THE EFFECT OF COGNITIVE REFLECTION ON THE EFFICACY OF IMPRESSION MANAGEMENT: AN ANALYSIS WITH FINANCIAL ANALYSTS

Ricardo Lopes Cardoso
PhD in Accountancy, University of São Paulo
Associate Professor – EBAPE/FGV & FAF/UERJ

Rodrigo de Oliveira Leite
MSc in Public and Business Administration, EBAPE/FGV
PhD Student in Public and Business Administration, EBAPE/FGV
rodrigo.leite2015@fgv.br – Praia de Botafogo 190, Sala 428, Rio de Janeiro/RJ

André Carlos Busanelli de Aquino
PhD in Accountancy, University of São Paulo
Associate Professor – FEARP/USP

ABSTRACT:

The study of impression management on disclosure of financial information did not consider, up to this point, the effect of the user's personal traits. We performed three randomized experiments with 525 financial analysts, who also answered the Cognitive Reflection Test developed by Frederick (2005). Results support impression management literature concerning the efficacy of graphical impression management techniques such as presentation enhancement, measurement distortion and selectivity. Moreover, results identify that presentation enhancement (through the usage of colors) affected only fully impulsive analysts. Measurement distortion and selectivity had significant effects on both groups of impulsive and reflective analysts, although impulsive analysts assigned smaller ratings than their reflective peers on the measurement distortion experiment. Thus, we show that the cognitive reflection traits play an important role on the efficacy of graphical impression management techniques.

Keywords: cognition, cognitive reflection test - CRT, impression management with graph, financial analysts.

Área Temática do evento: Contabilidade para Usuários Externos.

1 Introduction

Imagine that you will go on to an interview for a job that you really want. You start to think in what you should dress for this interview. What does pass in your mind? You probably will not think in dressing in your favorite clothing or the clothing that you dress in your day-to-day life, but in a cloth that passes a good impression to the interviewer, because you want to make a good impression of yourself and be hired (Wilhelmy et al., 2015).

Companies are not different. They want to pass a good image to the market, and they use different types of "clothing" in order to achieve this goal. One of these is the use of graphics, which are widespread used in non-audited financial information (such as the management commentary and chief executive's review). Indeed, the usage of graphs in the note to the financial statements has increased recently. In this regard, the International Accounting Standards Board (IASB) recently amended its accounting standard on financial

instruments disclosure[1] requiring that "if the quantitative data disclosed as at the end of the reporting period are unrepresentative of an entity's exposure to risk during the period, an entity shall provide further information that is representative" (IFRS 7.35). Whether this is the case, the implementation guidance exhorts the presentation of graphs: "[...] if an entity typically has a large exposure to a particular currency, but at year-end unwinds the position, the entity might disclose a graph that shows the exposure at various times during the period [...]" (IFRS 7.IG20).

Although the impression management literature has developed in the last two decades, no study has investigated the influence exercised by individual's characteristics on the effects of impression management on their perception of graphical disclosure; this paper fills this gap.

2 Literature Review and Hypothesis Development

Impression management can be described as "the process by which people control the impressions others form of them" (Leary & Kowalski, 1990). However, companies also use impression management, because they want to enhance their "respectability and impressiveness" in the eyes of the public (Highhouse et al., 2009). This process is very connected to signaling. Impression management consists in signals that are chosen to manipulate the impression that third parts have. Literature provides many examples whether certain signals such as press announcements (Standish & Ung, 1982) and dividend policies (Akhigbe et al., 1993) have an effect on the market's perception of a company.

Beattie & Jones (1992) already noted that in 1989 the average number of graphs in annual reports by large UK companies was 5.9 graphs per report. In addition, they found that, in average, 30% of graphs were distorted with data being exaggerated in about 10.7%. Godfrey et al. (2003) found similar results in Australia, concerning impression management done shortly after a CEO change; i.e., the new CEOs often use impression management in graphs during the following year.

Chernoff (1973) created a facial representation of multivariate data (a.k.a. Chernoff faces), in which sad/worried faces would represent negative information and happy faces would represent positive information. Moriarty (1979) and Stock & Watson (1984) applied such innovation in accounting research; they represented negative financial results by sad/worried faces and happy faces represented positive financial results. Both studies found that the use of this type of graphical representation of financial reports changed subjects' judgment about companies' performance. Smith & Taffler (1996) found similar results across undergraduate, MBA students and academics. Carter (1947) identified that the different ways to represent data in a graphical form alter the decision-making process of pilots and students. Tang et al. (2014) found that different ways of graphical user-interaction had significant effects on the user's perception of a company.

Jones (2011) classified impression management of graphs in three sub-domains: presentation enhancement, measurement distortion and selectivity. Each one them inspired the three experiments carried out with financial analysts. Presentation enhancement manipulates how a graph is presented, such as color or 2-D vs. 3-D graph. Measurement distortion defines the change of scale of a graph to give a better impression to the market. Finally, selectivity

[1] International Financial Reporting Standard 7 (IFRS 7) – Financial Instruments: Disclosures.

defines the selecting only the data that conveys positive information to be disclosed. However, there is a gap in the literature in which intrinsic individual characteristics are not addressed in the impression management literature, although those characteristics have a strong influence on the financial decision-making of individuals (Oechssler et al., 2009).

In order to assess the effects of those manipulations on the financial analysts' rating of a company, we designed three different experiments and applied them to the financial analysts, in the form of a web-based experiment. Aiming to "fill the gap" in the impression management literature about intrinsic individual characteristics, the CRT scale (Frederick, 2005) was applied to classify the financial analysts between impulsive and reflective, or, according to Kahneman (2011), those in which System 1 preponderates System 2, and those in which System 2 preponderates System 1.

