AN ASSESSMENT OF PROFESSIONAL ACCOUNTANTS’ COGNITIVE REFLECTION ABILITY

Ricardo Lopes Cardoso  
PhD in Accountancy  
Professor at Fundação Getulio Vargas, and at Universidade do Estado do Rio de Janeiro  
Praia de Botafogo, 190, sala 536, Botafogo, Rio de Janeiro, RJ, Brazil, 22250-900  
lopecardoso.ricardo@gmail.com +55-21-992080193

Leonardo Portugal Barcellos  
Student - PhD in Administration  
PhD Student at Fundação Getulio Vargas, and Professor at Universidade Federal do Rio de Janeiro  
Praia de Botafogo, 190, sala 536, Botafogo, Rio de Janeiro, RJ, Brazil, 22250-900  
lpbarcellos@gmail.com +55-21-981125115

André Carlos Busanelli de Aquino  
PhD in Accountancy  
Professor at Universidade de São Paulo  
Av. Bandeirantes 3900, Monte Alegre, Ribeirão Preto, SP, Brazil, 14040-905  
aa.bb.aquino@gmail.com +55-16-997642768

Patrick Del Bosco de Sales  
Student - Master in Administration  
Master Student at Fundação Getulio Vargas  
Praia de Botafogo, 190, sala 536, Botafogo, Rio de Janeiro, RJ, Brazil, 22250-900  
patrick_laser@yahoo.com.br +55-21-979871188

ABSTRACT
Cognitive Science literature attributes to the cognitive System 1 the intuitive decision, and attributes the reflective decision to the System 2 (Kahneman, Frederick, 2002). The intuitive System 1 is fast and straightforward, but it is more subject than System 2 to provide wrong decisions. On the other hand, the reflective System 2 is slower and more effort consuming. Therefore, it is not necessarily bad or good the usage of System 1, but it potentially affects the quality of judgment and decision-making. Frederick (2005) developed the cognitive reflection test (CRT), a simple three items test for cognitive abilities to assess individuals' tendency to override a preponderant response alternative that is incorrect and to engage in further reflection that leads to the correct response. We assessed the CRT scores for a large sample of 4,902 Brazilian accountants, and investigate which demographic characteristics are the most associated with a more reflective (System 2) or intuitive (System 1) behavior. For the multivariate analysis we used the ordered logistic regression model. Findings support the Cognitive Science literature, that higher test scores on the CRT are correlated with gender (male), age (young), educational level (post-graduate), and income (high-income, or well-succeeded). Results also show that Brazilian professional accountants scored lower than college students from Canada, Germany and the U.S.A. The findings invite accounting researchers to further investigate the impacts of such a behavior on judgment and decision-making in accounting related issues, such as the interpretation of accounting standards and its implementation on the recognition, classification and measurement of transactions.

Keywords: judgment; decision-making; cognitive reflection test; accountant; accounting.
1 INTRODUCTION

The exercise of judgment and decision-making are fundamental to assure the quality of financial reports; it has been widely recognized in the professional literature and accounting handbooks (Epstein and Jermakowicz, 2009; Pricewaterhousecoopers, 2014).

However, in most code law countries, listed companies were required to prepare financial statements in compliance with rules-based accounting standards, usually biased by the fiscal rules (Nobes, 1998; Gray, 1988), in such an environment, accountants were not trained to reflect deeply about their judgments and estimates, probably because more important than faithfully representing the substance of transactions, would be the compliance with standards and detailed codes (Arden, 1997; Colasse, 1997; Hulle, 1997). It is worth noting that the International Financial Reporting Standards (IFRS) require the disclosure of the judgments made in the process of applying the entity’s accounting policies and that have the most significant effects on the amounts presented in the financial statements (IASB, 2013a - IAS 1, item 122). Indeed, the training materials prepared by the IFRS Foundation Education Initiative emphasize the most significant judgments and estimates necessary for properly applying either the Full IFRSs (IASB, 2013a) or the IFRS for SMEs (IASB, 2009).

Even before the adoption of the IFRSs in Brazil, a few academics used to highlight the importance of judgment and decision making in accounting. Iudícibus (1998), for instance, teach that accountants should exercise the subjective allowed by accounting standards with responsibility. Despite of such an awareness about the significance of judgments and decision making in accounting, and despite of the increasing literature on behavioral accounting (Belkaoui, 1989; Ashton, Ashton, 1995; Bonner, 2008), until date, we do not know much about the accountants’ cognitive abilities.

This paper aims to identify the cognitive reflection abilities of professional accountants, and to explore which professional features a more associated with such a cognitive ability. Therefore, in a large-scale survey, we assessed the accountants’ tendency (i) to override a preponderant response alternative that is incorrect or (ii) to engage in further reflection that leads to the correct response. Based on Frederick (2005) cognitive reflection test (CRT), we investigated which (and how) demographic characteristics of Brazilian accountants are associated with a more reflective (or intuitive) behavior.

