD choice paradigm and intelligence type did not appear to moderate effect size.

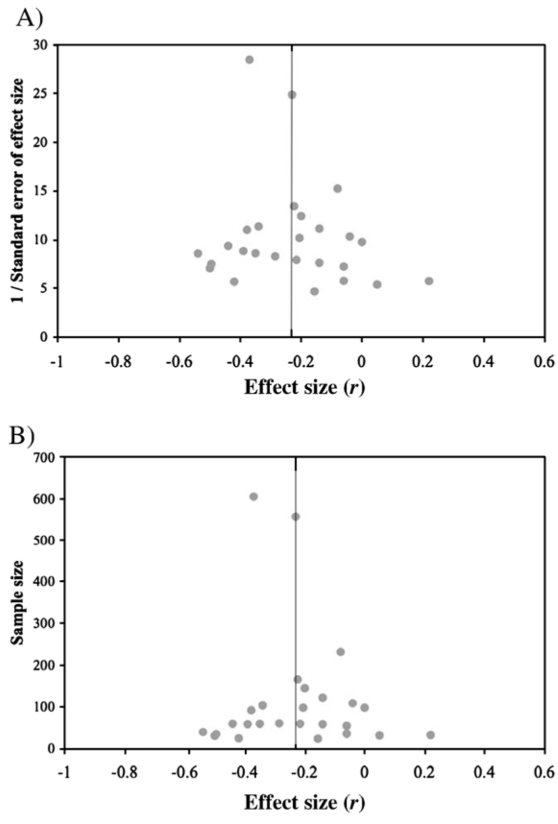

Fig. 1. Funnel plots of A) inverse of the standard error of the effect size against effect size and B) study sample size against effect size. The vertical line in both graphs at an effect size value of −0.23 represents the mean (unweighted) effect size across studies.

Notably, studies in which payoffs involved chance demonstrated weaker relations between DD and intelligence than did studies in which payoffs were all hypothetical or all real. This result is significant for two main reasons. First, it suggests that researchers should choose reward schemes carefully in future studies. Although some distinctions, such as that between hypothetical and real rewards, have been shown not influence DD preferences (Johnson & Bickel, 2002; Madden et al., 2003), the literature has all but ignored the potential significance of involving chance in compensating participants based on DD task performance. Presumably, chance payoff schemes are chosen as a compromise between the financial burden of paying out real rewards on every trial and presumed lack of ecological validity from using all hypothetical choices. The present findings, however, suggest that these potential practical benefits come at the cost of introducing additional psychological effects that may not be desired. Second, the result suggests that perceptions of risk can subtly influence DD behavior. Clues for the basis of such an influence come from the link between DD and another behavior, probability discounting (PD), which describes preferences for smaller, certain rewards over larger, risky ones. PD is similar to DD in that preferences vary according to similar functions: the value of riskier (larger) versus less risky (smaller) rewards decreases hyperbolically with risk (Green, Myerson & Ostaszewski, 1999; Prelec & Loewenstein, 1991). Because DD and PD are correlated within individuals, it is possible that the perception of risk through chance payoff schemes increases the influence of an overall "discounting trait" (Myerson et al., 2003), relative to intelligence, on DD behavior. Future work should aim to test more precisely how actual or implied risk can influence DD, and whether individual differences in cognitive capacity moderate this influence.

Please cite this article as: Shamosh, N. A., & Gray, J. R., Delay discounting and intelligence: A meta-analysis, *Intelligence* (2007), doi:10.1016/j.intell.2007.09.004

Other moderator tests that we conducted were significant only under liberal, fixed effects assumptions. We discuss these effects primarily to describe the studies within the present meta-analysis. Studies involving participants of lower socioeconomic status showed weaker associations between DD and intelligence. Given evidence that economic scarcity can influence discount rates (Ostaszewski, Green & Myerson, 1998), it makes sense that, for those under greater economic hardship, intelligence matters less in determining DD preferences than it does for people of greater financial means. It is possible that this effect would have held up in more conservative analyses if relevant information had been presented in more of the studies. Additionally, effect sizes were larger for samples of higher mean IQ, although the reason for such a relation is uncertain. Finally, publication year, sample mean age, and sample mean education level all demonstrated positive relations to effect size, though collinearity between these variables precluded clear interpretation of these results.

