

VALIDADE DE CONSTRUTO EM TESTES EDUCACIONAIS

Heraldo Marelim Vianna *

INTRODUÇÃO

O problema da validade de construto¹ é de grande relevância na área educacional, tendo em vista o fato de que a avaliação se vale, freqüentemente, de construtos, que, após sua definição operacional, são medidos por intermédio de testes. A *validade de conteúdo* e a *validade de critério* (concorrente e preditiva), atributos exigidos dos bons testes de escolaridade, não se preocupam, entretanto, com a compreensão dos construtos que os testes medem; conseqüentemente, impõe-se, conforme acentua Brown (1970), uma nova abordagem para análise dos instrumentos de medida da aprendizagem escolar.

A "validade" de muitos instrumentos de medida é, às vezes, inferida, mas tal procedimento é adotado à revelia da metodologia científica, não sendo, pois, um comportamento justificável. São necessárias provas insofismáveis de que um teste, construído para determinado fim, é efetivamente válido. Desse modo, se um teste visa a medir processos mentais complexos, como ocorre inúmeras vezes na área educacional, inclusive nos exames de acesso às universidades, é indispensável que existam provas irrefutáveis de que os instrumentos utilizados medem, efetivamente, o construto hipotetizado.

A validade de construto possibilita determinar qual a característica educacional que explica a variância do teste ou, então, qual o significado do teste (Kerlinger, 1973). Um avaliador, desse

* Do Departamento de Seleção de Recursos Humanos da Fundação Carlos Chagas.

¹ Construtos são traços, aptidões ou características supostamente existentes e abstraídos de uma variedade de comportamentos que tenham significado educacional (ou psicológico). Assim, fluência verbal, rendimento escolar, aptidão mecânica, inteligência, motivação, agressividade, entre outros, são construtos.

modo, poderá fazer perguntas do tipo: — o teste mede, efetivamente, a capacidade de expressão escrita? O teste é um reflexo do *status* sócio-econômico dos estudantes avaliados? A indagação, no caso, pretende esclarecer qual a proporção da variância total do teste que decorre desses construtos: — expressão escrita e *status* sócio-econômico. Ou, ainda, procura explicar as diferenças individuais nos escores desse instrumento de medida. O interesse, na validação de construtos, centraliza-se na característica ou traço que está sendo medido, mais do que no próprio teste.

A verificação da validade de construto, sua lógica e metodologia foram amplamente estudadas por Cronbach e Meehl (1955), que produziram documento fundamental para a compreensão desse tipo de validade, que é de particular importância sempre que um instrumento deva ser interpretado como proporcionando medidas de um atributo ou qualidade que, presumivelmente, as pessoas possuem. Outro documento básico para o estudo do problema é o ensaio de Campbell e Fiske (1959) sobre validade convergente e discriminante. Ambos os trabalhos são amplamente utilizados na fundamentação do presente estudo.

O CONCEITO DE VALIDADE DE CONSTRUTO

A validade de construto, ao contrário da validade empírica, não é expressa em termos de um coeficiente quantitativo, conforme ocorre no caso da validade preditiva e da validade concorrente. O conceito de validade de construto, por sua vez, é extremamente útil para explicar a natureza dos instrumentos que medem traços para os quais não se possuem critérios externos. Assim sendo, é necessário partir de uma variável logicamente definida. A variável, como um construto lógico, é inserida num sistema de conceitos, cujas relações são explicadas por uma teoria e a partir da qual certas conseqüências práticas, sob determinadas condições, podem ser extraídas e testadas (Magnusson, 1967). Se o resultado é o que se esperava em uma série de testes, o instrumento é considerado como possuindo validade de construto para a variável testada. Assim, a constatação da validade de construto resulta do acúmulo, por diferentes meios, de várias provas, que precisam ser analisadas em todos os seus detalhes, a fim de constatar, entre outros aspectos, quais as variáveis com as quais os escores do teste se correlacionam, quais os tipos de itens que integram o teste, o grau de estabilidade dos escores sob condições as mais variadas e o grau de homogeneidade do teste, com vistas a ter elementos que possam esclarecer o significado do instrumento.

