

INFLUÊNCIA DE CARACTERÍSTICAS DO ALUNO NA AVALIAÇÃO DO SEU DESEMPENHO *

CARMEN LÚCIA DE MELO BARROSO **

GUIOMAR NAMO DE MELLO **

ANA LÚCIA GOULART DE FARIA **

Entre as características do aluno que podem influir na avaliação do seu desempenho, interessa-nos particularmente o sexo. Isto porque investigações recentes a respeito de diferenças sexuais na escola (Rosemberg, 1975), e da participação e desempenho de moças e rapazes nos exames vestibulares (Barroso e Mello, 1975), constataram que:

- a) moças encaminham-se em sua maioria para carreiras tradicionalmente femininas;
- b) de um modo geral os rapazes tendem a obter maior sucesso que as moças no vestibular, independentemente da carreira escolhida; o desempenho das moças é inferior ao dos rapazes nas disciplinas de ciências exatas (física, química e biologia) e matemática, sendo igual ou superior nas disciplinas de maior conteúdo verbal, como português e inglês;
- c) no ensino médio (1º e 2º graus), as meninas tendem a obter notas mais altas que os meninos, são reprovados com menor frequência e apresentam menores atrasos de escolaridade.

DIFERENÇAS DE DESEMPENHO

A superioridade masculina em matemática e ciências exatas, quando se utilizam situações padronizadas de exame, parece ser fato bastante generalizado em vários países nos quais esses estudos têm sido realizados. No entanto, há indicações de que essas diferenças, grandes na adolescência, são pequenas ou inexistentes na infância. Maccoby e Jacklin (1974) realizaram uma revisão exaustiva de estudos realizados nos EEUU e assinalaram que a maioria dos estudos com amostras representativas não mostram diferenças entre os

* Pesquisa realizada com o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

** Do Departamento de Pesquisas Educacionais da Fundação Carlos Chagas

sexos antes dos 9 anos, em testes de conservação numérica e em testes de conceitos e operações numéricas.

Nos estudos de Comber e Keeves (1973) abrangendo 19 países e 3 níveis de idade (10, 14 e 18 anos), os meninos sempre obtiveram resultados mais altos que as meninas nos testes de ciências (química, física e biologia). As diferenças a favor dos meninos aumentaram com a idade. Husén (1967) estudando o desempenho de estudantes em testes de matemática em 12 países encontrou diferenças favoráveis aos meninos em todas as populações estudadas. Neste estudo, também, as diferenças aumentaram com a idade. Flanagan (1969) num estudo extensivo de caracterização da população escolar norte-americana obteve, para alunos de escolas médias, dados favoráveis aos meninos em testes de raciocínio matemático, mecânico e espacial e favoráveis às meninas em testes de capacidade verbal. Maccoby (1966) realizou uma revisão dos estudos que investigaram diferenças entre os sexos quanto a habilidades intelectuais. Os estudos que apontaram diferenças favoráveis ao homem referem-se, principalmente, a raciocínio matemático, raciocínio espacial e capacidade de solução de problemas. As diferenças a favor da mulher foram encontradas principalmente em estudos sobre habilidade verbal. Tyler (1969) revendo diversos estudos sobre diferenças de capacidade entre os sexos encontrou superioridade masculina em relações espaciais, solução de problemas e aptidão mecânica e superioridade feminina em fluência verbal, destreza manual e memória.

No Brasil, os Testes de Desenvolvimento Educacional (primeira bateria de testes de rendimento escolar criados e padronizados no país) aplicados a alunos de 2º ciclo da Guanabara revelaram superioridade masculina em vocabulário, compreensão de leitura, estudos sociais, matemática e ciências físicas e naturais. As meninas saíram-se melhor em uso da linguagem (Bessa, 1971).

A grande maioria dos estudos que constataram diferenças de desempenho intelectual entre os sexos utilizaram medidas objetivas e padronizadas. Parece haver certa inconsistência entre os resultados desses estudos e as medidas de rendimento escolar de meninos e meninas quando esse rendimento é expresso por avaliações ou notas dos professores. As vestibulandas estudadas por Barroso e Mello (1975), apesar de terem fracassado mais, relativamente aos seus colegas do sexo masculino, tiveram médias mais altas nas disciplinas da escola média. Os dados de Gouveia e Havighurst (1969) e de Rosemberg (1975) apontam no mesmo sentido, indicando menor número de reprovações e menores atrasos de escolaridade para as meninas. Carter (1952), estudando a validade das notas atribuídas pelos professores, conclui que, de um modo geral, controlada a variável de inteligência do aluno, as notas tendem a favorecer as meninas.

A ORIGEM DAS DIFERENÇAS DE DESEMPENHO

A explicação das diferenças de desempenho entre os sexos tem merecido a atenção de muitos estudiosos que não aceitam essas diferenças como sendo determinadas por fatores genéticos ou características biológicas inerentes. De um modo geral as explicações desses investigadores tendem a apontar como os principais responsáveis por essas diferenças os fatores ambientais presentes na socialização diferencial a que são submetidos meninos e meninas. Homens e mulheres desenvolvem suas potencialidades no sentido de corresponder às expectativas que a cultura possui quanto às características mais desejáveis do masculino e do feminino.

A prevalência de estereótipos e sua profunda incorporação nos autoconceitos de homens e mulheres foram documentadas por Broverman e colaboradores (1972). As mulheres são percebidas como relativamente menos competentes, menos independentes, menos objetivas e menos lógicas que os homens. No Brasil, constatações semelhantes foram feitas por Van Kolck (1969), com um grupo de alunos de Pedagogia e de Geologia, por Mello et al. (1976), com estudantes de pós-graduação de diversas áreas, e por Silva et al. (1976) com pais e mães de crianças de 4 a 9 anos em relação ao comportamento de seus filhos. Os traços estereotipicamente masculinos são percebidos como mais desejáveis que os femininos. As pesquisas de Broverman mostram que psiquiatras e psicólogos clínicos tendem a atribuir à mulher normal traços que não atribuem ao adulto normal: submissão, dependência etc.