The absence of some moderator effects is also potentially informative, notwithstanding the inferential limitations of null results. Effect sizes did not differ significantly in magnitude between both types of DD measures, time preference and discount rate. One possibility is that they equally measure an aspect of DD for which intelligence is critical. Alternatively, they could reflect separable aspects of DD that happen to require intelligence to the same degree, but for different reasons. Further work is needed to test these possibilities. Moreover, we hypothesized that discount rates would relate more strongly to intelligence because they partly reflect preference consistency, which is linked to intelligence in other contexts (Parker & Fischhoff, 2005). However, the tendency to demonstrate preference reversals on DD tasks may be too close to ceiling in the population (e.g., Kirby & Herrnstein, 1995) for intelligence to capture significant variability.

Additionally, the strength of association between DD and intelligence did not differ between commitment-choice and sustained-choice paradigms. Although both paradigms do measure DD in the sense of individuals' preferences for smaller, sooner versus larger, later rewards, they are separable to some extent, as Reynolds and Schiffbauer (2005) argue. Yet the absence of moderation in the present study speaks against these authors' implication that commitment-choice procedures are more cognitively oriented than sustained-choice procedures, because the DD × intelligenceassociation did not differ between the two types. It is possible that each procedure has particular requirements that tap intelligence, or that intelligence is involved in a process underlying both types of tasks (cf. Rachlin, 2000).

We hypothesized that the type of intelligence a study assessed would relate to its effect size. There were two classes of tests: those that measured both verbal and nonverbal intelligence and those that assessed verbal intelligence exclusively. Neither class was associated with larger effect sizes. A potential implication of this null result is that nonverbal abilities are not particularly critical for DD, as measuring them in addition to verbal ability did not strengthen the relation of the intelligence scores to DD. This implication offers some insight into the strategies people may deploy spontaneously on DD tasks. If participants were comparing rewards quantitatively or engaging in sophisticated calculations, a broader assessment of intelligence may have shown a stronger relation to DD for capturing the ability to execute these strategies successfully. Given that tests of verbal intelligence alone produced effect sizes that were equally large as those that also measured nonverbal abilities, our results suggest instead that facility with verbally intensive strategies (e.g., Metcalfe & Mischel, 1999) may crucially influence DD task performance. Another possibility, which remains to be explored, is that DD is most strongly associated with general intelligence rather than specific abilities (Jensen, 1998).

The present study has two main limitations. The first is that the number of eligible studies, though sufficient for meta-analysis, was low, which reduced power for moderator analyses. Relaxing the eligibility criteria would have increased the number of studies in the meta-analysis, but also would have rendered the primary variables ambiguous. We excluded many studies that measured DD and intelligence but did not report enough information to calculate an effect size (to prevent bias, we did not seek this information directly from the authors, as it is often unavailable for older articles). Hence, this limitation reflects the need for both more research and more thorough reporting of DD × intelligence associations, where applicable.

The second limitation is that the data permitted only limited conclusions about how different types of intelligence relate to DD. Based on the available information, our coding scheme only distinguished between intelligence tests that measured both verbal and nonverbal abilities and those that measured verbal abilities only. Future research should follow the lead of some researchers (e.g., Hongwanishkul et al., 2005; Kirby et al., 2005) and present the relation between DD and intelligence as broken down by subtests representing more specific types of intelligence so that future meta-analyses can compare the resulting effect sizes. Researchers should also aim to incorporate and contrast more diverse measures of intelligence, such as measures of fluid and crystallized intelligence. Doing so will enable a more nuanced look at which aspects of intelligence relate most strongly to DD.

Please cite this article as: Shamosh, N. A., & Gray, J. R., Delay discounting and intelligence: A meta-analysis, *Intelligence* (2007), doi:10.1016/j.intell.2007.09.004

The present report is, to our knowledge, the first to quantify the relation between DD and intelligence based on a comprehensive search of the literature. Thus, the key contribution of this meta-analysis is demonstrating that this relation not only exists, but also is consistent and substantial enough to constrain hypotheses about why individuals differ in DD. Although meta-analysis has the potential to end a line of inquiry into a psychological phenomenon by quantifying it and identifying relevant moderators, the present study has the potential to bring a noteworthy, but largely obscure, effect into the open for further investigation of its bases.