A compreensão de um instrumento que mede determinado construto ganha nova dimensão quando se conhece o grau de relacionamento com outros construtos. Exemplificando, os escores de um teste especialmente construído para medir "rendimento escolar" adquire novo sentido quando se estabelece o seu grau de associação com outros construtos, como "inteligência" e "criatividade". A validade de construto possibilita determinar se o teste é a melhor medida de algo diferente do que foi pretendido medir ou, então, se fornece elementos que não possuem efetiva conseqüência no processo de avaliação.

AS TEORIAS EDUCACIONAIS E A LÓGICA DA VALIDADE DE CONSTRUTO

É comum em educação (e em psicologia) o desenvolvimento de sistemas unificados de princípios, definições, postulados e observações, para explicar o relacionamento entre variáveis (Sax, 1980), ou seja, o desenvolvimento de teorias educacionais, a partir das quais são construídos instrumentos para a mensuração de um determinado traço ou característica. Somente por intermédio da constatação da validade de construto desses instrumentos é que se pode confirmar o significado dessas características ou traços apresentados pela teoria. E para alcançar esse objetivo, necessário se faz a aplicação dos procedimentos clássicos do método dedutivo: teoria, dedução, hipótese, experimentação e, finalmente, dados que confirmem ou neguem a hipótese, ou seja, o construto.

A lógica da validade de construto, assim como o seu processo, são, essencialmente, os do método científico. Parte-se de uma teoria sobre a natureza do construto e fazem-se predições sobre as relações entre os escores do teste e outras variáveis. A seguir, essas predições são verificadas empiricamente e, dependendo dos resultados, a teoria é aceita ou revista. O trabalho contínuo de fazer predições, testar hipóteses, através da experimentação, e rever a teoria são atividades que contribuem para aumentar a precisão da definição do construto.

A REDE NOMOLÓGICA, SEGUNDO CRONBACH E MEEHL

As teorias que procuram explicar um determinado fenômeno compreendem um conjunto interrelacionado de conceitos, proposições e leis. A esse sistema interligado Cronbach e Meehl (1955) deram o nome de rede nomológica². As leis estabelecem relações entre diferentes características, entre características e construtos, ou entre um construto e outro, conforme aqueles autores, os quais ressaltam ainda que as leis e os conceitos devem estar ligados a comportamentos observáveis. Às vezes, entretanto, isso não ocorre. Uma determinada lei ou conceito não se relaciona diretamente com as características, mas com outras leis e conceitos, que, por sua vez, são diretamente ligados a características observáveis. É condição necessária, na pesquisa da validade de construto, que a definição de conceitos ou formulação de leis estejam apoiados, direta ou indiretamente, em dados observáveis.

Cronbach e Meehl (1955) assinalam, ainda, que o significado de um conceito somente fica perfeitamente esclarecido quando se elabora uma rede de relações que mostre, claramente, que essas relações são específicas e definitivas. Dessa forma, na validação de um construto, há necessidade de um grande número de diferentes operações, inclusive de ordem qualitativa, para a mensuração de um conceito e a fim de mostrar que esse mesmo conceito está ligado a outro conceito por intermédio de uma rede nomológica.

A partir desse posicionamento, é preciso considerar as seguintes situações:

- 1 – o processo de inferência do significado de um construto exige que os dados sejam observáveis;
- 2 – o processo de inferência do significado de um construto, a partir de dados observáveis, deve ser explicitamente especificado, para que se possa verificar a precisão de uma inferência;
- 3 – diferentes usuários de um construto devem servir-se, essencialmente, de uma mesma rede nomológica, para que a concordância entre os pesquisadores seja possível, pois, freqüentemente, o nome de um mesmo construto é usado com referência a diferentes construtos (ex: – inteligência) ou diferentes nomes são empregados com relação ao mesmo construto (ex: – pensamento divergente e criatividade), conforme assinalou Brown (1970).

