Evaluating the Iowa Gambling Task as a direct assessment of impulsivity with low-income children

Jessica D. Burdick*, Amanda L. Roy, C. Cybele Raver

Department of Applied Psychology, New York University, 196 Mercer St. 8th Floor, New York, NY 10012, USA

ABSTRACT

This study examined performance on the Iowa Gambling Task (IGT; Bechara, Damasio, Damasio, & Anderson, 1994) as a measure of low-income school-aged children's affective decision-making and considered its utility as a direct indicator of impulsivity. One hundred and ninety-three 8–11 year olds performed a computerized version of the Iowa Gambling Task, a validated measure of decision-making. Multi-level modeling was used to examine children's performance over the course of the task, with age, gender, and teachers' ratings of child impulsivity (BIS-11; Patton, Stanford, & Barratt, 1995) used to predict children's Iowa Gambling performance. Higher impulsivity scores predicted a decrease in slope of Iowa Gambling performance, indicating students rated higher on impulsivity chose more disadvantageously across the task blocks. Results support evidence of the validity of the Iowa Gambling Task as a measure of impulsivity in low-income minority children.

Keywords: Decision-making Impulsivity Reward sensitivity Iowa gambling task

1. Evaluating the Iowa Gambling Task as a direct assessment of impulsivity with low-income children

Living in poverty is associated with the development of negative behavioral and emotional problems in children (Noble, McCandliss, & Farah, 2007; Takeuchi, Williams, & Adair, 1991). Research has shown that a key individual difference that distinguishes children's likelihood of avoiding costly negative outcomes such as delinquency, substance abuse, and school failure in the face of adversity may be their ability to control their impulses. That is, some low-income children may be at higher risk for experiencing problems in their schools and communities due to problems of behavioral disinhibition, while other children are better able to take advantage of learning opportunities because of higher levels of behavioral and emotional control (Raver et al., 2011). Emerging work in the fields of affective neuroscience and developmental science offer the prospect of a cohesive means by which to study impulse control (defined as inhibiting an automatic response in order to successfully complete a goal) (Bezdjian, Baker, Lozano, & Raine, 2009). Specifically, the Iowa Gambling Task (IGT; Bechara, Damasio, Damasio, & Anderson, 1994) has been viewed by many investigators as an excellent means of tapping “hot cognition” (affective decision-making) and holds potential as a direct assessment measure of impulsivity (Bubier & Drabick, 2008). Yet few studies have used assessments such as the Iowa Gambling Task to examine this key form of self-regulation among samples of low-income children (see Bubier & Drabick, 2008).

In the following study, we aim to address this empirical gap, given clear evidence that the optimal development of impulse control may be jeopardized by environmental stressors associated with chronic poverty (Blair & Raver, 2012; Noble et al., 2007). For example, several studies (de Wit, Flory, Acheson, McCloskey, & Manuck, 2007; Green, Myerson, Lichtman, Rosen, & Fry, 1996; Harrison, Lau, & Williams, 2002) have found lower income to be associated with greater impulsiveness and more delay-discounting, or preferring smaller rewards, sooner, over larger but delayed rewards. The development and validation of direct assessments of impulsiveness for use with children from lower income families may help researchers better understand both the correlates and consequences of individual differences in this key form of self-regulation for children facing economic adversity. In addition, little is known regarding children's performance on the IGT, particularly among samples of children facing higher environmental risk. The following paper seeks to address these questions, by assessing impulsiveness vs. impulse control using multiple methods (including teacher report and performance on the IGT) among a large, ethnic minority, low-income sample of students in urban elementary schools.

* Research reported in this publication was supported by Award R01HD046160 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

* Corresponding author. Tel.: +1 (212) 998 5647.
E-mail addresses: jdb513@nyu.edu (J.D. Burdick), alr260@nyu.edu (A.L. Roy), ccr4@nyu.edu (C.C. Raver).

