

A QUESTÃO DE REDAÇÃO NO CONCURSO VESTIBULAR À UNIVERSIDADE: VALIDADE E PODER DECISÓRIO

Fernando Lang da Silveira

Universidade Federal do Rio Grande do Sul – UFRGS

Carlos Eduardo da Cunha Pinent

Pontifícia Universidade Católica do Rio Grande do Sul – PUCRS, Faculdade de
Matemática, Universidade de Caxias do Sul – UCS,
Centro de Ciências Exatas e Tecnologia

Resumo

Apresentam-se análises de validade e do poder decisório da questão de Redação nos concursos vestibular da Universidade Federal do Rio Grande do Sul e da Pontifícia Universidade Católica do Rio Grande do Sul, envolvendo 6516 candidatos comuns aos dois concursos. O estudo mostrou que a fidedignidade dos escores na questão de Redação foi expressivamente inferior às fidedignidades dos escores nas demais provas dos concursos. O poder de decisão, isto é, o poder de decidir sobre o sucesso ou o fracasso na conquista de vagas da questão de Redação mostrou-se baixo; mais de noventa por cento dos que conquistaram uma vaga, teriam tido sucesso caso a questão de Redação não fosse considerada. O estudo aponta a necessidade das universidades desenvolverem projetos de pesquisa com o objetivo de aprimorar metodologias e critérios de atribuição de escores na questão de Redação, visando aumentar a sua validade.

Palavras chave: Concurso vestibular; questão de redação; validade; poder decisório.

Resumen

Este artículo presenta un análisis de la validez y del poder de decisión de los resultados de la Redacción conseguidos por 6.526 candidatos en el examen de ingreso en la Universidad Federal del Rio Grande del Sur y de la Pontifícia Universidad Católica del Rio Grande del Sur. El estudio demostró que el sistema de evaluación de la Redacción fue expresivamente menos fidedigno que el habido en los resultados de los demás exámenes. El poder de decisión, o sea, el poder de decidir sobre el suceso o el fracaso en la conquista de plazas en la cuestión de la Redacción, se mostró bajo: más de noventa por ciento de los que conquistaron una plaza habrían obtenido suceso aunque la Redacción no hubiera sido considerada. El estudio muestra la necesidad de que las universidades desenvuelvan proyectos de pesquisa con el objetivo de mejorar metodologías y criterios de evaluación de la Redacción para aumentar su validez.

Palabras-clave: Examen de ingreso; cuestión de la Redacción; valor; poder de decisión.

Abstract

This study presents an analysis of the validity and defining capacity of written test (essay) grades achieved by 6,516 common candidates to both Universidade Federal do Rio Grande do Sul and Pontifícia Universidade Católica do Rio Grande do Sul. When compared to objective test grades of the same group, reliability of essay grades is significantly lower than the reliability of the scores in the other tests of the exam. Similarly the power of decision, i.e., the effectiveness in defining candidates that should be approved, is also low when the essay is considered: over 90% of successful candidates would also enter university if there were no essay test at all. This investigation points out that university should develop research on evaluation methods and scoring criteria for the essay test to select university students, aiming at increasing its validity.

Keywords: Entrance examination; written test (composition); power of decision.

1. Introdução

Este trabalho constitui-se em um estudo empírico da validade e do poder decisório da questão de Redação em concursos de ingresso à universidade no Brasil, onde esse tipo de concurso recebe o nome de Concurso Vestibular (CV). Utilizamos os arquivos de dados dos participantes do CV de 1999 da Universidade Federal do Rio Grande do Sul (CV-UFRGS) e do CV de 1999 da Pontifícia Universidade Católica do Rio Grande do Sul (CV-PUCRS). Efetivamente participaram destes concursos 35.787 candidatos na UFRGS e 10.547 candidatos na PUCRS.

Ao cruzarmos as duas planilhas com os escores dos candidatos nos dois concursos, encontramos 6516 candidatos que participaram de ambos. Com este grupo foi possível realizar um estudo inédito, tendo por base os escores de um grupo expressivo de sujeitos que realizaram provas sobre os mesmos conteúdos duas vezes, num espaço de tempo de apenas duas semanas. Esses escores nos possibilitaram uma avaliação da validade e da fidedignidade da questão de Redação e de todas as provas que integraram os dois concursos. Adicionalmente, estudamos o poder decisório da questão de Redação no CV-UFRGS.

Os dois concursos vestibular constituíram-se de 9 provas: Língua Portuguesa, Língua Estrangeira, Literatura, História, Geografia, Matemática, Física, Biologia e Química. As 9 provas continham 30 questões de escolha múltipla e repostas únicas na UFRGS e 25 questões na PUCRS. Adicionalmente, a prova de Língua Portuguesa possuía uma questão de Redação.

As análises que apresentaremos a seguir foram possíveis graças à utilização do pacote estatístico "SPSS for Windows - Versão 10.05".