Daniel Kahneman also created the "Systems 1 and 2 Theory", in which he proposes that each and every one of us possess two "systems of thinking": one that is very impulsive (System 1) and the other that is more reflective (System 2) (Kahneman, 2011). While System 1 is always "turned on" and is "cognitively cheap" (requiring a low amount of effort), System 2 needs to be activated and is "cognitively expensive". As an example: how much is 2×2? Moreover, 137×363? Although it is the same operation (multiplication), the first question requires little effort, while the second requires a substantial amount of cognitive effort. That happens because we can answer the first question with the System 1, while we need to activate System 2 to answer the second question.

Frederick (2005) argued that there are people that are more reflective or more impulsive by nature. He adopted the Cognitive Reflection Test (CRT) in order to quantify how impulsive or reflective a person is. The CRT score can go from 0 to 3 with a 3 being fully reflective and a 0 being fully impulsive.

2.1 Cognitive Reflection Test

The original CRT, as proposed by Frederick (2005), consists of three questions:

1. "A bat and a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost?"[2]
2. "If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?"
3. "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?"

Each question has a wrong answer that is more intuitive, and a correct answer, that is more counter-intuitive. Thus, the person that correctly answered this question activated the System 2, while those that answered it in the wrong way did not activate the System 2 and used System 1 instead. Concerning the first question, for instance, the intuitive answer is $0.10, but if the ball costs $0.10, then the bat costs $1.00 that is only $0.90 more than the ball.

[2] In order to make the task more realistic for Brazilian respondents, we adjusted the costs to the price levels in Brazil, as suggested by Silva (2005). Therefore, in our questionnaire, we rewrote this question as follows: A bat and a ball cost $110. The bat costs $10 more than the ball. How much does the ball cost?

Therefore, the correct answer is: the bat costs $1.05 and the ball costs $0.05. For the second and third questions, the correct answers are: 5 minutes and 47 days, respectively.

Cokely & Kelley (2009) measured impulsiveness using the CRT scale and showed that impulsiveness harmed the decision-making of individuals. Koehler & James (2010) and Toplak et al. (2011) found similar results. In the behavioral economics field, Oechssler et al. (2009) proved that impulsiveness (measure with the CRT 3-items scale) affected in some cognitive biases that influenced economical decision-making.

Concerning the disclosure of information through graphics, Honda et al. (2015) founded that impulsiveness (measured with the CRT scale) had a negative effect on the experimental subjects decision-making regarding to food. However, there is no study about how impulsiveness may have an effect on the decision-making regarding graphical disclosure of financial information.

2.2 The Effect of Colors

Colors do have an important role in the "psychological functioning in humans" (Elliot & Maier, 2014). Mehta & Zhu (2008) identified that the red color (versus blue) creates more a motivation of avoidance, and enhances negative information. Indeed, Genschow et al. (2012) proved that the red color creates a sense of avoidance in real decision-making: red color diminished the snack food and soft drink consumption on the participants' choice. This "red vs. blue effect" is ostensibly used in marketing to try to manipulate the decision-making of consumers (Labrecque & Milne, 2012).

The color red, when compared to blue, harms decision-making in web-based tests of general knowledge (Gnambs et al., 2010). Even on animals, the red color is associated with negative signals, such as aggressiveness and intimidation (Pryke, 2009).

In the "corporate world", the red color has a negative connotation. In the English language, the term "red ink" is often related to a negative performance, and the term "blue ink" (or "black ink") is often related to a positive performance. Considering that the red color has a negative connotation and blue color has a positive, it is intuitive to present a positive performance with a blue graph and a negative performance with a red graph. Therefore, presenting a negative result in a blue graph is counter-intuitive and may be understood as an attempt to conceal a negative result with a "positive" color.

Thus, if an impression management through graphical presentation enhancement is considered to be effective, hence it should have an effect in both impulsive and reflective financial analysts. Therefore, we hypothesize that:

H1: *Presentation enhancement (color manipulation: blue vs. red) will have the same effect on impulsive and reflective financial analysts.*

2.3 Distorting Information

Companies often do use measurement distortion to enhance positive information and downplay the negative information (Beattie & Jones, 1993), aiming to portray a better image to the market. O'Reilly (1978) identified that, in the organizational field, the sender distorts

information presented to the receiver in order to portray a better image of himself. In addition, the desire to prove consistency is also a reason to distort the information to others (Russo et al., 2008).

In an experiment with undergraduate students, the subjects perceived companies whose graphics were distorted as performing better than other companies (Beattie & Jones, 2002). Such study replicated the research developed by Lawrence & O'Connor (1993). The results of measurement manipulation have effects on different areas such as contract formulations (Allen & Gale, 1992) and rating assignments (Beattie & Jones, 2002). Therefore, we hypothesize that:

H2: *Measurement distortion (use of different scales) will have the same effect on impulsive and reflective financial analysts.*

2.4 Information Disclosure

The selection process of information disclosure is an important factor in the development of financial reports (Healy & Palepu, 2001). According to Beattie & Jones (1996), companies often present graphs for "positive" information, but rarely present graphs for negative information.

If a graph does not present a negative information, only a positive one, then it may change the counterfactual that an analyst may have. If, for example, a company discloses that in the last 5 years its performance has increased, analysts would conclude that the company's performance prior to that period was even smaller. However, if a graph presents the information that six years ago the company's performance was better than it is today, than the counterfactual will change and also the analyst's conclusion. People often neglect the duration of information; hence, this effect could be amplified (Liersch & McKenzie, 2009). Medvec et al. (1995) clearly showed the importance of counterfactual comparison for Olympic medalists: those who were awarded the Silver medal were perceived as less happy than the ones who were awarded the Bronze one. This occurs because the Silver medalists compare themselves with the Gold ones, while the Bronze medalists compare themselves with those who got no medal. This effect is analogous to the prior example: by not disclosing negative information companies change the counterfactual of the analysts. Hence changes their perception.