© 2013 Elsevier Ltd. All rights reserved.
1.1. The Iowa Gambling Task

The IGT is a computerized card game commonly used to measure risky decision making tendencies or individuals' sensitivity to reward and loss. As a simulation of real-life decision-making, the IGT involves weighing “expected but uncertain rewards and penalties” (Franken, van Strien, Nijs, & Muris, 2008). Participants must choose from four decks of cards across 50 trials, with the goal of acquiring as much money as possible. Decks C and D consistently provide smaller wins of $50 and also lower levels of net loss over time, making these decks advantageous. In comparison, decks A and B consistently give out larger wins of $100 but have substantially higher levels of net loss over time, making these decks disadvantageous. Additionally, decks B and C produce more frequent, but smaller losses, while decks A and D produce infrequent but larger losses. Importantly, while rewards remain consistent, losses vary across trials, and the type of loss varies between decks. Analysis of participant choices across trials reveals their tendencies toward risk taking and abilities to weigh future outcomes.

Two types of indices can be calculated that measure participants’ (a) tendency to choose advantageously (where the choice of decks is likely to yield smaller rewards for each card drawn, but minimizes larger losses, over time) and their (b) tendency to select from decks that offer infrequent (though larger) loss. The first index, measuring long-term consequence, represents this tendency to play for “lower stakes” with the benefit of a more advantageous overall and indicates if a participant understands long-term effects of certain deck-choosing strategies. For this first index, an increase in positive values across the task indicates the participant's preference for “good” decks that promise a smaller gain for each card that is “dealt,” but that incrementally lead to greater total gain and lower loss, over time. The second index, a bias for infrequent loss, indicates the number of deck choices that lead to infrequent-but-larger losses relative to the number of choices that lead to frequent-but-smaller losses (for this second index, positive scores indicate more infrequent-loss choices) (Hooper, Luciana, Conklin, & Yarger, 2004). Both indices are calculated across five blocks of task trials.

Findings on the IGT long-term consequence index suggest that children may be sensitive to the anticipatory experience of reward, and consequently may be less good at playing the game in terms of avoiding losses (Crone, Jennings, & van der Molen, 2004). In short, children may struggle with a tendency to focus on immediate outcomes rather than future consequences (Hooper et al., 2004; Overman, 2004). This is in keeping with the landmark hypothesis proposed by (Damasio, Tranel, and Damasio, 1991), that the prospect of winning (or losing) a large sum of money can serve as a “primary inducer,” whereby a somatic state of pleasure (at the possibility of winning) or discomfort (at the prospect of losing) may non-consciously drive decision-making. While healthy adults and adolescents tend to learn to choose advantageously across the task (Overman, 2004; Sweitzer, Allen, & Kaut, 2008; Upton, Bishara, Ahn, & Stout, 2011), some research has shown that younger children tend to select disadvantageously (Blair, Colledge, & Mitchell, 2001; Hooper et al., 2004; Huizenga, Crone, & Jansen, 2007). One notable exception is a study of 8-year-olds employing a 280-trial child version of the IGT in which children learned to choose advantageously in later task blocks (Carlson, Zayas, & Guthornsen, 2009). Children also tend to have a bias for larger, infrequent-loss decks on the IGT, preferring those decks to the decks that provide smaller, but more frequent losses (Carlson et al., 2009; Hooper et al., 2004; Huizenga et al., 2007). The majority of studies that have examined gender differences on the IGT among children or adolescents have found no differences between boys and girls (e.g. Carlson et al., 2009; Hooper et al., 2004). Generally, these studies have taken place in laboratory settings where the strengths of high precision in measurement and experimental control have been balanced against constraints such as limited generalizability to broader samples or other populations (Enticott, Ogloff, & Bradshaw, 2006).

In stark contrast, large scale survey studies on risky behavior in American youth have tried to address empirical questions regarding individuals' predispositions towards impulsive or risky behavior, by using parent- and teacher report. One strength of that measurement approach is that it can easily be deployed in large school- and community-based studies. However, this approach lacks precision relative to direct assessments such as the IGT, and may also suffer from reporter bias or lack objectivity (Arnold & Feldman, 1981; Kroes, Veerman, & De Bruyn, 2003). One of the benefits of the IGT is that it offers a more empirically precise “lens” into processes of decision-making in the contexts of reward and loss, and as a direct assessment of child skill, may be more “objective” than teacher- or parent-report measures. In this study, we examine whether the IGT may provide a promising resolution to some of these concerns.