The remaining sections of this paper are organized as follows. The next section presents an overview of judgments and decision-making in accounting literature. Section three describes the method to access the cognitive reflection ability of individuals, data collection approach, presents the sample of analysis, defines variables, and the model used. The forth section presents the results. In the last section we discuss the results and present implications for future research.

2 JUDGMENT AND DECISION-MAKING IN ACCOUNTING

Psychology and Cognitive Science are mature fields of study, and their interactions with Finance and Economics are also mature (see Frederick, 2005; Kahneman, 2011; Mackenzie, 2009, 2010). There are also applications of the psychological research into Law

---

1 Please, refer to IASB (2013b); also refer to the IFRS Foundation continuing professional development training material (available at http://www.ifrs.org/Use-around-the-world/Education/CPD/Pages/CPD-training-material.aspx), and the training modules on the IFRS for SMEs (available at http://www.ifrs.org/IFRS-for-SMEs/Pages/Training-material.aspx).

2 Indeed, Daniel Kahneman received the Noble Prize in 2002 “for having integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty”(source: http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2002/kahneman-facts.html).
Behavioral Accounting can be described as the application of behavioral science to Accounting (Belkaoui, 1989, p. xi), consequently, it is open for many different researching methods, for example: interviews (Beattie et al, 2001; Beattie et al, 2012), surveys (Govindarajan, 1984; Hartmann, Maas, 2011), and experiments (Libby, 1975; Rosman, 2011). Behavioral Accounting and, more precisely, Judgment and Decision-Making (JDM) research, is of special interest in the context of the adoption of principles-oriented accounting standards, such as the IFRS, because the interpretation of such principles requires more intensive cognitive efforts than rules-based accounting standards. In the absence of any relevant reference or empirical evidence to support such a statement, we present three anecdotal evidences that compare the accounting policies generally accepted in Brazil until 2010 with those adopted as a consequence of the implementation of the IFRSs.\(^4\)

*Judgment on the recognition and measurement of the allowance for bad debts*

The Brazilian fiscal law 4,506/1964 established the maximum amount for the deductible expense (i.e., tax shield) measured as 3% of the ending balance of accounts receivable. Later, the law 8,541/1992 reduced such a percentage by a half. Finally, the law 9,249/1995 prohibited the deductibility of the allowance for bad debt expense. There is a common sense that, while permitted, entities used to accrual the allowance for bad debts in their general purpose financial statements at the maximum amount allowed by fiscal rules (i.e., 3% for the period 1965-1992, and 1.5% for the period 1993-1995), and that in 1996 many companies stopped recognizing the allowance for bad debts – in financial reports – due to the fiscal rules prohibition – for tax purposes (Cardoso, 2007, p. 152, note 4).

Nowadays, there is a debate in the IFRS arena, because the IASB decided to change the recognition criteria of the impairment of financial instruments, from an incurred loss approach to an expected loss approach. Which requires much more cognitive efforts than were required before, especially if such an effort were compared with those required to measure the allowance for bad debt as a certain percentage (i.e., 3% or 1.5%) of the ending balance of accounts receivable.

*Judgment on the measurement of depreciation*

The fiscal rules (i.e., acts issued by the Brazilian Internal Revenue Service, 162/1998 and 130/1999) establish a benchmark for estimating the useful life of items of property, plant and equipment. For instance, such rules suggest 25 years for buildings, 10 years for machinery, and 4 years for cargo dedicated motor vehicles. Until 2010, the Brazilian listed companies used to measure the depreciation expense of their buildings, machineries and vehicles based on a depreciation rate of 4%, 10% and 25%, respectively – on their “general purpose” financial statements.

\(^{3}\) Bonner (2008, p.2) defines *judgment* as “forming an idea, opinion, or estimate about an object, an event, a state, or another type of phenomenon”, and defines *decision* as “making up one’s mind about the issue at hand and taking a course of action”.

\(^{4}\) The first set of financial statements prepared by Brazilian listed companies in accordance with the IFRS was related to the fiscal year ended 31 December 2010, with data from 2009 presented for means of comparability. The convergence towards IFRS was mandatorily required by law 11,638/2007.
During 2011 and 2012, the Brazilian market regulator (Comissão de Valores Mobiliários, from now on CVM)\(^5\) required the management of listed companies to explain why they did not apply the fair value as the deemed cost of items of property, plant and equipment as allowed by IFRS 1.

*Judgment on the measurement of fair value*

The fair value was introduced in the Brazilian accounting literature with the implementation of the IFRSs.

In 2011, Alessandro Broedel Lopes, former director of the CVM, commented during a speech at an IFRS Conference\(^6\), that, in 2009, the management of a listed company asked him if the CVM would publish a table of fair values for items of property, plant and equipment similarly to the fiscal rules useful life time (see the previous anecdotal). Such an example presents anecdotal evidence that judgment and decision making in accounting are still incipient in Brazilian business context.

Despite the fact that anecdotal evidences are not theoretically appropriate, we are confident that studying the cognitive abilities of professional accountants is a necessary condition in order to improve the consistence of the IFRS implementation, especially in the context of code law jurisdictions. Indeed, we used Frederick’s cognitive reflection test (CRT) to assess the cognitive abilities of Brazilian accountants. The method to access the cognitive reflection ability is described on the next section.

