JIDR Journal of Intellectual Disability Research
Published on behalf of mencap and in association with IASSID

Journal of Intellectual Disability Research doi: 10.1111/j.1365-2788.2010.01260.x

VOLUME 54 PART 4 pp 380–391 APRIL 2010

Evaluation of the ability of people with intellectual disabilities to 'weigh up' information in two tests of financial reasoning

P. Willner,[1,2] R. Bailey,[1] R. Parry[1] & S. Dymond[2]

1 *Directorate of Learning Disability Services, Abertawe Bro Morgannwg University Health Board, Bridgend, UK*
2 *Department of Psychology, Swansea University, Swansea, UK*

Abstract

Background An assessment of mental capacity includes an evaluation of the ability to 'weigh up' information, but how to do this is uncertain. We have previously used a laboratory decision-making task, temporal discounting, which involves a trade-off between the value and the delay of expected rewards. Participants with intellectual disabilities (ID) showed very little evidence of 'weighing up' of information: only a third of participants showed consistent temporal discounting performance, and when present, consistent performance was usually impulsive; and the ability to perform consistently was more strongly related to executive functioning than to IQ. The aim of the present study was to replicate these observations and extend them to a more realistic financial decision-making task.
Methods We administered a temporal discounting task and a financial decision-making task, as well as tests of executive functioning and IQ, to 20 participants who attended day services for people with learning disabilities (mean Full-Scale IQ = 59), and to 10 staff members.
Results Performance in both decision-making tasks was related more strongly to executive functioning than to IQ. In both tasks, decisions by service users were made largely on the basis of a single item of information: there was very little evidence in either task that information from two sources was being 'weighed'.
Conclusions The results suggest that difficulty in 'weighing up' information may be a general problem for people with ID, pointing to a need for psycho-educational remediation strategies to address this issue. The importance of executive functioning in decision-making by people with ID is not recognized in the legal test for mental capacity, which in practice includes a possibly irrelevant IQ criterion.

Keywords decision-making, executive functioning, intellectual disability, IQ, mental capacity, temporal discounting

Correspondence: Professor Paul Willner, Community Support Team, The Laurels, 87 Lewis Road, Neath SA11 1DJ, UK (e-mail: p.willner@swansea.ac.uk).

Introduction

The Mental Capacity Act (2005), which applies in England and Wales, introduced a formal legal definition of mental capacity: the ability to understand and retain information relevant to a decision, weigh-up the information for the purposes of making the decision, and communicate the decision. The issue of communication does not usually arise (as this applies mainly to people with 'locked in' syndrome, it being assumed that others, if they are able to reach a decision, would also be able to communicate it by one means or another) but the other three aspects of capacity, understanding, retention and 'weighing up' need to be assessed when capacity is questioned. While it is relatively straightforward to assess understanding and retention of information, the same cannot be said of the ability to 'weigh up' information. Indeed, in relation to people with intellectual disabilities (ID), there is a profound dearth of research on reasoning abilities. The British Psychological Society's guidance on the assessment of capacity provides a useful summary of factors that may influence reasoning, and much specific advice on the assessment of understanding and retention of information, but there is almost nothing on the assessment of the ability to 'weigh up' information (British Psychological Society 2006).

The need to 'weigh up' information arises when it is necessary to evaluate two or more items of information relative to one another: for example, evidence in favour of a particular choice and evidence against. If the two elements are expressed in the same dimensions, then the problem of 'weighing up' is trivial: for example, provided the information is well understood, the choice between paying £10 or £100 for the identical outcome is almost automatic. It is when there is a need to integrate information from two or more sources that are expressed in different dimensions that weighing-up of information becomes a problem (Green & Myerson 2004, p. 569).

In an earlier study, we examined this issue using an adapted version of a popular laboratory decision-making task, temporal discounting (TD), in which decisions are based on both the magnitude and the delay of an expected reward. In TD tasks, participants are presented with a series of choices between small immediate rewards and large delayed rewards. The overwhelming majority of intellectually able adult participants produce an orderly trade-off between magnitude and delay, such that larger rewards are preferred at short delays, but smaller rewards are preferred when there is a long wait for the larger alternative (Chapman & Elstein 1995; Critchfield & Kollins 2001; Green & Myerson 2004; Wileyto *et al.* 2004; Reynolds 2006; Estle *et al.* 2007). There exists a small literature on reward choice in people with ID (Kopp 1990; Dixon *et al.* 1998; Vollmer *et al.* 1999; Cuskelly *et al.* 2001, 2003), but earlier use of the TD procedure amounts to two studies involving extensive training with a total of eight children (Ragotzy *et al.* 1988; Schweitzer & Sulzer-Azaroff 1988). Using a version of the TD task designed for young children to use without training (Scheres *et al.* 2006), we found that, unlike intellectually able participants, most participants with ID either behaved randomly, or used only a single source of information, almost always choosing the immediate alternative (or in one case, the larger alternative). There was very little evidence that participants were taking both sources of information into account and 'weighing' them (Willner *et al.* 2010a).