In sum, the following study examines ways that the IGT, developed as a measure of decision-making in adults, may offer promise for the assessment of affective decision-making among low-income, ethnic minority children, serving as a valid means of measuring impulsivity directly in field-based settings. Accordingly, this study aims to address the following questions:

1. How do low-income children perform on the IGT administered in a field-based setting?
2. Is IGT performance related to teachers' ratings of child impulsivity, after adjusting for age and gender?

Consistent with previous research demonstrating that children have difficulty weighing future outcomes of decks’ various reward and loss intervals (Crone et al., 2004; Hooper et al., 2004; Overman, 2004), we predict that children in this sample will choose in ways that highlight their somatic preferences for immediate reward but that may lead to disadvantageous long term outcome across the IGT. Further, we expect children will have a strong preference for infrequent-loss decks over frequent-loss decks, as found in prior studies. We base this hypothesis on the possibility that children may focus on the frequency of the loss as the most salient feature of the task. In so doing, children may have a more difficult time with mastering a key cognitive demand of IGT, namely that winning involves attending to two dimensions, rather than a single dimension of the task (e.g. both the loss's frequency and the magnitude; Huizenga et al., 2007). Finally, we predict that higher teacher-rated impulsivity scores will be associated with lower long-term consequence scores, or more disadvantageous choices, across the IGT. We base this prediction on our theory that the IGT can serve as a valid direct measure of impulsivity among this sample, and thus should be related to teachers' subjective reports of child impulsivity (assessed with the BIS, a well-known measure of impulsivity).

2. Method

2.1. Sample

Data for this study comes from the Chicago School Readiness Project (CSRP), a socioemotional intervention trial implemented in preschool programs located in high-poverty Chicago neighborhoods. The current study sample consists of 193 children who took part in a follow-up wave of data collection. Assessors administered the IGT to individual students in Chicago Public Schools using laptop computers. The majority of participants were African-Ameri-
can (48.9%) or Hispanic (43.3%), 48.2% were male, and the average age was 9.89 (S.D. = 0.72). 33.7% of the sample reported “income-to-needs risk” (see Table 1) at the time of data collection. A small proportion of participants (N = 29, 15%) were missing information on the BIS. In order to assess demographic differences as a function of missingness, the likelihood of having a missing value was predicted from a set of demographic characteristics. As the results revealed no differences, the decision was made to impute BIS values at their mean.

2.2. Measures

2.2.1. IGT

Prior piloting of the Hungry Donkey and 100-trial version of the IGT suggested that these versions were too long to keep children engaged. Therefore, we used a shortened version of the IGT consisting of five 10-trial blocks. The learning of long-term consequences (LTC) index was calculated by subtracting the number of disadvantageous deck choices from the number of advantageous deck choices ([C + D] − [A + B]). The bias for infrequent loss (IFL) index was calculated by subtracting the participants’ number of frequent-loss deck choices from the number of infrequent-loss deck choices ([A + D] − [B + C]). Indices were calculated within each 10-trial block.

2.2.2. Barratt Impulsiveness Scale

The Barratt Impulsiveness Scale Version 11 (BIS-11) is a 30 item self-report questionnaire of impulsivity in adults (Patton, Stanford, & Barratt, 1995). Measurement work in a previous time point of the study (McCoy, Raver, Lowenstein, & Tirado-Strayer, 2011) demonstrated the validity of the 7-item, teacher-reported version of the BIS used here. A total impulsiveness score was created by averaging the seven BIS items, rated on a four-point Likert scale (0–3) (M = 1.25, S.D. = 0.78, α = 0.91). BIS scores were grand-mean centered for analyses.