2. A equivalência dos escores obtidos pelos candidatos nos dois concursos vestibular

Nesta seção testaremos a hipótese da equivalência dos escores obtidos pelos candidatos nos dois concursos vestibular. Uma análise do programa de conteúdos dos dois concursos revelam a equivalência formal. Aliás, não poderia ser diferente na medida em que esses conteúdos são os do ensino médio. A equivalência de fato somente pode ser testada empiricamente se houver um grupo de candidatos que

participou dos dois concursos. Na seção anterior já acusamos a existência desse grupo (6516 candidatos).

Os arquivos com os dados do CV-UFRGS registram o escore total de cada sujeito em cada uma das nove provas (um escore entre zero e trinta), bem como o escore final da questão de Redação (um escore entre zero e vinte, resultante da média aritmética de duas avaliações independentes dessa questão). Os resultados da análise de consistência interna da cada prova com questões de escolha múltipla, efetuada com todos os participantes que a responderam (cerca de 35.000 respondentes por prova), estavam disponíveis¹. Faz parte desta análise, além dos coeficientes utilizados para avaliar cada item, o *coeficiente alfa de Cronbach* (Cronbach, 1967), que na versão para itens com apenas dois escores (zero ou um) também é denominado de *fórmula KR.20* (Kuder e Richardson, 1967). O *coeficiente alfa* constitui-se em uma estimativa, por consistência interna², do coeficiente de fidedignidade do escore total na prova.

Já os arquivos com os dados do CV-PUCRS registram apenas o escore total de cada sujeito em quatro pares de provas (Biologia e Química, Física e Matemática, História e Geografia, Língua Estrangeira e Literatura Brasileira), mais a prova de Língua Portuguesa e a questão de Redação. Esses escores variam entre zero e cinquenta para os pares de provas e entre zero e vinte e cinco para a prova de Língua Portuguesa. O escore da Redação, também obtido da média de duas avaliações independentes, é um escore entre zero e quatro. Utilizamos como estimativa do coeficiente de fidedignidade dos escores totais nas provas com questões de escolha múltipla a *fórmula KR.21*³ (Kuder e Richardson, 1967), calculada para a totalidade dos candidatos ao CV-PUCRS (10.547 sujeitos).

¹ Ao que nos conste, não tem sido aplicado qualquer tratamento estatístico à questão de Redação com o objetivo de avaliar a fidedignidade dos escores que dela resultam. Se as universidades investissem nisso, certamente teriam como retorno o aprimoramento da avaliação.

² As estimativas do coeficiente de fidedignidade baseadas em métodos de consistência interna necessitam uma única aplicação do instrumento e dependem das correlações entre os escores de suas partes (usualmente as partes do teste são os itens ou questões; podendo ser também conjuntos de itens).

³ A *fórmula KR.21* é uma subestimativa do *coeficiente alfa*. Para testes com cerca de 50 itens a diferença entre os coeficientes obtidos por essas duas fórmulas não excede 0,03 (Thorndike e Thorndike, 1997).

Conforme Vianna (1972; p. 166):

A fidedignidade de um teste é determinada para um grupo específico. Um mesmo teste aplicado a diferentes grupos possui, geralmente, várias fidedignidades. Conhecida a variabilidade de determinado grupo, pode-se estimar a fidedignidade para outro grupo, através da fórmula:

$$F_x = 1 - \frac{s^2(1-f_x)}{S^2} \quad (1)$$

F_x . estimativa da fidedignidade para outro grupo

s^2 . variância de certo grupo

f_x . fidedignidade de certo grupo

S^2 . variância do outro grupo

Com a equação acima, partindo dos coeficientes obtido para a totalidade dos candidatos em cada concurso, calculamos os coeficientes de fidedignidade de cada prova no grupo dos 6516 sujeitos comuns aos dois concursos. Computamos os escores totais em pares de provas do CV-UFRGS (Biologia e Química, Física e Matemática, História e Geografia, Língua Estrangeira e Literatura Brasileira), já que no CV-PUCRS (conforme relatado anteriormente) não existiam os escores individualizados por área de conteúdo. Para obter o coeficiente de fidedignidade do escore em cada par de provas no CV-UFRGS, valemo-nos da equação 2 apresentada por Nunnally (1978):

$$f_y = 1 - \frac{\sum_{i=1}^n [(1-f_i)s_i^2]}{s_y^2} \quad (2)$$

onde:

y – escore total ou $y = \sum_{i=1}^n x_i$

f_y – coeficiente de fidedignidade do escore obtido da soma de dois ou mais escores x_i .

f_i – coeficiente de fidedignidade do escore x_i .

s_i^2 – variância do escore x_i .

s_y^2 – variância do escore y .

