

UMA ANÁLISE DA FIDEDIGNIDADE DA TAXONOMIA DE OBJETIVOS EDUCACIONAIS (DOMÍNIO COGNITIVO)

MARIA AMÉLIA AZEVEDO GOLDBERG

1.0. Introdução

Em 1956, Bloom e seus colaboradores publicavam nos Estados Unidos, uma Taxonomia de Objetivos Educacionais para o domínio das operações intelectuais. Definida como "classificação dos comportamentos do aluno que representam resultados esperados do processo educacional" (Bloom, 1972, pág. 11), sua estrutura compreendia seis classes principais, dispostas numa hierarquia de complexidade crescente.

- 1.00 Conhecimento
- 2.00 Compreensão
- 3.00 Aplicação
- 4.00 Análise
- 5.00 Síntese
- 6.00 Avaliação

Já em 1956, Bloom relatava que o esquema inicial da Taxonomia tinha sido extensivamente aplicado tanto em pesquisa, quanto em planejamento educacional. A partir de então, vem crescendo o número de estudos envolvendo o emprego da Taxonomia. Só na década de 60, a *Encyclopedia of Educational Research* aponta a existência de estudos feitos por Sanders, Russell, Cox, Guszak, Harrison, Klein e Ellis aplicando a Taxonomia tanto na **análise de objetivos educacionais**, quanto na **análise de itens de avaliação educacional** e, até mesmo, na **análise da interação professor-aluno**. A esses, poderíamos acrescentar outros estudos como os de Brown, Ober e Soar

resultando numa modificação da Taxonomia de Bloom, com o fim de possibilitar a observação, em classe, de interações de natureza cognitiva. Essa modificação é hoje conhecida como a **Florida Taxonomy of Cognitive Behavior** de 1968.

No Brasil, por sua vez, a década de 70 está presenciando o despertar do interesse pelos trabalhos de Bloom.

Assim, registramos a presença da Taxonomia de Objetivos Educacionais, tanto em publicações recentes sobre testes e medidas educacionais, como em trabalhos destinados a treinamento de professores. Além disso, pelo menos três pesquisas de nosso conhecimento, empregaram a Taxonomia como instrumento auxiliar. Uma, realizada no Rio Grande do Sul, sobre **Avaliações e Operações Mentais**, outra em Minas Gerais, sobre previsão de sucesso ou fracasso do estudante de Medicina e outra, no Paraná, sobre ensino normal. (*) Além disso, nosso mercado bibliográfico acaba de receber a tradução portuguesa da obra de 1956 de Bloom, sobre a Taxonomia de Objetivos Educacionais.

O que chama a atenção, porém, em todas essas utilizações da Taxonomia de Bloom é que, se em todas elas a Taxonomia vem sen-

(*) Cf. Ott, Margot Bertolucci — *Avaliação e operações mentais*, RGS, CRPE-MEC/INEP, 1971.

Cf. Alvarenga, G.P.M. e Alvarenga, Clímene Silva M. — Previsão do sucesso ou fracasso do estudante de Medicina utilizando as fontes obtidas nas provas de Física, Química e Biologia do vestibular, in *Ciência e Cultura*, (23), junho 1971, Suplemento, p. 235-236.

Cf. FUNDEPAR — *O ensino normal no Paraná e os recursos humanos para o desenvolvimento*. Paraná, 1971 (brochura).

do utilizada como um **instrumento de mensuração**, poucos foram aqueles usuários que se preocuparam com a investigação de suas qualidades docimológicas de **fidedignidade e validade**. Todavia, esta é uma preocupação básica, pois antes de aplicar um instrumental é preciso decidir sobre sua **adequação**.

O presente trabalho se insere nesta linha de estudos e pretende retomar alguma das colocações feitas por Bloom acerca do que ele denominou de "comunicabilidade" da Taxonomia.

2.0. Fidedignidade de uma taxonomia — Pressupostos teóricos

Fidedignidade "é a **exatidão** ou **precisão** de um instrumento de medida" (Kerlinger, 1964, pág. 430). Embora aparentemente unívoca, a definição é do tipo redutivo, admitindo significados vários. "O coeficiente de fidedignidade é um termo genérico que engloba vários tipos de evidência; cada tipo de evidência sugere um significado diferente" (**American Psychological Association**, 1966, pág. 25). Embora isto seja verdade, qualquer tipo de evidência acerca da fidedignidade de um instrumento de medida envolve necessariamente um ponto comum: todos eles procuram determinar a extensão da influência dos **erros de medida** nas mensurações feitas. Um estudo de fidedignidade deve, pois, necessariamente, partir da **análise das fontes de variância de erro nas mensurações em causa**. Essas fontes de **imprecisão de medida**, por sua vez, variam de instrumento para instrumento. A taxonomia pode ser considerada uma escala ordinal de medida, pois consiste em **categorias** ordenadas. "Este é o problema básico da taxonomia — hierarquizar fenômenos, de forma que se manifestem algumas de suas propriedades essenciais e inter-relações dos próprios fenômenos" (Bloom, 1972, pág. 15). As operações básicas envolvidas são as de **categorização e ordenação**. De fato, aplicar a taxonomia significa atribuir cada objeto a uma categoria de acordo com a regra representada pela definição constitutiva da própria ca-

tegoria. (*) Além disso, considera-se que as categorias devem estar ordenadas, segundo uma relação de complexidade.