Therefore, as in the first two hypotheses, if this impression management technique is effective, then it must alter the perception of both impulsive and reflective financial analysts in a similar manner. Then, we hypothesize that:

H3: *Selectivity (through the disclosure of positive information only) will have the same effect on impulsive and reflective financial analysts.*

3 Sample and Procedures

The sample consists of 525 financial analysts regularly registered at the CFC (*Conselho Federal de Contabilidade*, Brazilian Accounting Association) at the time of data collection (from August 2012 until July 2013). They responded, for this study, three web-based experiments, and also the three CRT questions and their demographics (gender, age,

place of residence and monthly income). In total, 530 individuals responded the questionnaire, but four participants did not inform their monthly income, and one did not inform his age, hence we excluded them from the analysis. Thus, the final sample consists of 525 individuals.

3.1 Randomization Test

Each experiment had two conditions (i.e., a manipulated graph and a non-manipulated graph) randomly assigned to the participants. The expected result of randomization is that both groups are comparable, and the only difference is the treatment assignment. To test for randomization we performed a series of t-tests in the demographic questions: gender, age, place of residence and monthly income (low, medium and high). Table 1 presents the aforementioned results.

To avoid the Type I error in 18 tests across three experiments with the same sample, we used the Bonferroni correction. The Bonferroni correction is credited as being conservative, since it assumes there is no correlation between the groups. However, this assumption is fulfilled in this case, since each one of the groups is comprised by two random subgroups, the expected correlation between the experiments assignment (control and treatment) is zero, and in practice it was very small ($r_{1,2}$= .003, $r_{1,3}$= .016 and $r_{2,3}$= .027).[3]

The results presented no significant difference in the corrected 95% overall confidence level, and, therefore, we assume that the randomization worked in making the groups comparable in the three experiments. Another way to assess if there is randomization balance is the rule of thumb that the differences in means should be smaller than ¼ of the standard deviation (Rosenbaum & Rubin, 1985; Rubin, 2001). As Table 1 shows, all differences are smaller than the threshold of .25, therefore proving that the groups in all three experiments have randomization balance. Therefore, we can claim exogeneity for this study.

Table 1: Randomization tests

		Difference in means $[\Delta M/SD]^{\alpha}$		
Variable	**Mean (SD)**	**Exp. 1**	**Exp. 2**	**Exp. 3**
Gender (1 = Male)	.657 (.475)	0.078 [.164]	0.011 [.023]	0.019 [.040]
Age (Years)	39.60 (10.93)	1.247 [.114]	0.291 [.027]	0.635 [.058]
Residence (1 = Capital)	.675 (.469)	0.078 [.166]	0.047 [.100]	0.021 [.045]
Low Income (1 = Yes) (<USD 10,593 p.a.)$^{\beta}$	.132 (.339)	0.064 [.189]	0.008 [.024]	0.011 [.032]
Middle Income (1 = Yes) (USD 10,593–USD 35,310 p.a.)	.511 (.500)	0.019 [.038]	0.010 [.020]	0.024 [.048]
High Income (1 = Yes) (>USD 35,310 p.a.)	.352 (.478)	0.084 [.176]	0.006 [.013]	0.041 [.086]

Bonferroni corrected p-values: * p < .05, ** p < .01, *** p < .001. Std. Deviation in parenthesis.

[3] Notice that a 95% confidence level at each experiment produces only a 86% ($.95^3$) overall confidence level across all three experiments. Because the same comparison is performed three times (one for each experiment), and the sample is the same for all three experiments (although the sub-sample in each condition is different due to randomization) the Bonferroni correction was adopted. The confidence interval adopted was 98.3% ($100 - 5^3 = 98.3$) instead of the usual 95%, to secure a 95% ($.983^3 = .95$) confidence level across all three experiments.

[α] Values in brackets denote the mean difference divided by the std. deviation.
[β] Annual income was asked to participants in BRL (the Brazilian currency), we converted into U.S. dollars at the average rate for the period of data collection, i.e., USD 1 = BRL 2.29.

3.2 Data analysis criteria

The results were estimated using OLS regressions. The OLS regression requires continuous dependent variable, but rating is not continuous. However, since rating is a linear variable and OLS regressions are simpler to interpret and gives a true R^2, therefore, we used this estimation analysis in this study. In addition, the OLS regression with dummies gives the same p-values as the ANOVA test.

Three models were estimated in each experiment. The first tests the direct effect of the experiment, while the second also tests the CRT effect and the interaction. The third model is the same as the second, but using control variables (gender, age, residence and income).

Table 2 presents the descriptive statistics.

Table 2 – Descriptive Statistics

			Correlations			
Variables	**Mean**	**SD**	**Exp. 1**	**Exp. 2**	**Exp. 3**	**CRT**
Experiment 1 (Ratings)	3.72	2.15	1			
Treatment = blue color	*3.67*	*2.20*				
Control = red color	*3.76*	*2.11*				
Experiment 2 (Ratings)	8.42	1.84	-0.114	1		
Treatment = small scale	*8.90*	*1.72*				
Control = large scale	*7.88*	*1.81*				
Experiment 3 (Ratings)	6.85	1.57	-0.185	0.274	1	
Treatment = partial disclosure	*7.26*	*1.61*				
Control = full disclosure	*6.51*	*1.45*				
Cog. Ref. Test (CRT Score)	1.41	0.95	0.054	0.038	0.134	1
CRT Dummy	*0.41*	*0.49*				

The ratings from experiments 1, 2 and 3 range from 0 (very bad performance) to 10 (very good). For each experiment, it presents the mean and standard deviation for the entire sample, and for each condition (i.e., treatment and control). The CRT score range from 0 (fully impulsive) to 3 (fully reflective), and the CRT dummy is binary, where 0 (impulsive, i.e., CRT score 0 or 1) and 1 (reflective, i.e., CRT score 2 or 3).