Os estudos de Broverman incluem grupos variados quanto a sexo, idade, nível educacional etc. Apesar dessa variação, esses grupos apresentaram grande consenso a respeito das características típicas de homens e mulheres. Esse consenso parece ser o produto

de uma percepção estereotipada e também de uma constatação de diferenças realmente existentes entre os sexos. Engendra-se em parte um círculo vicioso no qual a existência do estereótipo contribui para o desenvolvimento de diferentes características no sexo masculino e no feminino, o que, por sua vez, reforça a manutenção do estereótipo.

Uma das formas mais diretas de manifestação do estereótipo é na avaliação do trabalho produzido por homens e mulheres. Goldberg (1968) fez um experimento no qual verificou que as mulheres recebem avaliação mais baixa ainda quando apresentam trabalho de igual qualidade e ainda que a avaliação seja feita por mulheres. Neste estudo, estudantes universitários deveriam avaliar alguns artigos. Cada estudante recebeu um livreto com artigos de lingüística, direito, história da arte, dietética, educação e planejamento urbano. Os livretos eram idênticos, apenas mudados os nomes dos autores, de modo que cada artigo era atribuído a um autor masculino em alguns livretos e a uma autora em outros. De acordo com sua hipótese, Goldberg verificou que até os artigos de áreas tipicamente femininas como dietética e educação receberam classificação mais baixa quando atribuídos a uma mulher. Outro experimento de Pheterson, Kiesler e Goldberg (1971), usando quadros que teriam sido pintados por artistas premiados numa exposição ou artistas desconhecidos, mostrou novamente a existência do preconceito contra as mulheres, exceto nos casos em que não havia ambigüidade em relação à competência do indivíduo julgado.

Torna-se de grande importância, portanto, a análise de como atuam os diferentes agentes socializadores — entre eles a escola — no sentido de orientar e reforçar diferentes habilidades nos meninos e nas meninas. É plausível supor que tais agentes disponham de formas mais ou menos sutis pelas quais transmitem as expectativas culturais quanto ao tipo de desempenho intelectual mais adequado para cada sexo e, concomitantemente, disponham de formas para manipular sanções e recompensas sempre que tais expectativas sejam ou não satisfeitas.

Os dados de alguns estudos sobre a influência do sexo do aluno na interação em sala de aula sugerem que as meninas recebem menor número de reforços positivos que os meninos e que, de um modo geral, os professores matêm maior número de contatos com os meninos (Meyer e Thompson, 1956; Brophy e Good, 1970; Biber et al., 1972). Feshbach (1969), estudando preferências de futuros professores quanto às características de personalidade de alunos de escola primária, concluiu que os traços mais desejados eram rigidez, conformismo, ordem, dependência, passividade e submissão e que havia uma ordem consistente de preferência quanto ao sexo do aluno: em primeiro lugar a menina com as características acima mencionadas, em segundo lugar o menino com as mesmas

características. Nessa ordem, o último lugar coube às meninas que apresentassem características de independência e auto-afirmação.

Gurwitz e Kenneth (1975) realizaram um experimento no qual estudantes universitários de ambos os sexos assistiram a um vídeo-tape com uma criança de 3 anos. Para metade da amostra foi dito que a criança era um menino e para a outra metade, uma menina. Após o filme, os sujeitos avaliaram a criança em algumas medidas de personalidade e capacidade. Os autores observaram que o sexo da criança interfere no julgamento dos sujeitos pois avaliaram a criança do sexo oposto mais positivamente que a do seu própria sexo. Um experimento de Dion (1974) sugere que as mulheres aplicam penalidades menos severas aos meninos, mas somente quando são fisicamente atraentes.

É possível que também haja diferença entre os sexos quanto ao efeito da avaliação do professor sobre a interpretação que lhe dá o aluno. Várias pesquisas observaram diferenças entre meninos e meninas quanto a «learned helplessness»: a atitude de desesperança que se segue a insucessos frequentes. Discutindo esses resultados, Dweck (1976) argumenta que eles refletem a similaridade da situação experimental à situação de ensino, na qual os meninos não se sentem tão ameaçados quanto as meninas por valorizarem menos a avaliação do professor.

Outras variáveis que podem influenciar a avaliação do professor são a raça e o nível sócio-econômico. Freijo e Jaeger (1976) estudaram o «efeito de halo» na avaliação dos professores e observaram que a precisão da avaliação é maior (o efeito de halo é menor) quando o professor percebe o aluno como sendo de nível sócio-econômico igual ao seu. O mesmo não ocorre com raça, o que se explica, segundo os autores, por nível sócio-econômico ser mais importante que raça na determinação de distância social. O estudo de Miller (1973) confirma que as expectativas do professor seriam influenciadas pela classe social do aluno, mas não pela sua raça.

Num estudo longitudinal de observação de uma classe de crianças de «gueto» durante o jardim da infância, o 1º e o 2º anos, Rist (1970) detectou formas pelas quais a escola ajuda a reforçar a estrutura de classes da sociedade. Os alunos de origem sócio-econômica mais alta, percebidos pelos professores como possuidores de características atitudinais e comportamentais desejáveis eram aqueles pelos quais os professores manifestavam maior interesse, com os quais passavam a maior parte do tempo, aos quais dirigiam menor controle, e que eram apresentados como modelos para o resto da classe etc.

A TEORIA DA ATRIBUIÇÃO

Uma contribuição que merece ser explorada para a análise das avaliações feitas por professores é a

oferecida pela teoria da atribuição elaborada por Weiner e seus colaboradores (1971), a partir dos pressupostos de Heider (1958). Esses autores postulam que o empenho das pessoas na realização de uma determinada tarefa depende em parte de suas expectativas, de sucesso ou fracasso nesta tarefa. Tais expectativas, por sua vez, são determinadas pela percepção que as pessoas adquirem das causas dos eventos que se seguem ao seu próprio comportamento e ao comportamento de outras pessoas que as cercam.

Segundo o modelo desenvolvido por Weiner, o sucesso ou fracasso podem ser atribuídos a quatro fatores assim categorizados:

	Internos	Externos
Estáveis	Habilidade	Dificuldade da tarefa
Instáveis	Esforço	Sorte

Para Weiner, a dimensão estabilidade-instabilidade dos fatores causais aos quais o sujeito atribui o resultado de sua ação é a mais importante na determinação de sua expectativa de sucesso ou fracasso, e conseqüentemente na probabilidade de que ele venha a repetir essa mesma ação em outras situações semelhantes. Se o sucesso é atribuído à habilidade ou à facilidade da tarefa, a expectativa de sucesso no futuro é aumentada; do mesmo modo diminui a expectativa de sucesso — ou aumenta a de fracasso — se este é atribuído à falta de habilidade ou à dificuldade da tarefa. A ocorrência de um evento (sucesso ou fracasso) atribuído a um fator instável (esforço ou sorte) não altera a expectativa de sucesso ou fracasso porque o esforço pode ser intensificado, do mesmo modo que a sorte pode melhorar.

