A LINGUAGEM FAZ DIFERENÇA? UM EXAME EXPERIMENTAL SOBRE *MOTIVATED SENSITIVITY* NO JULGAMENTO DO AUDITOR GOVERNAMENTAL BRASILEIRO

DOES LANGUAGE MAKE A DIFFERENCE? AN EXPERIMENTAL EXAM ON THE MOTIVATED SENSITIVITY IN THE JUDGEMENT OF BRAZILIAN GOVERNMENTAL AUDITING

ROSSANA GUERRA DE SOUSA
EDILSON PAULO
CÉSAR AUGUSTO TIBÚRCIO SILVA

ROSSANA GUERRA DE SOUSA
Doutoranda pelo Programa Multi-Institucional e Inter-Regional de Pós-Graduação em Ciências Contábeis (UnB/UFPB/UFRN) da Universidade Federal da Paraíba (UFPB).
E-mail: rossanagsousa@yahoo.com.br

EDILSON PAULO
Doutor em Ciências Contábeis pela Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (USP). Professor adjunto da Universidade Federal da Paraíba (UFPB). Docente do Programa Multi-Institucional e Inter-Regional de Pós-Graduação em Ciências Contábeis (UnB/UFPB/UFRN).
E-mail: epaulo@ccsa.ufpb.br

CÉSAR AUGUSTO TIBÚRCIO SILVA
Doutor em Ciências Contábeis pela Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (USP). Professor titular da Universidade de Brasília (UnB). Docente do Programa Multi-Institucional e Inter-Regional de Pós-Graduação em Ciências Contábeis (UnB/UFPB/UFRN).
E-mail: cesartiburcio@unb.br
Endereço: UnB/FACE/CCA - Campus Universitário Darcy Ribeiro - Prédio da FACE - Salas B1-02 - Asa Norte
Brasília - DF - CEP 70910-900

Recebido em: 13.08.2012.
Revisado por pares em: 15.10.2012.
Aceito em: 04.12.2012.
Publicado em: 30.04.2013.
Avaliado pelo sistema *double blind review.*

Resumo:

Este estudo busca examinar a ocorrência de *motivated sensitivity* na análise de informações econômicas, financeiras e patrimoniais das entidades realizada pelo auditor governamental brasileiro em decorrência da influência da preferência linguística em seu processamento. A metodologia utilizada foi o experimento de campo com a exposição de dois grupos aleatoriamente constituídos de auditores governamentais, de órgãos de fiscalização interna e externa dos estados e do governo federal do Brasil à manipulação da variável linguística entre textual e numérica. Os dados coletados foram analisados com a utilização da Anova e do Qui-Quadrado. Os resultados indicam que a preferência por estilos linguísticos guia a atenção e a interpretação das análises realizadas, podendo interferir na percepção de confiabilidade das informações postas sob exame dos auditores governamentais e na formulação do julgamento profissional. A pesquisa contribui para o entendimento de parte do processo de julgamento e dos efeitos de fatores motivacionais desse grupo e pode guiar a adoção de ações de mitigação do risco para os casos de julgamentos relevantes.

Palavras-chave: Auditoria governamental. Julgamento profissional. *Motivated sensitivity.*

Abstract: *This study aimed at examining the occurrence of motivated sensitivity in the analyses of economical, financial and patrimonial information carried out by Brazilian government auditors, which results from the influence of linguistic preference in their processing. Methodology used was field experiment: two groups formed at random were exposed to the manipulation of the linguistic variable (textual and numerical). Subjects were government auditors from an agency in charge of inspecting both Brazilian states and the federal government. Collected data was analyzed by using Anova and Chi-Squared test. Results indicate that the preference for linguistic styles guide the attention and interpretation of analyses carried out, which may interfere with the perception of reliability of the information being examined by government auditors, as well as with reaching professional judgment. This research contributes to the understanding of part of the judgment process, as well as of motivational factor effects acting on these professionals. It may also guide actions of risk mitigation in the case of relevant judgments.*

Keywords: *Governmental auditing. Professional judgment. Motivated sensitivity.*

1 INTRODUÇÃO

Os estudos sobre auditoria comportamental estão presentes na literatura científica focados especialmente na dimensão psicológica envolvida no julgamento e na tomada de decisões e utilizando como fonte central de investigação os auditores independentes no processo de análise das demonstrações financeiras.

A abordagem psicológica que integra a contabilidade, a auditoria e a pesquisa cognitiva (SHANTEAU, 1989), materializada no estudo do julgamento, envolto em sua subjetividade (HOGARTH, 1991) e presente em todas as fases do processo, tem apresentado alterações ao longo do tempo.

Desde a década de 1970, segundo Koch e Wustermann(2009), com os trabalhos de Tverskye Kahneman(1974), essa área de estudo foi estimulada e tornou-se pauta frequente nas pesquisas científicas, examinando prioritariamente as heurísticas da ancoragem e representatividade. Ainda na área da auditoria comportamental, destacam-se os trabalhos de HogartheEinhorn (1992) sobre *belief-adjustmentmodel*, evoluindo para o exame de vieses orientados a questões específicas de auditoria, investigados sob o enfoque do comportamento de auditores independentes.

O futuro da pesquisa em auditoria era direcionado por Shanteau (1989) a mover-se na direção de mais e melhores experimentos sobre os processos de decisão e julgamento, em oposição às demonstrações de ocorrências de heurísticas, baseadas, preferencialmente, em conceitos quantitativos e tendo por meta entender e não apenas descrever o processo de julgamento.