Criamos também, para os dois concursos, uma medida global de desempenho nas nove provas com itens de escolha múltipla, calculando um escore total (somatório do número de acertos) sobre todas as provas. Esse total resultou em um escore entre zero e duzentos e setenta (9x30) para um sujeito no CV-UFRGS e entre zero e duzentos e vinte e cinco (9x25) no CV-PUCRS. O coeficiente de fidedignidade das duas medidas globais foi obtido também com a equação 2. A tabela 1 apresenta os coeficientes de fidedignidade, bem como o coeficiente de correlação entre os escores na UFRGS e na PUCRS (todos os coeficientes de correlação são estatisticamente significativos em nível inferior a 0,001). Observa-se que os coeficientes de fidedignidade das provas que versam sobre os mesmos conteúdos são semelhantes, aproximadamente iguais ao coeficiente de correlação entre os escores nessas provas.

Tabela 1
Coeficientes de fidedignidade e correlação

Provas	Coeficiente de fidedignidade PUCRS	Coeficiente de fidedignidade UFRGS	Coeficiente de correlação	Coeficiente de correlação desatenuado
Biologia e Química	0,84	0,79	0,80	0,98
Matemática e Física	0,84	0,84	0,78	0,93
História e Geografia	0,83	0,79	0,80	0,99
Língua Estrangeira e Literatura Brasileira	0,81	0,84	0,79	0,96
Língua Portuguesa	0,66	0,69	0,62	0,92
Número total de acertos nas 9 provas	0,95	0,96	0,92	0,96

A última coluna da tabela 1 apresenta os coeficientes de correlação *desatenuados* ou *verdadeiros*⁴. Como eles são próximos da

⁴ O coeficiente de correlação observado entre dois escores é atenuado pelos coeficientes de fidedignidade de cada um deles. Ele é "menor ou igual à raiz quadrada do produto dos coeficientes de fidedignidade" (Silveira, 1993; p. 76). É possível estimar qual seria o coeficiente de correlação se ambos os escores tivessem fidedignidade perfeita (igual 1). O coeficiente de correlação *desatenuado* ou *verdadeiro* ($[r_{xy}]_{\text{Verdadeiro}}$) é dado pela seguinte equação (Wherry, 1984):

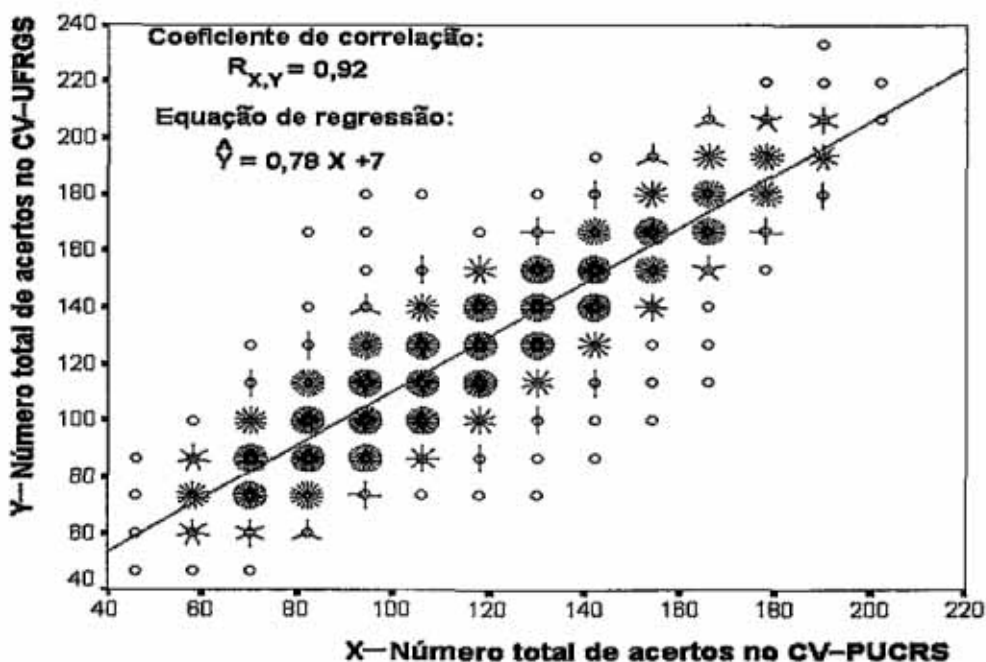
$$[r_{xy}]_{\text{Verdadeiro}} = r_{xy} / \sqrt{f_x f_y} \text{ onde } r_{xy} \text{ é o coeficiente de correlação observado, } f_x \text{ e } f_y \text{ são os coeficientes de fidedignidade.}$$

unidade, corrobora-se empiricamente que as provas do CV-UFRGS e do CV-PUCRS são medidas equivalentes dos mesmos conteúdos. Em outras palavras, medem a mesma coisa.