Embora, "estritamente falando, classificação e taxonomia não possuam o mesmo significado" (Bloom, 1972, pág. 15) a taxonomia não deixa de ser, no fundo, um esquema de classificação e assim sendo, o problema de sua fidedignidade converte-se no problema da fidedignidade de um sistema de classificação. Quais serão, pois as principais fontes de variância de erro de um sistema classificatório? Classificar significa **julgar** e todo julgamento, para ser digno de confiança, deve ser **objetivo**, isto é, **independe da pessoa que julga** e **estável**, isto é, **independe da ocasião em que a mesma pessoa julga**. Portanto, duas são as principais fontes de variância de erro numa classificação: **juiz** e **ocasião**. Cada uma delas é uma **faceta**, na terminologia da teoria da generalizabilidade. A cada faceta deverá corresponder, pois, um **coeficiente de fidedignidade**. Considerando a objetividade e estabilidade como dimensões necessárias da confiabilidade de um sistema classificatório e considerando, por outro lado, que essas dimensões respondem pelas características de **comunicabilidade** e **consistência** da classificação, podemos afirmar que estudar a fidedignidade da Taxonomia de Objetivos Educacionais de Bloom, equivale a procurar comprovar a afirmação de que ela é "**um sistema lógico**, no sentido de que seus termos estão definidos precisamente e são usados consistentemente" (Krathwohl, 1964, pág. 21).

Até aqui discutimos os pressupostos relativos à **fidedignidade** e à **taxonomia**. Restaria discutir, ainda, os pressupostos ligados aos **objetivos educacionais**. Os objetivos educacionais são "proposições sobre mudanças desejadas nos pensamentos, nas ações ou sentimentos dos alunos e que um determinado curso ou um programa educacional deve produzir" (Bloom, 1967, pág. 389). As proposições sobre as mudanças desejadas podem ser formuladas por sua vez em termos descritivos

(*) Definição constitutiva é aquela que define uma palavra ou construto, usando outras palavras ou construtos (cf. Kerlinger, 1964, p. 33).

ou em termos observáveis. Nesta última categoria, uma das possibilidades é a chamada **definição operacional**, ou seja, aquela que especifica em maior ou menor grau as operações necessárias à medida dos objetivos educacionais. Isso equivale a dizer que qualquer objetivo educacional pode ser definido num "continuum" de generalidade — especificidade. A tradução de objetivos educacionais em procedimentos de avaliação representa um "esforço para clarificar e dar uma definição operacional aos objetivos" (Bloom, 1967, pág. 389).

Foi a estreita relação entre objetivos educacionais e avaliação educacional que levou Bloom a considerar sua taxonomia do domínio cognitivo como aplicável também a instrumentos de medida educacional. Cada categoria recebeu, pois, três níveis de definição — uma descrição verbal, uma exemplificação de objetivos educacionais pertinentes e uma amostra de exemplos de questões de provas destinados a ilustrar tanto "o comportamento abrangido pela categoria, quanto o tipo de item capaz de medir essa objetivo" (Cecco, 1968, pág. 52).

Todavia, se considerarmos que a posição de objetivos educacionais e a avaliação desses mesmos objetivos devem ser mediadas, por experiências de aprendizagem, fácil será entender que os comportamentos envolvidos na solução de um dado tipo de item dependem das experiências anteriores de aprendizagem dos sujeitos. "Por exemplo, dois alunos solucionam um problema algébrico. Um deles pode resolvê-lo de memória, devido a uma experiência anterior com um problema idêntico em aula. O outro aluno deparou com o problema igual e necessita raciocinar, aplicando princípios gerais para encontrar a solução" (Bloom, 1972, pág. 14). Se empregássemos a Taxonomia, o mesmo item poderia ser classificado como pertencendo à categoria de **Conhecimento** e à categoria de **Aplicação**, por exemplo. Esse fator não poderá ser ignorado sempre que se pretender testar a fidedignidade da Taxonomia de Objetivos Educacionais, quando aplicada à classificação de itens de avaliação educacional.

3.0. Objetivos

O presente estudo propõe-se a analisar a fidedignidade da Taxonomia de Objetivos Educacionais de Bloom — Domínio cognitivo, fidedignidade essa definida em termos de **objetividade** e **estabilidade** na classificação de itens de avaliação educacional.

4.0. Metodologia

Para testar a objetividade e a estabilidade na classificação de itens de uma prova de escolaridade, definimos, respectivamente, duas evidências de fidedignidade: a) concordância entre julgamentos feitos por juizes independentes; b) concordância entre duas classificações feitas pelo mesmo juiz, com um certo intervalo de tempo.

Em vez de reduzir, simplesmente, a fidedignidade a porcentagens de concordância inter, ou intra-juizes, preferimos estimá-la através do coeficiente de Scott (cf. Ober, 1971, pág. 71 e segs.). Este coeficiente revela-se superior às simples medidas de porcentagem de concordância, pois estima, em vez disso, a relação entre concordância observada e concordância devida ao acaso. De fato, segundo o coeficiente de Scott.

$$\text{Fidedignidade} = \frac{\text{Concordância observada} - \text{concordância casual}}{\text{Maior concordância possível} - \text{concordância casual}}$$

O coeficiente de Scott foi empregado sempre que quisemos comparar **pares** de classificações. A comparação entre mais de duas classificações foi feita através de uma medida não paramétrica de correlação, tendo em vista a natureza ordinal da Taxonomia. Foi utilizado aqui o coeficiente de concordância W de Kendall, considerado útil na "determinação da concordância entre vários juizes" (Siegel, 1956, pág. 239).

Todos os julgamentos foram feitos sobre 30 questões de uma prova de escolaridade aplicada em um curso de educação integrada do programa MOBREAL de alfabetização de Adultos.