2.3. Data analysis

Multi-level models were run in HLM 7 (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011) were used to examine children’s IGT deck choice trajectories. Although a three-level model was initially considered because of clustering of students within schools, intraclass correlation coefficients revealed a minimal amount of variability in outcomes (3% of LTC, 1% of IFL) at the school level. Therefore, a two-level model was used, in which index scores across blocks were nested within individuals. At Level 1, the index score (LTC or IFL), Ybi, is a function of the linear model specified in Eq. (1).

\[ Y_{bi} = \pi_{o1} + \pi_{1b} X_{bi} + e_{bi} \]

Table 1
Sample demographic characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Analytic sample (N = 193)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>9.89 (0.72)</td>
</tr>
<tr>
<td>Male</td>
<td>48.2%</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td></td>
</tr>
<tr>
<td>African-American/Black</td>
<td>48.9%</td>
</tr>
<tr>
<td>Hispanic/Latino/Latina</td>
<td>43.3%</td>
</tr>
<tr>
<td>White/Non-Hispanic</td>
<td>4.5%</td>
</tr>
<tr>
<td>Bi-racial</td>
<td>2.8%</td>
</tr>
<tr>
<td>Other</td>
<td>0.6%</td>
</tr>
<tr>
<td>Income-to-needs risk</td>
<td>33.7%</td>
</tr>
</tbody>
</table>

*Income-to-needs risk is calculated using the ratio of participants’ family income to the Census Poverty Threshold for family size and number of children. At risk refers to families whose total income falls below poverty threshold for their household size.

for \( i = 1, \ldots, N \) subjects, \( \pi_{o1} \) is the intercept, or the index score at block 1 (the block variable was coded to range from 0 to 4) for subject \( i \), \( \pi_{1b} \) is the task block, \( \pi_{1i} \) is the growth rate in index scores across blocks for subject \( i \), and \( e_{bi} \) is random error.

Eq. (2) specifies the unconditional Level 2 model, without predictors entered.

\[ \pi_{o1} = B_{o0} + \gamma_{o1} \]
\[ \pi_{1i} = B_{i0} + \gamma_{1i} \]

In these equations, the individual growth parameters are predicted as a function of the average sample intercept (\( B_{o0} \)) and growth rate (\( B_{i0} \)) and their random variation (\( \gamma_{o1} \) and \( \gamma_{1i} \)).

To examine whether IGT performance differed by age, gender, or teacher-rated impulsivity, these individual-level characteristics were entered as predictors into the model at Level 2.

3. Results

3.1. Question 1: How do children perform over the course of the task?

On average, children had negative LTC scores in blocks two through five, although these scores were not significantly different from zero, with the exception of block five (t (192) = −2.126, p = 0.035) (Fig. 1). Additionally, the index showed a general decrease across blocks; participants chose more often from the disadvantageous, “larger short-term gain, with higher long-term loss” decks, and continued to choose more frequently from these “bad” decks as the task progressed. However, the lack of significance of these scores in blocks 1–4 signifies the number of disadvantageous choices is not statistically greater than advantageous choices until the end of the task. Participants had positive IFL scores, which tended to increase across task blocks (Fig. 2), signifying that participants chose more often from the decks with infrequent but larger losses than from decks with smaller, more frequent losses. IFL scores across each task block were all significantly different from zero at the p < 0.001 level; participants’ tendency to choose infrequent-loss decks was significantly greater than their tendency to choose from frequent-loss decks during each task block.

To determine whether children’s IGT performance changed over the course of the task and whether there was variation in response strategies, two separate multi-level unconditional growth models were run, in which time (i.e. block) was used to predict the LTC and the IFL indices. Results revealed a small negative but significant coefficient for the LTC slope (\( B = −0.13, SE = 0.06, p = 0.03 \)), indicating that on average children made more disadvantageous choices over time (Table 2, Model 1). The coefficient for the LTC intercept was non-significant indicating that on average, LTC was not different from 0 at block 1. There was significant random variation around both the LTC intercept (\( \chi^2 (192) = 237.99, p = 0.01 \)
and slope ($\chi^2 (192) = 240.22, p = 0.01$), demonstrating that children varied in how they chose in block 1 and how they continued to choose over the course of the task, variation that may be explained by individual-level differences. Results also revealed a positive growth rate ($B = 0.28, SE = 0.05, p = 0.000$) in IFL scores, indicating that participants increasingly favored infrequent-larger losses over frequent-smaller losses. However, only variance around the intercept of the IFL index was significant ($\chi^2 (192) = 329.62, p = 0.000$), meaning there was variation in how children chose on block one, but no significant variation across task blocks Table 3.