VALIDAÇÃO DE TESTES E VALIDAÇÃO DE TEORIAS

A validação de construtos não se limita apenas a validar um teste, o seu alcance é bem mais amplo, centrando-se o seu objetivo na validação da teoria em que se apoiou a construção do instrumento (Kerlinger, 1973); desse modo, o trabalho de validação de um construto é uma pesquisa científica empírica, porque, definidos os construtos que seriam responsáveis pelo desempenho no teste, o avaliador passa a formular hipóteses sobre a teoria dos construtos e, a seguir, testa empiricamente essas hipóteses. A testagem de hipóteses sobre construtos exige a verificação da *convergência* e da *discriminatividade*. A *convergência* mostra que as evidências coletadas de diferentes fontes e de diferentes modos indicam um significado igual ou semelhante para o

² Nomológico, que é relativo à nomologia, estudo das leis que presidem aos fenômenos naturais.

construto. A *discriminatividade* refere-se à possibilidade de diferenciar, empiricamente, um construto de outros construtos semelhantes, assim como a de constatar o que não está correlacionado aos construtos (Kerlinger, 1973). As idéias de convergência e de discriminatividade serão retomadas mais adiante, quando for discutida a metodologia de Campbell e Fiske (1959) relativa à matriz do multitraço-multimétodo. A validação de construtos ultrapassa, assim, os limites de uma validação empírica, pois, além de constatar a correlação com um critério ou em que medida o instrumento separa indivíduos que possuem alto ou baixo grau de uma determinada característica, explica, também, o porquê dessas ocorrências.

O PROCESSO DE VALIDAÇÃO E TESTAGEM DE HIPÓTESES

O processo de validação exige, essencialmente, o estabelecimento de hipóteses, a partir de leis e construtos definidos pela rede nomológica, e a coleta de dados para testar essas hipóteses. Ao discutir os dados sobre a validade de um construto, é necessário que sejam claramente especificados os seguintes aspectos (Brown, 1970):

- 1 - a interpretação proposta, ou seja, qual o construto que se tem em mente, como esse construto é definido e como a hipótese testada foi estabelecida a partir de uma teoria importante;
- 2 - a comprovação adequada da interpretação, oferecendo elementos que apoiaram ou rejeitaram as hipóteses;
- 3 - a argumentação sobre a concretização dos objetivos propostos (apresentar detalhes sobre os procedimentos experimentais e a linha de raciocínio que permitiu inferências sobre o significado do construto).

O que significam, efetivamente, os resultados da comprovação das hipóteses levantadas? Se as predições forem confirmadas empiricamente, pode-se acreditar que o teste meça o construto e ter maior confiança no conceito adotado. Um construto nunca pode ser comprovado como correto em termos absolutos (Cronbach e Meehl, 1955), mas somente adotado como a melhor definição de trabalho. Se, ao contrário, os resultados forem negativos e os dados não confirmarem as predições, três interpretações são possíveis: 1º - o teste não mede o construto; 2º - o referencial teórico não é correto, possibilitando inferências errôneas; e 3º - o planejamento experimental não possibilita a testagem de hipóteses.

O fracasso na confirmação de uma predição indica a necessidade de uma revisão na rede teórica ou no procedimento experimental (Brown, 1970). As interpretações ambíguas de resultados negativos constituem uma das desvantagens da validade de construto.

A METODOLOGIA DA PESQUISA DA VALIDADE DE CONSTRUTO

A validade de construto pode ser pesquisada por diferentes métodos, inclusive os que são empregados na validade de conteúdo e de critério. Ao utilizar diferentes métodos é importante que se estude (Magnusson, 1967), entre outros aspectos:

- a - as diferenças entre os grupos em função do que a teoria estabelece relativamente à variável pesquisada;
- b - como os resultados dos testes são influenciados por mudanças nos indivíduos ou no meio, as quais, segundo a teoria, devem, respectivamente, influenciar ou deixar de influenciar as várias posições dos indivíduos no *continuum*;
- c - as correlações entre diferentes testes que, supostamente, medem a mesma variável. É necessário cautela a fim de que as correlações entre as medidas não se elevem em virtude de similaridades nos métodos utilizados (Cronbach e Meehl, 1955). Isso pode acontecer (Magnusson, 1967) se as respostas dos testes exigirem alguma aptidão especial além da que está sendo considerada. Uma possível concordância entre as medidas poderia, simplesmente, provocar um aumento da correlação do efeito de diferenças individuais relativas a essa aptidão especial;

d – as correlações entre itens isolados ou diferentes partes do teste, a fim de verificar se possuem alta intercorrelação e o teste possa ser considerado como medindo uma variável unitária.