Os juizes foram alunos do Curso de Testes e Medidas ministrado em 1972, na Pontificia Universidade Catolica de São Paulo, dentro do Programa de Pós-Graduação em Psicologia da Educação. Esses juizes estudaram a Taxonomia de Bloom, como parte obrigatória do curso e seu domínio da mesma foi considerado satisfatório pela professora da área. Além disso, exercitaram sua habilidade classificatória, em um treinamento com itens de uma forma paralela àquela cujas questões deveriam classificar depois. Por outro lado, tendo em vista a relação entre itens de rendimento escolar e experiências anteriores de aprendizagem forneceu-se aos juizes uma relação de conteúdos programáticos que os itens a serem classificados pretendiam cobrir. Esses conteúdos eram tratados em três materiais didáticos, doravante designados **material didático A - B - C**. Cada material didático incluía tanto o livro do aluno quanto o guia metodológico do professor.

Embora os mesmos conteúdos estivessem presentes nos três materiais didáticos, uma inspeção do tipo de perguntas que os três guias do professor sugeriam fossem feitas aos alunos, fez-nos supor que os materiais didáticos distribuir-se-iam diferentemente, em termos da versão modificada da Taxonomia de Bloom, proposta por Sanders (1966), (cf. Ober, 1971, pág. 99). Assim sendo, o mesmo item de teste poderia ser resolvido por esta ou aquela operação mental, conforme se tratasse de um aluno que tivesse recebido este ou aquele material didático. Conseqüentemente, nossos juizes, ao classificarem os mesmos itens, colocar-se-iam no lugar dos alunos e refletiriam, através de discrepâncias nos respectivos julgamentos, a influência do material didático que tivessem consultado. Essa a razão pela qual nossos juizes foram distribuídos em grupos A - B e C, conforme tivessem lido os materiais didáticos A - B e C.

Após todo esse trabalho preliminar, cada juiz recebeu os 30 itens para categorizar, trabalhando independentemente. De posse dos resultados, procedemos então a uma seleção, em cada grupo, dos juizes cujas classificações tivessem menor número de erros. O erro foi

definido, estipulativamente, como **não concordância** com:

- a) classificação modal do grupo de juizes e/ou
- b) classificação da autora do teste e/ou
- c) classificação da professora de Testes e Medidas Educacionais.

Com isto, o grupo inicial de 12 juizes reduziu-se a 6, havendo um par para cada grupo. A partir das classificações dos juizes, eliminamos, também, 7 itens (N.^{os} 12, 14, 16, 25, 26, 28 e 30), cujas modas apresentavam porcentagens de freqüência inferiores a 50%. Todos os nossos resultados baseiam-se pois em dados de uma matriz de classificações feitas por 6 juizes para 23 itens de teste. A classificação dos itens, por sua vez, foi feita apenas a partir das seis grandes categorias da Taxonomia de Bloom.

5.0. Resultados

A primeira indagação que fizemos foi acerca do **grau de concordância** existente nas classificações dos seis juizes.

Para cálculo do coeficiente de concordância W de Kendall, as classificações foram convertidas em postos. O coeficiente obtido foi de 0,632, significativo ao nível de 0,001. Isto permite-nos afirmar que os juizes aplicaram, essencialmente, o mesmo padrão nas classificações dos 23 itens de teste. Esse resultado, embora esperado tendo em vista o processo final de seleção de juizes, não nos satisfaz porque, muito dificilmente nos emprego da taxonomia, poder-se-ia contar com um grupo tão grande de juizes. Decidimos, pois, refinar nosso nível de análise e investigar as concordâncias e discordâncias entre **pares de juizes**. Usando a fórmula de Scott para pares de juizes em três condições, correspondentes à leitura que haviam feito dos materiais didáticos A, B e C, obtivemos, respectivamente, coeficientes de fidedignidade de 0,59, 0,07 e 0,24. Considerando como aceitável, de acordo com Ober, um coeficiente de 0,60 ou mais, verificamos que apenas o primeiro deles preen-

che o critério. Obviamente, na medida em que cada par de juízes estava em igualdade de condições, a fonte de variação aqui deveria ser o juiz. Tanto isso é verdade que o exame das matrizes da figura 1 permite-nos detectar, imediatamente, o viés do juiz 4 e do juiz 6, exatamente nos pares em que os coeficientes de fidedignidade foram extremamente baixos.

Enquanto o juiz 4 carregou nas classificações da categoria 1, o juiz 6 saturou seus julgamentos com classificações na categoria 2. A análise dessas matrizes fez, porém, que indagássemos em que medida a concordância entre juízes não seria uma decorrência necessária do tipo de condição a que haviam sido submetidos em termos de leitura do material didático. Para investigar o grau de concordância entre juízes que leram materiais didáticos diversos, selecionamos de cada par de juízes aquele que havia tido menos erros segundo os mesmos critérios que fixamos anteriormente. Obtivemos, assim, 3 juízes e às suas classificações aplicamos o procedimento da **Análise de Variância, por postos**, de Friedman ("two-way") (cf. Siegel, 1956, págs. 166-172). Para cálculo do χ^2 , as classificações dos três juízes foram convertidas em postos. O χ^2 obtido foi de 1,0, não significativo ao nível de 0,001. Podemos pois concluir que a leitura dos diferentes materiais didáticos não teve efeito diferencial na classificação dos itens da prova, feita pelos três juízes que selecionamos. Uma das possíveis explicações é que, de fato, os materiais seriam equivalentes do ponto de vista das experiências que proporcionam. A outra, seria a de que, ao lerem os materiais didáticos, os juízes tenham se preocupado mais com o conteúdo propriamente dito, do que com as estratégias propostas para lidar com esse conteúdo.