A figura 1 apresenta um diagrama de dispersão do número total de acertos no CV-UFRGS contra o número total de acertos no CV-PUCRS. Os "girassóis" apresentados neste diagrama devem assim ser interpretados: apenas o "centro do girassol" (uma pequena circunferência) representa até 10 pares de escores; um "girassol" com n "pétalas" representa no mínimo $(10.n + 1)$ pares de escores e no máximo $10.(n+1)$ pares de escores. Ou seja, um "girassol com 4 pétalas" representa entre 41 e 50 pares de escores.

Figura 1

Diagrama de dispersão do número total de acertos no CV-UFRGS contra o número total de acertos no CV-PUCRS

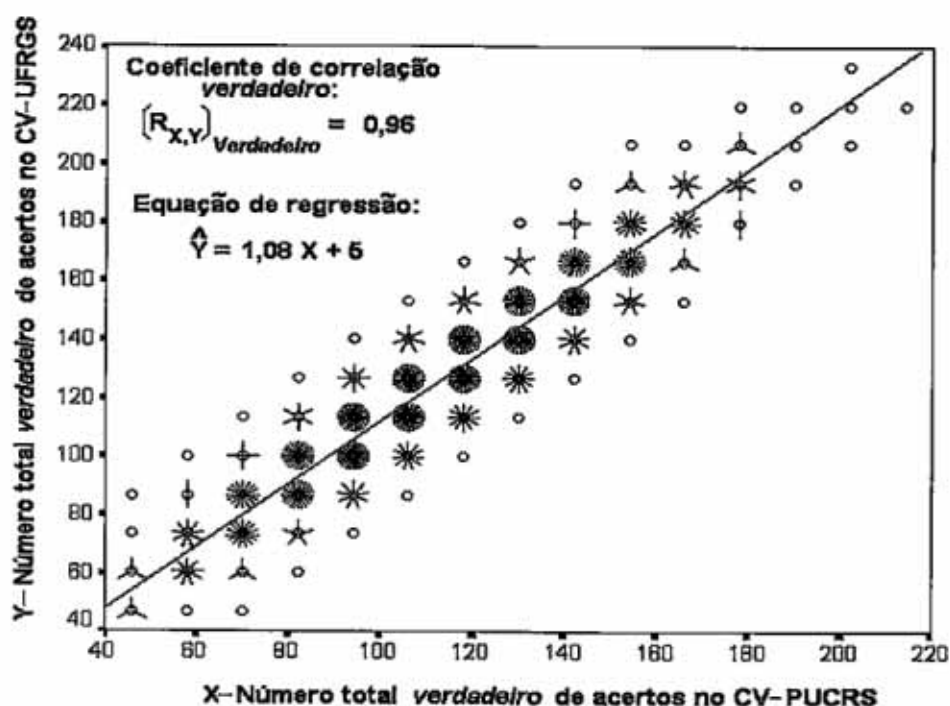


Na figura 1 visualiza-se a intensa correlação observada (0,92) entre os resultados nos dois concursos pela concentração dos escores em torno da reta de regressão. Por simulação, através do método de Monte Carlo (Sobol, 1983), é possível se construir um diagrama de

dispersão que represente a correlação *desatenuada* ou *verdadeira* entre os totais de acertos. A figura 2 apresenta esse diagrama.

Figura 2

Diagrama de dispersão dos números de acertos *verdadeiros*, obtidos pelo método de Monte Carlo



Na figura 2, como decorrência da forte correlação *desatenuada* ou *verdadeira*, os totais de acertos concentram-se mais próximos da reta de regressão do que na figura 1.

A tabela 2 apresenta os coeficientes de correlação entre os escores em todos os pares de provas dos dois concursos. As duas letras iniciais, tanto na primeira coluna quanto na primeira linha, identificam as provas (BQ – Biologia e Química; MF – Matemática e Física; HG – História e Geografia; LL – Língua Estrangeira e Literatura; LP – Língua Portuguesa). A terceira letra identifica a universidade (P – PUCRS; U – UFRGS). Todos os coeficientes são estatisticamente significativos em nível inferior a 0,001.

Tabela 2
Coeficientes de correlação entre os escores em todas as provas dos dois concursos

	BQ-P	MF-P	HG-P	LL-P	LP-P	BQ-U	MF-U	HG-U	LL-U	LP-U
BQ-P						0,80	0,75	0,71	0,65	0,59
MF-P	0,79					0,75	0,78	0,65	0,58	0,54
HG-P	0,75	0,69				0,71	0,67	0,80	0,69	0,63
LL-P	0,70	0,64	0,71			0,65	0,58	0,67	0,79	0,65
LP-P	0,53	0,48	0,58	0,62						
BQ-U										
MF-U						0,75				
HG-U						0,72	0,69			
LL-U						0,64	0,58	0,69		
LP-U						0,58	0,54	0,62	0,67	

Abaixo da diagonal da tabela 2 encontram-se os coeficientes de correlação entre os escores de um mesmo concurso (coeficientes intra-concurso); acima da diagonal estão os coeficientes de correlação entre os escores dos diferentes concursos (coeficientes entre-concursos). A média dos coeficientes de correlação intra-concurso na UFRGS é 0,65 e na PUCRS é 0,65; a média dos coeficientes de correlação entre-concursos é 0,68. Ou seja, os padrões de correlação intra e entre-concursos são quase os mesmos, corroborando novamente a hipótese da equivalência dos escores obtidos nos dois concursos⁵.