A validade em geral e, especialmente, a validade de construto são estimadas pela concordância de medidas obtidas por métodos tão diferentes quanto possível. As dissimilaridades dos métodos, no estudo da validade de construto, são importantes para que as intercorrelações obtidas possam ser interpretadas como expressando realmente esse tipo de validade.

A VARIÂNCIA NO PROCESSO DE VALIDAÇÃO: – MÉTODOS E INDIVÍDUOS

Deve-se assinalar que a variação entre os indivíduos é expressa por escores que foram obtidos por um determinado método pré-definido, podendo esses escores ser afetados por diferenças, ainda que irrelevantes, nas reações dos indivíduos às características do método ou, então, por diferenças nas posições ao longo do *continuum* que se pretende que o teste meça, ou, ainda, por ambas as situações (Cronbach e Meehl, 1955; Magnusson, 1967). Assim, uma parte da variância total da distribuição dos escores pode ser atribuída a aspectos especificamente característicos do método empregado na mensuração, enquanto outra parte da variância pode ser considerada como resultado de diferenças realmente verdadeiras entre os indivíduos no que diz respeito ao traço mensurado. Esta variância expressa a *variância verdadeira* e é a que se deseja determinar com o máximo de precisão na pesquisa da validade. Desse modo, pode-se decompor a variância sistemática em:

a – variância devida às propriedades do método empregado,

e

b – variância decorrente de características relevantes dos indivíduos testados.

A variância resultante de propriedades do método empregado, conforme acentua Magnusson (1967), gera uma espécie de efeito de halo³ metodológico. Desse modo, quando as medidas de um certo número de variáveis são baseadas em um único método, os coeficientes da matriz de intercorrelação são grandes, em geral. A medida de diferentes traços é, portanto, afetada pelas propriedades do método empregado, as quais contribuirão para que resulte uma certa quantidade de variância comum.

É preciso atentar para o fato de que, quando dois métodos diferentes mas com propriedades semelhantes são independentemente empregados na medida de certo traço, uma parte da variância das medidas baseadas em um método pode repetir-se, sistematicamente, na medida baseada no outro método (Magnusson, 1967). A variância comum, resultante de semelhanças entre os métodos empregados, redundará numa superestimação da validade de construto, quando esta é verificada pela intercorrelação de escores obtidos a partir de diferentes métodos.

MÉTODOS USADOS NA VALIDAÇÃO DE CONSTRUTOS

Os métodos usados no estudo da validade de construto podem ser classificados, de acordo com Brown (1970), em cinco categorias: métodos intratestes, métodos entretestes, estudos relacionados a critérios, estudos experimentais e estudos de generalizabilidade⁴.

Métodos intratestes – esses métodos usam técnicas que permitem o estudo da estrutura

3 O efeito de halo é um efeito sistemático do avaliador que deve ser levado em conta quando traços humanos estão sendo avaliados. Uma atitude positiva ou negativa do avaliador em relação ao avaliado afeta, geralmente, na direção da atitude, as medidas de cada traço sujeito à avaliação. O efeito poderá provocar um decréscimo nas diferenças individuais e um aumento na correlação entre as medidas de diferentes traços.

4 *Generalizabilidade* – propriedade que têm as coisas de se tornarem generalizáveis.

interna do teste — seu conteúdo, as interrelações entre os itens e os subtestes e os processos relacionados com as respostas aos itens. Essas técnicas não consideram variáveis externas, porque se preocupam, antes de mais nada, com a estrutura interna do teste. Assim, não se pode usar essa metodologia para obter amplas informações sobre a validade de construto do teste, quando muito o seu emprego possibilitaria algum conhecimento sobre a natureza do construto, mas não o relacionamento do construto com outras variáveis.

