

Estudo Comparativo dos Métodos Keller e Tradicional em Termos de Conhecimento Adquirido e Índice de Desistências

P. H. DIONISIO e M. A. MOREIRA

Instituto de Física, Universidade Federal do Rio Grande do Sul, Porto Alegre RS*

Recebido em 15 de Julho de 1974

The results of a Keller vs *lecture* comparative study, in an introductory College physics course, are presented. In the selection of the samples, a matching procedure has been used and the comparison has been made in terms of achievement and number of withdrawals. The results of the achievement tests were favorable to the "Keller group", whereas the number of withdrawals was not significantly different. It was also observed that score differences decreased during the course.

Apresentam-se resultados de uma comparação entre o Sistema Keller e o sistema convencional expositivo em um curso de Física Geral. Na amostragem, foi usado um procedimento de "matching" e a comparação feita em termos de conhecimento adquirido e índice de desistências. Os resultados obtidos, em testes de conhecimento, favoreceram o "grupo Keller", enquanto que, em relação ao índice de desistências, não houve diferença significativa. Observou-se ainda que a diferença entre os escores diminuiu ao longo do curso.

1. Introdução

Pela sua universalidade, o sistema convencional de ensino expositivo serve, via de regra, como sistema de referência para a avaliação de novos métodos de ensino. A eficiência e as vantagens de um novo método de ensino são frequentemente demonstradas (às vezes erroneamente) através de comparações com o método convencional.

O mesmo vem acontecendo com o Sistema Keller, ou "Método Keller", como é comumente conhecido esse sistema de ensino. As características básicas do Sistema Keller¹ são fundamentalmente diferentes mas a comparação é inevitável. Por mais que se argumente que os dois sistemas não são comparáveis, é muito difícil ao administrador, ao professor e ao estudante, acostumados com o sistema tradicional, aceitarem o Sistema Keller em termos absolutos. É, portanto, natural que, no processo de implantação desse Sistema, antes de se passar a uma análise intrínseca do mesmo, tomando o próprio aluno como sistema de

*Av. Luiz Englert, s/n, 90000 - Porto Alegre RS.

referência, se façam comparações do tipo "Método Keller x Método Convencional". Várias dessas comparações tem sido feitas^{2,3}; porém, considerando que o Sistema Keller representa uma mudança bastante radical no processo educacional e que os resultados obtidos com uma população não são necessariamente válidos para outra, existe ainda lugar para novas comparações dessa natureza que venham a acrescentar mais dados aos já existentes.

Um estudo desse tipo foi realizado na disciplina Física II (Termodinâmica, Eletricidade e Magnetismo) do Departamento de Física da UFRGS e os resultados serão aqui relatados. O objetivo do estudo foi comparar um grupo de alunos que fez o curso pelo sistema Keller com um grupo que o fez de forma convencional, em vários aspectos, dois dos quais são objeto deste trabalho: conhecimento adquirido e índice de desistências.

2. Descrição da Experiência

No 2.º semestre de 1973, matricularam-se, na disciplina Física II, 251 alunos distribuídos em 8 turmas. Trabalharam, na disciplina, 4 professores, cada um dos quais responsável por uma "turma Keller" e uma "Não Keller", na mesma faixa horária. Havia, portanto, 4 "turmas Keller" e 4 "Não Keller", totalizando, respectivamente, 122 e 129 alunos. Ante a impossibilidade de escolher as amostras aleatoriamente, utilizou-se um processo de "matching". Fez-se um pareamento entre os alunos dos dois grupos nas seguintes variáveis: resultado de um pré-teste, repetência ou não, ter emprego ou não e número de disciplinas cursadas simultaneamente, além das variáveis faixa horária e professor. Assim, supomos que um determinado aluno de uma "turma Keller" tivesse obtido um escore x no pré-teste, não fosse repetente, trabalhasse, além de estudar, e cursasse um número y de disciplinas. Procurou-se então na turma convencional, situada na mesma faixa horária (mas em dias diferentes) e sob a responsabilidade do mesmo professor, um aluno que não fosse repetente, trabalhasse, cursasse aproximadamente o mesmo número de disciplinas e tivesse obtido aproximadamente o mesmo escore no pré-teste. Com esse procedimento, formaram-se 63 pares, ou seja, foram obtidas duas amostras de 63 alunos equivalentes em relação as variáveis mencionadas.