3. A fidedignidade da questão de Redação

Na seção anterior corroboramos a hipótese de equivalência dos escores nos dois concursos vestibulares. Em conseqüência, tomaremos como estimativa da fidedignidade da questão de Redação o coeficiente de correlação entre os escores de Redação obtidos pelos sujeitos nos dois concursos.

⁵ Segundo Ghiselli (1964), uma das condições para que duas medidas possam ser consideradas equivalentes é a de apresentar os mesmos coeficientes de correlação com outras variáveis.

No CV-UFRGS nem todas as Redações são avaliadas; apenas aquelas dos candidatos que fizeram mais de 108 acertos (40%) do total das 270 questões de escolha múltipla. Entre os 6516 candidatos comuns aos dois concursos, 4184 tiveram a Redação avaliada no CV-UFRGS. O coeficiente de correlação entre os escores de Redação nos dois concursos, calculado no grupo de 4184 candidatos, resultou em 0,36 (estatisticamente significativo em nível inferior a 0,001). Portanto, estimamos o coeficiente de fidedignidade da Redação em 0,36 neste grupo restrito. A tabela 3 apresenta esse coeficiente, os desvios padrão dos escores de Redação no CV-PUCRS no grupo de 4184 sujeitos e no grupo de 6516 sujeitos; na última coluna aparece o coeficiente de fidedignidade, estimado pela equação 1, para o grupo de 6516 sujeitos.

Tabela 3
Coeficientes de fidedignidade e desvios padrão dos escores de Redação para 4184 e 6516 candidatos do CV-PUCRS

Coeficiente de fidedignidade (n = 4184)	Desvio padrão (n = 4184)	Desvio padrão (n = 6516)	Coeficiente de fidedignidade (n = 6516)
0,36	0,88	0,92	0,41

Assim como o coeficiente de fidedignidade de uma variável modifica-se de um grupo para outro, o coeficiente de correlação entre duas variáveis também. A equação abaixo, denominada de *correção para a restrição em variabilidade* (Ghiselii, 1964; Guilford e Fruchter, 1973), permite que se estime o coeficiente de correlação em um grupo maior (grupo total) partindo-se do coeficiente de correlação observado em outro grupo menor (grupo restrito):

$$(R_{xy})_{est} = \frac{r_{xy} S_x}{S_x \sqrt{1 - r_{xy}^2 \left(1 - \left[\frac{\sum x}{S_x} \right]^2 \right)}} \quad (3)$$

onde:

$(R_{XY})_{est}$ – coeficiente de correlação estimado para o grupo total.

r_{XY} – coeficiente de correlação calculado no grupo restrito.

S_x – desvio padrão da variável x no grupo total.

s_x – desvio padrão da variável x no grupo restrito.

A equação 3 tem nos seus pressupostos a ausência de erro na variável x (coeficiente de fidedignidade de x igual a 1). Ora, no domínio das medidas educacionais, ausência de erro de medida ou mesmo um pequeno erro de medida, é mais a exceção do que a regra. Conforme Silveira (1996, p. 365):

A equação 3, não levando em consideração a existência de erro de medida na variável x (isto é, não levando em consideração que o coeficiente de fidedignidade de x é menor do que um), produz um coeficiente de correlação estimado menor para todo o grupo do que na verdade ele é; havendo erro de medida na variável x, a equação 3 produzirá subestimativas cada vez mais afastadas do valor real ao decrescer o desvio padrão no grupo restrito. Apresentamos a seguir uma equação inédita na literatura, que leva em conta o erro de medida na variável independente.

A equação apresentada por Silveira (1996) é a seguinte:

$$(R_{XY})_{est} = \frac{r_{XY} F_x S_x}{s_x \sqrt{f_x^2 - r_{XY}^2 \left(f_x - F_x \left[\frac{S_x}{s_x} \right]^2 \right)}} \quad (4)$$

onde:

F_x – coeficiente de fidedignidade da variável x no grupo total.

f_x – coeficiente de fidedignidade da variável x no grupo restrito.

Com auxílio da equação 4, partindo-se do grupo composto por 4184 sujeitos, estimamos todos os coeficientes de correlação dos escores na questão de Redação do CV-UFRGS com as demais provas dos dois concursos (inclusive com a Redação do CV-PUCRS) no grupo dos 6516 candidatos. A tabela 4 apresenta esses coeficientes, bem

como os coeficientes de correlação efetivamente observados da questão de Redação do CV-PUCRS com as demais provas entre os 6516 candidatos.