A determinação da *validade de conteúdo* fornece informações sobre a validade de construto e é um tipo de estudo que pode ser incluído na categoria dos métodos intratestes. As especificações do conteúdo e do domínio comportamental “amostrado” no teste, condição essencial no estudo da validade de conteúdo, também servem para definir a natureza do construto que o teste mede.

A análise da *homogeneidade do teste*, método que também pode ser incluído na categoria dos intratestes, por intermédio de medidas de consistência interna (coeficientes de Kuder-Richardson), pela análise fatorial dos itens, entre outros estudos de homogeneidade, auxiliam na definição do construto, especialmente ao indicar se o teste mede um único traço ou se, ao contrário, mede diversos traços.

Ao analisar um teste, o interesse nem sempre se limita ao conhecimento do conteúdo dos itens, aprofunda-se e procura conhecer, também, o processo usado pelos examinandos nas suas respostas aos itens. Assim sendo, qualquer processo de análise que identifique capacidade e habilidades pode, em princípio, esclarecer o significado do construto que o teste mede, ao indicar as variáveis que estão sendo medidas pelos itens do instrumento. É preciso, entretanto, usar de cautela quando for empregado o processo de análise, pois, indiscutivelmente, diferentes pessoas poderão utilizar processos diversos, mas igualmente válidos, de resposta a um item, criando-se, desse modo, uma situação complexa que pode levar a falsas interpretações.

Métodos entretestes — os métodos incluídos nessa categoria consideram, simultaneamente, vários testes, mas não levam em consideração variáveis extratestes. Os métodos entretestes permitem indicar, geralmente, os aspectos comuns a vários testes, mas não possibilitam a realização de inferências diretas sobre a relação entre escores do teste e variáveis externas.

O método mais simples dessa categoria consiste em correlacionar um teste novo a um outro teste já amplamente estudado e conhecido nas suas diversas dimensões. É a chamada *validade congruente*. Se existe uma alta correlação entre os dois testes pode-se dizer que ambos medem o mesmo construto. Essa abordagem apresenta um aspecto que merece reflexão, pois a menos que os dois testes sejam altamente correlacionados (Brown, 1970), isto é, as correlações e as fidedignidades sejam da mesma magnitude, os fatores que influenciam no abaixamento da correlação podem ser importantes para determinar a relação entre o teste e a variável externa e, desse modo, invalidar as inferências realizadas com base na intercorrelação dos testes.

Uma outra abordagem, ainda nessa categoria, consiste em promover a *análise fatorial* em um grupo de testes. A análise fatorial é, no momento, o método mais promissor para estimativa da validade de construto (Kerlinger, 1970), pois objetiva reduzir um grande número de medidas a um número reduzido de fatores a fim de estabelecer quais os que medem as mesmas coisas e em que medida está, realmente, ocorrendo a mensuração que era esperada. A análise fatorial, ainda que seja um procedimento complexo, exigindo inclusive o emprego de computador, é um caminho fecundo na pesquisa de construtos, pois mostrará quais os testes que compartilham da variância comum e, assim, medem o mesmo construto. A pesquisa das cargas do mesmo fator (Brown, 1970), no conteúdo comum dos testes, possibilita inferir sobre a natureza do construto e até mesmo pode levar à identificação do fator. A análise fatorial permitirá verificar em que medida cada teste está saturado pela variância comum e em que medida os seus escores dependem da variância específica. A proporção da variância total dos escores do teste, que é variância comum, é um índice da validade de construto. O presente método visa a estabelecer a *validade fatorial*, denominação esta freqüentemente usada como sinônimo de validade de construto.

Outra abordagem possível nessa categoria refere-se às concepções de Campbell e Fiske

(1959) para estabelecimento da *validade convergente* e da *validade discriminante*. A validação, segundo estes autores, processar-se-ia por meio de métodos que visariam a estabelecer se duas técnicas de medida (ou testes) diferentes estariam medindo, efetivamente, o mesmo construto; daí serem esses métodos chamados de convergentes. A *validade congruente* e a *validade fatorial* são exemplos típicos do uso de métodos convergentes.