A fim de comparar os grupos, em termos de conhecimento adquirido, os alunos foram submetidos a três testes comuns durante o semestre,

um de Termodinâmica, um de Eletricidade e um de Eletromagnetismo, de acordo com a seguinte técnica: os alunos, do grupo de controle, faziam o teste em conjunto após terem estudado o conteúdo correspondente, enquanto os alunos, do grupo experimental, o faziam individualmente após concluírem a última unidade relativa ao conteúdo do teste. Todos os testes foram do tipo objetivo (as provas normais de ambos os grupos eram discursivas) e os alunos foram cientificados de que os resultados não seriam computados para fins de avaliação na disciplina. Registre-se que, no grupo experimental, os alunos fizeram o mesmo teste em oportunidade diferentes, podendo, portanto, ter ocorrido comunicação entre os alunos. Não havia, no entanto, razão para isso pois os alunos sabiam que o resultado do teste não seria computado. Ademais, os testes eram realizados sem consulta, não eram devolvidos aos alunos e eram suficientemente longos para que não pudessem ser memorizados.

Quanto ao índice de desistências, foram considerados como desistentes aqueles alunos que iniciaram o curso, ou seja, os que chegaram a participar das atividades didáticas em ambos os grupos mas abandonaram o curso nas primeiras semanas.

O grupo experimental teve o curso dividido em 17 unidades enquanto o grupo de controle teve o mesmo conteúdo avaliado em 6 verificações, aproximadamente quinzenais.

3. Resultados

As Tabelas 1, 2 e 3 apresentam os resultados obtidos pelos grupos nos três testes comuns, mantendo-se o pareamento inicial. Os coeficientes de fidedignidade estimados, a partir da fórmula de Spearman Brown⁴, para esses testes foram: Termodinâmica = .72, Eletricidade = .46 e Eletromagnetismo = .50. Embora os correspondentes coeficientes de correlação par-ímpar⁴, respectivamente, .56, .30 e .34, fossem todos significativos ao nível .01, esses coeficientes de fidedignidade (principalmente os dois últimos) são relativamente baixos. Os testes, portanto, encarados como instrumentos de medida apresentam uma certa margem de erro.

A fim de ser mantido o pareamento inicial, sempre que um aluno de um dos grupos deixou de apresentar-se para o teste, o score do seu par no outro grupo não foi considerado. A Tabela 4 é similar as anteriores.,

	N(N.º de alunos)	M(Média)	T	Z	Diferença significativa ao nível .05
Grupo Experimental	50	10,8	127	4,29	SIM
Grupo de Controle	50	8,0			

Tabela 1 – Resultados do teste de Termodinâmica (18 questões). Té a variável característica do teste de significância usado e, Z, a da distribuição normal. Para $N > 25$, a distribuição T tende a distribuição normal Z .

	N	M	T	Z	Diferença significativa ao nível .05
Grupo Experimental	43	9,8	264,5	2,51	SIM
Grupo de Controle	43	8,4			

Tabela 2 – Resultados do teste de Eletricidade (16 questões).

	N	M	T	Diferença significativa ao nível .05
Grupo Experimental	22	9,7	74,5	NÃO
Grupo de Controle	22	9,5		

Tabela 3 – Resultados do teste de Eletromagnetismo (16 questões)

	N	M	U	Z	Diferença significativa ao nível .05
Grupo Experimental	37	9,8	463	1,36	NÃO
Grupo de Controle	31	9,1			

Tabela 4 – Resultados do teste de Eletromagnetismo, não mantendo o pareamento inicial. U é a variável característica do teste de significância utilizado. Para $N > 20$, a distribuição C tende à distribuição normal Z .

porém apresenta os dados referentes ao teste de Eletromagnetismo, sem manter o pareamento inicial.

A Tabela 5 é uma tabela de contingência 2 x 2 apresentando as desistências ocorridas nos grupos inicialmente pareados. O valor de χ^2 , (Ref. 5), obtido para essa tabela, é de 3.84, não significativo ao nível .05.

	Não Desistentes	Desistentes	Total
Grupo Experimental	54	9	63
Grupo de Controle	48	15	63
Total	102	24	126

Tabela 5 – Desistências ocorridas nos grupos pareados para fins de comparação.