3 METHODOLOGY

3.1 Cognitive Reflection Test

CRT was introduced by Frederick (2005). It is criteria to measure how impulsively or reflectively people make decisions. The CRT consists of three questions to which the intuitive (i.e., impulsive and spontaneous) answer is wrong. Correct answers indicate a higher degree of reflectivity and deliberate thinking. The quintessential item from the CRT was first discussed by Kahneman and Frederick (2002) in an article that reframed the heuristics-and-biases literature in terms of the concept of attribute substitution.

The intuitive answers are attributed to the “System 1” process of decision-making, because it is the first answer that respondents’ cognition suggests; therefore, System 1 is considered ‘fast’. If such an answers is not identified being wrong, the “System 2” is not activated. However, if respondent reflectively tests whether the first answer is wrong and deliberately think about the problem again, she might find the correct answer; therefore, System 2 is considered ‘slow’ (Toplak, West, Stanovich, 2011). Notice that, it is not necessarily bad or good the usage of System 1, but it potentially affects the quality of judgment and decision-making.

In order to estimate the participants’ CRT scores, we asked them the following questions:

- A bat and a ball cost $110. The bat costs $10 more than the ball. How much does the ball cost?
- If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

---

\(^5\) CVM has powers and responsibilities similar to those carried out by the U.S.A. Securities and Exchange Commission.

\(^6\) IFRS Conference in Latin America and the Caribbean, São Paulo, 27 and 28 October 2011.
• In a lake, there is a patch of lily pads. Every day, patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

The impulsive and correct answers are:

• Bat & ball: Impulsive answer: $100; correct answer: $50.7
• Widgets: Impulsive answer: 100 minutes; correct answer: 5 minutes.
• Lily pads: Impulsive answer: 24 days; correct answer: 47 days.

Frederick (2005); Toplak, West and Stanovich (2011) and other authors from the Cognitive Science literature have shown evidences that the CRT is a more potent predictor of performance on a wide sample of tasks from the heuristics-and-biases literature than measures of cognitive ability, thinking dispositions, and executive functioning. Therefore, in this paper, we rely exclusively on the CRT to assess accountants’ cognitive abilities.

We assessed the CRT scores for a large sample of 4,902 Brazilian accountants, and investigate which demographic characteristics are the most associated with a more reflective (System 2) or intuitive (System 1) behavior. For the multivariate analysis we used the ordered logistic regression model, as described in the next section.

3.2 Data collection and sampling

Data were collected via a questionnaire applied by Brazilian Federal Accounting Council (Conselho Federal de Contabilidade, hereafter CFC). Questions were classified into two groups: (i) demographic questions (such as age, gender, formal instruction level, income, professional segment, and size of the firm to which she works), and (ii) Frederick’s (2005) CRT questions.

Respondents were recruited via CFC’s publications (e.g., newsletter, professional magazine and academic journal) and e-mail market. The target audience was the entire population of professional accountants. In June 2012, the closest data for which the CFC has accrued information, there were 302,697 professional accountants registered as bachelor accountants at CFC8.

A total of 9,355 bachelor professional accountants answered the questionnaire. As a check against respondents answering randomly, we eliminated the observations of respondents who presented unreasonable answers for the three CRT questions, who did not answer their income and who did not fit in any professional segment of interest.

After eliminations, the analyzed sample was comprised by 4,902 professional accountants. Table I presents the sample selection steps.

---

7 Frederick’s originally worded the bat & ball question as follows: “A bat and a ball cost 110 cents. The bat costs 100 cents more than the ball. How much does the ball cost?”. Therefore, the impulsive answer is 10 cents and the correct answer is 5 cents (Frederick, 2005). Silva (2005), who first submitted the CRT questions to Brazilian respondents, adapted the values to became more realistic in comparison to the general prices in the Brazilian market, but the gist were keep in the way that do not influence on the result. We decided to apply the Silva (2005) version.

8 Notice that in Brazil professional accountants are required by law to register themselves at the CFC in order to properly provide accounting services.
Table 1 – Sample selection.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total CRT survey answers:</td>
<td>9,355b</td>
</tr>
<tr>
<td>(-) CRT inconsistent answers (A1 U A2 U A3):</td>
<td>(516)</td>
</tr>
<tr>
<td>(-) Answers lower than 10 or higher than 110 for Question 1 (bat and ball) (A1):</td>
<td>(228)</td>
</tr>
<tr>
<td>(-) Answers lower than 2 or higher than 96 for Question 2 (lily pads) (A2):</td>
<td>(222)</td>
</tr>
<tr>
<td>(-) Answers lower than 1 or higher than 500 for Question 3 (machines) (A3):</td>
<td>(210)</td>
</tr>
<tr>
<td>(-) Respondents who did not answer the income question:</td>
<td>(49)</td>
</tr>
<tr>
<td>(-) Respondents who did not fit in any of the professional segments:</td>
<td>(3,888)</td>
</tr>
<tr>
<td>(=) Total of answers considered:</td>
<td>4,902b</td>
</tr>
</tbody>
</table>

a. The initial number of respondents was 13,084, but we had to exclude 3,142 answers from accounting technicians and other 587 inconsistent answers.
b. The population of interest was 302,697 in June 2012, according to CFC database. It means that final sample is comprised by 1.6% of the population (CFC, 2013).
c. A1 U A2 U A3 represents the union of inconsistent answers, once some respondents answered inconsistently more than one question.