A teoria da atribuição pode ser estendida para uma melhor compreensão dos fatores que interferem na avaliação que fazemos do desempenho dos outros. Neste sentido, alguns experimentos já foram realizados. Weiner e Kukla (1970) relatam uma série de investigações nas quais, replicando os procedimentos utilizados por Lanzetta e Hannah (1969), mediram os efeitos das causas de sucesso e fracasso sobre a avaliação do rendimento. Nesses estudos pediu-se aos sujeitos que se imaginassem como professores que tinham dado um exame a uma classe, e que, então, deveriam avaliar os alunos. Para essa avaliação, foram fornecidas três informações: resultado no exame, esforço expendido estudando para o exame e nível de habilidade. Os sujeitos distribuíam recompensas ou punições em cada uma de 20 condições simuladas (5 níveis de resultado x 2 níveis de esforço x 2 níveis de habilidade). Em três experimentos independentes, observou-se que o sucesso era recompensado mais do que o fracasso, grande esforço mais do que pequeno esforço, e, em geral, pequena habilidade mais do que grande habilidade. As maiores recompensas eram dadas aos estudantes com baixa habilidade que se esforçaram

muito e obtiveram sucesso, e as maiores punições aos estudantes com grande habilidade que não se esforçaram e apresentaram um desempenho pobre. Rest e colaboradores (1973) confirmaram esses dados e, além disso, demonstraram que a avaliação do rendimento também é influenciada por esforço como **traço**, isto é, o grau de empenho que o estudante exibe comumente (e não somente o empenho em relação a uma tarefa específica).

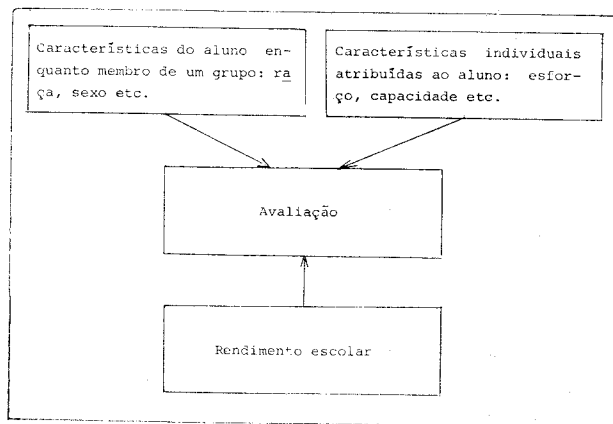
Parece claro que os diferentes pesos atribuídos aos fatores postulados por Weiner sofrem a influência de valores culturais. É possível que os valores predominantes na escola brasileira determinem pesos diferentes dos obtidos nas amostras estudadas (professores suíços e estudantes secundários e universitários americanos e alemães). É também possível supor que o sexo da pessoa que está sendo avaliada, se controlada essa variável, possa levar a resultados ou interações diferentes das obtidas nos estudos mencionados.

Dessa forma, foram realizados dois experimentos para investigar a influência de características do aluno na avaliação que o professor faz do seu desempenho.

O MODELO EXPLICATIVO

Em suma, o modelo de explicações no qual esse estudo se insere pode ser esquematizado no Quadro 1.

QUADRO I



É preciso ter em mente que se trata de um esquema em que relações extremamente complexas são reduzidas a sua forma mais simples a fim de possibilitar a análise de alguns dos seus aspectos. Neste sentido, não somente foram excluídas variáveis que podem desempenhar papel importante como também foram ignoradas algumas das influências recíprocas que as variáveis consideradas exercem entre si.

Saliente-se portanto que, embora a motivação inicial da pesquisa tenha sido uma preocupação com as diferenças de realização escolar entre os sexos, essa não é diretamente estudada. Partindo de uma

revisão ampla da literatura achamos razoável supor que essas diferenças estariam ligadas, entre outras coisas, a diferenças na interação que o professor desenvolve com alunos de cada um dos sexos. Também com base na literatura revista considerou-se que um dos elementos importantes dessa interação era exatamente a avaliação dos trabalhos dos alunos que o professor está a realizar constantemente. Além disso, o estudo da avaliação é importante em si mesmo pois aquelas diferenças de realização escolar, que constituíram a preocupação inicial de nossa pesquisa, talvez sejam explicadas em parte, não apenas pela influência que a avaliação tem sobre a realização posterior, mas também pelo próprio viés da avaliação, uma vez que as duas medidas — realização e avaliação — estão inextricavelmente confundidas na maioria dos estudos que medem realização utilizando indicadores escolares tais como taxa de aprovação e notas, que dependem diretamente da avaliação realizada pelo professor.

Assim, procuramos estudar as características do aluno que influiriam na avaliação que o professor faz de seu trabalho escolar. Essas influências são concebidas não só como vieses involuntariamente introduzidos no processo, como resultado de adoção não-crítica de estereótipos vigentes, como também medidas deliberadamente adotadas pelo professor em virtude da função ambígua que é atribuída à avaliação. De fato, parece não haver uma ideologia clara e explicitamente definida em relação ao papel da avaliação. Se, de um lado ela se insere num sistema meritocrático com a finalidade de julgar a qualidade do produto apresentado e servir como árbitro de um processo competitivo onde os «fortes» devem ser recompensados e os «fracos» punidos, de outro reconhece-se o papel que pode desempenhar na dinâmica motivacional do aluno utilizando-a como instrumento incentivador de maior empenho por parte de todos os estudantes. Esse segundo papel não chega a contradizer o sistema competitivo pois, se este exige que apenas alguns vençam, também requer que a maioria procure fazê-lo. No entanto, a situação da avaliação na escola permanece contraditória pois, no caso de um resultado pobre obtido através de grande esforço, contrapõem-se as exigências de objetividade e de tentativa de manutenção de uma atitude considerada desejável.