The main aim of the present study was to replicate these observations and to examine whether a similar outcome would be seen in a second, more realistic, financial decision-making task (FDMT). The task chosen was developed for use with people with ID (Suto *et al.* 2005a). It consists of five scenarios, of increasing complexity, describing choices that people identified in the scenario need to make, each followed by a structured interview designed to elucidate five aspects of the decision: identification of the relevant information, understanding, reasoning, appreciation of who is affected by the decision and communication. We focused here on the two elements most relevant to 'weighing up' of information, understanding and reasoning.

In the previous study, we also found that two aspects of performance in the TD task, the ability to perform non-randomly and the ability to benefit from training to increase the consistency of performance, were both related more strongly to performance on tests of executive functioning (EF) than to IQ (Willner *et al.* 2010a). Therefore, a second aim of the present study was to replicate this observation in the TD task and to establish whether reasoning ability in the FDMT is also more dependent on EF than on IQ. A replication of this observation across tasks would suggest a more general hypothesis that EF, rather than IQ, underpins reasoning abilities in people with ID.

Method

Participants

The participants were 20 people [55% male, with a mean (SD) age of 39.9 (9.6)], who attended day services for people with mild to moderate learning disabilities. (The term 'learning disability' is used in the UK to refer to people with significant impairments of both intellectual and functional abilities, acquired in childhood. Participants' disabilities were of mixed etiology, and the etiology was typically unknown.) A control group comprised 10 staff members [20% male, with a mean (SD) age of 38.6 (12.4)]. All participants provided informed consent and the study was approved by the Local NHS Research Ethics Committee. Potential service user participants were screened using a simple test of financial knowledge (Coins and Costs: Willner *et al.* 2010a), with a threshold value of 4. All potential participants met this criterion.

Intellectual ability

Service users were assessed for intellectual ability using the Wechsler Abbreviated Scale of Intelligence (WASI), so as to confirm that they met the IQ criterion for a diagnosis of 'learning disability', and for receptive language ability using the British Picture Vocabulary Scale (2nd edition) (BPVS). Participants' mean (SD) scores were: Full-Scale IQ, 58.8 (3.6; range 50–67) and BPVS raw score, 81.6 (20.5). Controls were assessed using the BPVS: the mean (SD) raw score was 152.2 (6.4).

Executive functioning

Executive functioning was evaluated, in service users only, using two batteries of tests, the Children's version of the Behavioural Assessment of the Dysexecutive Syndrome (BADS-C: Emslie *et al.* 2003), which includes six tests of EF, and the Cambridge Executive Functioning Assessment for people with Intellectual Disability (CEFA-ID: Ball *et al.* 2008), which consists of eight EF tests (including two tests of 'executive memory') and four memory tests, the results of which are not reported (Ball *et al.* 2008). The procedure for the BADS-C was as described in the user guide (Emslie *et al.* 2003) with some minor modifications: in addition to the written versions, the rules for the Playing Cards and Six-Part tests were also provided in a visual format; and in the Six-Part Test, participants were asked to make their responses verbally rather than in writing. These modifications were made in order to increase accessibility and decrease dependence on literacy. The CEFA-ID EF tests were administered as described by Ball *et al.* (2008).

Temporal discounting

The TD task was programmed in E-Prime and presented on a laptop computer. The task consisted of a series of choices between small immediate and large delayed rewards, which were presented on the screen, while the experimenter also provided the same information verbally. Feedback was also presented both visually and verbally.

On each trial the participant saw a picture of a long road winding into the distance up the screen and was offered a choice between a truck at the bottom of one side of the screen and a second truck some way up the road at the opposite side of the screen. The closer truck carried either '£2', '£4', '£6' or '£8' to be delivered 'Now', while the more distant truck carried '£10' to be delivered after '1 day', '2 days', '4 days', '1 week' or '2 weeks', with increasing times represented by a smaller picture higher on the screen (i.e. further away). If the participant chose the immediate alternative, the message 'Well done! You have won £x' was presented immediately. If the delayed alternative was chosen, the message 'Well done! You have won £10 but you will have to wait for it' was presented, after a symbolic delay of 1–6 s that increased as a function of the actual delay. The test was preceded by a first block of practice trials in which the participant was asked to choose between immediate delivery of different amounts of money (£10 vs. £2/4/6/8), a second practice block where the choice was between delayed deliveries of £10 (Now vs. 1/2/4/7/14 days), and a third practice block of 20 choice trials where the two alternatives varied in both value and delay. On the first two blocks of practice trials (which unlike later choice trials have a right and a wrong answer), feedback, and in the case of incorrect responses, an explanation, were provided after each trial, as this has been found to improve choice performance in the main task (Willner *et al.* 2010a).

Scoring

We have previously reported that a proportion of participants tested in this task produce an essentially random pattern of choices, while others behave in a more orderly manner, and that participants who behave non-randomly vary greatly in the extent to which their choices reflect the use of both sources of available information (money and time). Two measures were developed to capture these features, *Complexity*, which measures the extent to which participants use both choice dimensions and *Inconsistency*, which measures the extent to which choices are made in a disorderly fashion (Willner *et al.* 2010a). These two measures were used to categorize participants as performing randomly or consistently (see below: Statistical Analysis).