3.3 The model and variables description

In order to achieve the purposes of this paper, we apply the Cognitive Reflection Test (CRT) scores and some demographic features as dependent and independent variables, respectively. Prior to entering in modeling matters, we present the variables descriptions in the Table 2.

Table 2 – Variables descriptions.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Score</td>
<td>Sum of Cognitive Reflection Test (CRT) scores, where each correct answer add one point in a scale that varies from 0 to 3. It's the response (dependent) variable of our model.</td>
</tr>
<tr>
<td>Gender</td>
<td>Categorical predictor (independent) variable which assumes 1 (one) if the participant is a woman and 0 (zero) otherwise.</td>
</tr>
<tr>
<td>Age</td>
<td>Quantitative independent variable that represents the age of the participants in years.</td>
</tr>
<tr>
<td>Highest Education Level</td>
<td>Categorical predictor variable which assumes 1 (one) if the participant completed a MBA or a Ph.D. and 0 (zero) otherwise, i.e., if participant completed only the bachelor level.</td>
</tr>
<tr>
<td>Monthly income intervals</td>
<td>Categorical predictor variable with seven levels (until 1,866.00; 1,866.01 - 3,110.00; 3,110.01 - 6,220.00; 6,220.01 - 12,440.00; 12,440.01 - 18,660.00; 18,660.01 - 31,100.00, more than 31,100.00) in Brazilian currency (Real). For each level we apply a dummy variable assuming 1 (one) if the respondent belongs to that category, and 0 (zero) otherwise, so that each participant is classified in only one category.</td>
</tr>
</tbody>
</table>

As at July 2012 the exchange rate was US$ 1 = BRL 2.29. Therefore, those ranges in U.S. dollars (US$) as annual income, based on 13 salaries per year, are: (i) until 10,593.01; (ii) 10,593.02 – 17,655.02; (iii) 17,655.03 – 35,310.04; (iv) 35,310.05 – 70,620.09; (v) 70,620.10 – 105,930.10; (vi) 105,930.11 – 176,550.20; and (vii) more than 176,550.20.
To assess the relationship among dependent and independent variables stated above the most suitable technic of multivariate data analysis seems to be ordered logistic regression (OLOGIT), in which, for each unit increased in a given independent variable, the CRT score (dependent variable) is expected to change by its respective coefficient, if the other variables are held constant. On the other hand, it provides another interpretation approach through the odds ratio, which is especially useful if the independent variable is a dummy. In our specific case, the probability of a higher cognitive score (3) versus lower (0) is odds ratio value times higher (if it is greater than 1) or lower (if it is less than 1) for the group that assumes value 1 (one) according to the categorical variable design than for the other.

For instance, if the odds ratio result for gender is 0.5, then the probability of an individual owning a high CRT score (equal to 3) is 0.5 times lower if it is a woman (if participant is woman, the gender variable assumes the value 1, as described in Table 2), while the other variables are held constant. Still, if the odds ratio result for gender is 1.5, then the probability of an individual owning a high CRT score (3) is 1.5 times higher (not lower) if it is a woman.

We developed three ordered logistic regression models to explore the professional features association with the CRT score. First, we only analyzed the control variables (model 1). Then, we included the (expected) main effect variables (model 2). Models 1 and 2 were analyzed based on the entire sample. Model 3 is similar to model 2, except but the fact that sample was split by the professional segment variable, therefore this variable was omitted. Hence, the following Equations (models) were developed and assessed.

\[
\text{Cognitive Score} = \alpha + \alpha \text{Gender} + \alpha \text{Age} + \alpha \text{Formal Education Level} + \varepsilon \quad (\text{Model 1})
\]

\[
\text{Cognitive Score} = \alpha + \alpha \text{Gender} + \alpha \text{Age} + \alpha \text{Formal Education Level} + \alpha \text{Monthly income intervals} + \alpha \text{Professional segment} + \alpha \text{Larger Company’s Professional} + \varepsilon \quad (\text{Model 2})
\]

\[
\text{Cognitive Score} = \alpha + \alpha \text{Gender} + \alpha \text{Age} + \alpha \text{Formal Education Level} + \alpha \text{Monthly income intervals} + \alpha \text{Larger Company’s Professional} + \varepsilon \quad (\text{Model 3})
\]

Moreover, it is also important to highlight that there are a large body of statistical studies presenting an OLOGIT equivalent to the R-square value found in OLS regression. Our main purpose is not to report how much of the cognitive score variance is explained by the

---

\(^{10}\) Odds ratio = \(e^\alpha\), where \(\alpha\) is the variable coefficient.
predictors, but if and how the predictors are related with the response variable, thus, we present the McFadden’s pseudo R-square for each estimation, but do not discuss it.