Tabela 4
Coeficientes de correlação dos escores de Redação com outras provas

Prova	Coeficiente de correlação da Redação no CV-UFRGS com outra prova	Coeficiente de correlação da Redação no CV-PUCRS com outra prova
Biol.- Quím. PUCRS	0,29	0,39
Biol.- Quím. UFRGS	0,28	0,38
Mat.- Fís. PUCRS	0,24	0,36
Mat.- Fís. UFRGS	0,20	0,34
His.- Geo. PUCRS	0,32	0,40
His. - Geo UFRGS	0,29	0,38
L. Est. - Lit. Br. PUCRS	0,47	0,47
L. Est. - Lit. Br. UFRGS	0,49	0,46
L. Port. - PUCRS	0,49	0,44
L. Port. UFRGS	0,55	0,46
Redação PUCRS	0,41	
Redação UFRGS		0,41

Se comparamos os valores da tabela 4 com os da tabela 2, constatamos que as correlações da Redação com as outras dez provas são significativamente menores do que as correlações entre essas dez provas. Isto poderia indicar que, de acordo com o usualmente admitido, a questão de Redação avalia aspectos independentes dos que são medidos em questões de escolha múltipla. Entretanto, essa conclusão é questionável: como os escores de Redação se correlacionam mais fracamente entre si (0,41) do que qualquer par de provas sobre o mesmo conteúdo nos dois concursos, não pode haver correlação intensa da Redação com qualquer outra prova mesmo que os conteúdos e habilidades avaliadas fossem idênticas. Ou seja, neste estudo há fortes indícios de que a fidedignidade dos escores de

Redação é pequena. Assim sendo, qualquer correlação da Redação com outra prova não poderá ser grande (vide a nota de rodapé 4).

4. O poder decisório da questão de Redação no CV-UFRGS

Nesta seção é estudado o poder decisório da questão de Redação – isto é, o poder que a questão de Redação efetivamente apresentou para decidir se um candidato conseguiu ou não uma vaga. Este estudo é realizado apenas com os candidatos ao CV-UFRGS pois o poder decisório da Redação cresce com as taxas de concorrência por curso (razão candidato/vaga). Como essas taxas são maiores no CV-UFRGS, sabemos que o efetivo poder decisório constatado neste estudo é maior do que no CV-PUCRS. Destaque-se que a análise aqui desenvolvida não tem por objetivo argumentação contra a Redação nos vestibulares, pois cremos que o conhecimento de nossa língua é *conditio sine qua non* de acadêmicos.

No CV-UFRGS a questão de Redação valeu 40% do escore total da prova de Língua Portuguesa; os demais 60% do escore total foram constituídos pelo escore da prova de escolha múltipla com 30 itens. A prova de Língua Portuguesa foi a única dentre as 9 provas que teve peso 3 independentemente do curso ao qual o candidato concorre; portanto, com peso 1,2 (40% de 3) para a Redação e 1,8 (60% de 3) para as questões objetivas. Para cada curso havia mais uma prova com peso 3, 2 provas com peso 2 e 5 provas com peso 1. Desta forma, a soma dos pesos foi 15; como o peso da Redação é 1,2, então ela contribuirá com, no máximo, 8% (1,2 de 15) do argumento de concorrência (escore composto final do candidato)⁶.

⁶ Na verdade o peso efetivo da questão de Redação no escore composto de um concorrente no CV-UFRGS, denominado *argumento de concorrência*, é mais complicado de ser calculado com precisão pois é obtido da seguinte forma: A) Cada um dos nove escores brutos (escore total em cada prova) é transformado linearmente em um escore padronizado com média 500 e desvio padrão 100. B) De acordo com o curso ao qual o candidato concorre é atribuído um peso (1, 2 ou 3) a cada um dos nove escores. C) O *argumento de concorrência* é a média harmônica ponderada dos nove escores padronizados. Um dos autores do presente trabalho (Silveira, 1997), demonstrou que a complicada média harmônica ponderada produz praticamente os mesmos resultados da média aritmética. Por isso, grosso modo, podemos estimar *a priori* um peso pequeno para a questão de Redação.