Comunale, Sexton e Sincich (2005) apontam mais um caminho a percorrer nas alterações das abordagens de pesquisa em auditoria, quando indicam a ocorrência de limitados estudos na literatura sobre a questão relacionada à comunicação e julgamento, especificamente no que tange ao estilo linguístico utilizado, que pode, segundo os autores, influenciar o julgamento do auditor.

Na busca por novos caminhos nas investigações em auditoria, identifica-se, adicionalmente uma categoria representativa de auditores, que em sua diversidade, não tem sido contemplada pela pesquisa em auditoria comportamental: os auditores governamentais. Atuantes em todos os países, com suas peculiaridades locais, as atividades desse corpo de auditores difere do tradicional enfoque de emissão de opiniões e julgamentos sobre instrumentos financeiros, não prescindindo, entretanto, do julgamento profissional no exercício de seu mister.

No Brasil, em decorrência da estrutura de controle governamental, os auditores atuam no âmbito do controle externo, ligado ao legislativo, e do controle interno, exercendo suas atribuições junto ao poder executivo nas três esferas de governo. Esses profissionais são rotineiramente submetidos à realização de julgamento, além do foco tradicional de análise de demonstrações contábeis e da documentação a ela correlacionada, quando instados a se posicionarem sobre a regularidade ou adequação de processos públicos, como compras, contratações, questões legais e organizacionais. Esses julgamentos são realizados a partir de material enviado aos órgãos de controle para exame, cujo conteúdo e formato não apresentam a uniformidade de estrutura e conteúdo em geral presentes nas análises financeiras realizadas pelos auditores independentes.

Materializadas em forma de pareceres ou relatórios de auditoria, a opinião do auditor pode ter forte impacto nas questões de interesse coletivo relacionadas à mitigação de ocorrências de corrupção, favorecimento ou desvios de recursos públicos, como destacado nos estudos de Schelker e Eichenberger (2010), que indicam que a auditoria governamental aumenta a transparência das políticas públicas e reduz o desperdício dos gastos. Ferraz e Finan (2011, 2008), ainda em estudos sobre o impacto das auditorias governamentais, apontam achados no sentido de sua influência no processo eleitoral no cenário brasileiro, enquanto que Liu e Lin (2012)trazem evidências de sua importância no controle da corrupção.

O contexto em que é exercida a auditoria governamental conduz a diferenciações entre as abordagens e julgamentos dos auditores governamentais e corporativos (independentes), conforme relatado nos estudos de Zhou (2012). A diferenciação existente entre o auditor independente e o governamental em relação ao ambiente e à abordagem do trabalho, aos objetivos ou material analisado para realização do julgamento, remetem à necessidade de realização de investigações específicas sobre essa categoria profissional que permitam identificar os aspectos que norteiam esse processo no contexto próprio dessa área, possibilitando, pelo melhor entendimento das limitações e vieses inerentes, a atuação na melhoria de sua qualidade.

Nesse sentido, este estudo busca investigar o processo de julgamento do auditor governamental a partir da seguinte questão central: a percepção de credibilidade dos auditores governamentais nas informações prestadas para análise é afetada pela preferência linguística textual ou numérica? O objetivo é examinara influência da preferência linguística no processamento das informações submetidas para análise, verificando sua interferência na percepção de confiabilidade e, como consequência, na definição do julgamento do auditor governamental brasileiro.

Baseada na noção em que a confiabilidade nas informações está relacionada à consistência em relação a preferências (O'KEEFE, 1990; KUNDA, 1987;

MAYSELESS; KRUGLANSKI, 1987[1] *apud* HAMILTON, 2005), um experimento de campo foi conduzido, com a manipulação da variável linguística entre textual e numérica, para examinar a seguinte hipótese: a apresentação de informações para análise de forma consistente com as preferências linguísticas dos auditores governamentais são examinadas menos criticamente e menos extensivamente, exigindo menos informações para chegar à conclusão e, portanto, aceitas como mais confiáveis que quando inconsistentes com suas preferências.

O experimento coletou dados através de utilização de questionários de internet, o que possibilitou a cobertura de todas as cinco regiões brasileiras e teve a participação de 114 auditores governamentais nos dois grupos de tratamento. O resultado do experimento, analisado sob uma ótica quantitativa, através da Anova e do Qui-Quadrado, buscou identificar a existência de diferentes influências e interações de variáveis motivacionais e demográficas sobre os resultados.

O estudo e entendimento do impacto da preferência linguistica na confiabilidade do auditor governamental brasileiro é importante, do ponto de vista teórico, contribuindo para o crescimento da literatura relacionada ao tema, e prático, através da indicação de vieses de julgamento que podem contribuir com a ação preventiva de gestores e reguladores na busca pela redução do impacto desse fator no produto final do trabalho desse grupo.

O restante deste trabalho está dividido da seguinte forma: a seção 2 trata das teorias relacionadas ao estudo e à construção das hipóteses; os detalhes quanto ao método experimental constam da seção 3; os resultados são apresentados e analisados na seção 4; e as considerações finais são expressas na seção 5.

2 TEORIA E HIPÓTESES

Auditores governamentais são expostos em seu trabalho à avaliação e emissão de julgamento quanto a uma diversidade de questões que envolvem desde o exame de contas públicas e aspectos de conformidade legal a posicionamentos subjetivos quanto à adequação de situações à probidade e moralidade pública. Essa atuação para a além dos formatos numéricos e normatizados e dos testes estruturados, quase exclusivos nas auditorias de demonstrações contábeis, expõe esse profissional aos mais diferentes formatos linguísticos na apresentação das informações para exame e emissão de julgamento.