4 RESULTS
4.1 Descriptive statistics

Aiming to provide an accurate description of the analyzed data, Table 3 summarizes each variable in details, in the sense that, for each demographic classification of respondents, it presents CRT score mean, standard deviation, the frequency by number of hits and the absolute number of observation per condition.

Among the reviewed literature, the most previous study assessed the CRT scores for samples comprised almost exclusively by college students (Frederick, 2005; Hope, Kusterer, 2011; Oechssler et al., 2009; Toplak et al., 2011). Thus, Moritz et al. (2013) is the only research we found that assess the CRT scores for professionals, instead of students. They reported a mean CRT score of 1.51 for 313 supply chain managers employed at one of three Fortune 500 supply-chain-intensive firms from the U.S.A. Therefore, the mean CRT score of Brazilian professional accountants (i.e., 1.55) is slightly higher than the mean CRT score of American supply chain managers (Moritz et al., 2011). Notice that on average, the Brazilian professional accountants had higher mean CRT scores than the 3428 individuals reported in Frederick (2005, Table 1); specifically, of the eleven sample populations in Frederick (2005), only the student population from MIT and Princeton University had higher average CRT scores than the practitioners in our study.

On the other hand, Oechssler (2009) found a 2.05 CRT mean for a sample comprised by 1,250 individuals (90% of college students) whose mean age was 24 years, whereas we have only professionals accountants whose mean age is 40.35 years old and we do not control for being a student or not. The further analysis shows that our result (mean CRT score equal to 1.55) corroborates the literature prediction that age tend to be negatively associated with CRT scores.
### Table 3 – Descriptive Statistics: CRT Scores for each group of accountants.

<table>
<thead>
<tr>
<th>Accountants groups</th>
<th>CRT Score</th>
<th>Percentage Scoring 0, 1, 2 or 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>St. Dev.</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1.66</td>
<td>0.98</td>
</tr>
<tr>
<td>Female</td>
<td>1.31</td>
<td>0.92</td>
</tr>
<tr>
<td>Age intervals (years old)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 - 35</td>
<td>1.59</td>
<td>0.98</td>
</tr>
<tr>
<td>36 - 50</td>
<td>1.55</td>
<td>0.98</td>
</tr>
<tr>
<td>51 - 88</td>
<td>1.45</td>
<td>0.96</td>
</tr>
<tr>
<td>Highest education level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor</td>
<td>1.50</td>
<td>0.94</td>
</tr>
<tr>
<td>MBA/PhD</td>
<td>1.59</td>
<td>0.98</td>
</tr>
<tr>
<td>Monthly income intervals (BRL 1,000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>until 1.8</td>
<td>1.37</td>
<td>1.00</td>
</tr>
<tr>
<td>1.8 – 3.1</td>
<td>1.39</td>
<td>0.96</td>
</tr>
<tr>
<td>3.1 – 6.2</td>
<td>1.52</td>
<td>0.93</td>
</tr>
<tr>
<td>6.2 – 12.4</td>
<td>1.70</td>
<td>0.98</td>
</tr>
<tr>
<td>12.4 – 18.6</td>
<td>1.74</td>
<td>1.01</td>
</tr>
<tr>
<td>18.6 – 31.1</td>
<td>1.80</td>
<td>1.01</td>
</tr>
<tr>
<td>&gt; 31.1</td>
<td>1.67</td>
<td>1.07</td>
</tr>
<tr>
<td>Professional Segment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparers</td>
<td>1.50</td>
<td>0.97</td>
</tr>
<tr>
<td>Auditors</td>
<td>1.67</td>
<td>1.02</td>
</tr>
<tr>
<td>Analysts</td>
<td>1.54</td>
<td>0.97</td>
</tr>
<tr>
<td>Managers</td>
<td>1.59</td>
<td>0.97</td>
</tr>
<tr>
<td>Professors</td>
<td>1.54</td>
<td>0.99</td>
</tr>
<tr>
<td>Larger Cia’s Professional</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.57</td>
<td>0.97</td>
</tr>
<tr>
<td>Yes</td>
<td>1.54</td>
<td>0.99</td>
</tr>
<tr>
<td>Total</td>
<td>1.55</td>
<td>0.98</td>
</tr>
</tbody>
</table>

*a. The age intervals were applied as additional classification range. It’s a quantitative variable in the ordered logistic model.*
Next section presents the multivariate outputs of the OLOGIT model estimation and we also apply the descriptive statistics from the Table 3 as support for our conclusions.

4.2 Multivariate analysis

Table 4 presents the estimations of ordered logistic regression model settled in Equation 1. Model 1 presents the estimation only for control variables, Model 2 contains the
estimation outputs for the entire set of predictors and, in order to provide a detailed analysis of the cognitive abilities of each professional segment, we run the Model 3 for each one of those segments.