4 The Experiments

The tasks in all three experiments are the same: giving a rating (0-10 range) to a hypothetical company based on a graph (a different graph in each condition) that represents its performance (net profit graphs). After the three experiments, participants responded the 3-questions CRT designed by Frederick (2005).

4.1 First Experiment: Presentation Enhancement

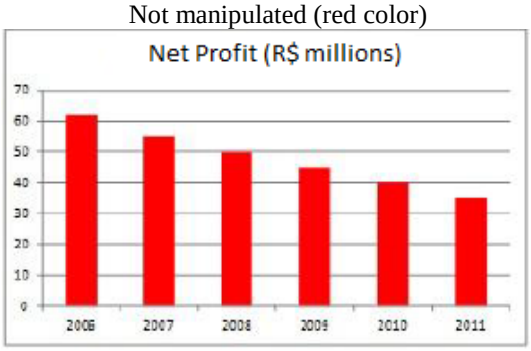
4.1.1 Experimental Setup

The first experiment consisted in randomly assigning the participants to one of two different bar graphs showing a decrease in the profits for a given company each year. The manipulation consisted in one graph being red, and the other being blue. Figure 1 shows both graphs that were presented to subjects, and the question they answered concerning such a graph. Impulsiveness was measured using a 3-questions test (CRT score), those that correctly answered 2 or 3 questions were coded as 1 (reflective) and those that correctly answered 0 or 1 of the questions were coded as 0 (impulsive).

This experiment explores the presentation enhancement sub-domain of impression management, as expressed by Jones (2011). The red color in graphics is often related to a sense of avoidance (Mehta & Zhu, 2008), what may enhance the negativity that is already present in the graph (decrease in profits). Companies often do use different colors to draw attention to the last year's results (Beattie & Jones, 1993). Therefore, we expect that the subjects who received a blue graph attributed a better rating than those that received the red graph.

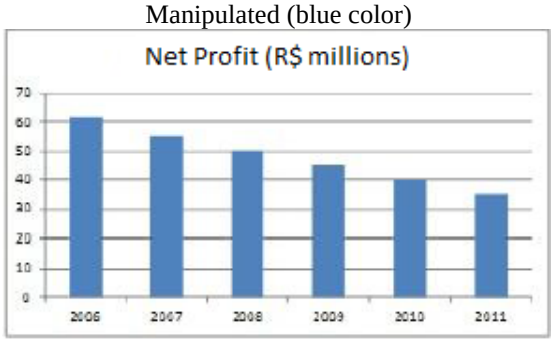
Figure 1: First Experiment conditions

Not manipulated (red color)



Respondents: 50%

Manipulated (blue color)



Respondents: 50%

Based exclusively on this information, how would you evaluate the performance of this company? Please, attribute a rating from 0 (very bad) to 10 (very good).

4.1.2 Results

On average, respondents that received the red graph attributed a rating of 3.76, and those that received the blue graph attributed 3.67. Parting the sample between impulsive subjects (i.e., those that correctly answered no more than 1 CRT question) and reflective subjects (i.e., those that correctly answered at least 2 CRT questions); impulsive (reflective) analysts attributed 3.43 (3.99) for the red graph and 3.83 (3.66) for the blue graph, on average.

We calculated two t-tests to assess if gender was a confounder in this study, i.e. if men and women perceived differently the colors. There was no significant difference between the rating assessed by both genders in regard to the red graph ($Mean_{Women}$= 3.76, $Mean_{Men}$= 3.61; t(259) = 0.52; p = .606) and the blue one ($Mean_{Women}$= 3.78, $Mean_{Men}$= 3.75; t(267) = 0.09, p = .925). Thus, gender is not a confounder in this study.

Table 3 presents the results of OLS regression. Although the main effect of the color manipulation was not significant (standardized β = .99, n.s.), being impulsive or reflective had a significant impact on graph interpretation, with impulsive financial analysts giving a lower score (0.6 point in average) than reflective ones (standardized β = .137, p = .026) in the

not manipulated graph. The interaction between the color manipulation and the CRT dummy was marginally significant (standardized β = .137, p = .053).

Figure 2 shows the difference in rating for the fully impulsive (CRT score = 0) versus the fully reflective (CRT Score = 3). While the manipulation increased the rating of the fully impulsive ($Mean_{Red}$= 2.98, $Mean_{Blue}$= 4.33; t(87) = 2.75, p = .007), it did not affect the rating of the fully reflective ($Mean_{Red}$= 4.23, $Mean_{Blue}$= 3.82; t(84) = .87, n.s.). This gives evidence that an impression management of financial information with color only have an effect on impulsive financial analysts, but not on reflective ones. Therefore, results reject the first hypothesis, because the presentation enhancement only had an effect on the fully impulsive group.