Analisando as pesquisas em finanças e contabilidade, Hales, Kuang e Venkataraman (2011) citam que as pessoas não aplicam uma atenção uniforme sobre toda a informação sob exame, apresentando o que chamam de *motivated sensitivity*, que significa, segundo os autores, que as pessoas tornam-se sensíveis e tendem a um maior escrutínio de informações inconsistentes com suas preferências e aceitam, com um exame menos minuncioso, informações consistentes com tais preferências, sugerindo que essa relação é decorrente de diferenças na linguagem. Essa conclusão é similar a estudos anteriores como os de Ditto e Lopez (1992) e Dawson, Gilovich e Regan (2002). Os trabalhos de Lord, Ross e Liepper (1979[2] *apud* HAMILTON, 2005), Kunda (1987) e Mayseless e Kruglanski (1987[3] *apud* HAMILTON, 2005) apontam para que se aceite mais facilmente conclusões e resultados que sejam consistentes com as metas ou preferencias, realizando uma análise menos acurada, consistente e extensiva de argumentos e evidências.

Os estudos de Kunda (1990) demonstram que o processamento de informações é influenciado por diversos fatores, situações ou motivações. A preferência, elencada como uma motivação pelo autor, similarmente ao desejo, pode afetar a tarefa de raciocínio e constituir-se, segundo Ditto e Lopez (1992), em um dos itens que podem afetar profundamente a forma como uma dada quantidade de informações é processada, atuando na utilização seletiva dos recursos cognitivos, um dos principais objetos da auditoria comportamental.

A influência dos aspectos relacionados à linguagem tem sido estudada nos mais diversos segmentos, e na área contábilé dirigida para investigar questões relacionadas ao impacto da linguística na comunicação financeira, como as relacionadas à reação de investidores a informações não numéricas nos estudos de Hutton, Miller e Skinner (2003), ou a resposta do mercado de ações a estímulos pessimistas ou otimistas ou a divulgações de gestão, como presentes nos estudos de Feldman *et al.* (2010).

No âmbito do julgamento profissional de auditores a questão da linguagem é introduzida por Comunale, Sexton e Sincich (2005), que examinam a influência do estilo de linguagem e da credibilidade do cliente no julgamento do auditor.A percepção de credibilidade é composta, segundo

1 MAYSELESS, O.; KRIGLANSKI, A.W. Accuracy of estimates in the social comparison of abilities. *Journal of Experimental Social Psychology*, v. 23, p. 217-219, 1987.

2 LORD, C. G.; ROSS, L.; LIEPPER, M. R. Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidency. *Journal of Personality and Social Psychology*, v. 37, p. 3098–2109, 1979.

3 MAYSELESS, O.; KRIGLANSKI, A.W. Accuracy of estimates in the social comparison of abilities. *Journal of Experimental Social Psychology*, v. 23, p. 217-219, 1987.

Berlo, Lemert e Mertz (1969[4] *apud* COMUNALE; SEXTON; SINCICH, 2005) por três fatores: competência, fidedignidade e objetividade, sendo incluído pelos autores, como resultado de seu estudo, um quarto fator: o estilo de 'entrega' da linguagem, que em conjunto afetam a sensibilidade do auditor.

Resumidamente, a literatura investigada aponta para a existência de influência da preferência linguística no processamento da informação com reflexo na percepção de confiabilidade das pessoas, e, como consequência, em seu julgamento em relação à aceitação de conclusões ou resultados, levando a um exame mais ou menos extensivo das informações disponíveis. A partir desse conceito, será possível afirmar que a variável preferência linguística - textual ou numérica - ligada a outro contexto que não conclusões e resultados promoveriam no raciocínio a mesma *motivated sensitivity*? A preferência por estilos linguísticos na apresentação das informações para análise poderia induzir a um exame em graus diferenciados de profundidade e influenciar na percepção de credibilidade das pessoas?

Essas questões aplicadas ao contexto do ambiente público de auditoria, em que há entrega de informações para análise em formatos linguísticos variados e realização de alguns julgamentos sem seguir procedimentos normatizados de testes e evidências, poderia indicar fatores motivadores de vieses de julgamento nesse segmento.

Para identificar a alteração na percepção de credibilidade dos auditores governamentais, as informações disponíveis para análise e sua interação com as preferências linguísticas, textual ou numérica, foi testada a seguinte hipótese de pesquisa através de um estudo experimental de campo: as informações apresentadas para exame de auditores governamentais de forma consistente com suas preferências quanto à variável linguística - textual ou numérica - são analisadas menos profundamente e a elas é atribuído um grau de credibilidade maior, evidenciando o fenômeno da *motivated sensitivity*.

O entendimento desse aspecto do processo de julgamento dos auditores governamentais poderá ser aplicado para a melhoria de sua qualidade, do planejamento de treinamentos e do apoio à decisão de gestores dessas equipes. Cabe ressaltar, entretanto, que apesar da busca da casualidade do viés estudada pelo experimento, segundo Sellitz (1959), esta não é determinística e, assim, outros aspectos relacionados ao julgamento e à percepção de confiabilidade não podem ser negligenciados na obtenção da compreensão dos vieses existentes no processamento de informações para realização de julgamentos.

3 PROCEDIMENTOS METODOLÓGICOS

O estudo foi delineado como exploratório devido à própria natureza do tema da pesquisa, sobre a qual há pouco acumulado e sistematizado (VERGARA, 2003) no cenário brasileiro, buscando-se, assim, maior conhecimento sobre o problema investigado.