Um dos problemas centrais da análise da validade de um construto consiste em que os testes, além de apresentarem uma alta correlação com outros testes que medem o mesmo construto, devem, também, demonstrar que não apresentam correlação com testes que medem claramente construtos diferentes. Sabe-se, por exemplo, que há uma associação entre inteligência e criatividade; dessa forma, segundo o ponto de vista de Campbell e Fiske (1955), somente podemos aceitar um teste de criatividade se o mesmo não apresentar qualquer correlação com os resultados de testes de inteligência, quando se terá certeza de que aquele construto está sendo medido e não se confunde com este outro. Esse é, sem dúvida, o ponto discriminante na validação de construto, pois, conforme acentuou Brown (1970), um teste somente é válido para medir um determinado construto quando, sem sombra de dúvida, é independente de testes que medem outros construtos perfeitamente definidos.

Estudos relacionados a critérios — A natureza e o tipo de critérios que os escores de um teste predizem são indicativos do construto que o teste mede; isso posto, é perfeitamente possível obter importantes informações para o estabelecimento da validade de construto a partir dos dados de estudos sobre validade de critério.

Uma das maneiras de coletar evidências sobre um construto é a partir de escores de um teste que seja capaz de separar grupos naturalmente existentes de grupos organizados experimentalmente. Os escores para esse fim devem estabelecer a diferenciação entre grupos. Suponhamos, a título de exemplificação, que foi construído um teste para identificar criatividade literária. Os itens desse teste devem ter sido elaborados de forma a identificar dois grupos bem distintos: — os que são capazes de produzir textos de valor artístico nos vários setores da literatura e aqueles que apenas possuem o domínio da língua, como qualquer pessoa comum. O instrumento assim construído teria validade de critério (concorrente), o que representa importante informação para a caracterização do construto.

Uma outra abordagem, ainda segundo essa perspectiva, seria a formação de dois grupos distintos com base nos escores do teste (quartil superior e inferior, por exemplo) e o estabelecimento das características de cada um desses grupos, o que permitiria estabelecer uma definição, tão completa quanto possível, sobre a natureza do construto.

Destaca-se, nessa categoria, o método baseado no emprego de *coeficientes de validade*, o qual consiste em verificar o êxito de um instrumento na predição de um determinado comportamento. Um teste de aptidão escolar deve ser um bom preditor do desempenho acadêmico, pois há uma clara associação entre essas duas variáveis. Ora, na medida em que isso ocorre, o teste critério estaria medindo se o construto é realmente aptidão escolar.

Estudos experimentais — Outros métodos usados para determinação da validade de construto exigem a manipulação de algumas variáveis e a observação dos efeitos consequentes nos escores do teste. Tomemos o exemplo da ansiedade, conforme a colocação de Brown (1970). Os estudantes, em época de exames, costumam demonstrar um aumento na sua ansiedade. A ansiedade durante a realização de um teste refletiria medo de fracasso no exame e o comprometimento do auto-conceito da pessoa. A partir dessa definição de ansiedade, pode-se estabelecer a hipótese de que o desempenho no exame é inversamente correlacionado à ansiedade durante a realização de um teste, se o teste for de grande importância para a pessoa. Igualmente, pode-se hipotetizar que ansiedade e exame não se correlacionam, se o teste for estruturado de forma a não constituir uma ameaça ao indivíduo. Se, nesse contexto, os escores num teste de ansiedade não apresentarem relação de predição com os escores de um exame, ter-se-ia, então, evidência de que o teste mede realmente o construto ansiedade.

Um conceito de grande importância em psicometria — o de fidedignidade — pode ser usado

no estudo da validade de construto (Cronbach e Meehl, 1955). Se o construto estabelece que o traço a ser medido é grandemente estável ao longo do tempo e resiste a mudanças, o coeficiente de estabilidade do teste será um indicador do construto que foi hipotetizado. Se o construto estabelece que os escores, em certas circunstâncias, aumentam com a idade, na medida em que isso é constatado, tem-se uma prova da validade de construto de instrumento.