Our findings also have possible implications for phenomena with broad social impact. For example, because burdensome levels of credit card debt and insufficient retirement savings may stem in part from higher DD (Angeletos et al., 2001; Bar-Gill, 2004), individuals with lower intelligence may be more prone to these financial hardships. Forthcoming work by Benjamin, Brown, and Shapiro (2006) supports this possibility with evidence that individuals with lower intelligence tend to have lower levels of financial asset accumulation, controlling for income. Therefore, DD may represent a critical variable through which intelligence influences important economic outcomes (cf. Heckman, 2007).

In summary, individual differences in the tendency to prefer smaller, sooner rewards to larger, later ones are associated with intelligence. This result also suggests that DD reflects not only motivational processes, but also cognitive abilities that critically influence the formation or integration of preferences. The consistency and magnitude of the association evince the necessity for further work to determine its bases. Although we explored several moderators, the range of possible reasons for the relation underscores the need for empirical tests. We hope that the present investigation will prompt such research, guided by the findings that we have presented.

Acknowledgments

The authors thank Marvin M. Chun and Marcia K. Johnson for comments on a previous draft. This research was supported by a National Science Foundation Graduate Research Fellowship to the first author.

Appendix A

Coded values for each study included in the meta-analysis

Study	Sample size	Raw effect size, *r* (SE)	Mean age†	% Male	Mean/median years education†	Mean/median SES†
Baker et al. (2003)	60	−0.215 (0.126)	31.95	57	13.85	Higher
Crean et al. (2000)	24	−0.156 (0.213)	33.5	50	14.5	Higher
Cuskelly et al. (2001)	31	−0.500 (0.142)	19.8	58	–	–
Cuskelly et al. (2003)	25	−0.420 (0.176)	10.29	44	–	–
de Wit et al. (2007)	606	−0.370 (0.035)	45.18	50	16	Higher
Dolan and Fullam (2004)	40	−0.539 (0.117)	42.2	100	12.25	Lower
Dom et al. (2007)	92	−0.378 (0.091)	42.3	74	12.7	–
Funder and Block (1989)	104	−0.340 (0.088)	14	48	–	–
Hongwanishkul et al. (2005)	98	−0.205 (0.098)	4	51	–	–
Kendall et al. (1981)	98	0.000 (0.103)	9.4	46	3.5	–
Kirby et al. (1999)	59	−0.140 (0.131)	35.8	–	12.6	Lower
Kirby and Petry (2004)	145	−0.200 (0.081)	39.97	72	12.51	Lower
Kirby et al. (2005)‡	232	−0.080 (0.066)	20	44	14.5	–
Krietler and Zigler (1990) (1)	60	−0.350 (0.116)	5.45	50	–	Higher
Krietler and Zigler (1990) (2)	60	−0.440 (0.107)	11.56	50	–	Higher
Lessing (1969)	558	−0.230 (0.040)	–	–	9.5	Lower
LeSure (1977)	61	−0.285 (0.121)	–	–	4.5	–
Macbeth (1974)	109	−0.040 (0.097)	–	44	14.5	–
Mischel and Metzner (1962)‡	122	−0.140* (0.090)	7.5	55	3.5	Lower
Monterosso et al. (2001)	32	0.050 (0.185)	39	78	–	–
Paulsen and Johnson (1980)	55	−0.060 (0.138)	4.56	49	0	–
Petry (2002)	166	−0.2235 (0.074)	39.21	67	12.41	Lower
Rodriguez et al. (1989)	59	−0.390 (0.113)	10.19	100	–	Lower
Shoda et al. (1990) (1)‡	35	−0.495 (0.133)	4.33	44	0	Higher
Shoda et al. (1990) (2)‡	33	0.220 (0.174)	4.33	44	0	Higher
Tweedie (1967)	36	−0.060* (0.173)	9	56	3.5	Lower

*Converted from d using Cohen's (1988) formula $r = d/\surd(d2+4)0.5$.

†Contains imputed values.

‡Sample characteristics apply to a broader sample.