3.2. Question 2: Is teacher-rated impulsivity related to performance?

Child impulsivity, age, and gender were entered into the Level 2 model as predictors of LTC and IFL indices (Table 2). After adjusting for gender and age, impulsivity was not significantly related to the LTC intercept. However, impulsivity was negatively related to the LTC slope ($B = -0.159, r = 0.078, p = 0.044$); children rated as more impulsive by teachers scored more disadvantageously across the course of the IGT. Predicted values graphed +/- 1 standard deviation for impulsivity are displayed in Fig. 3. In addition, there continued to be significant variation around both the slope ($\chi^2(189) = 234.01, p = 0.014$) and intercept ($\chi^2 (189) = 231.48, p = 0.019$), suggesting additional individual differences may need to be considered to understand variation in LTC performance. After controlling for age and gender, BIS scores did not predict the IFL intercept or slope and there continued to be variation around the intercept, ($\chi^2(189) = 326.40, p = 0.000$), but not the slope.

4. Discussion

This work demonstrates the utility of using the IGT as a field-based, direct assessment of children’s sensitivity to reward and loss, as well as a good indicator of impulsivity. In our sample of low-income 8–11 year-olds, our results are in keeping with a key component of the somatic marker hypothesis proposed by Damasio et al. (1991), in that the children in our sample tended to make choices based on the prospect of more immediate, larger rewards even when this choice led to more disadvantageous outcomes over the course of the task. This finding parallels previous research that has found that children in this age range do not learn to choose from the “safer,” more advantageous decks, but instead tend to decrease their selection from those more advantageous decks, across the task (Hooper et al., 2004; Huizenga et al., 2007). We also found that children’s IFL scores increased over the course of the task, indicating that children in our sample continued to select cards from the decks that promised bigger winnings, even though those decks also favored infrequent-but-larger losses, over time. This is consistent with previous findings (Carlson et al., 2009; Hooper et al., 2004).

These findings are thought-provoking: Why would children choose less well over time, making consecutively fewer advantageous deck choices? To answer this question, we highlight that there are three different aspects of the IGT that can affect performance: reward amount, loss amount, and the schedule of rewards and losses. To perform advantageously on the task, children must consider both decks’ losses and reward amounts. Advantageous decks accumulate winnings over time, but have smaller immediate rewards ($50) compared to disadvantageous decks, which have larger immediate rewards ($100) but lead to losses over time. As would be predicted by the somatic marker hypothesis in conditions where connectivity between orbitofrontal cortex, amygdala, and prefrontal cortex may not be fully mature, children may have been particularly attracted to the high reward, infrequent loss decks, while failing to process the overall net score that accumulates over time, therefore producing a negative LTC and positive IFL slope across blocks. This hypothesis is congruent with previous work suggesting that younger children focus on immediate outcomes and less on future consequences (Hooper et al., 2004; Huizenga et al., 2007; Overman, 2004). These findings also suggest that the IGT task places significant demand on children’s ability to balance multiple dimensions of a cognitive problem in mind. Children were aware of the differential loss schedule and were actively avoiding frequent-loss schedule decks, but may have had trouble realizing that the infrequent loss decks actually led to fewer total winnings over time.