Os procedimentos metodológicos serão apresentados a seguir de forma a possibilitar um melhor entendimento do experimento de campo realizado, na seguinte ordem:

a) visão global e desenho do experimento;
b)detalhes sobre participantes, materiais e procedimentos utilizados; e
c) uma breve análise dos dados coletados.

3.1 VISÃO GLOBAL E DESENHO

A partir da definição do problema, foi conduzido um experimento de campo observando as orientações de Smith (2003) e da Resolução nº 196 do Conselho Nacional de Saúde do Brasil (BRASIL, 1996)[5]. Mediante a exposição de dois grupos de auditores governamentais aleatoriamente constituídos a partir da manipulação experimental da variável linguística, por meio de uma simulação de análise que solicita o posicionamento do auditor quanto a adequação de uma justificativa apresentada por um órgão governamental X para aquisição de produtos, buscou-se coletar a variação na percepção de confiabilidade do grupo. Adicionalmente, informações demográficas quanto a sexo, experiência profissional e estado da federação foram incluídas no estudo para possibilitar a melhor categorização dos respondentes.

O desenho do experimento foi definido seguindo a notação padrão[6] (CAMPBELL; STANLEY, 1963; MATTAR, 1997), como segue: grupo 1: R X1 O1 e grupo 2: R X2 O2. As variáveis de análise foram determinadas pela questão de pesquisa e hipóteses e definidas como:

a) y: variável dependente (VD) referente à confiabilidade do auditor na informação apresentada, medida em escala métrica discreta, com variação de 1 a 10, sendo 1 designado como pouco confiável e 10 como muito confiável;

4 BERLO, D. K.; LEMERT, J. B.; MERTZ, R. J.P. Dimensions for evalutating the acceptability of messagem sources. *Public Opinion Quarterly*, v. 33, n. 4, p. 563-376, 1969.

5 Resolução CNS 196/96– Diretrizes e normas regulamentadoras de pesquisa envolvendo seres humanos no Brasil. Regulamenta a pesquisa envolvendo seres humanos de forma direta ou indireta, incluindo o manejo de informações, e aplica-se (item III) às pesquisas realizadas em qualquer área do conhecimento.

6 Pela notação padrão: *X* representa um tratamento cujo efeito sobre a variável dependentes deverá ser determinado; O representa a saída ou observação e *R* indica que os grupos foram distribuídos de forma aleatória para receber os diferentes tratamentos.

b) x_n: variáveis independentes (VI), sendo x1 (variável linguístico-textual) e x2 (variável linguístico-numérica) o tratamento aplicado em dois níveis; duas variáveis categóricas nominais: x3, referente à preferência da variável linguística, e x4, que mediu a experiência do respondente, e, por último, uma variável categórica de validação da confiabilidade x5, na qual era declarado a tipificação do parecer a ser emitido (adequação ou inadequação).

A validade interna foi analisada com base nos estudos de Cook e Campbell (1979[7] *apud* SMITH, 2003) e demonstram a presença de risco não controlado de desmoralização, ou seja, de diferentes tratamentos agirem em diferentes níveis de motivação nos participantes com consequente impacto nos resultados. Quanto à validade externa, tratada a partir das preocupações apontadas por Christensen (1994[8] *apud* SMITH, 2003) indicou o controle sobre a validade da população e validade ecológica, não sendo possível controlar, pela natureza do experimento conduzido, a validade temporal.

O teste de entendimento do instrumento foi realizado com quatro auditores de instituições diferentes previamente à aplicação efetiva do instrumento e possibilitou identificar e corrigir questões relacionadas à clareza das questões. Os auditores envolvidos no pré-teste não participaram da pesquisa.

3.2 PARTICIPANTES, MATERIAIS E PROCEDIMENTOS

Seguindo as posições de Shanteau (1989) e Smith (2003), que entendem que os estudos que envolvam habilidades profissionais devem ser realizados com participantes qualificados e sem a utilização de incentivos, o experimento foi conduzido exclusivamente com auditores governamentais em exercício da profissão nos órgãos de controle governamental do Brasil (controladorias gerais dos estados, tribunais de contas estaduais, Tribunal de Contas da União e Controladoria Geral da União), com participação voluntária, não tendo sido realizada nenhuma outra restrição à participação.

A coleta de dados foi realizada a partir da disponibilização do instrumento de pesquisa em coletor eletrônico de dados com operação em ambiente de internet, sendo essa configuração definida para proporcionar a captura de dados de participantes institucionalmente e geograficamente dispersos e seus efeitos similares aos experimentos presenciais(KRANTZ; DALAL, 2000). O instrumento de pesquisa constou com seis questões elaboradas de forma a possibilitar a coleta de todas as variáveis e aquelas cujo conteúdo poderia influenciar na resposta de outras mantidas em páginas separadas, buscando garantir a validade interna do experimento no que tange ao efeito teste.

O formato de coleta favoreceu a distribuição aleatória dos participantes pelas condições experimentais, assegurando que as diferenças apontadas pelo experimento são causadas pelo tratamento e não por diferenças pré-existentes entre os grupos. A formação dos grupos deu-se a partir do envio ou disponibilização de dois *links* da pesquisa que eram escolhidos livremente pelos participantes, sendo o *software* utilizado na coleta programado para não aceitar respostas do mesmo IP nos dois *links* e,adicionalmente, este alerta feito aos participantes.