Notice that the following interpretation relies on the odds ratios ($e^c$, where $c$ is the coefficient) at most, once it allows a clear and more understandable analysis, as stated in the previous section. The results interpretation also takes into account the large sample size, especially when dealing with the Model 1, in which the significance level may be due the relatively large number of observations (i.e., $n = 4,902$). The further divided analysis mitigates this problem.

**Gender**

In line with previous research (Frederick, 2005; Hoppe, Kusterer, 2011; Oechssler et al., 2009), descriptive statistics suggests that men scored significantly higher than women on the CRT, as shown in Table 3. Inferential tests, more precisely ordered logistic regression estimations, strongly confirm such a tendency for Brazilian accountants as shown in Table 4. The odds ratio value indicates that the probability of an individual owning a high CRT score (3) is 0.54 times lower if it is a woman, held constant all other variables.

It is important to mention that, with a slight variation in the coefficients, such a tendency kept constant for all the professional groups, unless for Professors. For this group, comprised by 135 men and 79 women, in spite of the difference in the average cognitive score between men and women (1.64 versus 1.37, respectively), the statistical results confirm that this difference may be only by chance in the sample process\(^\text{11}\). Perhaps, the teaching activity, which is marked by a high level of reflection in every day routine, reduces the gender distortions in accountants’ cognitive abilities.

Nevertheless, further research seems necessary to confirm this, once Frederick’s (2005) results kept constant for gender even after the establishment of controls for sampling problems, differences in attention, effort expended to answer and mathematical ability. Frederick (2005), in face of a mean CRT score of 1.47 for males versus 1.03 for females, suggests that men are more likely to reflect on their answers and less inclined to go with their intuitive responses. As shown, our results are strong in this sense.

**Age**

The ordered logistic model estimations shown in Table 4 strongly confirm the evidences from descriptive statistics that CRT scores correlate negatively with age. While Table 3 show that younger groups of accountants demonstrate higher average CRT levels and also tend to present lower proportion of individuals in the “low” CRT group, the multivariate outputs reinforce such an evidence in order that, for each additional year old, the CRT rates of Brazilian accountants tend to decrease by 0.03, if all other variables are kept constant. Note that there is only a slight variation among the coefficients for professional segments, but in all cases they are statistically significant.

Those results are as expected by Kahneman (2011) and Moritz et al. (2013), which coherently noted that some individuals may have skills and experience in making repeated decisions over the years. In such settings, the intuitive system may play a decisive role in the judgment and decision making.

In this sense, facing age as a proxy for experience and taking into account the Frederick (2005) conclusion that CRT is a measure to avoid acting based on the initial response to a problem, then, our results suggest that the more mature accountants tend to

---

\(^{11}\) ANOVA test ($F = 3.75, p = 0.054$) corroborates the result.
judge and decide more intuitively than the younger ones. Then, in unexpected new situations in which no previous knowledge can be retrieved, older accountants have greater chance of simply underestimate the complexity involved, taking wrong decisions. It seems likely to occur specially in an environment marked by accounting complexity\textsuperscript{12} as has been witnessed in Brazil since the IFRS adoption.

**Education level**

In spite of the two groups substantial differences in the CRT mean (Table 3 shows 1.59 for who conclude MBA or PhD and 1.50 for others), the only significant coefficient for education in multivariate data analysis was shown by Model 1 ($\beta = 0.16, p < 0.01$). Further investigation also demonstrates that education variable is significant only when the dummies for income are not included in the Equation 1. It probably means that income variables exert a mediation effect on the relationship between education level and cognitive score.

In such a case, mediation may occurs because there is an obvious causal influence of education on income, then, if income is included in the proposed model, education level has two pathways to influence CRT as outlined in the Figure 1. Unfortunately, Sobel test and other techniques are not applicable when sequential dummies are in mediator role, then, there is no known statistical tool to confirm this hypothesis\textsuperscript{13}. However, it is not a problem, once the focus of this research is not the specific mediation effects exerted by income on the relationship between education level and CRT score, but the cognitive abilities of accountants according to professional categories and other features.

Figure 1 – Intuitive design of the pathways to Education Level influence on CRT Score.

Thus, when considering accountants in general, it’s possible to assume that the differences in education level have a positive effect on cognitive score if we disregard the income effects\textsuperscript{14}. In this sense, the probability of owning a high CRT score (3) is 1.18 times higher if the accountant completed the MBA or the PhD level, holding constant the other variables and excluding income ones.

Based on a similar analysis for each professional category, i.e., also disregarding the income effects, education level is influent on cognitive score only among managers, and running the OLOGIT model for the 813 postgraduate managers together with the others 495 who completed only the bachelor level, we found that the probability of a manager owning a high CRT score is 1.45 times higher if he (she) completed the MBA or the PhD level. Once again, these analyses are made considering constants all other variables.

\textsuperscript{12} Beattie et al. (2011, p. 14-16) present a large body of studies that supports the IFRS complexity.