Table 2

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Unconditional growth</th>
<th>Age &amp; gender</th>
<th>Age, gender, BIS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
<td>B</td>
</tr>
<tr>
<td>Model for initial level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean initial level</td>
<td>0.04</td>
<td>0.14</td>
<td>0.21</td>
</tr>
<tr>
<td>Male</td>
<td>-0.37</td>
<td>0.29</td>
<td>-0.54</td>
</tr>
<tr>
<td>Age</td>
<td>0.12</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>BIS score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model for growth rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean growth rate</td>
<td>-0.13*</td>
<td>0.06</td>
<td>-0.07</td>
</tr>
<tr>
<td>Male</td>
<td>-0.14</td>
<td>0.12</td>
<td>-0.06</td>
</tr>
<tr>
<td>Age</td>
<td>-0.05</td>
<td>0.08</td>
<td>-0.06</td>
</tr>
<tr>
<td>BIS score</td>
<td>-0.16*</td>
<td>0.08</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effects</th>
<th>Variance component</th>
<th>$\chi^2$ (df)</th>
<th>Variance component</th>
<th>$\chi^2$ (df)</th>
<th>Variance component</th>
<th>$\chi^2$ (df)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level-1 error</td>
<td>5.39</td>
<td>5.39</td>
<td>Initial status</td>
<td>237.99 (192)*</td>
<td>0.78</td>
<td>235.60 (190)*</td>
</tr>
<tr>
<td>Growth rate</td>
<td>0.14</td>
<td>0.14</td>
<td>Growth rate</td>
<td>240.22 (192)**</td>
<td>0.14</td>
<td>238.13 (190)**</td>
</tr>
</tbody>
</table>

*p < 0.05, **p < 0.01, ***p < 0.001.
Lastly, we found clear evidence to support our hypothesis that individual differences in children's performance on the IGT would be significantly associated with teachers' reports of their impulsivity in the "real world" setting of the classroom. Our results indicated that teachers' reports on the BIS indeed negatively and significantly predicted children's performance. While overall, children were making disadvantageous choices, children rated as more impulsive by their teachers were making more disadvantageous choices than less impulsive children. This may suggest that over the course of the task less impulsive children were beginning to process more of the task's informational cues. Rather than focusing solely on immediate rewards, less impulsive children may have begun to understand the importance of considering both rewards and losses, while more impulsive children continued to focus on only one dimension of the task. These results indicate that the IGT may be a valid direct measure of impulsivity.

There are several potential limitations to our study. First, we used a 50-trial, rather than the standard 100-trial, version of the IGT. It is possible that the brevity of the task does not enable enough time for children to learn deck payoffs and make informed choices. However, in piloting the 100-trial IGT with similarly-aged children, we found that children failed to stay engaged throughout the entire task. This, in combination with concerns about the feasibility of completing an extended battery of assessments with children in school-settings, prompted us to use a shorter version. A second limitation was the use of a shortened version of the BIS. This concern is somewhat assuaged by measurement work demonstrating the validity of a shortened BIS and BRIEF (Behavior Rating Inventory of Executive Function; Gioia, Isquith, Guy, & Kenworthy, 2000) composite using data from a prior wave of data collection with this sample (McCoy et al., 2011). Although concerns about generalizability remain, the items used here were selected as salient dimensions of classroom impulsivity and arguably still capture the construct well. Another limitation is the lack of a higher-income comparison sample. However, one of our study strengths is the examination of the IGT, typically tested with higher-income, white samples, in a sample of low-income, ethnic-minority children. Our findings parallel prior work demonstrating an increase in disadvantageous choices across the task (declines in LTC), and as such extend the generalizability of the IGT among more diverse samples.

Several important questions should be explored further: for example, significant variation in the LTC slope remained after adjusting for age, gender, and impulsivity. This suggests that other individual characteristics, such as personality or temperament, may further explain variation in performance. In addition, triangulation through multiple methods of assessment is needed to more accurately measure key dimensions of children's self-regulation. The present study demonstrates the utility and validity of using the IGT in field-based settings and with low-income, minority children. Improving the measurement of impulsivity among children from diverse backgrounds in "real-world" contexts is necessary in order to better inform research-based policy and practice decisions that directly affect children in low-income communities.

References