Ao respondente foi inicialmente apresentado um termo de consentimento livre e esclarecido para a realização da pesquisa, atendendo a Resolução nº 196 do Conselho Nacional de Saúde (BRASIL, 1996)[9] e, ainda, na mesma tela, coletadas informações demográficas quanto a sexo, experiência profissional e estado da federação (sigla), de modo a determinar a abrangência do estudo.

Em sequência, e em tela exclusiva, acessada após as respostas anteriores, o tratamento do experimento foi introduzido com a apresentação para cada grupo de uma situação fictícia em que um dado órgão público solicita ao órgão de controle parecer quanto à adequação de justificativa apresentada para viabilizar uma compra de material. A descrição do ambiente de trabalho com informações restritas foi obtida a partir da utilização, na situação fictícia, de nomes e características de materiais médicos pouco comuns, buscando a não interferência de referências passadas na confiabilidade reportada e da informação da indisponibilidade, devido a um problema técnico de acesso a internet naquele momento. A situação foi expressa para os dois grupos com dimensões, conteúdos e quantitativo informacional idêntico, com variação da forma linguística de apresentação, ou seja, ao grupo 1 as informações foram prestadas exclusivamente em forma textual e ao grupo 2 em forma numérica. Os textos utilizados no experimento constam no Apêndice A.

Após o texto, foi solicitado ao participante que expressasse a confiabilidade nas informações analisadas (VD), através de marcação em escala métrica, e declarasse

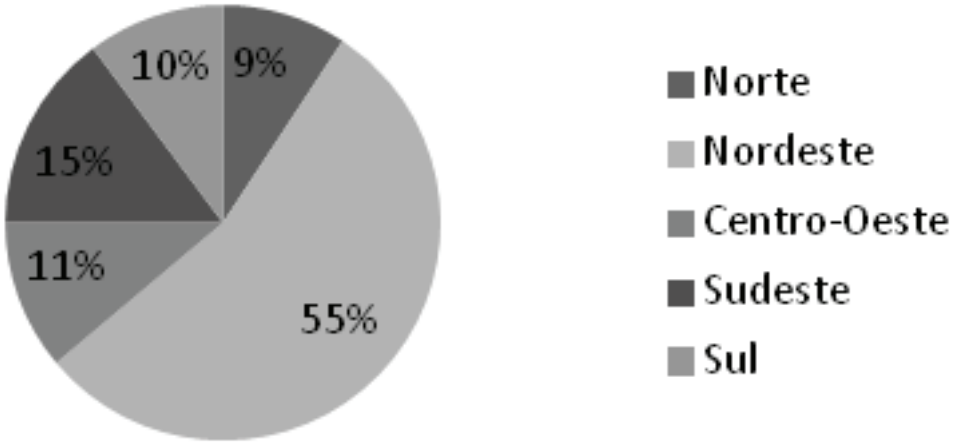
7 COOK, T. D.; CAMPBELL, D.T. *Quasi-Experimentation*: design and analysis issues for field settings. Boston: Houghton Mifflin, 1979.
8 CHRISTENSEN, L. B. *Experimental Methodology*, 6th Ed. Allyn, 1994.
9 Termo de consentimento livre esclarecido visa permitir ao sujeito uma decisão autônoma e voluntária de participara ou não da pesquisa.

sobre que tipo de parecer emitiria ao órgão solicitante (adequação ou inadequação). Na tela seguinte, foi realizada a declaração de sua preferência entre as variáveis linguístico-textuais ou linguístico-numéricas para fins de análise.

3.3 ANÁLISE DOS DADOS

O experimento coletou informações de 114 respondentes, distribuídos nos dois grupos, sendo 108 dados válidos, configurando 5,1% de dados perdidos em decorrência de respostas incompletas. Os grupos tiveram distribuição de respondentes de 62 (57,4%) para o grupo 2 (numérica) e 46 (42,6%) para o grupo 1 (textual), com prevalência do sexo masculino (63%). A distribuição dos respondentes, nas cinco regiões geográficas do país, consta na Figura 1.

Figura 1 - Distribuição dos respondentes por região geográfica



Fonte: Elaboração própria a partir de dados da pesquisa.

Para verificar a existência de diferenças entre os grupos a partir dos efeitos do tratamento (VI) sobre a confiabilidade (VD), foi utilizada a análise estatística *oneway* Anova, com nível de significância de 5%. Os pressupostos para a sua aplicação foram testados e apresentaram os seguintes resultados: ausência de normalidade (Shapiro-Wilk, t=0,919; gl=108; p=0,000; e Kolmogorov-Smirnov, t=0,130; gl=108; *p*=0,000), a hipótese nula para homogeneidade das variâncias foi rejeitada (Levene, $f_{1,106}$=1,784; *p*=0,185), indicando variâncias homogêneas e apresentam observações independentes.

Apesar da violação do pressuposto de normalidade, a utilização da Anova para o estudo foi mantida, considerando os estudos de Donaldson (1968), Lunney (1970) e Reis e Ribeiro Júnior (2007), que apontam a robustez do teste F em relação aos desvios de normalidade. Para corroborar os resultados encontrados com utilização da Anova com a violação da normalidade, foi realizado adicionalmente o teste não paramétrico de Wilcoxon-Mann-Whitney para amostras independentes, equivalente ao teste t para duas amostras, pois foram utilizados neste estudo dois grupos de teste.

A busca de correlação entre as variáveis independentes foi testada com Qui-Quadrado, com significância de 5%, na qual os pressupostos de quantidade mínima de frequência foram alcançados.