Quanto à atribuição, deve-se notar que, conforme mostra o Quadro 1, a presente pesquisa limita-se a investigar o que ocorre na avaliação, uma vez dadas certas atribuições. Ou seja, não foi nossa preocupação determinar como se processam essas atribuições. É isto constitui importante limitação pois é bem possível que os estereótipos vigentes levem à atribuição de certas características individuais a todos os membros de um certo grupo racial ou de um sexo, independentemente do comportamento do indivíduo a quem a característica é atribuída. No presente estudo, as

atribuições são manipuladas igualmente para ambos os sexos, e esse controle permite analisar a influência dessas atribuições em si mesmas e a influência do sexo do aluno independentemente dessas atribuições; porém não é possível detectar a influência do sexo, que seria mediada pela atribuição de esforço ou capacidade a um dos sexos.

Quase não é necessário mencionar que, subjacente às nossas hipóteses, há o pressuposto de que os estereótipos, já constatados por inúmeras pesquisas anteriores, estejam atuantes e sejam incorporados nas representações os professores tem de alunos e de alunas de diferentes raças. Isto não implica em negar a existência de diferenças reais de comportamento entre os diferentes grupos. É bem provável que essas diferenças existam e contribuam, de certa forma, para a manutenção dos estereótipos. Mas, ao acentuar as diferenças e considerá-las absolutas e irreversíveis os estereótipos, em sua inúmeras formas de manifestação, inclusive a avaliação, reforçam a existência dessas diferenças e dificultam sua eliminação. Sendo esse, portanto, o processo que nos interessa investigar, o rendimento escolar foi manipulado em um experimento e controlado em outro a fim de analisar a influência do sexo e das atribuições de esforço e capacidade em si mesmas, independentemente do efeito que o rendimento escolar possa ter sobre essas mesmas atribuições (ver Beckman, 1970, 1973) e sobre as representações de alunos e alunas de cada grupo racial.

O modelo esquematizado no Quadro 1 também não inclui «expectativas». Essas são concebidas como mediando raça, sexo e características individuais atribuídas e a avaliação. Isto é, o professor desenvolve certas expectativas para cada aluno, em função de sua raça, seu sexo e o grau de esforço e capacidade que lhe atribui. Essas expectativas podem ser confirmadas ou contrariadas pelo desempenho do aluno mas, em ambos os casos, influenciam a avaliação resultante. Este é um processo dinâmico, no qual a desconformação pode em certas circunstâncias levar à alteração das expectativas e até mesmo das atribuições das quais decorrem. O processo é complexo conforme têm demonstrado inúmeras pesquisas, algumas das quais citamos anteriormente. No entanto, é estudado aqui apenas indiretamente uma vez que não constituía nosso foco principal de interesse.

Uma última observação a respeito do Quadro 1 é que se procurou investigar o seu grau de generalização verificando a sua aplicabilidade a diferentes condições. Em relação à variável sexo, julgamos de particular relevância variar a natureza da tarefa específica a que a avaliação se referia, uma vez que, mesmo dentro da escola, parece persistir uma «divisão sexual de atividades».

Procurou-se, assim, investigar sob forma rigorosamente controlada, algumas variáveis individuais que

interferem na avaliação. Assim, pretende-se contribuir indiretamente para esclarecimento das seguintes questões:

- a obtenção de notas mais altas por parte das meninas na escola média estaria ligada a vieses do professor na avaliação?
- havendo viés, ele seria independente da atribuição de maior esforço ou de maior habilidade às meninas?
- havendo viés, seria generalizado ou se referiria especificamente às tarefas que a cultura costuma considerar como típica de um ou outro sexo?

Foram realizados dois experimentos para investigar essas hipóteses. Um deles replica o procedimento experimental de Weiner e Kukla (1970), e introduz duas outras variáveis independentes: sexo do aluno e disciplina.

O estudo do problema por meio de um procedimento que permita controlar e aleatorizar as variáveis investigadas é sobretudo vantajoso, na medida em que grande parte da literatura que investigou diferenças sexuais na escola utiliza uma metodologia correlacional e de observação no contexto natural, com as grandes limitações que esse tipo de metodologia apresenta (Campbell e Stanley, 1966).

Tal como no estudo citado, os procedimentos consistiram em dar a professores protocolos individuais de alunos hipotéticos contendo dados a respeito de: capacidade e esforço (variáveis manipuladas por Weiner e Kukla), acrescentando-se a essas informações o sexo de cada aluno. No primeiro experimento, manipulou-se também o resultado que cada aluno obteve em determinada prova e a matéria sobre a qual a prova foi realizada (Português e Matemática).

Era então solicitado aos sujeitos que utilizassem, para comunicar a cada aluno sua avaliação, de 0 a 10 estrelas douradas ou vermelhas, as quais significavam recompensa e punição, respectivamente, em diferentes graus.

No segundo experimento, o professor deveria corrigir uma redação. Além das variáveis capacidade, esforço e sexo, que já haviam sido manipuladas no primeiro experimento, foram também manipulados o tipo das redações (masculino ou feminino) e raça do aluno (branca ou preta).

Em ambos os experimentos, para maior realismo, procurou-se construir uma situação na qual os sujeitos acreditaram estar avaliando crianças realmente existentes, as quais receberiam os resultados atribuídos pela avaliação realizada.

No primeiro experimento, não se procurou controlar o sexo dos sujeitos avaliadores porque a maioria dos estudos não tem constatado influência do sexo do professor (p. ex: Good, 1973).

EXPERIMENTO I¹

PROCEDIMENTO

Este experimento foi adaptado de Rest et al. (1973). Cada sujeito recebeu 24 fichas individuais de alunos de 1º grau, sendo 12 de meninos e 12 de meninas, nas quais constavam: o nome, a série e o sexo da criança, o seu grau de esforço e de capacidade em Português ou em Matemática e o resultado por ela obtido numa prova objetiva dessa matéria em termos de porcentagem de acertos. Com base nessas informações solicitou-se que os sujeitos avaliassem cada aluno e lhe atribuissem uma nota em termos de recompensa (0 a 10 estrelas douradas) ou de punição (0 a 10 estrelas vermelhas).