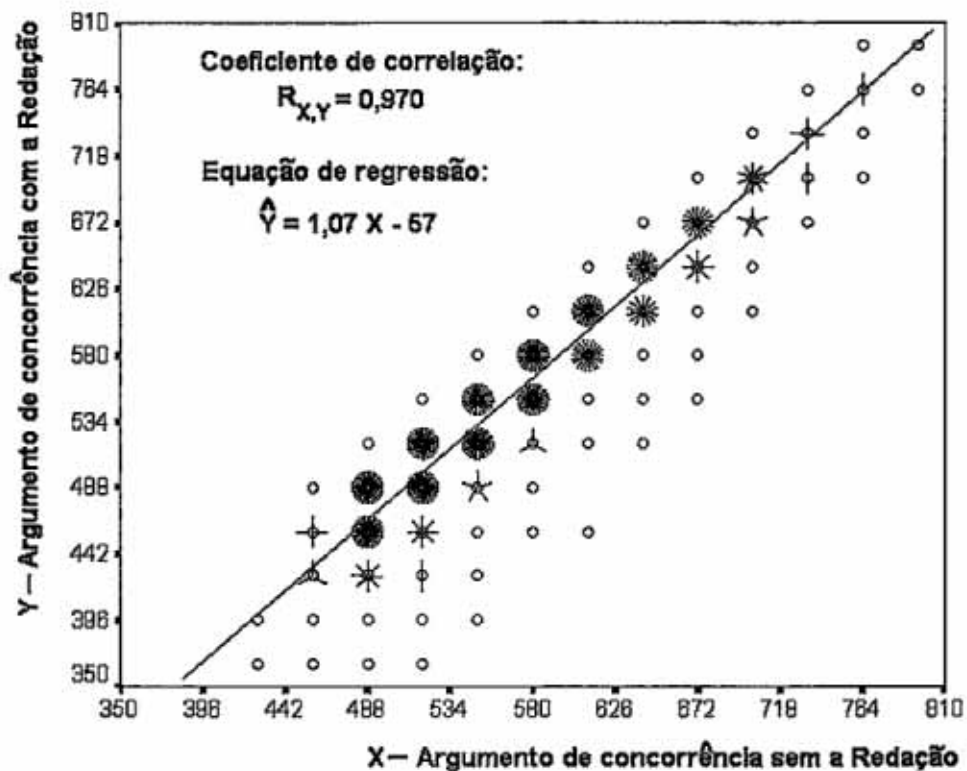
No CV-UFRGS nem todas as redações elaboradas foram avaliadas, pois, candidatos com número de acertos igual ou inferior a 40% (108 acertos) do número máximo possível de acertos (270 acertos) nas 9 provas, estavam automaticamente eliminados. Efetivamente foram avaliadas 17501 das 35787 redações produzidas. A cada redação foram atribuídos dois escores de forma independente por dois avaliadores, sendo a média de ambos o escore final. O processo de avaliação tomou cerca de uma semana, envolvendo quase uma centena de avaliadores.

Há os que entendem que a universidade deveria avaliar todas as redações produzidas. Argumentam que a eliminação de candidatos pelo critério dos 40% poderia prejudicar alguém que tivesse realizado uma excelente redação: obtendo um escore elevado em Redação, um candidato poderia terminar com um argumento de concorrência superior ao de outro candidato que efetivamente conquistou uma vaga.

A priori, estimamos que o poder do escore de Redação para decidir se um candidato conquistou ou não uma vaga foi pequeno, pois, conforme discutido anteriormente, esse escore contribuiu com aproximadamente 8% do argumento de concorrência. Entretanto, tal fato só se define mais precisamente com base nos resultados dos candidatos.

Partindo dos escores obtidos pelos candidatos ao CV-99, calculamos um novo argumento de concorrência. A única diferença entre este novo argumento e o oficialmente adotado pela UFRGS era que o escore de Língua Portuguesa não incluía a questão de Redação. Esse novo argumento de concorrência se correlacionou fortemente (0,97) com o argumento de concorrência oficial da UFRGS, conforme pode-se verificar no gráfico da figura 3. Os "girassóis" apresentados neste gráfico devem assim ser interpretados: apenas o "centro do girassol" representa até 30 pares de escores; um "girassol" com n "pétalas" representa no mínimo $(30.n + 1)$ pares de escores e no máximo $30.(n + 1)$ pares de escores. Ou seja, um "girassol com 4 pétalas" representa entre 121 e 150 pares de escores.

Figura 3
 Relação entre o argumento de concorrência com e sem a questão de Redação no CV-UFRGS



Baseado nesse novo argumento de concorrência produzimos um novo "listão" (uma nova lista de candidatos classificados). Depois, calculamos a taxa de coincidência dessa nova lista com a lista oficial do CV-99. Dos 3860 candidatos presentes na lista oficial, 3480 estavam presentes na lista construída sem o escore de Redação. Desta forma, a taxa de coincidência foi de 90,2%. Quando calculamos tais taxas por curso, 59,2% dos cursos tiveram taxas iguais ou superiores a 90,0% (alguns dos quais com taxa de 100%), sendo a taxa mínima de coincidência igual a 83% (ocorrida na Medicina da FFFCMPA).