Appendix A (*continued*)

Psychological disturbance (for majority of participants)	Mean IQ
Addiction/substance abuse	–
Heterogeneous diagnoses	118
Down Syndrome	–
Down Syndrome	–
None/no data reported	117.2
Personality disorders	98.2
Addiction/substance abuse	96.7
None/no data reported	115.75
None/no data reported	–
None/no data reported	113.6
Addiction/substance abuse	98.5
Addiction/substance abuse	94.32
None/no data reported	106.57
None/no data reported	107.17
None/no data reported	107.03
None/no data reported	104.87
None/no data reported	106.44
None/no data reported	96.97
None/no data reported	102.51
Addiction/substance abuse	93.4
None/no data reported	112.5
Addiction/substance abuse	94.2
None/no data reported	–
None/no data reported	–
None/no data reported	–
None/no data reported	110

Intelligence test(s) referenced	IQ type	DD task referenced
Quick Test (Ammons & Ammons, 1962)	Verbal	Richards et al. (1999)
National Adult Reading Test (Grober & Sliwinski, 1991)	Verbal	Richards et al. (1999)
Expressive Vocabulary Test (Williams, 1997)	Verbal	Mischel (1974)
Peabody Picture Vocabulary Test, 3rd ed. (Dunn & Dunn, 1997)	Verbal	Mischel (1974)
Wechsler Abbreviated Scale of Intelligence, Vocabulary and Matrix Reasoning Subtests (Wechsler, 1999)	General	Mitchell (1999)
National Adult Reading Test (Nelson, 1982)	Verbal	Newman et al. (1992)
Wechsler Adult Intelligence Scale, 3rd ed., Dutch version (Swets Test Publishers, 2000)	General	Richards et al. (1999)
Wechsler Intelligence Scale for Children—Revised (Wechsler, 1974); Wechsler Adult Intelligence Scale—Revised (Wechsler, 1981a)	General	Original
Stanford–Binet Intelligence Scale, 4th ed., Bead Memory and Pattern Analysis subtests (Thorndike et al., 1986)	General	Prencipe and Zelazo (2005)/ Thompson et al. (1997)
Peabody Picture Vocabulary Test (Dunn, 1965)	Verbal	Original
Wechsler Adult Intelligence Scale—Revised (Wechsler, 1981b)	General	Kirby and Marakovic (1996)
Shipley Institute for Living Scale (Zachary, 1991)	General	Kirby and Marakovic (1996)
Scholastic Aptitude Test (no reference given)	General	Kirby (1997)
Wechsler–Bellevue Test, Israeli version (Lieblich, Ben-Shahar & Ninio, 1975)	General	Mischel (1961)
Wechsler–Bellevue Test, Israeli version (Lieblich et al., 1975)	General	Mischel (1961)
Assorted IQ tests (no specific tests or references given)	General	Original
Otis–Lennon Mental Ability Test (no reference given)	General	Mischel and Staub (1965)
Scholastic Aptitude Test (no reference given)	General	Original
Lorge–Thorndike, Form A; Pintner General Ability Test, Verbal Series, Forms A or B (no references given)	General	Mischel (1961)
Wechsler Adult Intelligence Scale, 3rd ed. (Wechsler, 1997)	General	Kirby and Marakovic (1996)
Peabody Picture Vocabulary Test (no reference given; see above)	Verbal	Original
Shipley Institute for Living Scale (Zachary, 1991)	General	Original
Peabody Picture Vocabulary Test (no reference given; see above)	Verbal	Mischel (1974)
Scholastic Aptitude Test (no reference given)	General	Mischel (1974)

(continued on next page)

Appendix A (*continued*)

Intelligence test(s) referenced	IQ type	DD task referenced
Scholastic Aptitude Test (no reference given)	General	Mischel (1974)
Kuhlmann–Anderson Intelligence Test, form C–D (no reference given)	General	Mischel (1961)

Choice paradigm	DD measure	Reward material	Reward scheme
Commitment-choice	Time preference	Assorted, no money	Hypothetical
Commitment-choice	Time preference	Money	Chance payoff
Sustained-choice	Discount rate	Assorted, no money	All choices paid
Sustained-choice	Discount rate	Assorted, no money	All choices paid
Commitment-choice	Time preference	Money	Hypothetical
Commitment-choice	Discount rate	Money	All choices paid
Commitment-choice	Time preference	Money	Hypothetical
Commitment-choice	Discount rate	Money	All choices paid
Commitment-choice	Discount rate	Assorted, includes money	All choices paid
Commitment-choice	Discount rate	–	All choices paid
Commitment-choice	Time preference	Money	Chance payoff
Commitment-choice	Time preference	Money	Chance payoff
Commitment-choice	Time preference	Money	Chance payoff
Commitment-choice	Discount rate	Assorted, no money	All choices paid
Commitment-choice	Discount rate	Assorted, no money	All choices paid
Commitment-choice	Discount rate	Assorted, includes money	Hypothetical
Commitment-choice	Discount rate	Assorted, includes money	Chance payoff
Commitment-choice	Discount rate	Assorted, includes money	Chance payoff
Commitment-choice	Discount rate	Food(s)	All choices paid
Commitment-choice	Time preference	Money	Chance payoff
Commitment-choice	Discount rate	Food(s)	All choices paid
Commitment-choice	Time preference	Money	Hypothetical
Sustained-choice	Discount rate	Food(s)	All choices paid
Sustained-choice	Discount rate	Assorted, includes money	All choices paid
Sustained-choice	Discount rate	Assorted, includes money	All choices paid
Sustained-choice	Discount rate	Food(s)	All choices paid