Table 3: OLS Results for Experiments 1, 2 and 3

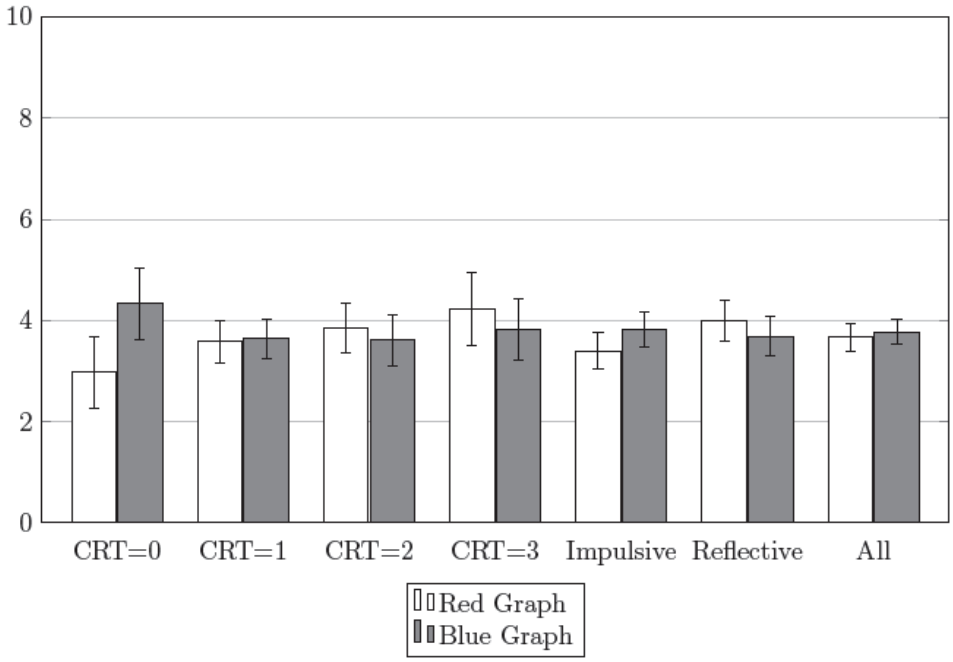
	Experiment 1 (Presentation Enhancement)			**Experiment 2 (Measurement Distortion)**			**Experiment 3 (Selectivity)**		
	Model 1	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
Experiment	0.110	0.425	0.388	1.005***	0.950***	0.957***	0.739***	0.886***	0.933***
(1 = manipulated)	(0.188)	(0.245)	(0.246)	(0.155)	(0.203)	(0.204)	(0.134)	(0.173)	(0.173)
CRT score		0.601*	0.594*		0.121	0.118		0.538**	0.539**
(0 = low, dummy)		(0.270)	(0.273)		(0.227)	(0.234)		(0.182)	(0.183)
Interaction		-0.742	-0.770*		0.161	0.138		-0.312	-0.378
		(0.382)	(0.381)		(0.315)	(0.319)		(0.271)	(0.272)
Control variables	No	No	Yes	No	No	Yes	No	No	Yes
Observations	525	525	525	525	525	525	525	525	525
Adjusted R^2	0.001	0.005	0.009	0.073	0.073	0.073	0.053	0.068	0.072
F	0.342	1.858	1.570	42.176	14.708	6.143	30.502	13.663	6.055

Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

While this experiment only had a "soft" impression management (color manipulation), the next experiment will have a "stronger" impression management: changing the scales of a graph.

Figure 2: Rating results for Experiment 1


10
8
6
4
2
0
CRT=0
CRT=1
CRT=2
CRT=3
Impulsive
Reflective
All
Red Graph
Blue Graph

4.2 Second Experiment: Measurement Distortion

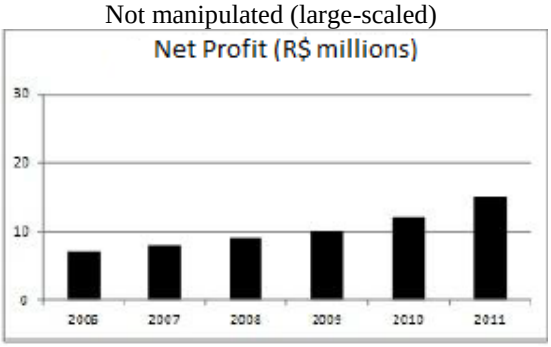
4.2.1 Experimental Setup

In the second experiment the participants were randomly assigned to one of two different types of bar graphs of a time series from a given company's profits that were increasing, one with a large scale (0 - 30 million) and the other with a small scale (0 – 15 million), as presented in Figure 3. Compared to the large scale, the small-scaled graph had bigger bars and a bigger slope. Then the participants were asked to rate the company's performance in a 0-10 range, in the same way as in the first experiment.

This second experiment explores the measurement distortion sub-domain (Jones, 2011). Therefore, in theory, a company would use a small-scaled graph to convey positive information, but a large-scaled graph to convey negative information.

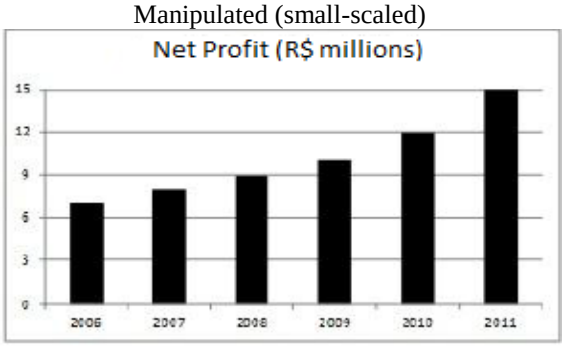
Figure 3: Second Experiment Conditions

Not manipulated (large-scaled)



Respondents: 50%

Manipulated (small-scaled)



Respondents: 50%

Based exclusively on this information, how would you evaluate the performance of this company? Please, attribute a rating from 0 (very bad) to 10 (very good).