As variáveis independentes foram portanto: sexo (masculino ou feminino), esforço (pouco ou muito), capacidade (alta ou baixa), resultado (alto ou acima de 70% de acertos, médio ou entre 40% e 70% de acertos, ou baixo ou menos de 40% de acertos) e matéria (Português ou Matemática). A matéria variou entre blocos, isto é, cada sujeito fez todas as 24 avaliações na área de Português ou na área de Matemática. As variáveis sexo, esforço, capacidade e resultado foram aleatorizadas para cada sujeito avaliador. Assim, cada uma das 24 fichas que o sujeito recebeu apresentava uma combinação aleatória de sexo, esforço, capacidade e resultado ($2 \times 2 \times 2 \times 3 = 24$).

A variável dependente foi o número de estrelas douradas ou vermelhas atribuídas a cada aluno. Quando o avaliador atribuiu estrelas douradas a nota variou de 0 a 10 e no caso de estrelas vermelhas de 0 a -10.

Planejamento: o plano de Análise de Variância adequado para esses dados é o «split-plot», descrito por Kirk (1972). Trata-se de um modelo no qual as combinações de variáveis são distribuídas entre blocos de tal forma que apenas um nível de uma das variáveis (Matéria) aparece em cada bloco, mas todos os níveis das demais variáveis aparecem em todos os blocos. Assim, usando a nomenclatura mais comum nos experimentos de psicologia, diríamos que o «tratamento» «Matéria» foi distribuído **entre** sujeitos enquanto os «tratamentos» «sexo», «esforço», «capacidade» e «resultado» variavam **dentro** dos sujeitos. Neste caso, o poder associado com o teste de Matéria é menor que o dos demais testes.

Aplicação: o experimento foi realizado simultaneamente com dois grupos independentes de sujeitos. O primeiro grupo era composto de 39 alunas e 1 aluno de curso de Pedagogia de uma faculdade oficial do interior do Estado de São Paulo. Todos os 40 sujeitos possuíam experiência anterior de magistério na pri-

meira etapa do 1º grau (1ª a 4ª séries). A cada sujeito foi atribuída aleatoriamente uma matéria para fazer suas avaliações — Português ou Matemática — de modo que 20 avaliaram uma matéria e 20 a outra.

O segundo grupo foi constituído de 40 voluntários dos cursos de Licenciatura para o Magistério da Pontifícia Universidade Católica de São Paulo, 20 da área de Português e 20 da área de Matemática, os quais fizeram avaliações em suas respectivas especialidades. Neste grupo não se controlou a experiência anterior de magistério.

Os experimentadores se apresentaram para os sujeitos como orientadores de uma suposta escola de 1º grau e disseram que estavam solicitando de pessoas alheias a essa escola que avaliassem seus alunos com base em alguns dados objetivos. Alegou-se que se queria investigar se a subjetividade do professor que conhecia pessoalmente os alunos chegava a afetar muito as suas avaliações. Esse procedimento teve por finalidade levar os sujeitos a pensar que estavam avaliando crianças realmente existentes e não simplesmente participando de um experimento.

Depois dessa explicação pediu-se ao sujeito que tentasse ser o mais objetivo possível e entregou-se a cada um um envelope contendo o seguinte material (Anexo 1):

- carta do «diretor» da escola, mimeografada em papel timbrado com o nome da escola, tudo de modo a fazer crer que ela realmente existia; nessa carta se repetia o que já fora oralmente dito e se colocavam as instruções sobre como o sujeito deveria proceder para realizar sua avaliação;
- conjunto das 24 fichas individuais dos alunos, cada uma representando uma combinação diferente dos fatores sexo, esforço, capacidade e resultado, e ordenadas aleatoriamente para cada sujeito; a série dos alunos a serem avaliados foi a 4ª para o primeiro grupo e a 8ª para o segundo grupo;
- lista com os nomes dos 24 alunos (12 meninos e 12 meninas), que foi a mesma para cada sujeito embora o esforço, capacidade e resultado associados a cada nome tenham variado de sujeito para sujeito; ao lado de cada nome havia o espaço para o sujeito escrever quantas estrelas, e de que cor, ele atribuíra a cada aluno.

Embora os sujeitos tenham participado coletivamente do experimento foi possível controlar para que não houvesse comunicação entre eles. Uma vez completadas as avaliações, cada um recolocou seu material no envelope e o devolveu ao experimentador, e aguardou.

Quando todos haviam entregue o material completado, contou-se aos sujeitos o objetivo real do estudo e agradeceu-se sua colaboração.

(1) Realizado com a colaboração dos alunos do curso de Pós-Graduação em Psicologia Educacional da Pontifícia Universidade Católica de São Paulo.

Resultados e discussão

As Tabelas 1 e 2 resumem análises de variância feitas em separado para as duas amostras. A semelhança dos dados obtidos, apesar das diferenças existentes entre os dois grupos de avaliadores, permite-nos considerar as duas aplicações como réplicas e aumenta a confiança nos resultados.

A capacidade e o esforço do aluno, bem como o resultado alcançado por ele, foram as variáveis que tiveram efeito principal significativo sobre as notas atribuídas pelos avaliadores. O sexo do aluno avaliado e a matéria não tiveram influência.

TABELA 1
ANÁLISE DE VARIÂNCIA
PROFESSORES DA CAPITAL

Fonte	gl	QM	F
A (matéria)	1	7,53	0,09
B (sexo)	1	0,46	0,36
C (capacidade)	1	1.573,38	36,12**
D (esforço)	1	3.515,18	111,91**
F (resultado)	2	2.317,62	116,41**
A x B	1	0,30	0,24
A x C	1	0,00	
A x D	1	1,28	0,04
A x F	2	47,40	2,38
B x C	1	0,46	0,21
B x D	1	4,96	1,46
B x F	2	0,16	0,08
C x D	1	46,38	4,38**
C x F	2	1,43	0,22
D x F	2	43,85	6,27*
A x B x C	1	1,28	0,58
A x B x D	1	0,05	0,01
A x B x F	2	3,05	1,48
A x C x D	1	14,75	1,39
A x C x F	2	3,02	0,45
A x D x F	2	7,27	1,04
B x C x D	1	4,68	2,27
B x C x F	2	0,71	0,33
B x D x F	2	2,96	1,11
C x D x F	2	11,27	1,29
A x B x C x D	1	6,18	3,00
A x B x C x F	2	0,88	0,40
A x B x D x F	2	0,04	0,75
A x C x D x F	2	3,30	0,38
B x C x D x F	2	1,30	0,49
A x B x C x D x F	2	1,43	0,54