References[1]

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*(1), 30−60.

Ammons, R. B., & Ammons, C. H. (1962). The Quick Test (QT): Provisional manual. *Psychological Reports*, *11*, 111−161.

Angeletos, G. -M., Laibson, D., Repetto, A., Tobacman, J., & Weinberg, S. (2001). The hyperbolic consumption model: Calibration, simulation, and empirical evaluation. *Journal of Economic Perspectives*, *15*(3), 47−68.

*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*(3), 382−392.

Bar-Gill, O. (2004). Seduction by plastic. *Northwestern University Law Review*, *98*(4), 1373−1434.

Barkley, R. A., Edwards, E., Laneri, M., Fletcher, K., & Metevia, L. (2001). 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*, *29*(6), 541−556.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1–3), 7−15.

Beier, M. E., & Ackerman, P. L. (2005). Working memory and intelligence: Different constructs. Reply to Oberauer et al. (2005) and Kane et al. (2005). *Psychological Bulletin*, *131*(1), 72−75.

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2006, April). Who is "behavioral"? Cognitive ability and anomalous preferences. *Working Paper.* Harvard University. Retrieved August 10, 2007 from http://www.dklevine.com/archive/refs4122247000000001334.pdf

Bickel, W. K., & Marsch, L. A. (2001). Toward a behavioral economic understanding of drug dependence: Delay discounting processes. *Addiction*, *96*(1), 73−86.

Bickel, W. K., Odum, A. L., & Madden, G. J. (1999). Impulsivity and cigarette smoking: Delay discounting in current, never, and ex-smokers. *Psychopharmacology*, *146*(4), 447−454.

Bishai, D. M. (2004). Does time preference change with age. *Journal of Population Economics*, *17*, 583−602.

Borenstein, M. (2005). *Comprehensive meta-analysis (Version 2).* Englewood, NJ Biostat.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in

[1] Those included in the meta-analysis are preceded by an asterisk.

the Raven Progressive Matrices test. *Psychological Review*, *97* (3), 404−431.

Coffey, S. F., Gudleski, G. D., Saladin, M. E., & Brady, K. T. (2003). Impulsivity and rapid discounting of delayed hypothetical rewards in cocaine-dependent individuals. *Experimental and Clinical Psychopharmacology*, *11*(1), 18−25.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Second ed.). Hillsdale, N.J.: Lawrence Erlbaum.

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547−552.

*Crean, J. P., de Wit, H., & Richards, J. B. (2000). Reward discounting as a measure of impulsive behavior in a psychiatric outpatient population. *Experimental and Clinical Psychopharmacology*, *8*(2), 155−162.

Critchfield, T. S., & Kollins, S. H. (2001). Temporal discounting: Basic research and the analysis of socially important behavior. *Journal of Applied Behavior Analysis*, *34*(1), 101−122.

*Cuskelly, M., Einam, M., & Jobling, A. (2001). Delay of gratification in young adults with Down Syndrome. *Down Syndrome Research and Practice*, *7*(2), 60−67.

*Cuskelly, M., Zhang, A., & Hayes, A. (2003). A mental age-matched comparison study of delay of gratification in children with down syndrome. *International Journal of Disability, Development, and Education*, *50*(3), 239−251.

*de Wit, H., Flory, J. D., Acheson, A., McCloskey, M., & Manuck, S. B. (2007). IQ and nonplanning impulsivity are independently associated with delay discounting in middle-aged adults. *Personality and Individual Differences*, *42*, 111−121.

*Dolan, M., & Fullam, R. (2004). Behavioral and psychometric measures of impulsivity in a personality disordered population. *The Journal of Forensic Psychiatry and Psychology*, *15*(3), 426−450.