\textsuperscript{13} We regressed income as a quantitative variable (assuming values between one and seven according to the level stated in the Table 3) against education level applying an OLOGIT model for the five professional groups and for the entire sample (see a path on Figure 1). We applied an analogous process to regress CRT score against income (see b path on Figure 1). In both cases the coefficients were statistically significant for all groups. It reveals evidences of mediation effect, but we are not able to assure this, once the indirect effects were not tested.

\textsuperscript{14} ANOVA test ($F = 10.71, p = 0.001$) with CRT score as dependent and education level as independent variable shows that the differences are not by chance when considering the entire sample (n=4,902). Nevertheless, the same test for each professional group confirms that the difference is not by chance only for managers.
Our intuition suggests that this result can be assigned to the fact that education level becomes insignificant among accountants after years of training programs within companies, which probably make education abilities more homogeneous along the years as academic knowledge are forgotten and prevails what is being acquired over the career. On the other hand, managers are probably not affected by training within firms once they use not to be direct involved in technical issues and, then, attending to less training courses.

**Larger Companies’ Professional**

Although we have found no previous study applying this analysis, our intuition suggests that larger companies’ professionals use to work inside high departmentalized administrative structures. In such an environment they tend to develop repetitive and detailed duties connected to specific points of companies’ process, thus, their capacity to act reflectively should decrease due to the routine mental idleness.

On the other hand, we expect that accountants from small companies to work in an environment completely different, because the lower number of employees must force them to develop activities connected to the entire administrative process, which requires a more comprehensive cognitive reflection.

The results are as we expected. Descriptive statistics in Table 3 present the first evidences, once the CRT score mean is slightly higher for accountants who do not work at larger companies. Inferential statistics also corroborate such an assumption, in the sense that for the entire sample the probability of owning a high CRT score (3) is 0.85 times lower if the accountant works at a larger firm, holding constant the other variables.

Analyses for professional segments (Model 3) present analogous results for preparers and for managers. Among the 1,850 preparers, 730 larger co. workers and 1,120 not, the probability of owning a high CRT score (3) is 0.80 times lower if the accountant belongs to the group of larger co. workers. In the same sense, among 1,308 managers (642 larger co. workers and 666 not), the probability of getting the highest cognitive score (3) is 0.81 times lower for accountants who work for larger companies. Once again, we consider constant all other variables in these analyses.

Thus, those results for preparers and managers make sense with the aforementioned assumptions, once these professionals seem to be more exposed to the “high detailed and repetitive duties effect” in comparison with auditors, analysts and professors, whose work assignments tend to be broader even in larger companies.

Among the professional segments, this result holds for preparers and for managers (odds ratio equal to 0.80 and 0.81, respectively), what makes sense in light of the aforementioned assumptions, because preparers and managers seem to be more exposed to the “high detailed and repetitive duties effect”, once the work assignments of auditors, analysts and professors tend to be broader even in larger companies.

Considering that the present constitutes an exploratory analysis, we suggest further investigation in order to test the proposed theoretical approach.

**Monthly income**

The descriptive results presented in Table 3 point out a consistent growth of the mean CRT score analyzed from the lower income levels to the higher ones, except for the last income interval (i.e., more than BRL 31,100), in which we observe a slight decrease. A pronounced positive association between cognitive reflection score and monthly income was in fact found through inferential tests, as shown in Table 4.

The estimations of the models were made with the first monthly income level (from zero to BRL 1,866) as reference, then, the interpretation must take into account that the coefficients for income levels are relative to a shift from the first range to the analyzed one.
Hence, Model 1 demonstrates that there are no significant CRT score differences between the first and the second income levels, but the differences are growing along the remaining levels in the sense that the probability of getting the highest cognitive score (3) is 1.50 times greater for accountants who earn between BRL 3,110.01 and 6,220 when compared with professionals in the first range. The odds ratio values for the other income levels are 2.22, 2.47, 3.05 and 2.51, respectively. The only exception for such a tendency is for the last income level.

When considering only preparers and only professors (see Model 3 for each of these groups), the results are quite similar to the others, but any professor amongst the respondents earn more than BRL 18,660, than we have no data for the last two income levels. Considering exclusively managers (Model 3), the only shift in relation to the two previous cases is that the difference between the first and the second levels is statistically significant, whereas between the first and the last is not. The coefficients keep the tendency of increasing along the income levels increases, as we can observe in Table 4.

The evidences presented for the entire sample (i.e., Model 2) are in the sense that accountants with higher monthly incomes tend to make reflective decisions more frequently than the lower income professionals. Despite we have no further theoretical argument, these results can be connected with the fact that positions of greater responsibility tend to be better rewarded, as well as the judgments and decisions demanded from well-succeeded professionals may contain greater complexity, thus probably requires more reflective abilities. On the other hand, accountants in the lower income levels must develop less complex activities that also require less responsibility. Hence, they must be less used to reflect deeply on their decisions.