Por outro lado, verificamos a existência de um alto coeficiente de estabilidade, no caso de um dos juízes, a quem pedimos que classificasse duas vezes os mesmos itens. O coeficiente de Scott obtido, foi da ordem de 0,89 enquanto que a figura 4 dá uma boa idéia da consistência nos julgamentos de um mesmo juiz. Sua matriz, é, predominantemente, uma

matriz de concordâncias com os julgamentos distribuindo-se, sobretudo, na diagonal principal (fig. 5). O juiz "retestado" foi, diga-se de passagem, aquele que menos "erros" (1) cometeu na primeira classificação dos itens. Para ele, pelo menos o decurso de cerca de dois meses entre uma e outra aplicação da Taxonomia, praticamente nenhum efeito teve. O espaço de dois meses, por sua vez, foi suficiente para eliminar o "efeito memória".

Portanto, parece-nos que até agora só os juízes, individualmente tomados, apareceram como fonte de imprecisão ou infidedignidade na classificação. Duas indagações se colocam agora:

1.º) Que fatores poderiam ser responsáveis por essa variação inter-individual nas classificações?

2.º) Toda a variância atribuída aos juízes seria, necessariamente, **variância erro** e, portanto, fonte de infidedignidade?

Respondendo à primeira indagação, diríamos que esses fatores são inúmeros, desde condições emocionais até um domínio insuficiente da taxonomia. Mas, dentre esses prováveis fatores, um deles merece algumas considerações: é o fator que o próprio Bloom denominou de **imprecisões e ambiguidades na definição das categorias**. Tentando verificar uma possível interferência desse fator, construímos nas figuras 2 e 3, matrizes de **concordâncias e discordâncias** para os nossos 3 pares de juízes. As porcentagens de concordância oscilam ao redor de 50%. No caso das discrepâncias (fig. 3), procuramos verificar se haveria concentração nos nichos que assinalamos, nichos esses que seriam indicativos de confusões entre categorias limítrofes. Pode-se verificar que essa não foi a regra geral. As discordâncias não se distribuíram como seria de se esperar, caso houvesse indefinição de limites entre elas. Observando a figura 3 verificamos, por exemplo, a existência de discrepâncias entre as categorias 1, de Conhecimento, e 3, de Aplicação. Curioso notar que, no caso do mesmo juiz, aplicando duas vezes

(1) No sentido já definido anteriormente.

a Taxonomia, também encontramos a maioria de suas poucas discrepâncias concentrada na intersecção das categorias 1 e 3 (fig. 6). Por outro lado, se examinarmos a fig. 3, veremos que as categorias 1, 2, 3 e 4 aparecem confundidas com várias outras. Todavia, se examinarmos detidamente a definição que Bloom oferece para essas categorias verificaremos que a maior dificuldade talvez resida na categoria 4, já que Análise, no sentido mais usual do termo, estaria incluída, inclusive, na solução de problemas, atividade típica, segundo Bloom, da categoria 3 (Aplicação). “As linhas limítrofes entre Análise e Compreensão, por um lado, e entre Análise e Avaliação, por outro, não estão suficientemente esclarecidas” (Bloom, 1972, pág. 123). Além disso, procurando tornar mais discriminativos certos limites, Bloom introduziu recentemente uma nova categoria, entre as categorias de Conhecimento e Compreensão. Trata-se da “habilidade no uso de processos e procedimentos” (Bloom, 1971, pág. 119). Por outro lado, se examinarmos o gráfico da figura 2, veremos que a categoria 6 (Avaliação) foi a que registrou menor ambigüidade. Evidentemente é preciso não esquecer que todas essas considerações sobre as **discrepâncias e concordâncias** devem ser consideradas como **meramente sugestivas** de vez que é relativamente pequeno o número de observações de que dispomos.

Mas, voltemos a indagar: toda discordância entre classificações dos juízes deverá ser sempre considerada na variância de erro? As intersecções de julgamentos que observamos na categoria 1, com todas as demais categorias, poderão ser explicadas apenas por ambigüidades na definição das classes? Não será legítima essa variação, ou seja, não haverá condições em que essa variância será variância verdadeira? A resposta é afirmativa, pois, como o próprio Bloom afirma, “na avaliação de comportamentos que transcendem as categorias de Conhecimento, o contexto em que os termos ou símbolos aparecem deve ser,

de algum modo, um novo contexto” (Bloom, 1972, pág. 82).

Por essa razão, exercícios idênticos poderão ser categorizados ou na categoria de Conhecimento ou em uma das outras, dependendo do **contexto das experiências anteriores** dos sujeitos, contexto esse que determinará, por exemplo, a **novidade** ou não da situação de avaliação.

Para tentar investigar mais esta possibilidade, elaboramos a tabela 1, onde aparecem as classificações individuais de nossos 6 juízes, sua moda, a classificação da professora do Curso de Testes e Medidas e a classificação da autora do teste. Procuramos então, verificar, para cada item, as discordâncias entre a moda das classificações dos juízes e as classificações da professora e da autora. A fig. 7 ilustra esses resultados.