*Dom, G., De Wilde, B., Hulstijn, W., & Sabbe, B. (2007). Dimensions of impulsive behaviour in abstinent alcoholics. *Personality and Individual Differences*, *42*, 465−476.

Dunn, L. M. (1965). *Expanded manual for the Peabody picture vocabulary test.* Minneapolis, MN American Guidance Services.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test*, Third Edition. Circle Pines, MN American Guidance Service.

Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*(12), 939−944.

Duckworth, A. L., & Seligman, M. E. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, *98*(1), 198−208.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629−634.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309−331.

Franco-Watkins, A. M., Pashler, H., & Rickard, T. C. (2006). Does working memory load lead to greater impulsivity? Commentary on Hinson, Jameson, and Whitney (2003). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*(2), 443−447 discussion 448–450.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25−42.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or *g*? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, *15*(6), 373−378.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172−179.

*Funder, D. C., & Block, J. (1989). The role of ego-control, ego-resiliency, and IQ in delay of gratification in adolescence. *Journal of Personality and Social Psychology*, *57*(6), 1041−1050.

Funder, D. C., Block, J. H., & Block, J. (1983). Delay of gratification: Some longitudinal personality correlates. *Journal of Personality and Social Psychology*, *44*(6), 1198−1213.

Gottfredson, L. S. (1997). Why *g* matters: The complexity of everyday life. *Intelligence*, *24*(1), 279–132.

Green, L., Fry, A. F., & Myerson, J. (1994). Discounting of delayed rewards: A life-span comparison. *Psychological Science*, *5*(1).

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 and Aging*, *11*(1), 79−84.

Green, L., Myerson, J., & Ostaszewski, P. (1999). Amount of reward has opposite effects on the discounting of delayed and probabilistic outcomes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *25*(2), 418−427.

Grober, E., & Sliwinski, M. (1991). Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *Journal of Clinical and Experimental Neurophysiology*, *13*, 933−949.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capacity formation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(33), 13240−13255.

Hinson, J. M., Jameson, T. L., & Whitney, P. (2003). Impulsive decision making and working memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*(2), 298−306.

Hinson, J. M., & Whitney, P. (2006). Working memory load and decision making: A reply to Franco-Watkins, Pashler, and Rickard (2006). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*(2), 448−450.

*Hongwanishkul, D., Happaney, K. R., Lee, W. S., & Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Developmental Neuropsychology*, *28*(2), 617−644.

Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, *8*(4), 275−292.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA Sage Publications.

Jaroni, J. L., Wright, S. M., Lerman, C., & Epstein, L. H. (2004). Relationship between education and delay discounting in smokers. *Addictive Behaviors*, *29*, 1171−1175.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, Conn.; London: Praeger.

Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior*, *77*(2), 129−146.

Kahana, B., & Kahana, E. (1975). The relationship of impulse control to cognition and adjustment among institutionalized aged women. *Journal of Gerontology*, *30*(6), 679−687.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review*, *9*(4), 637−671.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology General*, *132*(1), 47−70.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 66−71.

Kendall, P. C., & Wilcox, L. E. (1979). Self-control in children: Development of a rating scale. *Journal of Consulting and Clinical Psychology*, *47*(6), 1020−1029.

*Kendall, P. C., Zupan, B. A., & Braswell, L. (1981). Self-control in children: Further analyses of the self-control rating scale. *Behavior Therapy*, *12*, 667−681.

Kirby, K. N. (1997). Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General*, *126*, 54−70.

Kirby, K. N., & Herrnstein, R. J. (1995). Preference reversals due to myopic discounting of delayed reward. *Psychological Science*, *6*(2), 83−89.

Kirby, K. N., & Marakovic, N. N. (1995). Modeling myopic decisions: Evidence for hyperbolic delay-discounting within subjects and amounts. *Organizational Behavior and Human Decision Processes*, *64*(1), 22−30.

Kirby, K. N., & Marakovic, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic Bulletin and 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*(4), 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*(1), 78−87.

*Kirby, K. N., Winston, G. C., & Santiesteban, M. (2005). Impatience and grades: Delay-discount rates correlate negatively with college GPA. *Learning and Individual Differences*, *15*(3), 213−222.

*Krietler, S., & Zigler, E. (1990). Motivational determinants of children's probability learning. *The Journal of Genetic Psychology*, *151*(3), 301−316.