For participants who performed consistently, a TD function was constructed by determining the indifference point at each time delay, defined as the monetary value at which the immediate alternative was chosen on 50% of trials (for further details, see Willner *et al.* 2010a). The indifference points were summed to provide an 'area under the curve' (AUC: Myerson *et al.* 2001). AUC values were used to categorize participants as 'impulsive' (AUC = 0–10), 'orderly' (AUC = 11–49) or 'acquisitive' (AUC = 50–60).

Financial decision-making

The FDMT was presented and scored as described by Suto *et al.* (2005a). It consisted of five vignettes, of the same form but graded in difficulty, which describe a situation where a choice has to be made ('In the supermarket'; 'Buying jeans'; Going to work'; 'The noisy car'; 'Selling shares'). For example, the third scenario explains that John, who works as a gardener earning £5 per h, does not feel like going to work today because it is raining. After each scenario, a series of questions is asked, which are scored for Identification (of the choice to be made and who has to make it), Understanding, Reasoning, Appreciation (of who is affected) and Communication. Each question is scored out of two, making a maximum score of 10 for each of the five vignettes or for each of the five types of decision-making ability, and a maximum total score of 50.

Procedure

Participants first provided informed consent; which was witnessed by a staff member. Tests were conducted over five sessions, in the order: session 1: WASI, BPVS, Coins & Costs; session 2, TD; session 3, FDMT; session 4: BADS-C; session 5: CEFA-ID. The order of testing was not randomized in order to simplify the analysis, and because the potential for carry-over effects was considered to be minimal.

After each session, participants were asked to rate their experience of the session on a 3-point scale ('I didn't like it'; 'I liked it a bit'; 'I liked it a lot'). In total, 83 of the 100 sessions were rated as 'liked a lot' and 16 as 'liked a bit'; only one session was rated as 'not liked'.

Controls were tested over two sessions on the BPVS, the TD task and the FDMT.

Statistical analysis

A cluster analysis of complexity and inconsistency scores in the TD task was first applied to identify consistently and randomly performing groups of service users (see Willner *et al.* 2010a). Subsequently, independent *t*-tests (one-tailed, as this was a replication of earlier results) and stepwise logistic regression were used to compare the test performances of these two groups on other experimental measures. A χ^2 test was used to compare profiles of performance between service users and controls. In order to increase the size of the data set for the profile analysis, the data were combined with those from a comparable study (Willner *et al.* 2010a) involving 20 service users and 20 controls with very similar demographic characteristics to the present participants (service users: 50% male, mean age 40.7, mean FSIQ 59.8, mean BPVS score 84; staff: 25% male, mean age 41.7, mean BPVS score 149).

For the FDMT, on most comparisons between service users and controls, service users were significantly more variable. Therefore, these comparisons were made using non-parametric Mann–Whitney *U*-tests. Further tests on data from service users included one-way analysis of variance to analyze performance across the five decision-making scenarios, Pearson product–moment correlation coefficients and stepwise multiple regression to

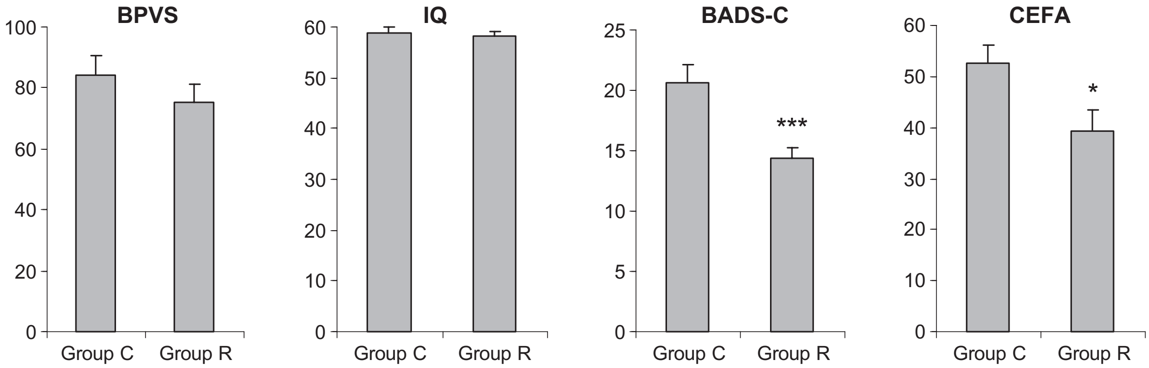

Figure 1 Comparison of performance on cognitive measures for the groups performing consistently (C) and randomly (R) on the temporal discounting task. $*P < 0.05$; $***P < 0.001$. BPVS, British Picture Vocabulary Scale; BADS-C, Children's version of the Behavioural Assessment of the Dysexecutive Syndrome; CEFA, Cambridge Executive Functioning Assessment.

examine the relationship between scores on the FDMT and cognitive variables, and χ^2 tests to compare the quality of financial reasoning between service users and controls.

Results

Temporal discounting

Cluster analysis identified a set of 15 service users with low complexity and inconsistency scores, and a smaller set of five service users with high scores on both measures, whose performance was essentially random. One service user was transferred from the consistent to the random group on the basis that this participant's low scores were achieved by always pressing the same response key. With the exception of this participant, all participants performed well above chance (= 18 errors) on unidimensional practice trials. Largely because of this participant, the random group made more practice errors than the consistent group, but the difference was in any case not significant [mean (±SEM) = 7.3 (±2.6) vs. 3.2 (±1.0), $t(18) = 1.5$, NS].