Para identificar a relação entre confiabilidade e preferência foi realizada a verificação dos escores médios de confiabilidade através da segmentação dos grupos originais em dois outros, que foram constituídos por aqueles que receberam a informação coerente com suas preferências e pelos que as receberam em desacordo com a preferência.

4 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Os dados do experimento, tratados através de *oneway* Anova, com nível de significância de 5%, indicam a existência de diferença significante entre os grupos de tratamento relacionado à variável de tratamento (VD) – confiabilidade e a variável independente (VI) preferência ($f_{1,106}$= 5,920; *p*=0,017)–, similarmente ao teste adicional nãoparamétrico de Wilcoxon-Mann-Whitney (w= 998,5; *p*=0,02148), não tendo apresentado significância estatística a nenhuma outra relação entre as demais variáveis independentes (sexo e experiência).

A magnitude do efeito experimental foi mensurada utilizando-se, segundo classificação de Cooper (1997[10] *apud* CONBOY, 2003), o tipo *d* de Cohen (1988), equivalente às diferenças padronizadas, e seguindo-se a mesma escala de classificação (COHEN, 1988), com *d*=0,47, que indica que o tratamento explica em torno de 33% da variância observada.

Considerando a indicação estatística quanto à explicação da variância da confiabilidade em decorrência da preferência pela variável linguística, os grupos iniciais (1 e 2) foram segmentados por tipo de informação recebida e por preferência, ou seja, os que receberam informações para análise de acordo com sua preferência linguística declarada (A) e os que receberam informações para análise em formato linguístico não consistente com sua preferência declarada (B). A partir dessa classificação, foram calculados os valores médios atribuídos à VD confiabilidade pelos grupos segmentados, conforme a Tabela 1.

Tabela 1 - Confiabilidade média

Informação recebida para análise	Preferência	Grupo segmentação	Quantidade	Confiabilidade média
Textual	Textual	A	21	4,05
Numérica	Textual	B	41	3,38
Textual	Numérica	A	25	4,40
Numérica	Numérica	B	21	4,90

Fonte: Dados da pesquisa.

A confiabilidade média dos que recebem informações consistentes com sua preferência é de 4,48, superior àquela apurada quando as informações analisadas não são consistentes com sua preferência, que é de 3,78.

Esse resultado sugere a não rejeição da hipótese de pesquisa e indica que a percepção de confiabilidade do auditor governamental é influenciada pela preferência da variável linguística, textual ou numérica, na apresentação das informações para análise, sendo a variável entendida como mais confiável àquela de sua preferência, apontando para a ocorrência no processamento das informações de *motivated sensitivity* e na mesma linha com os estudos de Ditto e Lopez (1992) e Dawson, Gilovich e Regan (2002).

A análise dos resultados permite evidenciar que os auditores governamentais que têm preferência por linguagem numérica, quando expostos à variável textual, apresentam um nível crítico em relação à confiabilidade superior aos dos que preferem a variável textual e recebem a numérica, tendo sido esse fato o responsável direto pela redução na confiabilidade média percebida quando da exposição para análise a uma variável linguística contrária a sua preferência. O resultado é coerente com a crença geral de que números apresentam uma maior objetividade e, consequentemente, maior confiabilidade.

A busca de correlação entre as variáveis independentes foi testada com Qui-Quadrado, com significância de 5%, e demonstrou existir correlação entre as variáveis tipo de tratamento e preferência (Paerson, $\chi 2$= 4,528; gl=1; p= 0,33). O teste de contagem realizado pela tabulação cruzada indicou que 54% dos respondentes que receberam informações com a variável textual preferem a variável numérica, e 66% dos que receberam informações numéricas declaram a preferência pela textual, o que pode sugerir duas interpretações: a ocorrência do efeito desmoralização do experimento, no qual o tratamento empregado induz ao comportamento do respondente, ou o chamado "o jardim do vizinho é sempre mais verde", em que as pessoas consideram que a alternativa não disponível seria a melhor.

Os dados indicam que 57,4% dos respondentes preferem a variável textual em detrimento da numérica, que contou com 42,6% da preferência, indo de encontro à visão do senso comum deque auditores em geral são mais afetos a números.

Quanto ao tipo de parecer emitido para a avaliação solicitada, 29,6% emitiram parecer de adequação da justificativa para a compra e 70,4% emitiram parecer de não adequação da justificativa. Esses dados, coerentes com o baixo nível de confiabilidade expressado pelos respondentes, refletem o perfil conservador dos auditores governamentais, motivado, especialmente, pelas condições de incerteza geradas pelo experimento (desconhecimento do tipo de material e impossibilidade de consultas adicionais).

10 COOPER, H. Some finer points in meta-analysis. In: HUNT, M. (Ed.). *How science takes stock: the story of meta-analysis. New York: Russell Sage Foundation.* 1997. p. 169-181.

5 CONSIDERAÇÕES FINAIS

Este estudo experimental foi conduzido para investigar a ocorrência de *motivated sensitivity* no processamento das informações pelo auditor governamental brasileiro a partir dainterferência da preferência pela variável linguística – textual ou numérica – na apresentação das informações para análise, na percepção de confiabilidade.

A hipótese inicial da pesquisa foi formulada a partir dos indicativos do referencial teórico e aponta para a ocorrência de *motivated sensitivity* a partir da preferência pela variável linguística. O resultado do experimento é consistente com a predição inicial. A percepção de confiabilidade nas informações analisadas foi maior no grupo em que era consistente com as preferências linguísticas declaradas e menor no grupo no qual eram incompatíveis.