Como vemos, nove itens apresentam discordâncias internas. Verifiquemos se seria razoável supor que, para cada um deles, há mais de um caminho para se chegar à solução. Ou, por outras palavras, se o mesmo item pode ser resolvido através de diferentes operações mentais, categorizáveis em classes diferentes conforme a Taxonomia. Esses nove itens constam do ANEXO.

a) item 8 — a divergência se coloca entre categorias 2 e 4 e, pois, entre níveis de **compreensão**. Quer-nos parecer que, neste caso, a moda do grupo representaria a categorização mais correta.

b) item 21 — a discrepância está entre as categorias 1 e 3. Essa discrepância só seria legítima se os alunos já tivessem conhecimento anterior das mesmas frases.

c) item 3 — a discrepância se coloca entre as categorias 1, 2 e 5. Aqui, ela parece legítima, pois, tanto o sujeito pode chegar à solução certa apenas recordando o que sabia, como substituindo uma das afirmações pela alternativa correta, como fazendo um jogo mais complicado de comparações.

d) item 11 — a discordância ocorre entre categorias 1, 5 e 3. — Também é válida

essa categorização diferencial, pois, tanto o sujeito pode ter aprendido já a generalização implícita, como poderá chegar a ela, através das operações de aplicação, análise e síntese.

e) item 27 — as mesmas observações do item anterior.

f) item 13 — a discordância ocorre entre as categorias 2 e 5. Tanto um sujeito pode traduzir uma ou algumas das pistas dadas (“sensível à luz, aperta os olhos”) pelo seu correspondente — “visão”, como poderá fazer um jogo mais complicado de comparações até descobrir uma relação implícita entre os elementos propostos.

g) item 4 — a discrepância ocorreu entre as categorias 2 e 3. De fato, a solução tanto pode ser conseguida fazendo uma substituição rápida de “1 800” por “**século passado**” como fazendo uma subtração de 1972 — 1822 e chegando à mesma conclusão.

h) item 19 — as discordâncias ocorreram, aqui, em torno de 3 categorias: 1, 2 e 3. Aqui temos mais um item, para o qual nenhuma das categorizações apresentadas parece totalmente satisfatória. Talvez o melhor fosse classificá-lo em termos da **habilidade no uso de procedimentos e processos**, a categoria que Bloom introduziu, recentemente, entre as de Compreensão e Conhecimento.

i) item 20 — a divergência ocorreu, aqui, em termos das categorias 1 e 2. Aqui, também parece legítima a variação, na medida em que um sujeito pode acertar com a resposta, simplesmente porque aprendeu a relação entre definição — termo pedido, enquanto outro pode chegar aos termos, explorando o significado das definições.

Parece válido concluirmos que, na maioria dos itens analisados, eles não apresentam uma, e uma só, possibilidade de categorização. Ao contrário, cada um comporta uma distribuição de categorizações possíveis e o

que vai determinar a predominância de uma ou outra será, exatamente, o contexto das experiências anteriores de **aprendizagem do sujeito**. “Estes fatos sugerem que, em geral, os itens de um teste podem ser satisfatoriamente categorizados por meio da Taxonomia, quando o contexto em que as provas são aplicadas é conhecido ou considerado” (Bloom, 1972, pág. 19).

6.0. Implicações

Este estudo encerra, como é óbvio, várias limitações de natureza metodológica e, por isso mesmo, para que seus resultados pudessem ser considerados como generalizáveis, seria indispensável submetê-los a novos estudos, por exemplo, em termo de uma validação cruzada das classificações obtidas por nosso grupo de juízes.

Apesar disso, nossa análise sugere que a fidedignidade das classificações de itens de teste educacional, feitas por juízes a partir da Taxonomia de Objetivos Educacionais de Bloom (Domínio Cognitivo), depende do **contexto** das experiências anteriores de aprendizagem dos juízes. “Uma experiência educacional, ou de aprendizagem pode ser pensada como uma interação entre alunos e seu ambiente — professor, outros alunos, material didático e matéria” (Bloom, 1967, pág. 386). Nesse sentido, a mera informação sobre um dos aspectos — por exemplo material **didático** — pode ser considerada insuficiente em termos de redução do contexto anterior de experiências dos sujeitos, a um denominador comum.

Essa advertência é fundamental para todos aqueles que pretenderem classificar itens de avaliação educacional e calcar sobre essa classificação seja decisões relativas à pesquisa educacional, seja decisões relativas ao planejamento educacional.

FIG. 1 - MATRIZES DE RESULTADOS: 3 pares de juízes usando Taxonomia de Bloom (domínio cognitivo), para classificação de itens em seis categorias.

MATERIAL DIDÁTICO A							MATERIAL DIDÁTICO B						MATERIAL DIDÁTICO C										
Juiz 2			Juiz 1				Juiz 4			Juiz 3			Juiz 6			Juiz 5			Totais				
1	2	3	4	5	6	Totais	1	2	3	4	5	6	Totais	1	2	3	4	5	6	Totais			
•	•		•			4	•	•	•				5	•	•	•					8		
		•				3	•	•	•	•			4			•					4		
•	•	•	•			8	•	•	•	•			4	•	•	•	•				7		
		•	•	•		4				•			2								0		
						0	•	•					4								0		
				•	•	4					•	•	4						•	•	4		
Totais	6	5	6	2	0	4	23	Totais	15	3	0	1	0	4	23	Totais	3	10	6	1	0	3	23

- CÓDIGO: 1 = Conhecimento
 2 = Compreensão
 3 = Aplicação
 4 = Análise
 5 = Síntese
 6 = Avaliação

FIG. 2 - MATRIZES DE CONCORDÂNCIAS ENTRE TRÊS PARES DE JUÍZES

MATERIAL DIDÁTICO A

Juiz 2 \ Juiz 1	1	2	3	4	5	6	Totais
1	2						
2	2	2					
3			2	2			
4				2			
5					2		
6						2	
Totais	2	2	2	2	2	2	13