4.2.2 Results

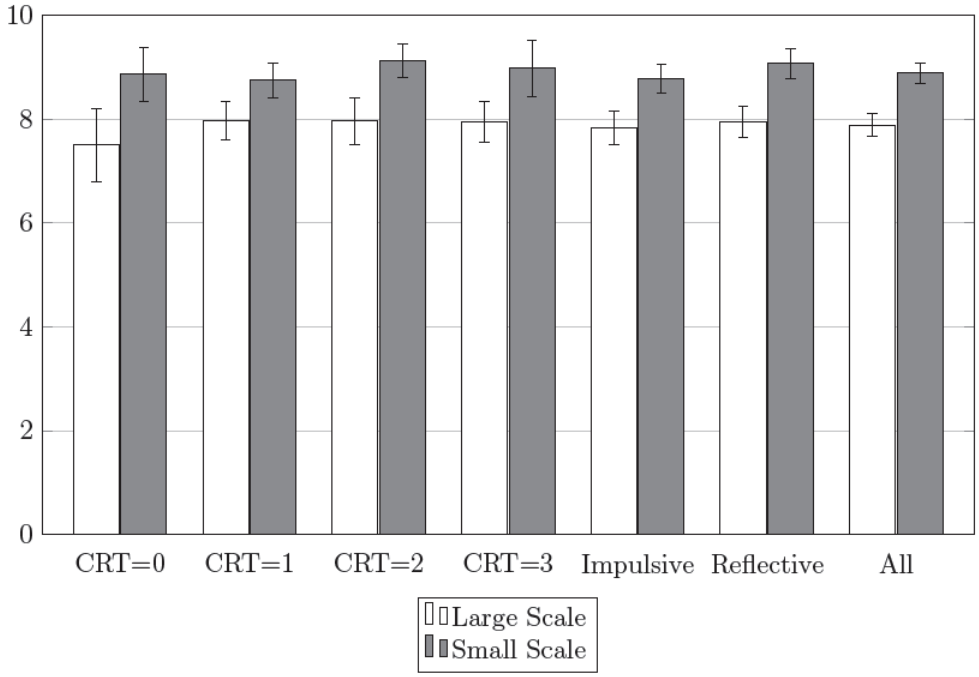
On average, respondents that received the large-scaled graph attributed a rating of 7.88, and those that received the small-scaled graph attributed 8.90. Parting the sample between impulsive and reflective subjects; impulsive (reflective) analysts attributed 7.82 (7.95) for the large-scaled graph and 8.78 (9.08) for the small-scaled graph, on average.

Using OLS regression, we estimated the results for the second experiment, in the same way as in the first experiment. The manipulation had a significant positive effect (1 point increase in average) on the rating given by the analysts (standardized β = .258, p<.001), being impulsive or reflective did not have an effect on the rating (standardized β = .032, n.s.), and the interaction was not significant (standardized β = .035, n.s.).

Figure 4 shows the experiment's results. Changing the scale of the graph had a large impact in the overall rating, and successfully misled both impulsive and reflective financial analysts. Changing the scale of the graph increased both bar height and slope. Impulsive financial analysts may be influenced by: (i) the height of the bar, and (ii) the fact that the last bar reached the top of the small-scaled graph (suggesting that the

company achieved its goal). On the other hand, in the large-scaled graph the last bar reached only the middle of the graph, suggesting that the company failure in achieving its goal. Additionally, the slope increase may suggest a greater progress of the company; hence, a better performance. Therefore, results support the second hypothesis.

Figure 4: Rating results from Experiment 2



In these two experiments, the same information was present in both graphs, and the manipulation only changed how information was disclosed. However, how do financial analysts perceive graphs with different information being disclosed? This is the objective of the third experiment.

4.3 Third Experiment: Selectivity

4.3.1 Experimental Setup

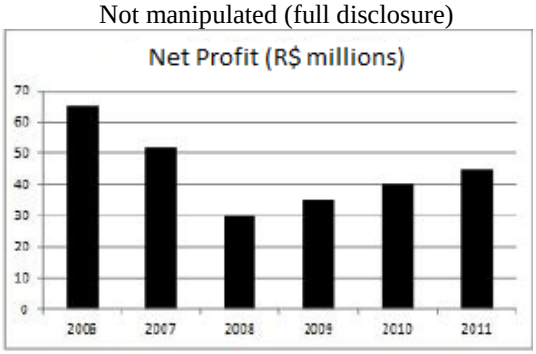
Participants were randomly assigned to one of the two graphical disclosures. One showing a profit increase from BRL 30 million to BRL 42 million (between 2008 and 2011). The other showing the same information from the previous graph, but with the pre-crisis figures (2006 and 2007) of BRL 62 million and BRL 51 million (much better than the performance from 2008 to 2011). Therefore, the graph that had an impression management showed less information to the financial analyst than the other, as denoted in Figure 5. Then, participants were asked to rate the company's performance in a 0-10 range scale, in the same way as in prior experiments.

This third experiment explores the concept of selectivity, i.e., the act of a company selecting only positive information to disclose in its graphs (Jones, 2011). In the case of this experiment a company selected only positive information to disclosure,

but the control had negative (drastically reduction of profits in 2008) and positive information (gradual recover in the three following years).

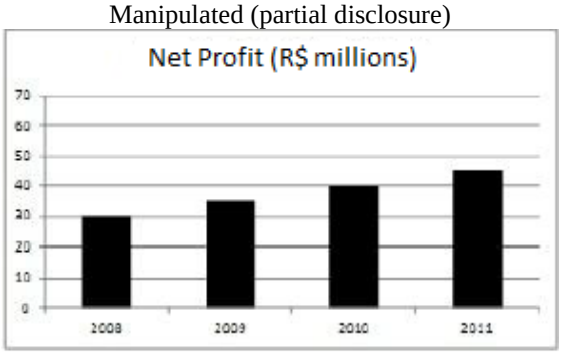
Figure 5: Third Experiment Conditions

Not manipulated (full disclosure)



Respondents: 50%

Manipulated (partial disclosure)



Respondents: 50%

Based exclusively on this information, how would you evaluate the performance of this company? Please, attribute a rating from 0 (very bad) to 10 (very good).