Role of executive functioning

Figure 1 compares the performance of the two groups on measures of intellectual ability and EF. There was no significant difference in IQ ($t(18) = 0.75$, NS) or receptive language ability ($t(18) = 0.91$, NS). However, the consistent group had significantly higher scores on both the BADS-C ($t(18) = 3.74$, $P < 0.001$) and the CEFA ($t(18) = 1.95$, $P < 0.05$).

In a logistic regression analysis to examine the ability of IQ, BPVS and EF to predict group membership, the EF score was retained as a significant predictor (χ^2 if removed: BADS-C = 7.09, $P < 0.01$; CEFA = 3.68, $P = 0.055$), but the IQ and BPVS scores were not (χ^2 when removed < 0.5). As this analysis included fewer than the recommended 10 cases/variable, a further set of analyses was performed in which only two variables were entered at a time. These analyses confirmed that BADS-C and CEFA scores were retained as significant predictors of group membership, but IQ and BPVS scores were not.

Quality of reasoning

Figure 2 shows the breakdown of performance between the randomly performing group and the three types of consistent performance: impulsive, orderly and acquisitive. Figure 2 includes data from a total of 40 service users (the 20 tested in the present study and the 20 tested previously by Willner *et al.* 2010a), and compares these data with the performance of 30 control participants, who were all staff working in learning disability services (the 10 controls from this study and the 20 tested previously by Willner *et al.* 2010a). It is apparent that there are marked differences between the performance of service users and controls ($\chi^2 = 42.7$, $P < 0.001$). In total, 17/40 of the service users (43.6%) performed randomly, compared with only a

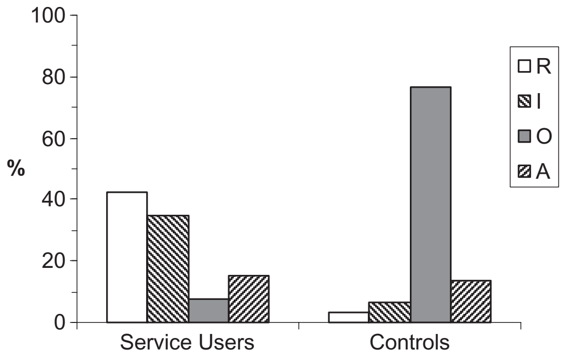

Figure 2 Proportion of participants showing each type of performance on the temporal discounting task. R, random; I, impulsive; O, orderly; A, acquisitive.

single control participant. Of the participants who performed consistently, in the control group there was a normal distribution of AUC scores, with 77% of participants performing in an 'orderly' manner, adjusting their choices according to both the value of the immediate option and the waiting time for the delayed option; of the remainder, almost twice as many were acquisitive than impulsive (13.3% vs. 6.7%). In contrast, the distribution of AUC scores in those service users who responded consistently was bimodal, with 61% of this group almost always choosing the immediate option, a further 26% almost always choosing the larger option, and only three participants (13% of the consistent group, or 7.5% of the total, all of whom were from the previous cohort) using both information sources. The difference between service users and controls in the proportion using both information sources was highly significant ($P < 0.001$) in both cohorts.

Financial decision-making

Scores on the individual decision-making tasks and decision-making abilities are shown in Table 1. As expected, control participants scored significantly higher than service users ($P < 0.001$) on all measures. Comparison between the performances of services users across the five vignettes found, as reported also by Suto *et al.* (2005a), that performance on the three simpler problems was similar, with a decrease on problem 4 and a very poor performance on problem 5 ($F(4,76) = 59.6$, $P < 0.001$). The present participants were similar to those tested by Suto *et al.* (2005a) in age and IQ, but their performance was worse by around 1 point on all tasks, and whereas more than 40% of the earlier sample had a maximum score of 10 on at least one scenario, only three (15%) of the present sample did so. From inspection of the two sets of data, it appears that the discrepancy arises almost entirely from a large difference between the two studies in 'Appreciation'; there was no marked difference between scores on the other four decision-making abilities.

Role of executive functioning

We first examined the correlations between FDMT total scores and four measures of cognitive ability: IQ, BPVS and the two EF scores. Performance on the FDMT was highly correlated with EF scores (BADS-C: $r(18) = 0.64$, $P < 0.002$; CEFA: $r(18) = 0.55$, $P < 0.02$). There was also a significant correlation between FDMT and BPVS scores ($r(18) = 0.69$, $P < 0.001$), but the correlation with IQ was non-significant ($r(18) = 0.41$, NS).