Krikorian, R., & Bartok, J. A. (1998). Developmental data for the Porteus Maze Test. *The Clinical Neuropsychologist*, *12*(3), 305−310.

*Lessing, E. E. (1969). Racial differences in indices of ego functioning relevant to academic achievement. *The Journal of Genetic Psychology*, *115*, 153−167.

*LeSure, G. E. (1977). Relationship between intelligence and preferences to work for delayed rewards. *Psychological Reports*, *40*, 493−494.

Lieblich, A., Ben-Shahar, S., & Ninio, A. (Eds.). (1975). *The Wechsler test for children*. Jerusalem: The Hebrew University, The Institute for Developmental Psychology and The Ministry for Education and Culture, The Psychological Counseling Service.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*, *Vol. 49*. Thousand Oaks: Sage Publications.

*Macbeth, L. (1974). The ability to delay gratification: A trait–or not a trait. *Multivariate Behavioral Research*, *9*(1), 3−19.

Madden, G. J., Begotka, A. M., Raiff, B. R., & Kastern, L. L. (2003). Delay discounting of real and hypothetical rewards. *Experimental and Clinical Psychopharmacology*, *11*(2), 139−145.

Madden, G. J., Petry, N. M., Badger, G. J., & Bickel, W. K. (1997). Impulsive and self-control choices in opioid-dependent patients and non-drug-using control participants: Drug and monetary rewards. *Experimental and Clinical Psychopharmacology*, *5*(3), 256−262.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*(5695), 503−507.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*(1), 3−19.

Mischel, W. (1961). Delay of gratification, need for achievement, and acquiescence in another culture. *Journal of Abnormal and Social Psychology*, *62*, 543−552.

Mischel, W. (1974). Processes in delay of gratification. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, *Vol. 7* (pp. 249–292). New York: Academic Press.

Mischel, W., & Baker, N. (1975). Cognitive appraisals and transformations in delay behavior. *Journal of Personality and Social Psychology*, *31*(2), 254−261.

*Mischel, W., & Metzner, R. (1962). Preference for delayed reward as a function of age, intelligence, and length of delay interval. *Journal of Abnormal and Social Psychology*, *64*, 425−431.

Mischel, W., Shoda, Y., & Rodriguez, M. I. (1989). Delay of gratification in children. *Science*, *244*(4907), 933−938.

Mischel, W., & Staub, E. (1965). Effects of expectancy on waiting and working for larger rewards. *Journal of Personality and Social Psychology*, *2*, 625−633.

Mitchell, S. (1999). Measures of impulsivity in cigarette smokers and non-smokers. *Psychopharmacology*, *146*(4), 455−464.

*Monterosso, J., Ehrman, R., Napier, K. L., O'Brien, C. P., & Childress, A. R. (2001). Three decision-making tasks in cocaine-dependent patients: Do they measure the same construct? *Addiction*, *96*(12), 1825−1837.

Myerson, J., & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of the Experimental Analysis of Behavior*, *64*(3), 263−276.

Myerson, J., Green, L., Hanson, J. S., Holt, D. D., & Estle, S. J. (2003). Discounting delayed and probabilistic rewards: Processes and traits. *Journal of Economic Psychology*, *24*, 619−635.

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior*, *76*(2), 235−243.

Nelson, H. E. (1982). *The National Adult Reading Test (NART): Test manual.* Windsor: NFER-Nelson.

Newman, J. P., Kosson, D. C., & Patterson, D. S. (1992). Delay of gratification in psychopathic and nonpsychopathic offenders. *Journal of Abnormal Psychology*, *101*(4), 630−636.

Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H. (2005). Working memory and intelligence—Their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 61−65.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. E. (2002). Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, *14*(8), 1215−1229.

Ohmura, Y., Takahashi, T., Kitamura, N., & Wehr, P. (2006). Three-month stability of delay and probability discounting measures. *Experimental and Clinical Psychopharmacology*, *14*(3), 318−328.

Orwin, R. G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics*, *8*(2), 157−159.

Ostaszewski, P. (1996). The relation between temperament and rate of temporal discounting. *European Journal of Personality*, *10*, 161−172.

Please cite this article as: Shamosh, N. A., & Gray, J. R., Delay discounting and intelligence: A meta-analysis, *Intelligence* (2007), doi:10.1016/j.intell.2007.09.004

Ostaszewski, P., Green, L., & Myerson, J. (1998). Effects of inflation on the subjective value of delayed and probabilistic rewards. *Psychonomic Bulletin & Review*, *5*(2), 324−333.