4.3.2 Results

On average, respondents that received the full disclosure graph (pre- and post-crisis information) attributed a rating of 6.51, and those that received the only post-crisis information (partial disclosure graph) attributed 7.26. Additionally, impulsive analysts attributed 6.28 for the full disclosure graph and 7.17 for the partial disclosure graph, on average. On the other hand, reflective analysts attributed on average 6.81 and 7.40, respectively.

As like as in the previous experiments, we used OLS regression to estimate results. The manipulation was successful in increasing (.8 point in average) the company's rating (standardized β = .281, p < .001). There was a significant effect of impulsiveness in the rating, with impulsive financial analysts attributing a lower rating (.538 point in average; standardized β = .169, p = .003) for the non-manipulated graph. However, the interaction was not significant (standardized β = -.076, n.s.).

Figure 6: Rating results from Experiment 3

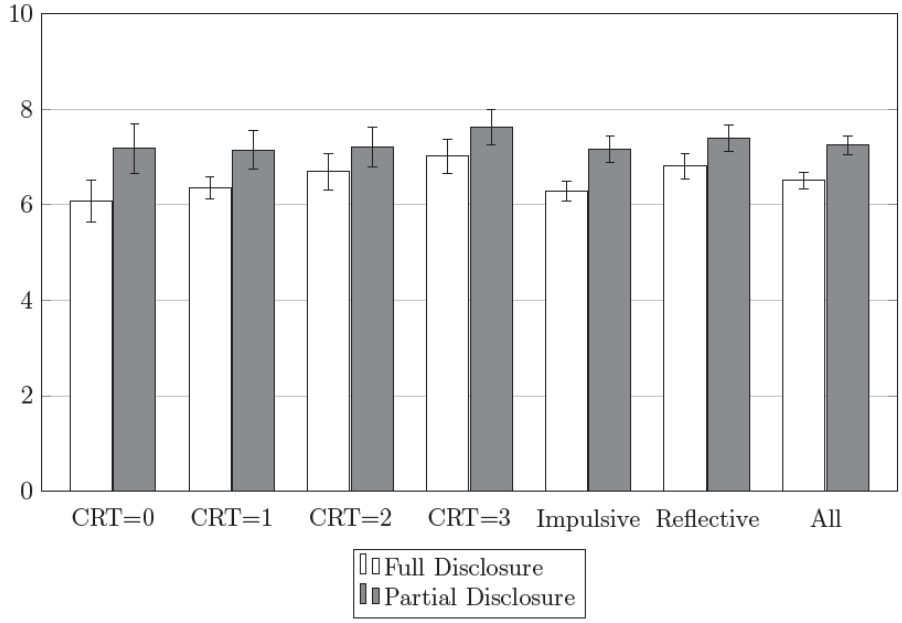


Figure 6 shows the comparison between the impulsive and reflective financial analysts in both the treatment and control groups. It clearly shows the two main effects as significant, while the interaction is not significant. Differently from the second experiment, the third experiment had the main effect of the CRT dummy as significant; it supports the theory that the manipulated graph from the second experiment suffered an impression management that manipulated both impulsive and reflective individuals. Therefore, results partially support the third hypothesis. The manipulation increased the rating of both impulsive and reflective financial analysts, but impulsive financial analysts still attributed a smaller rating than reflective financial analysts did. Table 3 summarizes the results of this study.

Table 3: Hypotheses summary

Hypoth.	Description	Supported?
H1	Presentation enhancement will have the same effect on impulsive and reflective financial analysts.	No
H2	Measurement distortion will have the same effect on impulsive and reflective financial analysts.	Yes
H3	Selectivity will have the same effect on impulsive and reflective financial analysts.	Partially

5 Concluding remarks

The three experiments provide many conclusions. The first is that, although different ways of graphical disclosure and presentation have different effects on the financial analysts' perception of a company's performance, not all types of presentation and disclosure have the same effect. Changing the color of a graph had an effect on the rating of fully impulsive financial analysts, while it had no effect on the others. Changing the scale of a graph had an effect on both impulsive and reflective financial

analysts, while homogenizing the rating between both groups (no CRT score effect). Disclosing only positive information had a positive effect when comparing to the disclosure of positive and negative information, but impulsive financial analysts assigned a lower rating to the company when compared to their reflective peers.

The experimental approach used in this study allowed for analyzing different ways of impression management with a high degree of internal validity. The sample was comprised by professional financial analysts; enhancing the external validity of this research. However, the ratings attributed by the analysts were theoretical, with no consequence for the participants, what hinders the external validity of this study.

The effect of the manipulations were small (the R^2 for the first experiment was .08 in the fully impulsive subgroup and .07 for the $AdjR^2$ of the second and third experiments), however, since we only manipulated minor graphical disclosure details, it is expected that those manipulations do not have tremendous power of explanation.

This study analyzed three different ways of impression management with graphs. However, it did not tested how the interaction of these different ways of manipulation can affect the financial analysts' perception. Does color moderates the effect of measurement distortion? How the scale may change the perception of different amounts of information that are disclosed? Those are questions were not investigated in this study; notwithstanding, they are interesting topics for further research.