* $\alpha < 0,05$

** $\alpha < 0,01$

TABELA 2
ANÁLISE DE VARIÂNCIA
PROFESSORES DO INTERIOR

Fonte	gl	QM	F
A (matéria)	1	6,18	0,10
B (sexo)	1	1,93	0,50
C (capacidade)	1	1.267,30	65,21**
D (esforço)	1	3.940,65	113,03**
F (resultado)	2	1.978,35	82,60**
A x B	1	25,10	0,65
A x C	1	52,73	2,71
A x D	1	0,46	0,01
A x F	2	41,23	1,72
B x C	1	17,88	3,77
B x D	1	0,23	0,05
B x F	2	0,95	0,17
C x D	1	181,13	20,89**
C x F	2	14,80	1,70
D x F	2	27,37	4,01*
A x B x C	1	3,63	0,77
A x B x D	1	0,88	0,20
A x B x F	2	8,88	1,59
A x C x D	1	34,88	4,02
A x C x F	2	4,57	0,53
A x D x F	2	4,93	0,72
B x C x D	1	2,50	0,74
B x C x F	2	5,26	1,48
B x D x F	2	0,15	0,03
C x D x F	2	23,38	3,03
A x B x C x D	1	5,55	1,64
A x B x C x F	2	1,14	0,32
A x B x D x F	2	0,80	0,16
A x C x D x F	2	6,34	0,82
B x C x D x F	2	8,06	2,10
A x B x C x D x F	2	5,82	1,52

* $\alpha < 0,05$

** $\alpha < 0,01$

As médias apresentadas na Tabela 3 mostram a direção dos efeitos principais. Quanto mais alta a capacidade, quanto maior o esforço e quanto mais alto o resultado, maior a recompensa atribuída. No que tange a esforço e resultado esses dados confirmam os de outros experimentos que manipularam essas variáveis (Weiner e Kukla, 1970 e Rest et al., 1973). Entretanto o efeito de capacidade difere do observado em alguns desses experimentos já realizados. Em Weiner e Kukla (1970) e em um dos experimentos realizados por Rest et al. (1973) a capacidade baixa foi mas recompensada que a capacidade alta. Esse fenômeno, designado por esses autores de «efeito de capacidade baixa», seria explicado pela tendência que teriam os avaliadores de incentivar indivíduos menos dotados. Todavia, controlando o grau de dificuldade de tarefa Rest et al. verificaram que o efeito de

capacidade desaparece. Ele não ocorreu também quando esses autores mudaram o planejamento experimental e apresentaram, para cada avaliador, os dados de apenas um indivíduo para ser avaliado. Uma das hipóteses levantadas para explicar o efeito de baixa capacidade que ora ocorre e ora não ocorre é a de que não seria a baixa capacidade «per se» o que os avaliadores valorizam. Estes tenderiam a inferir que a baixa capacidade exigiria um incremento no esforço para obter um mesmo resultado. Assim, o esforço seria o mediador na recompensa atribuída aos menos dotados.

De qualquer modo, o fato de que o «efeito de capacidade baixa» não tenha ocorrido no presente estudo corrobora a idéia de que esse fenômeno mereceria mais investigação. Além disso sugere que a ideologia do esforço pessoal, nos sujeitos estudados, estaria sendo minimizada pelo valor atribuído também a capacidade.

TABELA 3

MÉDIAS DAS VARIÁVEIS CUJOS EFEITOS PRINCIPAIS FORAM SIGNIFICATIVOS

Variáveis	Interior	Capital
Alta capacidade — C ₁	5,04	4,86
Baixa capacidade — C ₂	2,74	2,30
Pouco esforço — D ₁	1,87	1,67
Muito esforço — D ₂	5,92	5,49
Baixo resultado — F ₁	1,21	0,66
Médio resultado — F ₂	4,35	4,12
Alto resultado — F ₃	6,12	5,96

Nos Gráficos 1 e 2 estão representadas as duas interações significativas: de capacidade com esforço e de esforço com resultado.

A interação de capacidade com esforço mostra que este último é um fator mais poderoso na determinação da recompensa quando associado a capacidade baixa do que quando associado a capacidade alta: a diferença entre recompensa de muito e pouco esforço é maior para capacidade baixa do que para capacidade alta. Isso sugere que, no processo de avaliação, tenderá a ser especialmente recompensado o aluno que, embora pouco capaz, seja muito esforçado. Entretanto essa recompensa não será tão grande quanto a atribuída àquele que combine alta capacidade e muito esforço, o que reforça os comentários já feitos quando discutimos o efeito principal de capacidade. Assim é que a combinação de pouco esforço e baixa capacidade é mais punida do que a de pouco esforço e alta capacidade, como a sugerir que pouco esforço seria mais tolerável em pessoas mais capazes. Ocorre exatamente o inverso

GRÁFICO 1

PADRAO DOS DADOS RESULTANTES DA INTE- RAÇÃO CAPACIDADE × ESFORÇO, PARA AS DUAS AMOSTRAS.

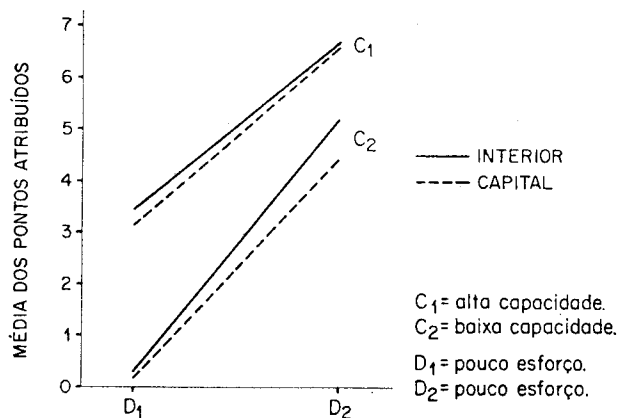
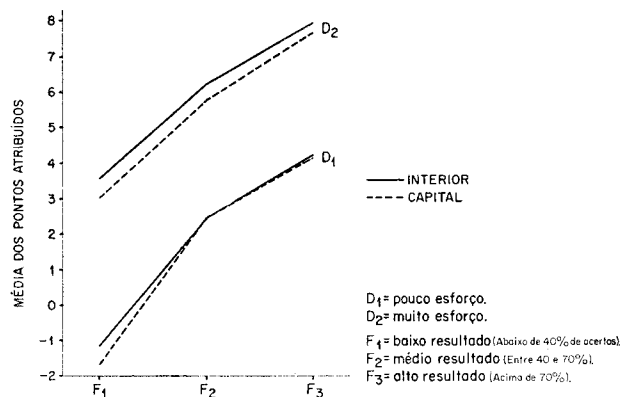