Notwithstanding, this pattern completely changes when considering the professional segments formed by auditors and analysts. According to the results stated in Table 4 for the Model 3, monthly income exerts a marginal role on the CRT Scores of those groups. Actually, the coefficient is negative for the second income range when considering only auditors, in the sense that the probability of getting the highest cognitive score (3) is 0.46 times lower for those who earn between BRL 1,866.01 and 3,110, when compared with those in the lowest income level. On the other hand, the tendency observed for other professional segments comes back for both, auditors and managers, when framed in the income range between BRL 18,660.01 and 31,100. In this case, the probability of scoring 3 at the CRT is 2.79 times higher for auditors who earn between BRL 1,866.01 and 3,110 when compared with auditors in the lowest income level. If we isolate the analysts, such a probability is 2.61 higher.

These results are aligned with our predictions, once auditors and analysts are expected to reflect heavily in their work routine regardless their income level because of the specificities inherent to their jobs.

Professional segment: further analysis

Considering that previous sections already discussed the main aspects of accountants’ professional segments relating it to CRT score and also to other demographic features, this last topic is a complementary one, useful for the addressing aspects not took into account yet.

Hence, according to the descriptive statistics presented in Table 3, auditors have the higher mean CRT score followed by managers, professors, analysts and preparers, in this order, but professors and analysts are tied with 1.54. Still, inferential outputs from Table 4 point out that, holding the preparers as the reference level (omitted dummy variable), the only statistically significant coefficient is for auditors. Notice that, the odds ratio value indicates that the probability of an accountant owning a high CRT score (3) is 1.20 times higher if he (or she) is an auditor, held constant all other variables. All other differences in mean CRT
scores among professional categories are only by chance as supported by the OLOGIT model (Model 2) estimations.

5 CONCLUSIONS

Judgment and decision making in accounting is an increasing avenue of research (Bonner, 2008), however, in its 40 decades of relevant publication, we did not have a large sample assessment of accountants’ cognitive abilities. This paper provides such an assessment in regard to the cognitive reflection test developed by Frederick (2005). Indeed, we analyze CRT scores in relation to demographic variables such as gender, age, formal education level, and income.

Either based on the analysis of proportions or on the ordered logistic regression’s odds ratios, the results suggest that (i) young, (ii) male, (iii) post-graduated, (iv) high-income and (v) auditors make more reflective decisions than (i) mature, (ii) female, (iii) not post-graduate, (iv) low-income and (v) preparers. According to Kahneman and Frederick (2002), Frederick (2005) and Kahneman (2011), intuitive answers are attributed to the fast ‘System 1’, while reflective answers are attributed to the slow ‘System 2’. Individuals that are more intuitive use to rely more frequently on the answers provided by System 1. Therefore, the results show that mature, female, low-income professionals that do not audit financial statements tend to be more confident on their System 1, than the other group of accountants.

Even if the working sample represents only 1.6% of the Brazilian population of accountants, it is reasonably balanced in terms of accountants’ gender, age, educational level, income and professional segment. It is also well balanced in terms of the geographic distribution over the Brazilian territory. Notice that Brazil has continental dimensions and its territory is politically divided among 27 federal states. Our sample has respondents from each state (data not presented in the paper, see CFC, 2013). Therefore, the working sample fairly and qualitatively represents the population of Brazilian accountants.

Every demographic variable analyzed in this paper is suitable for classifying accountants from any jurisdiction. Therefore, the results presented in this paper may be of interest of a wide audience, not only those interested in the Brazilian context.

Further international research could investigate if the findings presented in this paper are similar to those that could be obtained in other jurisdictions, especially in countries which environment are similar to the Brazilian environment. For such an assessment, the frameworks presented by Gray (1988) or Nobes (1998) could be of great use.

Further research on judgment and decision-making in accounting could investigate if the CRT score is associated with heuristic-bias and specific decision in the accounting context. For example: (i) is the interpretation of verbal probabilities (Doupink, Richter, 2003, 2004; Doupink, Riccio, 2006) associated with the accountants’ cognitive abilities?; (ii) is the classification of items in accordance with vague standards (Penno, 2008; Cardoso, Aquino, 2010) associated with the accountants’ cognitive abilities?; (iii) is the identification of impression management (Merkl-Davies, Brennan, 2011; Jones, 2011) associated with the accountants’ cognitive abilities?

Researchers interested in ‘Accounting and Gender’ could investigate the association of CRT scores and the fact that women hired into public accounting were not progressing up the ranks as expected (Dalton, Hill, Ramsay, 1997).

Researchers interested in ‘Accounting Education’ could further investigate the association of CRT scores and the formal education; and those interested in ‘Accounting Career’ could investigate the association of CRT scores and individual’s income.

Additionally, researchers not interested in Accounting, but in Cognitive Sciences, could analyze the impact of respondents’ recruitment strategy on the CRT scores. Notice that,
different from the majority of the researches that assessed CRT score recruited college students and offered them a monetary rewards; we recruited professional accountants and did not offer them any monetary reward. The CRT scores reported in this study are lower than the score presented in the reviewed literature (e.g., Canada: Toplak, West, Stanovich, 2011; Germany: Oechssler, Roider, Schmitz, 2009; Hoppe, Krusterer, 2011; and the U.S.A.: Frederick, 2005).

REFERENCES