Considering specifically the Reasoning component of the FDMT, the correlation with EF scores was even higher than for the overall FDMT score (BADS-C: $r(18) = 0.66$, $P < 0.002$; CEFA: $r(18) = 0.74$, $P < 0.001$). Reasoning was less strongly associated with BPVS and IQ scores (BPVS: $r(18) = 0.55$, $P < 0.02$; IQ: $r(18) = 0.42$, NS). Entering all four variables into a stepwise multiple regression analysis produced a significant model ($F(2,19) = 15.58$, $P < 0.001$) accounting for 64.7% of the variance in Reasoning scores, in which both BADS-C and CEFA scores were retained as significant predictors (BADS-C, $t = 2.24$, $P < 0.05$; CEFA: $t = 3.21$, $P < 0.005$), but BPVS and IQ scores were not (BPVS: $t = 0.05$, NS; IQ: $t = 1.13$, NS). As this analysis included a relatively low ratio of cases/variables, a further set of analyses was performed in which only two variables were entered at a time. These analyses confirmed that BADS-C and CEFA scores were retained as significant predictors of Reasoning scores, but IQ and BPVS scores were not.

Quality of reasoning

On the Reasoning component of the FDMT, a score of 1 is given for stating a choice and a score of 2 is awarded if a 'sufficient explanation' is given

Table 1 Performance on the financial decision-making task

Vignettes	Service Users	Controls	Abilities	Service Users	Controls
In the supermarket	7.05 (0.32)	9.70 (0.15)	Identification	6.35 (0.33)	9.20 (0.33)
Buying jeans	6.15 (0.46)	9.70 (0.15)	Understanding	4.30 (0.30)	9.70 (0.15)
Going to work	7.15 (0.51)	9.90 (0.10)	Reasoning	5.20 (0.30)	9.50 (0.17)
The noisy car	5.20 (0.47)	9.60 (0.16)	Appreciation	5.05 (0.78)	10.0 (0)
Selling shares	1.35 (0.53)	9.30 (0.30)	Communication	6.00 (0.44)	9.80 (0.13)

Values are mean scores (standard error).

(Suto *et al.* 2005a), which we interpreted as a valid reason for making the choice stated. This component of the FDMT should in principle assess the ability to 'weigh' information. However, the scoring does not adequately capture this process because full marks could be achieved without demonstrating any ability to weigh pros and cons (i.e. by providing a single piece of information). We addressed this issue by extending the rating scale from 3 points to 4 (0–3). Four types of valid information could be provided to justify a decision, reflecting the pros and cons of each of the two options. A score of 3 was awarded if reasons for the decision were provided from at least two different categories of evidence, indicating that information was being 'weighed'.

We also reexamined the Understanding component of the FDMT. On this component, a score of 2 is awarded if the participant can provide information for and against each alternative, and a score of 1 is awarded for partial information (Suto *et al.* 2005a). For our purposes, the relevant question was whether participants had been able to provide valid information from more than one category of evidence (without which there is insufficient information to 'weigh'). This was always the case for control participants. For service users, most could provide information from at least two categories of evidence for vignettes 1, 2 and 3 (n = 19, 17 and 18, respectively), but this fell to 13 on vignette 4 and only 3 on vignette 5. The analysis of Reasoning was therefore restricted to vignettes 1–3 and the 17 service users who had at least two categories of information available for all of these vignettes.

Service users were significantly less likely than controls to demonstrate that information was being 'weighed' by providing evidence from more than one category (Fig. 3): only 5/17 service users achieved a Reasoning score of 3 for any of the three problems, as compared with 9/10 of the controls ($\chi^2 = 9.3$, $P < 0.001$), and only 2/17 service users achieved a Reasoning score of 3 on two of the three problems, as compared with 6/10 controls ($\chi^2 = 7.0$, $P < 0.001$).

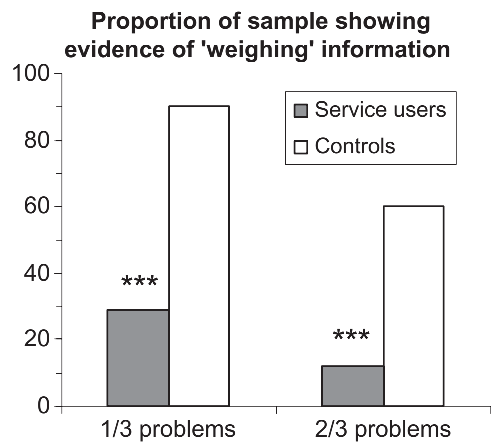

Figure 3 Proportion of service users (n = 17) and controls (n = 10) providing at least two pieces of information to support their decision in three financial reasoning problems. All participants were shown to have at least two pieces of information available to them on all problems. $^{***}P < 0.001$.

Discussion

Relationship of decision-making performance to EF

The groups of service users tested here and in our earlier study (Willner *et al.* 2010a) were very similar demographically, but differed markedly in their performance on the TD task, with 14/20 of the present participants performing consistently in the present study, compared with 14/20 performing inconsistently in the previous study. This difference arises because in the present study, participants were provided with feedback on practice trials, which has been shown to elicit a consistent performance in the main TD task in some participants who previously had performed inconsistently (Willner *et al.* 2010a). Despite the large difference between the two studies in the distribution of consistent and inconsistent performance, we replicated the earlier observations that consistent performers had higher EF scores than inconsistent performers, and that EF scores were significantly stronger than IQ or BPVS scores as predictors of TD performance.