GRÁFICO 2

PADRAO DOS DADOS RESULTANTES DA INTE- RAÇÃO ESFORÇO × RESULTADO, PARA AS DUAS AMOSTRAS.



no citado estudo de Weiner e Kukla (1970), no qual a maior dose de punição coube aos muito capazes que não se esforçaram. Essa diferença seria talvez explicada pelo fato dos sujeitos americanos serem mais intensamente adeptos da ideologia do esforço.

Na interação de esforço com resultado observa-se que a diferença na nota atribuída para pouco e muito esforço é maior no caso de resultado baixo do que no caso de resultado médio e alto. Ou seja, a recompensa para os que se esforçaram mas fracassaram é relativamente maior do que para os que se esforçaram mas conseguiram atingir pelo menos um resultado médio. Aqui aparece de certo modo, o valor do esforço para os avaliadores, que o utilizam como critério para um tipo de «prêmio de consolação».

EXPERIMENTO II

PROCEDIMENTO

Neste estudo 16 professores avaliaram 16 redações sobre o tema «Minhas aspirações para o futuro» — oito redações tipicamente feminina e oito tipicamente masculinas — variando-se sistematicamente o sexo, a raça, a capacidade e o grau de esforço do autor ou autora de cada redação. O procedimento constou de três partes: 1) seleção das redações; 2) seleção de fatos de identificação de crianças a quem as redações foram atribuídas e 3) avaliação de redações identificadas com as fotografias.

I. Seleção das Redações

Nosso objetivo era selecionar um conjunto final de 16 redações que satisfizessem a dois critérios: a) ser relativamente homogêneo quanto à qualidade, para que este fator não interferisse na avaliação; e b) possuir igual número de redações cujo conteúdo fosse típico do sexo feminino, mas plausível para o masculino, e redações cujo conteúdo fosse típico do sexo masculino mas plausível para o feminino, para que cada redação pudesse ser atribuída a um ou outro sexo para avaliadores diferentes¹.

Através dos professores de Português obtivemos redações de 280 crianças matriculadas em três escolas de bairros de periferia da cidade de São Paulo. Essas 280 redações sofreram várias seleções até se obter as 16 redações finalmente usadas na pesquisa de acordo com o procedimento que se descreve a seguir.

Inicialmente selecionamos 32 redações que nos pareceram satisfazer os critérios mencionados e cujos conteúdos consideramos mais interessantes.

A seguir, submetemos essas 32 redações a quatro juízes pedindo-lhes que fizessem uma seleção de acordo com os mesmos critérios e que sugerissem o que deveria ser alterado nas redações que não satisfizessem esses critérios. Das 32 redações foram selecionadas 26.

Tomamos essas 26 redações remanescentes, introduzimos pequenas modificações tendo em vista adaptá-las ainda mais aos critérios, e eliminamos palavras ou expressões que levassem a identificar o sexo do autor. Com essas modificações as redações foram mimeografadas e submetidas ao julgamento de um grupo de 19 estudantes de Letras da cidade de São Paulo, todos eles também professores de 1º grau. Cada estudante recebeu o conjunto de 26 redações e avaliou cada redação numa escala de 0 a 100, respondendo também às seguintes questões:

(1) «típico» e «plausível» aqui são definidos em função dos estereótipos prevalecentes.

Você acha mais provável que esta redação tenha sido escrita:

- por menino
- por menina

Foi difícil decidir se a redação era de menino ou de menina?

- muito difícil
- difícil
- fácil
- muito fácil

Se você recebeu esta redação com identificação da criança que a escreveu, e o sexo dessa criança não fosse o que você julgou, você acharia este fato:

- impossível
- pouco comum
- perfeitamente comum

As notas dadas por este grupo não apresentaram diferença significativa, tendo-se portanto assegurada a equivalência de qualidade entre as redações. Então, das 26 redações avaliadas, foram escolhidas oito tipicamente femininas (embora plausivelmente masculinas) e oito tipicamente masculinas (embora plausivelmente femininas). Esta última seleção obedeceu aos seguintes critérios por ordem de importância:

- 1 — mínimo de 15 juízes concordarem a respeito do sexo do autor da redação
- 2 — máximo de quatro juízes considerarem «difícil» ou «muito difícil» decidir sobre o sexo do autor
- 3 — mínimo de nove juízes considerarem «impossível» ou «pouco comum» que fosse de outro sexo.

Todos os critérios foram estabelecidos no sentido de eliminar as redações neutras e não as «absurdas» pois o número máximo de juízes que considerou impossível que o sexo do aluno-autor fosse diferente daquele que lhe havia atribuído foi quatro, e mesmo nesses casos, um número maior de juízes (cinco ou mais) considerou «perfeitamente normal» que fosse de outro sexo. (É verdade que as realmente «absurdas» — dona de casa, jogador de futebol, etc. — já haviam sido eliminadas pelas próprias pesquisadoras, na primeira seleção das 32 redações iniciais).

Todas as redações masculinas satisfizeram os três critérios de seleção. Das redações femininas só foi possível obter duas que satisfiziam os três critérios. Três delas atenderam ao critério de concordância a respeito do sexo do autor (critério 1) e ao critério de pequena dificuldade desta decisão (critério 2). Houve ainda três que só satisfizeram o primeiro critério. Os dados sobre todas as 26 redações estão no Quadro I e Ia.