Foi escolhido o Matched-Pairs Signed-Ranks *Test*⁵ de Wilcoxon como teste de significância para a diferença de escores e .05 como nível de significância. Esse é um teste estatístico não paramétrico apropriado para o caso que a amostragem é feita através de “matching” e que dá maior peso a um par que apresente grande diferença do que a um que apresente pequena diferença. Observou-se que, para os testes de Termodinâmica e de Eletricidade, a diferença foi estatisticamente significativa. Para o teste de Eletromagnetismo não houve diferença significativa ao nível escolhido. A primeira vista, poder-se-ia pensar que, por exemplo, a maior mortalidade ocorrida no grupo de controle tivesse causado uma melhoria nos escores desse grupo, fazendo com que diminuísse a diferença em relação ao grupo experimental. Tal fato parece não ter ocorrido, pois o teste de significância χ^2 , aplicado a Tabela 5, não permite rejeitar a hipótese de não diferença (hipótese nula) entre as desistências ocorridas em ambos os grupos. Além disso, na Tabela 4, confirma-se, com maior número de alunos, a impossibilidade de ser rejeitada a hipótese nula para a diferença de escores. Nesse caso foi usado o Mann-Whitney *U Test*⁵ como teste de significância, o qual é também não paramétrico, apropriado para o caso em que o “matching” não é considerado. A regressão estatística não parece, também, ser uma explicação satisfatória porquanto o conteúdo do pré-teste e dos testes posteriores foi sempre diferente.

Tabelas similares as Tabelas 1 e 2 foram também organizadas sem manter o pareamento inicial. Não houve, no entanto, diferença em relação as Tabelas 1 e 2.

Quanto ao índice de desistências, foi também organizada uma tabela, similar a Tabela 5, para as desistências ocorridas em toda a população. Também nesse caso não foi possível rejeitar a hipótese de não diferença ao nível .05.

4. Conclusão

Respeitadas as limitações da experiência, poder-se-ia dizer que os resultados obtidos fornecem evidências de que o sistema Keller dá melhores resultados em termos de conhecimento adquirido pelo aluno do que o sistema convencional, expositivo. Observou-se, no entanto, que essa diferença tendeu a diminuir ao longo do curso, o que sob certo ponto de vista é paradoxal, pois seria de se esperar que uma programação, pelo simples fato de ser individualizada, fornecesse melhores resultados. Uma vez rejeitada a hipótese de que a diminuição da diferença tivesse sido causada por um efeito de mortalidade ou por regressão estatística, poder-se-ia mencionar, entre outras, as seguintes razões para explicar esse efeito:

1. O segundo e o terceiro testes tiveram coeficientes de fidedignidade bastante baixos e, portanto, as medidas carecem de precisão e o efeito observado pode não ser verdadeiro.
2. A pressão de fim de semestre, decorrente de uma solicitação simultânea por parte de várias disciplinas, poderia ter tornado o desempenho dos alunos nos últimos testes inferior e conduzido a uma diferença não significativa entre os grupos.
3. A programação do grupo experimental, apesar de ser individualizada, não atentou especificamente as diferenças individuais, pois as unidades eram as mesmas para todos. Dessa forma, enquanto o conteúdo do curso era relativamente simples e até certo ponto já conhecido, o grupo experimental teria obtido melhores resultados em decorrência de um estudo melhor orientado e de maior motivação. Porém, a medida que o conteúdo foi se tornando mais complexo e a dificuldade aumentou, esses fatores não mais teriam sido suficientes para dar vantagem ao grupo experimental.

Precisamente esta última hipótese levou a uma pergunta que atualmente está sendo investigada em continuação a presente experiência: consi-

derando que em nenhum dos casos (programação individualizada e em grupo) foram levadas em conta, especificamente, as diferenças individuais, seria de se esperar que os grupos apresentassem diferença em termos de retenção do conhecimento adquirido?

Para tentar responder a essa questão, estão sendo aplicados, a esses mesmos grupos, testes de retenção versando sobre todo o conteúdo do curso.

Referências

1. F. S. Keller, *Good-Bye Teacher ...*, Journal of Applied Behavior Analysis, **1**, 79 (1968).
2. M. A. Philippas and R. W. Sommerfeldt, Keller vs. Lecture Method in General Physics, American Journal of Physics, **40**, 1300 (1972).
3. S. M. Austin and K. E. Gilbert, *Student Performance in a Keller-Plan Course in Introductory Electricity and Magnetism*, American Journal of Physics, **41**, 12 (1973).
4. F. N. Kerlinger, *Foundations of Behavioral Research*, Holt Rinehart and Winston Inc., New York, 1964.
5. S. Siegel, *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill Book Co., New York, 1956.