6. Bibliography

Akhigbe, A., Borde, S. F., & Madura, J. (1993). Dividend policy and signaling by insurance companies. *Journal of Risk and Insurance*, 60(3), 413-428.

Allen, F. & Gale, D. (1992). Measurement distortion and missing contingencies in optimal contracts. *Economic Theory*, 2(1), 1-26.

Beattie, V. & Jones, M. (1996). *Financial graphs in corporate annual reports: A review of practice in six countries*. Institute of Chartered Accountants in England and Wales.

Beattie, V. & Jones, M. J. (1992). The Use and Abuse of Graphs in Annual Reports: Theoretical Framework and Empirical Study. *Accounting and Business Research*, 22(88), 291-303.

Beattie, V. & Jones, M. J. (2002). Measurement distortion of graphs in corporate reports: an experimental study. *Accounting, Auditing & Accountability Journal*, 15(4), 546-564.

Beattie, V. A. & Jones, M. J. (1993). Information design and manipulation: Financial graphs in corporate annual reports. *Information Design Journal*, 7(3), 211-226.

Carter, L. F. (1947). An experiment on the design of tables and graphs used for presenting numerical data. *Journal of Applied Psychology*, 31(6), 640-650.

Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*, 68(342), 361-368.

Cokely, E. T. & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, 4(1), 20-33.

Elliot, A. J. & Maier, M. A. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology*, 65, 95-120.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4), 25-42.

Genschow, O., Reutner, L., & Wänke, M. (2012). The color red reduces snack food and soft drink intake. *Appetite*, 58(2), 699-702.

Gnambs, T., Appel, M., & Batinic, B. (2010). Color red in web-based knowledge testing. *Computers in Human Behavior*, 26(6), 1625-1631.

Godfrey, J., Mather, P., & Ramsay, A. (2003). Earnings and Impression Management in Financial Reports: The Case of CEO Changes. *Abacus*, 39(1), 95-123.

Healy, P. M. & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, 31(1), 405-440.

Highhouse, S., Brooks, M. E., & Gregarus, G. (2009). An Organizational Impression Management Perspective on the Formation of Corporate Reputations. *Journal of Management*, 35(6), 1481-1493.

Honda, H., Ogawa, M., Murakoshi, T., Masuda, T., Utsumi, K., Park, S., Kimura, A., Nei, D., & Wada, Y. (2015). Effect of visual aids and individual differences of cognitive traits in judgments on food safety. *Food Policy*, 55, 33-40.

Jones, M. (2011). Impression Management. In *Creative Accounting, Fraud and International Accounting Scandals* (pp. 97-114).

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Koehler, D. J. & James, G. (2010). Probability matching and strategy availability. *Memory & Cognition*, 38(6), 667-676.

Labrecque, L. I. & Milne, G. R. (2012). Exciting red and competent blue: the importance of color in marketing. *Journal of the Academy of Marketing Science*, 40(5), 711-727.

Lawrence, M. & O'Connor, M. (1993). Scale, variability, and the calibration of judgmental prediction intervals. *Organizational Behavior and Human Decision Processes*, 56(3), 441-458.

Leary, M. R. & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, 107(1), 34-47.

Liersch, M. J. & McKenzie, C. R. (2009). Duration neglect by numbers and its elimination by graphs. *Organizational Behavior and Human Decision Processes*, 108(2), 303-314.

Medvec, V. H., Madey, S. F., & Gilovich, T. (1995). When less is more: counterfactual thinking and satisfaction among Olympic medalists. *Journal of Personality and Social Psychology*, 69(4), 603-610.

Mehta, R. & Zhu, R. (2008). Blue or Red? Exploring the Effect of Color on Cognitive Task Performances. *Science*, 323(Feb.), 1226-1229.

Moriarty, S. (1979). Communicating Financial Information through Multidimensional Graphics. *Journal of Accounting Research*, 17(1), 205-224.

Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization*, 72(1), 147-152.

O'Reilly, C. A. (1978). The intentional distortion of information in organizational communication: A laboratory and field investigation. *Human Relations*, 31(2), 173-193.

Pryke, S. R. (2009). Is red an innate or learned signal of aggression and intimidation? *Animal Behaviour*, 78(2), 393-398.

Rosenbaum, P. R. & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1), 33-38.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2(3-4), 169-188.

Russo, J. E., Carlson, K. A., Meloy, M. G., & Yong, K. (2008). The goal of consistency as a cause of information distortion. *Journal of Experimental Psychology: General*, 137(3), 456.

Silva, J. S. (2005). Se um taco e uma bola custam R$110, e o taco custa R$100 a mais do que a bola, quanto custa a bola? *Thesis: Master in Business Administration*. Fundação Getulio Vargas. Rio de Janeiro (Brazil).

Smith, M. & Taffler, R. (1996). Improving the communication of accounting information through cartoon graphics. *Accounting, Auditing & Accountability Journal*, 9(2), 68-85.

Standish, P. E. & Ung, S.-I. (1982). Corporate signaling, asset revaluations and the stock prices of British companies. *Accounting Review*, 57(4), 701-715

Stock, D. & Watson, C. J. (1984). Human judgment accuracy, multidimensional graphics, and humans versus models. *Journal of Accounting Research*, 22(1), 192-206.

Tang, F., Hess, T. J., Valacich, J. S., & Sweeney, J. T. (2014). The Effects of Visualization and Interactivity on Calibration in Financial Decision-Making. *Behavioral Research in Accounting*, 26(1), 25-58.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, 39(7), 1275-1289.

Wilhelmy, A., Kleinmann, M., König, C. J., Melchers, K. G., & Truxillo, D. M. (2015). How and why do interviewers try to make impressions on applicants? A qualitative study. *Journal of Applied Psychology*, forthcoming.