MATERIAL DIDÁTICO B

Juiz 4 \ Juiz 3	1	2	3	4	5	6	Totais
1	2	2					
2	2	2					
3			2				
4				2			
5					2		
6						2	
Totais	2	2	2	2	2	2	11

MATERIAL DIDÁTICO C

Juiz 6 \ Juiz 5	1	2	3	4	5	6	Totais
1	2	2					
2	2	2					
3			2				
4				2			
5					2		
6						2	
Totais	2	2	2	2	2	2	10

Porcentagem de concordância: 57%

Porcentagem de concordância: 48%

Porcentagem de concordância: 43%

FIG. 3 - MATRIZES DE DISCORDÂNCIAS ENTRE TRÊS PARES DE JUÍZES

MATERIAL DIDÁTICO A MATERIAL DIDÁTICO B MATERIAL DIDÁTICO C

		MATERIAL DIDÁTICO A						MATERIAL DIDÁTICO B						MATERIAL DIDÁTICO C											
		Juiz 2	1	2	3	4	5	6	Totais	Juiz 4	1	2	3	4	5	6	Totais	Juiz 6	1	2	3	4	5	6	Totais
Juiz 1	1																								
	2																								
	3																								
	4																								
	5																								
	6																								
	Totais																								
	1																								
	2																								
	3																								
	4																								
	5																								
	6																								
	Totais																								

FIG. 4 - MATRIZ DE RESULTADOS:
o mesmo juiz, aplicando
duas vezes a Taxonomia de
Bloom, (domínio cognitivo),
para classificação de itens
em seis categorias.

II vez / I vez	1	2	3	4	5	6	Totais
1	•••••		•				8
2	••	••					4
3	•	•	•••				7
4							0
5							0
6					••	••	4
Totais	8	5	6	0	0	4	23

FIG. 5 - MATRIZ DE CONCOR-
DÂNCIAS:
o mesmo juiz, aplicando
duas vezes a Taxonomia
de Bloom, para classi-
ficação de itens em seis
categorias.

II vez / I vez	1	2	3	4	5	6	Totais
1	•••••						
2		••••					
3			••••				
4				••••			
5					••••		
6						••••	
Totais							20

FIG. 6 - MATRIZ DE DISCORDÂN-
CIAS:
o mesmo juiz, aplicando
duas vezes a Taxonomia
de Bloom, para categori-
zação de itens.

II vez / I vez	1	2	3	4	5	6	Totais
1		•					
2							
3	•						
4							
5							
6							
Totais							