Adicionalmente, investigamos qual era o posto (escore de ordem dos candidatos por curso), dentro da lista oficial do CV-99, mais alto do argumento de concorrência que resultou sem considerar a Redação (isto é, o posto do pior classificado sem a Redação entre

aqueles que integraram a lista oficial). Em todos os cursos o posto mais alto não atingiu o valor correspondente ao dobro do número de vagas. Exemplifiquemos com o curso de Medicina da UFRGS, um dos mais concorridos e onde efetivamente foram avaliadas 1971 redações: dos 140 candidatos que conquistaram vagas, 119 as conquistariam mesmo que não houvesse a Redação (portanto, a taxa de coincidência das duas listas acima referidas foi 85,0%). Dentre os restantes 21 candidatos que constaram da lista oficial de classificados, o posto mais elevado do argumento de concorrência sem a Redação (isto é, o posto do pior classificado) foi o 228º (portanto, inferior a 2 vezes o número de vagas). Desta forma, partindo-se da classificação sem considerar a Redação, bastaria serem avaliadas as redações dos 228 primeiros colocados para garantir que o último oficialmente classificado estivesse incluído. Portanto, para produzir a lista oficial de candidatos classificados, a rigor não haveria necessidade em qualquer um dos cursos de se avaliar mais redações do que 2 vezes o número de vagas.

Assim sendo, fica comprovado, como admitíamos *a priori*, que a questão de Redação possuiu um pequeno poder decisório sobre fato de um candidato conseguir ou não uma vaga na UFRGS.

5. Conclusão

Neste trabalho estudamos a validade e o poder decisório da questão de Redação em concursos de ingresso à universidade. Mostramos que os escores produzidos na avaliação da questão de Redação possui fidedignidade inferior a qualquer um dos escores obtidos nas demais provas dos concursos vestibulares da UFRGS e da PUCRS. Como a validade de um escore tem como condição necessária (mas não suficiente) a fidedignidade, o nosso trabalho aponta no sentido de uma baixa validade para a questão de Redação. Esse fato não é novo; a literatura especializada em medidas educacionais, por exemplo Vianna (1973) ou Ebel (1977), faz referência à baixa fidedignidade dos escores obtidos em questões de resposta livre.

Estudamos também o efetivo poder que a questão de Redação teve em decidir se um candidato conseguiu ou não uma vaga na universidade. Mostramos que esse poder foi muito pequeno pois, se a questão de Redação não existisse, praticamente os mesmos candidatos ingressariam na universidade. Acreditamos que o baixo poder decisório observado é extremamente desejável, em face dos problemas de

validade constatados na questão de Redação, pois, se o escore de Redação tem baixa fidedignidade, seu poder decisório deve ser pequeno.

Finalmente, os autores deste trabalho de maneira alguma advogam a eliminação da questão de Redação em concursos de ingresso à universidade. É nossa convicção que um acadêmico deva saber redigir e possuir bons conhecimentos da língua nacional, e isto pode ser estimulado mantendo-se a obrigatoriedade da Redação e da prova de Língua Portuguesa nos concursos de ingresso. Creemos que a universidade deveria providenciar na realização de estudos que visassem o aprimoramento do processo de avaliação, especialmente no da questão de Redação.

Referências Bibliográficas

CRONBACH, Lee J. Coefficient alpha and the internal structure of tests. In: MEHRENS, William A. e Ebel, Robert L. *Principles of educational and psychological measurement*. Chicago: Rand McNally, 1967.

EBEL, Robert L. *Fundamentos de la medicion educacional*. Buenos Aires: Guadalupe, 1977.

GHISELLI, Edwin E. *Theory of psychological measurement*. Nova Delhi: Tata McGraw-Hill, 1964.

GUILFORD, Joy P. e FRUCHTER, Benjamin. *Fundamental statistics in psychology and education*. New York: McGraw-Hill, 1973.

KUDER, Frederic G., RICHARDSON, Marion W. The theory of test reliability. In: MEHRENS, William A., EBEL, Robert L. *Principles of educational and psychological measurement*. Chicago: Rand McNally, 1967.

SILVEIRA, Fernando L. Comparação entre três argumentos de concorrência para o concurso vestibular da UFRGS. *Estudos em avaliação educacional*, São Paulo, 16: 43-57, 1997.

SILVEIRA, Fernando L. Correlação entre avaliações por testes de múltipla escolha e por provas analítico-expositivas: crítica e proposta metodológica. *Revista Brasileira de Ensino de Física*, São Paulo, 18(4): 362-371, 1996.

SILVEIRA, Fernando L. Validação de testes. In: MOREIRA, Marco A., SILVEIRA, Fernando L. *Instrumentos de pesquisa em ensino e aprendizagem*. Porto Alegre: EDIPUCRS, 1993.

SOBOL, Igor. *O método de Monte Carlo*. Moscou: MIR, 1983.

THORNDIKE, Robert L., THORNDIKE, Richard M. Reliability. In: KEEVES, John P. (org.) *Educational research, methodology, and measurement*. Cambridge: Pergamon, 1997.

VIANNA, Heraldo M. *Testes em educação*. São Paulo: IBRASA, 1972.

WHERRY, Robert J. *Contributions to correlational analysis*. Orlando: Academic Press, 1984.