Because inconsistent responding in TD tasks is so rarely seen in intellectually able participants (Green & Myerson 2004; Wileyto *et al.* 2004), there is little in the literature with which to compare these observations. We are aware of only a single study that included a proportion of inconsistent responders large enough to support an analysis. This study, in able children and adolescents, reported that IQ was lower in inconsistent responders than in consistent responders (Olson *et al.* 2007). However, the study involved a much wider range of IQ values than the present study (SD = 10.9 vs. 3.6), and the discrepancy with the present results could be accounted for by the significant correlation between IQ and EF (Ardila *et al.* 2000; Duff *et al.* 2005; Shamosh *et al.* 2008).

Performance in the FDMT was very similar to that reported by Suto *et al.* (2005a), in a demographically similar population of service users, with a single exception, that scores on one sub-scale, appreciation of who is affected by the decision, were lower. The reason for this difference is unclear, but Suto *et al.* (2005a) reported that this measure was significantly less reliable than the other four abilities measured in the FDMT. Consistent with what was observed in the TD task, reasoning performance in the FDMT was significantly correlated with EF, but was unrelated to either IQ or BPVS scores when EF abilities were taken into account. We are unaware of any previous studies of the role of EF in decision-making by people with ID. However, several studies have reported that impairments of EF are predictive of limitations of financial capacity in older adults (Earnst *et al.* 2001; Farias *et al.* 2003; Okonkwo *et al.* 2006; Sherod *et al.* 2009).

Suto *et al.* (2005b) reported a correlation between IQ and FDMT performance ($r = 0.42$) which was almost identical to the correlation that we observed ($r = 0.41$). However, as in the present study, the relationship became non-significant when other factors were taken into account: specifically, basic financial knowledge, and to a lesser extent, experience of decision-making. The only details of the financial knowledge task provided in the paper were that participants were required 'to order sets of items along (a) scale, as well as recognize and recall their values'. Aspects of EF may be involved in both of these sub-tasks (ordering items and recent memory), implying that, as with the FDMT itself, performance in the financial knowledge task may also depend to some extent on EF abilities. It is also possible that perceptions of a person's EF abilities may influence the extent to which others make opportunities available to gain experience of decision-making.

Interestingly, both BADS-C and CEFA scores were retained as independent predictors of reasoning performance in the FDMT, suggesting that these two batteries of EF tests measure different abilities. This is consistent with our earlier observation that there was very little cross loading in a factor analysis of subtests from these two batteries (Willner *et al.* 2010b). It is unclear how best to describe the differences between the BADS-C and the CEFA (see Willner *et al.* 2010b). However, there is a suggestion in the data that performance on the TD task was somewhat more closely related to scores on the BADS-C than the CEFA (see Fig. 1: as also seen in Willner *et al.* 2010a, Table 1), while conversely, the CEFA appears to be more closely related than the BADS-C to reasoning performance on the FDMT (see correlation and regression analyses).

'Weighing up' of information

In order to examine the quality of performance in the TD task, the present data were combined with a comparable set of earlier data (Willner *et al.* 2010a) because the larger data set provides a clearer picture of the distribution of the different types of response style. We had previously concluded, on the basis of the earlier data set, that service users' responses were typically either random or impulsive (always choosing the immediate reward), with very little evidence of the orderly responding that is typically shown by intellectually able control participants. This conclusion was confirmed. However, with the larger data set, it is now apparent that there was also a significant minority of service users who responded acquisitively (always choosing the larger reward). Therefore it may be more accurate to describe the typical responding of service users as either random or extreme: usually impulsive, but sometimes acquisitive. The common feature of these two extreme response styles is that they are both based on the use of a single class of information: delay (impulsive responding) or amount (acquisitive responding). Fewer than 10% of the service users (3/40) showed evidence that they were 'weighing up' both sources of information, in contrast to the control group, where an orderly trade-off between delay and amount of reward was evident in more than 75% (23/30) of participants.

A similar conclusion was reached in the FDMT. We first constructed a measure to confirm that service users had at least two sources of information available to them, and restricted the analysis of reasoning performance to service users and scenarios where this was demonstrated. We then constructed a second measure to assess whether two or more sources of information were being used to reach a decision. This showed that decisions were typically justified on the basis of a single piece of evidence. It cannot be excluded that service users were 'weighing up' different sources of evidence that they were not reporting. However, from their verbal responses, there was very little evidence that information was being 'weighed'.

In both tasks, therefore, it was clear that intellectually able control participants were typically 'weighing up evidence for the purpose of making a decision'. However, it appeared that service users typically were not 'weighing up' evidence, but rather, basing decisions on a single item of information. Such decisions may appear rational, if a justification is provided, but may actually not be rational if other important items of information are ignored. This can sometimes be obvious in relation to real life decisions. (Consider, for example, the common practice dilemma of patients who are determined to return home from hospital while ignoring their inability to self-care.) More commonly, decisions that appear rational but are actually irrational (because of a lack of 'weighing up' of evidence) might not be readily apparent in real life because the identification of such situations requires a specialized questioning technique.