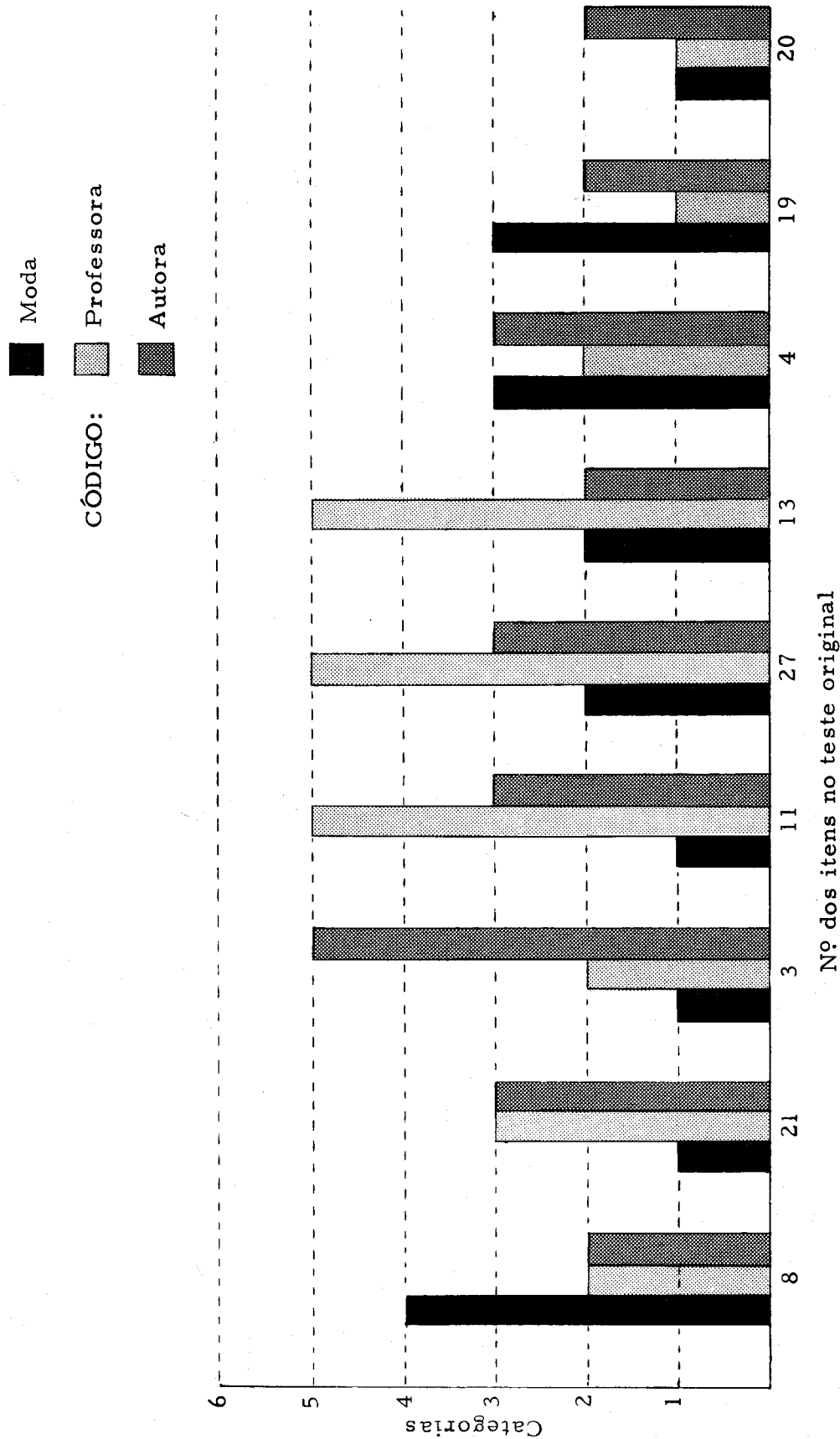
TABELA I

Classificação dos 23 itens feita: a) pelos seis juizes seleccionados
 b) pela professora de Testes e Medidas
 c) pela autora do teste

ITENS	Classificação dos Juizes							Classif. da Profess.	Classif. da Autora
	MATERIAL A		MATERIAL B		MATERIAL C		MODA		
	Juiz 1	Juiz 2	Juiz 3	Juiz 4	Juiz 5	Juiz 6			
1 (1)	1	1	2	1	1	1	1	1	1
2 (5)	3	1	3	1	3	3	3	3	3
3 (7)	2	2	4	2	2	2	2	2	2
④ (8)	4	2	4	4	3	4	4	2	2
5 (9)	3	3	1	1	1	3	3/1	3	3
6 (10)	1	3	1	1	1	2	1	1	1
⑦ (21)	3	1	1	1	1	3	1	3	3
8 (29)	3	1	3	1	1	3	1/3	3	3
9 (2)	6	6	6	6	6	6	6	6	6
⑩ (3)	1	4	1	1	1	2	1	2	5
⑪ (11)	1	1	5	1	1	1	1	5	3
12 (17)	6	6	6	6	6	6	6	6	6
13 (18)	6	6	6	6	6	6	6	6	6
⑭ (27)	4	4	2	2	2	2	2	5	3
⑮ (13)	2	2	5	2	2	3	2	5	2
16 (15)	6	6	6	6	6	3	6	6	6
⑰ (4)	3	2	3	1	3	2	3	2	3
18 (6)	2	2	2	1	2	2	2	2	2
⑱ (19)	3	3	1	1	3	2	3	1	2
⑳ (20)	3	1	2	1	1	1	1	1	2
21 (22)	3	3	5	1	3	2	3	3	3
22 (23)	4	3	3	1	3	2	3	3	3
23 (24)	4	3	5	1	3	2	3	3	3

Nota: os números entre parêntesis são os números do item no teste original.

FIG. 7 - GRÁFICO DAS DISCORDÂNCIAS DE CLASSIFICAÇÕES DE 8 ITENS: TOMANDO POR BASE A MODA DAS CLASSIFICAÇÕES DOS SEIS JUÍZES SELECIONADOS.



REFERÊNCIAS BIBLIOGRÁFICAS

- ALVARENGA, G. P. M. e ALVARENGA, Clímene Silva M. — Previsão do sucesso ou fracasso do estudante de Medicina, utilizando as fontes obtidas nas provas de Física, Química e Biologia do vestibular — Resultados de dois anos durante o curso básico, Minas Gerais, Departamento de Fisiologia do Instituto de Ciências Biológicas da UFMG (notícia in *Ciência e Cultura* (23), junho 1971, Suplemento, págs. 235-236).
- American Psychological Association — *Standards for Educational and Psychological Tests and Manuals*, 1966.
- BAUZER MEDEIROS, Ethel — *As provas objetivas*. RJ, FGV, 1971.
- BLOOM, B. S. e outros — *Taxonomy of Educational Objectives, Handbook I — Cognitive Domain*, N. Y. Longmans Green and Co., s/d.
- BLOOM, B. S. — Testing Cognitive Ability and Achievement in Gage, N. L. (editor) *Handbook of Research on Teaching* — Chicago, Rand McNally, 1967, cap. 8, págs. 379-397.
- BLOOM, B. S. e outros — *Handbook on Formative and Summative Evaluation of Student Learning*, N. Y., McGraw-Hill Co., 1971.
- BLOOM, B. S. e outros — *Taxionomia de Objetivos Educacionais*, [tradução de Flávia Maria Sant'Anna] P. Alegre, Globo, 1972.
- BROWN, B. B. e outros — *The Florida Taxonomy of Cognitive Behavior*, Florida Institute for Development of Human Resources, January, 1971, (mim.).
- CECCO, J. P. de — *The Psychology of Learning and Instruction* — N. Jersey, Prentice Hall, 1968.
- EBEL, R. L. (editor) — *Encyclopedia of Educational Research*, London, The MacMillan Co., 1969, verbete — «Objectives and outcomes», págs. 908-914.
- GAGNÉ, R. M. — Educational objectives and human performance in Krumboltz, D. (editor) — *Learning and the Educational Process*, Chicago, Rand McNally and Co., 1965, cap. 1, págs. 1-24.
- GRONLUND, N. E. — Stating behavioral objectives for classroom instruction in Biehler, F. R. — *Psychology Applied to Teaching, Selected Readings* — Boston, Houghton Mifflin Co., 1972, cap. 22, págs. 363-369.
- HAYS, W. I. — *Quantificação em Psicologia*. SP, Herder, 1970.
- KERLINGER, F. — *Foundations of Behavioral Research*, N. Y., Holt, Rinehart and Winston, 1964.
- KRATHWOHL, D. R. — The Taxonomy of Educational Objectives — Use of the Cognitive and Affective Domains — in Lindvall, C. M. (ed.) — *Defining Educational Objectives* — Pittsburgh, University of Pittsburgh Press, 1964, págs. 19-36.
- LOURENÇO FILHO (apresentação) — *Testes e Medidas na Educação*, RJ, FGV, 1970.
- MCDONALD, F. J. — *Educational Psychology*, California, Wadsworth, 1966.
- MCGAW, B. and WARDROP, J. L. and BUNDA, Mary Anne — Classroom Observation Schemes: Where are the errors. *American Educational Research Journal*, IX (1) 13-26, Winter, 1972.
- MEDLEY, D. M. and MITZEL, H. E. — Measuring Classroom Behavior by Systematic Observation in Gage, N. L. (editor) — *Handbook of Research on Teaching*, Chicago, Rand McNally, 1967, cap. 6, págs. 247-319.
- OBER, R. L.; BENTLEY, E. L.; MILLER, E. — *Systematic Observation of Teaching*, N. Jersey, Prentice Hall, 1971.
- OTT, B. Margot — *Avaliação e Operações Mentais*. P. Alegre, CRPE — MEC/INEP, 1971.
- Secretaria do Estado dos Negócios da Educação — Publicações GEPE Dr. Edmundo de Carvalho — Caderno II. *Planejamento de Ensino*, 1971.
- SIEGEL, S. — *Non parametric Statistic for the Behavioral Sciences*, N. Y., McGraw Hill Book Co., 1956.
- WASIK, J. L. — *The Construct and Concurrent Validation of the Taxonomy Processes of Synthesis and Evaluation*. (mim.).