Estudos de generalizabilidade: método do multitraço-multimétodo — Os estudos de generalizabilidade são aqueles em que o teste, cujo construto se deseja validar, é analisado sob diferentes condições, como, por exemplo, a aplicação do instrumento a diferentes amostras da população, a utilização de métodos diversos na sua aplicação etc. A abordagem mais frequentemente utilizada para estudos desse tipo é o da matriz do *multitraço-multimétodo*, conforme a proposta de Campbell e Fiske (1959). Esses autores estabeleceram algumas condições fundamentais para o sucesso de um processo de validação. Essas condições não se restringem apenas à validade de construto, incluindo, também, a verificação da validade preditiva e concorrente.

Um teste é uma unidade traço-método (Campbell e Fiske, 1959), ou seja, um teste mede determinado traço usando o único método. O interesse, portanto, no processo de validação, centra-se no conhecimento das contribuições relativas do traço e do método para o escore do teste, o que obriga a estudar mais de um traço e mais de um método. Isso significa que se deseja, na verdade, estabelecer a *validade convergente* e a *validade discriminante*.

A *validade convergente* será determinada por intermédio da correlação entre os *mesmos traços* medidos por *diferentes métodos*, esperando-se que essa correlação seja significativamente alta. A *validade discriminante*, por sua vez, será estabelecida comprovando-se que *diferentes traços*, mesmo quando medidos pelo *mesmo método*, não possuem uma alta correlação.

A Tabela 1.0 apresenta uma matriz relativa ao emprego do *multitraço-multimétodo*, em que temos três traços hipotéticos (A, B, C) medidos por três métodos diferentes (1, 2, 3), que geram nove variáveis separadas. Ou mais claramente, a título de exemplificação, os traços seriam: — compreensão de textos, raciocínio abstrato e capacidade de identificar elementos secundários numa informação, traços esses que seriam medidos por um teste de papel-e-lápis, um teste individual e uma escala de classificação. O método pode ser usado para o estudo de n traços, usando m métodos, não havendo necessidade de $n = m$. O número de traços é igual ao número de métodos, no exemplo citado, apenas por conveniência na discussão da metodologia.

Os coeficientes de correlação apresentados na Tabela 1.0 representam o grau de associação entre três *variáveis* (A, B, C), medidas por três *métodos diferentes* (1, 2, 3). Os escores para cada uma das variáveis são correlacionados com os escores de cada uma das outras variáveis, independentemente do método pelo qual os escores foram obtidos. Os valores na diagonal da matriz completa — 0,89; 0,89; 0,76;; 0,94; 0,92; 0,85 — representam as *fidedignidades das medidas*, valores esses que representam os resultados da medida do *mesmo traço* pelo *mesmo método*: portanto, são os valores *monotraço-monométodo*.

Os triângulos representados por linhas cheias, ao longo da diagonal da matriz completa, contêm coeficientes que dão a relação entre as medidas de *diferentes variáveis* pelo *mesmo método*, são, pois, valores *heterotraço-monométodo*. Considerando-se que o mesmo método foi usado para a medida de diferentes variáveis, as propriedades do método dão origem a uma variância comum para as diferentes variáveis, na medida em que as propriedades do método concorrem para a variância sistemática quando variáveis individuais são mensuradas.

Os triângulos em linhas interrompidas contêm coeficientes de correlação entre medidas de *diferentes variáveis* obtidas por *métodos diferentes*; dessa forma, são valores *heterotraço-heterométodo*. As diagonais entre esses triângulos apresentam os coeficientes de correlação entre as medidas da *mesma variável* por *diferentes métodos* e são *coeficientes de validade* (validade convergente). Esses valores em diagonal devem ser substanciais, porque refletem a correlação entre as mesmas variáveis medidas diferentemente. A variância comum que resulta, em virtude das semelhanças nos métodos, afetará o tamanho desses coeficientes na medida em que os

métodos têm propriedades iguais e lhes são oferecidas oportunidades de afetar as medidas de maneira sistemática.

TABELA 1.0.