Overton, R. C. (1998). A comparison of fixed-effects and mixed (random-effects) models for meta-analysis tests of moderator variable effects. *Psychological Methods*, *3*(3), 354−379.

Parker, A. M., & Fischhoff, B. (2005). Decision-making competence: External validation through an individual-differences approach. *Journal of Behavioral Decision Making*, *18*, 1−27.

*Paulsen, K., & Johnson, M. (1980). Impulsivity: A multidimensional concept with developmental aspects. *Journal of Abnormal Child Psychology*, *8*(2), 269−277.

*Petry, N. M. (2002). Discounting of delayed rewards in substance abusers: Relationship to antisocial personality disorder. *Psychopharmacology*, *162*, 425−432.

Pratt, T. C., & Cullen, F. T. (2000). The empirical status of Gottfredson and Hirschi's general theory of crime: A meta-analysis. *Criminology*, *38*(3), 931−964.

Prelec, D., & Loewenstein, G. (1991). Decision making over time and uncertainty: A common approach. *Management Science*, *37*(7), 770−786.

Prencipe, A., & Zelazo, P. D. (2005). Development of affective decision-making for self and other: Evidence for the integration of first- and third-person perspectives. *Psychologica Science*, *16*, 501−505.

Rachlin, H. (2000). *The science of self-control.* Cambridge, Mass: Harvard University Press.

Raudenbush, S. W. (1994). Random effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301−321). New York: Russell Sage Foundation.

Reynolds, B., & Schiffbauer, R. (2005). Delay of gratification and delay discounting: A unifying feedback model of delay-related impulsive behavior. *The Psychological Record*, *55*, 439−460.

Richards, J. B., Zhang, L., Mitchell, S. H., & de Wit, H. (1999). Delay or probability discounting in a model of impulsive behavior: Effect of alcohol. *Journal of the Experimental Analysis of Behavior*, *71*, 121−143.

*Rodriguez, M. L., Mischel, W., & Shoda, Y. (1989). Cognitive person variables in the delay of gratification of older children at risk. *Journal of Personality and Social Psychology*, *57*(2), 358−367.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, *86*(3), 638−641.

*Shoda, Y., Mischel, W., & Peake, P. K. (1990). Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions. *Developmental Psychology*, *26*(6), 978−986.

Sonuga-Barke, E. J., Dalen, L., & Remington, B. (2003). Do executive deficits and delay aversion make independent contributions to preschool attention-deficit/hyperactivity disorder symptoms? *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(11), 1335−1342.

Swets Test Publishers. (2000). *WAIS-III, Dutch version: Manual.* Lisse: Swets Test Publishers.

Tang, J. L., & Liu, J. L. Y. (2000). Misleading funnel plot for detection of bias in meta-analysis. *Journal of Clinical Epidemiology*, *53*, 477−484.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, *72*(2), 271−324.

Thompson, C., Barresi, J., & Moore, C. (1997). The development of future-oriented prudence and altruism in preschoolers. *Cognitive Development*, *12*, 199−212.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale*, (Fourth ed.) Chicago Riverside.

*Tweedie, J. (1967). Relationship of memory processes to delay of gratification. *Psychology in the Schools*, *4*(4), 324−328.

U.S. Census Bureau Historical income tables – people. Retrieved Jan 9, 2007, from http://www.census.gov/hhes/www/income/histinc/p07ar.html

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children—Revised.* New York: The Psychological Corporation.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale—Revised.* New York: The Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale: Administration and scoring manual.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence manual.* San Antonio, TX: Harcourt Assessment.

Weisz, J. R. (1978). Choosing problem-solving rewards and Halloween prizes: Delay of gratification and preference for symbolic reward as a function of development, motivation, and personal investment. *Developmental Psychology*, *14*(1), 66−78.

Williams, K. T. (1997). *Expressive vocabulary test.* Circle Pines, MN: American Guidance Service.

Zachary, R. A. (1991). *Shipley Institute of Living Scale: Revised manual.* Los Angeles: Western Psychological Services.

Please cite this article as: Shamosh, N. A., & Gray, J. R., Delay discounting and intelligence: A meta-analysis, *Intelligence* (2007), doi:10.1016/j.intell.2007.09.004