ANEXO

- 3) Adivinhe quem são eles;
- saíram de São Paulo e entraram no sertão
 - descobriram ouro e pedras preciosas
 - iam deixando povoados por onde passavam
 - por causa deles o Brasil ficou maior do que era
- a) os jesuítas
 - b) os bandeirantes
 - c) os governadores gerais

- 4) O Brasil tornou-se independente de Portugal no dia 7 de setembro de 1822. Isto aconteceu:
- a) neste século
 - b) no século passado
 - c) faz mais de dois séculos

- 8) Zezinho começou a trabalhar na enxada quando tinha:
- a) menos idade que seu pai
 - b) a mesma idade que seu pai
 - c) mais idade que seu pai

Mostre porque

.....

.....

.....

.....

- 11) O que há de comum entre eles?

Figura de um PRETO	Figura de um BRANCO	Figura de um ÍNDIO
-----------------------	------------------------	-----------------------

- a) invadiram o Brasil
- b) formaram o povo brasileiro
- c) descobriram o Brasil

- 13) “Esfrega os olhos com freqüência. É sensível à luz. Pisca excessivamente. Confunde letras, linhas e números. Aperta os olhos”. Essa pessoa deve estar com problemas de:

- a) audição
- b) visão
- c) olfação

- 19) Um compadre disse para o outro:

Adivinhe a chapa de meu caminhão

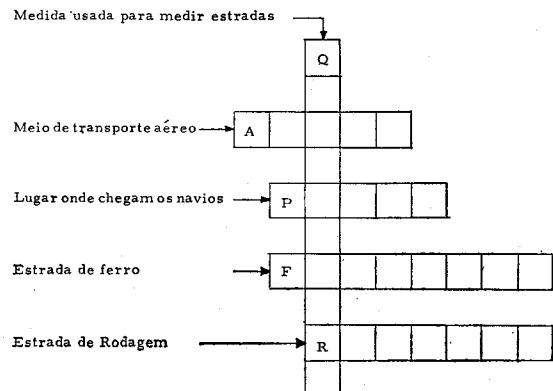
- o 1.º algarismo é igual a $56 \div 8$;
- o 2.º algarismo é igual a $64 \div 16$;
- o 3.º algarismo é igual a $144 \div 18$;
- o 4.º algarismo é igual a $56 \div 14$.

Escreva a chapa abaixo:

1º	2º	3º	4º
<input style="width: 30px; height: 20px;" type="text"/>	<input style="width: 30px; height: 20px;" type="text"/>	<input style="width: 30px; height: 20px;" type="text"/>	<input style="width: 30px; height: 20px;" type="text"/>

Faça as contas aqui:

- 20) Preencha as palavras cruzadas:



21) Complete com o verbo indicado no fim de cada frase:

O conjunto das estradas rodoviárias

 uma rede rodoviária. (FORMAR)

O verão e o inverno

 ao conjunto das estações do ano.
 (PERTENCER).

O conjunto de jovens

 parte de nossa classe. (FAZER).

Os caminhos
 escorregadios. (ESTAR).

Eu
 arroz e verdura todo dia. (COMER).

27) Vamos olhar bem para este mapa do Brasil.

Desenho do Mapa do Brasil com os limites estaduais, e com os traçados das rodovias:

TRANSAMAZÔNICA
 BELÉM-BRASÍLIA

RIO-BAHIA
 VIA DUTRA

Você será capaz de descobrir o que existe de comum entre os caminhos que marcamos no mapa?

Faça linha ligando a palavra CAMINHOS apenas aos que indicarem o que existe de comum.
 Veja como fizemos com