O entendimento do impacto da preferência linguística na confiabilidade do auditor governamental brasileiro é importante do ponto de vista teórico e prático. Teoricamente, o estudo permite iniciar, expandir e contribuir com o crescimento da literatura sobre o entendimento dos efeitos de diversos fatores motivacionais no julgamento dos auditores governamentais e possibilitar a condução inicial para outros estudos com esse segmento profissional.

Em relação ao aspecto prático, o conhecimento de vieses de julgamento, como os empiricamente indicados por este estudo, podem possibilitar a gestores e reguladores a adoção de medidas que busquem minimizar seu efeito sob o julgamento do auditor governamental, como treinamentos específicos, padronização de apresentação de informações para análise, dupla análise para casos relevantes, entre outros, contribuindo com a qualidade do julgamento final e com o seu consequente impacto nas questões sociais envolvidas.

Devido à natureza do estudo realizado, um experimento controlado, sua relevância para o mundo real está sempre sujeita a questionamentos, não devendo os resultados ser generalizados para outros segmentos da população. Essa é uma limitação que deve ser ressaltada. Limitações próprias desse tipo de estudo podem ser supridas por futuras pesquisas, como as decorrentes da demonstração dos efeitos do tratamento e diferenças entre grupos de tratamento e as decorrentes da própria natureza comportamental da análise e da influência de fatores diversos no julgamento, como, por exemplo, as questões culturais, que devem ser consideradas na condução da interpretação dos resultados.

REFERÊNCIAS

BRASIL. Ministério da Saúde. Conselho Nacional de Saúde. *Resolução CNS 196, de 16 de outubro de 1996*. Aprova as diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos. Disponível em: <http://conselho.saude.gov.br/resolucoes/1996/Reso196.doc>. Acesso em: 01 jul. 2011.

CAMPBELL, D. T.; STALNLEY, J. C. *Experimental and quasi experimental desing for research on teaching*. Chicago: Rand McNally, 1963.

COHEN, J. *Statistical power analysis for the behavioral sciences*. Hillsdale:Erlbaum, 1988.

COMUNALE, C. L.; SEXTON, T. R.; SINCICH, T. L. Linguistic delivery style, client credibility and auditor judgment. *Advances in Accounting Behavioral Research*, Stamford, n. 8, p. 59-86, 2005.

CONBOY, J. E. Algumas medidas típicas univariadas da magnitude do efeito. *Análise Psicológica*, Lisboa, v. 21, n. 2, p. 145-158, abr. 2003.

DAWSON, E.; GILOVICH, T.; REGAN, D. T. Motivated reasoning and performance on the wason selection Task. *Personality and Social Psychology Bulletin*, Beverly Hills, v. 28, n. 10, p. 1379-1387, Oct. 2002.

DITTO, P. H.; LOPEZ, D. S. Motivadet spektism: use of differential decision criteria for preferred and non preferred conclusion. *Journal of Personality and Social Psycology*, Washington, v. 63, n. 4, p. 568-584, 1992.

DONALDSON, T. S. Robustness of the F-test to errors of both kinds and the correlation between the numerator and denominator of the F-rati. *Journal of the American Statistical Association*, Alexandria, v. 63, n. 322, p. 660-676, 1968.

FELDMAN, R. *et al.* Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, Boston, v. 15, n. 4, p. 915-953, Dec. 2010.

FERRAZ, C.; FINAN, F. Exposing corrupt politicians: the effect of Brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics*, Cambridge, USA, v. 123, n. 2, p. 703-745, 2008.

_____. Electoral accountability and corruption: evidence from the audits of local government. *American Economic Review*, Nashville, v. 101, n. 4, p. 1274-1311, 2011.

HALES, J.; KUANG, X.; VENKATARAMAN, S. Who belives the hype? An experimental examination of how language affects investor judgment. *Journal of Accounting Research*, Chicago, v. 49, n. 1, p. 223-255, Mar. 2011.

HAMILTON, D. L. *Social cognition*. New York: Psychology Press, 2005.

HOGARTH, R. M. A perspective on cognitive research in accounting. *The Accounting Review*, Sarasota, v. 66, n. 2, p. 277-290, Apr. 1991.

HOGARTH, R. M.; EINHORN, H. J. Order effects in belief updating: the belief-adjustment mode. *Cognitive Psychology*, Amsterdam, v. 24, n. 1, p. 1-55, Jan. 1992.

HUTTON, A. P.; MILLER, G. S.; SKINNER, D. J. The role of supplementary statements with managements earnings forecast. *Journal of Accounting Research*,Chicago, v. 41, n. 5, p. 867-889, Dec. 2003.

KAHNEMAN, D. A perspective on judgment and choice: mapping bounded rationality. *American Psychologist*,Washington, v. 58, n. 9, p. 697-720, Sept. 2003.

KERLINGER, F. N.; LEE, H. B. *Foundations of behavioral research*. New York: Harcourt College, 2000.

KOCH, C.; WUSTEMAN, J. *A review of bias research in auditing*: opportunities for combining psychological and economic research. St. Petersburg, US: American Accounting Association: 2009.

KRANTZ, J. H.; DALAL, R.Validity of web-based psychological research. In: BIRNBAUM, M. H. (Ed.). *Psychological experiments on the internet*. New York: Academic Press, 2000. p. 35-60.

KUNDA, Z. The case for motivated reasoning. *Psychological Bulletin*,Washington, v. 108, n. 3, p. 480-498, Nov. 1990.