It is possible that the difficulties experienced by participants in 'weighing up' information in both tasks could reflect limitations of their understanding of the task requirements. While this possibility cannot be entirely excluded, the data suggest that this is unlikely: in the TD task, all bar one of the participants performed well on the unidimensional decision-making tasks used in the practice component, while in the FDMT, reasoning performance was only examined for participants and tasks where a sufficient degree of understanding was demonstrated. The fact that performance in both tasks was related to EF but not to IQ also argues against a lack of understanding of the task requirements as a significant factor.

Implications

Relationship of decision-making performance to EF

As we have pointed out previously (Willner *et al.* 2010a), the fact that the ability to 'weigh up' information appears to be much more strongly associated with EF than with IQ has some important practical implications.

The Mental Capacity Act (2005) defines incapacity as the inability to make a decision 'because of an impairment of, or a disturbance in the functioning of, the mind or brain'. Therefore, the assessment of capacity is a two-stage process, where it must first be established that there is 'an impairment of, or a disturbance in the functioning of, the mind or brain', before proceeding to the second stage of an assessment of decision-making abilities. In the context of ID, this means that people who have a diagnosis of 'learning disability' could be assessed as lacking capacity, but people who are more able and lack this diagnosis cannot (unless they have some other diagnosis that implies 'an impairment or disturbance in the functioning of mind or brain', such as a mental illness or an autistic spectrum disorder). However, the diagnosis of 'learning disability' is based, inter alia, on an IQ criterion (Full-Scale IQ < 70) (British Psychological Society 2000), which, as the present results demonstrate, may be of limited relevance to reasoning ability.

We have previously reported that EF abilities in people with a diagnosis of 'learning disability' are very much poorer than would be expected, on the basis of the published norms for people in the IQ range 70–130, suggesting that impairment of EF may be an important factor distinguishing those people with IQ < 70 who are known to learning disability services from those who are not (Willner *et al.* 2010b). But it is likely that there are also many people with IQ > 70 whose decision-making may be compromised by impairments of EF. The legal definition of capacity implies that such individuals could not, at present, be legitimately assessed as lacking capacity, since, in the absence of any diagnosis, there is no clear basis for reaching a conclusion, in the first stage of the capacity test, that they have 'an impairment of, or a disturbance in the functioning of, the mind or brain'. Therefore, these individuals are denied the protection that the MCA (Mental Capacity Act) affords to individuals who are assessed as lacking capacity. (They are also denied the protection from sexual exploitation provided by the Sexual Offences Act (2003), which uses a similar two-stage definition of capacity.) It may be that this issue will only be resolved by the development of case law around individuals who have been assessed as lacking capacity on the basis that their EF abilities are impaired, despite not meeting the criteria for any diagnosis.

'Weighing up' of information

Our earlier study of performance in the TD task identified two difficulties in service users: many were unable to make consistent decisions, while those who were able to respond consistently were highly impulsive (Willner *et al.* 2010a). There have been a number of studies in which adults or children with intellectual impairments were taught strategies to decrease impulsiveness (Dixon *et al.* 1998, 2003; Vollmer *et al.* 1999; Cuskelly *et al.* 2001), but these operated over short time-scales (min) and would be difficult to implement over longer real-life time-scales of days or weeks. However, the fact that, in addition to the impulsive group, there is also a small proportion of service users who behave acquisitively suggests that it may be more appropriate to think of the TD results as reflecting a single difficulty in 'weighing up' information, which is expressed as either random behaviour or extreme behaviour (usually impulsive but occasionally acquisitive), depending on the level of EF ability.

Temporal discounting is a laboratory procedure of questionable relevance to real life, but a similar difficulty in 'weighing up' information was also seen in the more realistic FDMT. The fact that difficulty in 'weighing up' information was observed in these two very different procedures suggests that this may be a general problem in the lives of many people with ID. This raises the question of whether it might be possible to develop psycho-educational strategies to improve reasoning ability, and the possibility that a psycho-educational strategy aimed at improving the ability to 'weigh up' information might also be effective in helping people with ID to suppress the tendency to behave impulsively.

Acknowledgement

This study was supported by a grant from the Welsh Office for Research and Development in Health and Social Care.

References

Ardila A., Pineda D. & Rosselli M. (2000) Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology* **15,** 31–6.

Ball S. L., Holland A. J., Treppner P., Watson P. C. & Huppert F. A. (2008) Executive dysfunction and its association with personality and behavior changes in the development of Alzheimer's disease in adults with Down syndrome and mild to moderate learning disabilities. *British Journal of Clinical Psychology* **47,** 1–29.

British Psychological Society (2000) *Learning Disability: Definitions and Contexts*. British Psychological Society, Leicester.

British Psychological Society (2006) *Interim Guidelines on the Assessment of Mental Capacity*. British Psychological Society, Leicester.

Chapman G. B. & Elstein A. S. (1995) Valuing the future: temporal discounting of health and money. *Medical Decision Making* **15,** 373–86.

Critchfield T. S. & Kollins S. H. (2001) Temporal discounting: basic research and the analysis of socially important behavior. *Journal of Applied Behavior Analysis* **34,** 101–22.

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

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,** 239–51.

Dixon M. R., Hayes L. J., Binder L. M., Manthey S., Sigman C. & Zdanowski D. M. (1998) Using a self-control training procedure to increase appropriate behavior. *Journal of Applied Behavior Analysis* **31,** 203–10.