QUADRO I

NÚMERO DE PROFESSORES, SEGUNDO O SEXO QUE ATRIBUÍRAM AO AUTOR DE CADA REDAÇÃO

REDAÇÃO	SEXO ATRIBUÍDO				DIFICULDADE DE ATRIBUIÇÃO					POSSIBILIDADE DE OUTRO SEXO			
	♀	♂	♀/♂	S/R	1	2	3	4	5	1	2	3	4
1	3	16	-	-	1	1	11	6	-	2	7	9	1
2	1	17	1	-	-	1	16	1	1	2	9	7	1
3	3	16	-	-	-	3	13	3	-	2	7	10	-
4	3	16	-	-	-	2	11	6	-	2	7	10	-
5	1	18	-	-	-	1	15	3	-	3	10	6	-
6	3	16	-	-	1	5	10	2	1	2	5	12	-
7	19	-	-	-	-	1	14	4	-	2	7	10	-
8	16	3	-	-	-	3	12	3	1	1	-	18	-
9	1	18	-	-	-	2	13	4	-	2	12	5	-
10	1	18	-	-	1	1	15	2	-	4	10	5	-
11	19	-	-	-	-	2	11	6	-	4	7	8	-
12	16	3	-	-	2	4	10	2	1	4	3	12	-
13	15	3	1	-	-	4	12	2	1	-	5	13	1
14	17	1	-	1	-	6	11	1	1	2	4	12	1
15	16	3	-	-	-	5	12	2	-	2	3	14	-
16	14	5	-	-	-	6	10	3	-	-	3	16	-
17	14	4	-	1	-	7	8	3	1	3	4	11	1
18	14	5	-	-	-	4	12	3	-	1	6	12	-
19	5	14	-	-	-	6	10	3	-	1	7	11	-
20	-	19	-	-	1	2	13	3	-	4	4	11	-
21	3	16	-	-	1	3	12	3	-	2	3	14	-
22	-	19	-	-	-	2	15	2	-	4	8	7	-
23	5	14	-	-	1	7	9	2	-	2	2	15	-
24	19	-	-	-	-	3	13	3	-	2	3	14	-
25	11	8	-	-	1	8	8	2	-	-	1	18	-
26	14	5	-	-	-	5	11	3	-	-	3	16	-

Sexo atribuído

- ♀ - mulher
- ♂ - homem
- ♀/♂ - mulher ou homem
- S/R - sem resposta

Dificuldade de atribuição

- 1 - muito difícil
- 2 - difícil
- 3 - fácil
- 4 - muito fácil
- 5 - sem resposta

Possibilidade de outro sexo

- 1 - impossível
- 2 - pouco comum
- 3 - perfeitamente comum
- 4 - sem resposta

QUADRO Ia

ANALISE DE VARIANCIA DA NOTAS DE CADA REDAÇÃO

Fonte	g1	QM	F
Redação	25	96,37	0,78
Erro	442	122,76	
Total	467		

Como se vê, um número maior de redações masculinas satisfaz todos os três critérios, o que parece confirmar que o papel profissional masculino é definido com menor ambiguidade.

II. Seleção de Fotografias

Era necessário obter fotografias que fossem típicas de crianças de cada raça e sexo, e que não apresentassem grandes diferenças entre si, exceto quanto a esses dois fatores. O processo de seleção dessas fotos foi o seguinte:

Em primeiro lugar, fomos a uma escola da periferia de São Paulo e tiramos fotografias 3x4 de 80 crianças de 12 anos de idade, sendo 20 meninos pretos, 20 meninos brancos, 20 meninas pretas e 20 meninas brancas.

Selecionamos seis fotos de cada sexo e raça das crianças mais «bonitinhas», segundo o julgamento de três pesquisadoras. Depois submetemos estas fotos a três outras pesquisadoras para julgarem quão típicas de cada sexo e raça eram as fotografias escolhidas. Levantou-se a suspeita de que as crianças brancas seriam de nível sócio-econômico superior ao das crianças pretas. Então, voltamos à escola e colhemos informações sobre a ocupação dos pais das crianças fotografadas. Daí selecionamos quatro fotos de cada sexo e raça segundo os critérios:

- sexo bem definido
- cor bem definida
- filhos de operários qualificados
- aparência agradável.

A seguir selecionamos, paralelamente, quatro redações: uma de um aluno que queria ser militar, outra de um aluno que queria ser engenheiro, outra de uma aluna que queria ser artista e outra que queria ser pediatra. Quatro crianças de boa letra fizeram, cada uma, dezesseis cópias de uma dessas redações. Solicitamos a 16 alunas de um curso de pesquisa que as corrigissem. Estas alunas eram formadas, na sua maioria, em Sociologia e Psicologia. Eram professoras, algumas professoras universitárias, de São Paulo, Rio de Janeiro, Bahia e Rio Grande do Sul.

Atribuímos cada redação a um sexo e uma raça, mas variamos a fotografia que representava este sexo e esta raça para cada uma das 16 avaliadoras.

A distribuição de fotos e redações pelas avaliadoras encontra-se no Quadro II.

QUADRO II
DISTRIBUIÇÃO DE FOTOS E REDAÇÕES PELOS AVALIADORES

Avaliadores	REDAÇÕES			
	Militar	Pediatra	Engenheiro	Artista
A	F	F	F	F
1	111	121	211	221
A	F	F	F	F
2	111	121	211	221
A	F	F	F	F
3	111	121	211	221
A	F	F	F	F
4	111	121	211	221
A	F	F	F	F
5	112	122	212	222
---	---	---	---	---

onde A_i = avaliadores $i = 1 \dots 16$

F_{ijk} = fotos $i = 1$ (preto)
2 (branco)

$j = 1$ (masculino)
2 (feminino)

$k = 1 \dots 4$ (quatro crianças de cada cor e raça)

Para cada sexo e raça fizemos uma análise de variância e verificamos que, em nenhum subgrupo, havia qualquer criança que se destacasse por sua aparência a tal ponto que influenciasse a correção. Isto é, não houve diferença significativa entre as fotos. Os dados aparecem no Quadro III. A ausência de

QUADRO III
ANALISES DE VARIANCIA

Fonte	g1	QM	F
Fotos de meninos pretos	3	0,85	0,26
Erro	12	3,32	
Total	15		
Fotos de meninos brancos	3	0,60	0,58
Erro	12	1,03	
Total	15		
Fotos de meninas pretas	3	2,18	1,62
Erro	12	1,35	
Total	15		
Fotos de meninas brancas	3	0,43	0,17
Erro	12	2,56	
Total	15		