LUNNEY, G. H. Using analysis of variance with a cichotomous dependent variable: an empirical study. *Journal of Education Measurement*, New York, v. 7, n. 4, p. 263-269, 1970.

LIU, J.; LIN, B. Government auditing and corruption control: evidence from China's provincial panel data. *China Journal of Accounting Research*, Hong Kong, v. 5, n. 2, p. 163-186, Jun. 2012.

MATTAR, F. N. *Pesquisa em marketing*. São Paulo: Atlas, 1997. v. 1.

O'KEEFE, D. J. *Persuasion*: theory and research. Newbury Park: SAGE, 1990.

REIS, G. M.; RIBEIRO JÚNIOR, J. I. *Comparação de testes paramétricos e não paramétricos aplicados em delineamentos experimentais*. 2007. Disponível em : <http://www.saepro.ufv.br/Image/artigos/SA03.pdf>. Acesso em: 02 jul. 2011.

SCHELKER, M.; EICHENBERGER, R. Auditors and fiscal policy: empirical evidence on a little big institution. *Journal of Comparative Economics*, Amsterdam, v. 38, n. 4, p. 357-380, Dec. 2010.

SELLTIZ, C. *Research methods in social relations*. New York: Holt, 1959.

SHANTEAU, J. Cognitive heuristics and biases in behavioral auditing: review, comments and observations. *Accounting, Organizations and Society*, Amsterdam, v. 14, n. 1-2, p. 165-177, 1989.

SIMON, H. *Models of man social and rational*. New York: Wiley, 1957.

SMITH, M. *Research methods in accounting*. London: Sage, 2003.

TVERSKY, A.; KAHNEMAN, D. Judgment under uncertainty: heuristics and biases. *Science*, New Series, Washington, v. 185, n. 4157, p. 1122-1131, Sept. 1974.

WOLK, H. I. *Accounting theory*: conceptual issues in a political end economic environment. Los Angeles: Sage, 2008.

ZHOU, Y. Government audit materiality (part one): how and why is it different from corporate audit materiality – a theoretical matrix on three materiality differences and corresponding contextual reasons. *International Journal of Economics and Finance*, Toronto, v. 4, n. 1, p. 80-89, Jan. 2012.

APÊNDICE A – INSTRUMENTO DE PESQUISA

Texto inicial comum aos dois instrumentos de coleta.

Imagine a seguinte situação:

O processo em suas mãos trata do questionamento de um Órgão X quanto à adequação da justificativa para fundamentar a aquisição de equipamentos que foi elaborada pelo corpo técnico do citado órgão. O objetivo da consulta é a obtenção de uma análise prévia da justificativa, de forma a não ter problemas futuros com a compra que será realizada, tendo em vista tratar-se da aquisição de OxímetrosFingertip SB 220 para monitorização (*sic*) e spot cehks, cujo valor unitário é muito elevado. O seu parecer deve restringir-se à análise da justificativa para a quantidade requerida e escolha do modelo.

Hoje é o último dia para emissão deste parecer, falta apenas 1 hora para encerrar o turno de trabalho, a internet para consultas está sem operação devido a um problema técnico e somente agora você pode se deter na análise da questão. A sua avaliação de desempenho depende, dentre outros fatores, do cumprimento de prazos na entrega dos trabalhos.

a) Abordagem Numérica

Justificativa Técnica para a compra de 52 unidades.

Serviço de atendimento móvel de emergência em 01.09.2011: 20 unidades remotas de atendimento (URA).

Necessidade de OxímetroFingertip SB 220 (OFSB) para monitorização e spot cehks para leitura de saturação de oxigênio no sangue e frequência de pulso.

Cálculo Necessidade:

01 p/URA x 20 URA = 20 OFSB

Substituição de 01 OFSB p/dia para recarga x 20 URB = 20OFSB

Substituição por avaria por mês = 2%*20 OFSB*30 dias= 12 OFSB

Necessidade estimada para uso- 52 OFSB

Escolha do Modelo:

Economia de energia: desligamento em 8s sem utilização.

Peso: 50g

Medida de Intervalo de saturação: 0% a 99%

Variação de frequência máxima: +/- 3%

Medida de intervalo de frequência de pulso: 30-250 pulsos por minuto.

b) Abordagem Textual

O serviço de atendimento móvel de emergência no primeiro dia de setembro deste ano dispõe de vinte unidades remotas de atendimento que necessitam cada uma de um OxímetroFingertip SB 220 para monitorização (*sic*) e spot cheks. A utilização do equipamento para leitura de saturação de oxigênio no sangue e frequência de pulso precisa de substituição para recarga a cada dia de operação durante vinte horas. Os dados de acompanhamento da utilização mostram que a cada dia dois por cento deles são avariados necessitando de substituição. O modelo escolhido foi devido a sua economia de energia, pois tem como característica o desligamento automático após oito segundos sem operação, seu peso de cinquenta gramas e intervalo de saturação medido completamente e a frequência de variação de no máximo três por cento, medindo intervalos de frequência de pulso entre trinta e duzentos e cinquenta pulsos por minuto.

Considerando a justificativa acima e a impossibilidade de deixar de emitir o julgamento e consequente parecer naquele momento, responda as questões que seguem: (todas apresentadas em páginas separadas).

O quanto você considera a justificativa apresentada confiável (medida em escala Linkert)?

Que tipo de parecer você emitira a partir da análise da justificativa? Adequação ou inadequação?

Caso você precise fundamentar uma opinião e possa optar por uma abordagem numérica ou textual, qual seria sua escolha?