Dixon M. R., Horner M. J. & Guercio J. (2003) Self-control and the preference for delayed reinforcement: an example in brain injury. *Journal of Applied Behavior Analysis* **36,** 371–4.

Duff K., Schoenberg M. R., Scott J. G. & Adams R. L. (2005) The relationship between executive functioning and verbal and visual learning and memory. *Clinical Neuropsychology* **20,** 111–22.

Earnst K. L., Wadley V., Aldridge T., Steenwyk A., Hammod A., Harrell K. *et al.* (2001) Loss of financial capacity in Alzheimer's disease: the role of working memory. *Aging, Neuropsychology and Cognition* **8,** 109–19.

Emslie H., Wilson F. C., Burden V., Nimmo-Smith I. & Wilson B. A. (2003) Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C). *Child Neuropsychology* **13,** 539–42.

Estle S. J., Green L., Myerson J. & Holt D. D. (2007) Discounting of monetary and directly consumable rewards. *Psychological Science* **18,** 58–63.

Farias S. T., Harrell E., Neumann C. & Houtz A. (2003) The relationship between neuropsychological performance and dail functioning in individuals with Alzheimer's disease: ecological validity of neuropsychological tests. *Archives of Clinical Neuropsychology* **18,** 55–672.

Green L. & Myerson J. (2004) A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin* **130,** 769–92.

Kopp C. (1990) The growth of self-monitoring among young children with Down Syndrome. In: *Children with Down Syndrome: A Developmental Perspective* (eds D. Cicchetti & M. Beeghly). pp. 231–51.Cambridge University Press, Cambridge.

Mental Capacity Act (2005) Office for Public Sector Information, London.

Myerson J., Green L. & Warusawitharana M. (2001) Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior* **76,** 235–43.

Okonkwo O. C., Wadley V. G., Griffith H. R., Ball K. & Marson D. (2006) Cognitive correlates of financial abilities in mild cognitive impairment. *Journal of the American Geriatrics Society* **54,** 1745–50.

Olson E. A., Hooper C. J., Collins P. & Luciana M. (2007) Adolescents' performance on delay and probability discounting tasks: contributions of age, intelligence, executive functioning, and self-reported externalizing behaviour. *Personality and Individual Differences* **43,** 1886–97.

Ragotzy S. P., Blakely E. & Poling A. (1988) Self-control in mentally retarded adolescents: choice as a function of amount and delay of reinforcement. *Journal of the Experimental Analysis of Behavior* **49,** 191–9.

Reynolds B. (2006) A review of delay discounting research with humans: relations to drug use and gambling. *Behavioural Pharmacology* **17,** 651–67.

Scheres A., Dijkstra M., Ainslie E., Balkan J., Reynolds B., Sonuga-Barke E. *et al.* (2006) Temporal and probabilistic discounting of rewards in children and adolescents: effects of age and ADHD diagnosis. *Neuropsychologia* **44,** 2092–103.

Schweitzer J. B. & Sulzer-Azaroff B. (1988) Self-control: teaching tolerance for delay in impulsive children. *Journal of the Experimental Analysis of Behavior* **50,** 173–86.

Sexual Offences Act (2003) Office for Public Sector Information, London.

Shamosh N. A., DeYoung C. G., Green A. E., Reis D. L., Johnson M. R., Conway A. R. A. *et al.* (2008) Individual differences in delay discounting: relation to intelligence, working memory, and anterior prefrontal cortex. *Psychological Science* **19,** 904–11.

Sherod M. G., Griffith H. R., Copeland J., Belue K., Krzywanski S., Zamrini E. Y. *et al.* (2009) Neurocognitive predictors of financial capacity across the dementia spectrum: normal aging, mild cognitive impairment, and Alzheimer's disease. *Journal of the International Neuropsychological Society* **15,** 258–67.

Suto W. M. I., Clare I. C., Holland A. J. & Watson P. C. (2005a) Capacity to make financial decisions among people with mild intellectual disabilities. *Journal of Intellectual Disability Research* **49,** 199–209.

Suto W. M. I., Clare I. C., Holland A. J. & Watson P. C. (2005b) The relationship among three factors affecting the financial decision-making abilities of adults with mild intellectual disabilities. *Journal of Intellectual Disability Research* **49,** 210–7.

Vollmer T. R., Borrero J. C., Lalli J. S. & Daniel D. (1999) Evaluating self-control and impulsivity in children with severe behavior disorders. *Journal of Applied Behavior Analysis* **32,** 451–66.

Wileyto E. P., Audrain-McGovern J., Epstein L. H. & Lerman C. (2004) Using logistic regression to estimate delay-discounting functions. *Behavior Research Methods* **36,** 41–51.

Willner P., Bailey R., Parry R. & Dymond S. (2010a) Performance in temporal discounting tasks by people with intellectual disabilities reveals difficulties in decision-

making and impulse control. *American Journal on Intellectual and Developmental Disability* **54,** 366–79.
Willner P., Bailey R., Parry R. & Dymond S. (2010b) Evaluation of executive functioning in people with intellectual disabilities. *Journal of Intellectual Disability Research* **54,** 366–79.

Accepted 29 January 2010