Matriz multitraço-multimétodo, segundo Campbell e Fiske (1959)

Traço	Método 1			Método 2			Método 3		
	A ₁	B ₁	C ₁	A ₂	B ₂	C ₂	A ₃	B ₃	C ₃
Método 1	A ₁ 0,89								
	B ₁ 0,51	0,89							
	C ₁ 0,38	0,37	0,76						
Método 2	A ₂ 0,57	0,22	0,09	A ₂ 0,93					
	B ₂ 0,22	0,57	0,10	B ₂ 0,68	0,94				
	C ₂ 0,11	0,11	0,46	C ₂ 0,59	0,58	0,84			
Método 3	A ₃ 0,56	0,22	0,11	A ₃ 0,67	0,42	0,33	A ₃ 0,94		
	B ₃ 0,23	0,58	0,12	B ₃ 0,43	0,66	0,34	B ₃ 0,67	0,92	
	C ₃ 0,11	0,11	0,45	C ₃ 0,34	0,32	0,58	C ₃ 0,58	0,60	0,85

O MÉTODO DO MULTITRACO-MULTIMÉTODO: - CONDIÇÕES PARA APLICAÇÃO

Campbell e Fiske (1959) estabeleceram algumas condições para, usando o método do multitraço-multimétodo, desenvolver o processo de validação:

1º - os coeficientes de correlação entre medidas da *mesma variável* com *diferentes métodos* (coeficiente de validade convergente) devem ser significativamente maiores do que zero. (Este critério é, normalmente, considerado suficiente para caracterizar a validade.);

2º — as medidas de uma variável devem apresentar uma correlação mais estreita com medidas do *mesmo tipo*, e que foram obtidas por um *outro método*, do que com medidas de *outro tipo* que foram estabelecidas pelo *mesmo método*. Os coeficientes de validade para uma certa variável devem, assim, ser maiores do que os coeficientes para a mesma variável nos triângulos delimitados por linhas contínuas;

3º — o coeficiente de validade para uma determinada variável deve ser maior do que a correlação entre as medidas dessa variável e as medidas de todas as outras variáveis, obtidas por qualquer outro método. Um coeficiente de validade, desse modo, será maior do que os correspondentes coeficientes na mesma linha e coluna no interior do triângulo representado por linhas interrompidas;

4º — se os mesmos métodos ou métodos diferentes forem usados, as magnitudes dos coeficientes para as correlações entre diferentes variáveis devem ter a mesma configuração.

Se a primeira condição for satisfeita, os métodos possuem *validade convergente*: — concordância significativa entre medidas do mesmo tipo com diferentes métodos. A concretização dessa condição não é suficiente para satisfazer o processo de validação. É necessário que a segunda e a terceira condição também ocorram. Se ambas ocorrerem, dir-se-á que as medidas têm *validade discriminante*.

É preciso, no caso das condições estabelecidas por Campbell e Fiske (1959), considerar a fidedignidade dos métodos, pois se ocorrer a falta desse atributo, a validade discriminante será afetada. A fidedignidade, na presente abordagem, refere-se à concordância entre duas medidas do *mesmo traço* usando o *mesmo método* enquanto que a *validade* é definida em termos da concordância entre duas medidas do *mesmo traço* usando *diferentes métodos*. Assim sendo, cumpre ressaltar que a diferença fundamental entre fidedignidade e validade está na similaridade dos métodos de medida. A quarta condição apresentada por Campbell e Fiske (1959) é irrealista (Magnusson, 1967) e impossível de ser obedecida rigorosamente. Se fosse seguida, a validade discriminante dificilmente seria estabelecida, em face da dificuldade de julgar o efeito da falta de fidedignidade em matrizes complexas, com inúmeras variáveis, como exigem os estudos de validação de construtos.

REFERÊNCIAS BIBLIOGRÁFICAS

- BROWN, F.G. (1970) *Principles of Educational and Psychological Testing*. The Dryden Press Inc. Hinsdale, Illinois.
- CAMPBELL, D. T. e FISKE, D. W. (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 59.
- CRONBACH, L. J. e MEEHL, P. E. (1955) Construct validity and psychological tests. *Psychological Bulletin*, 52.
- KERLINGER, F.M (1973) *Foundations of Behavioral Research*. Second edition. Holt, Rinehart and Winston, Inc. Nova Iorque.
- MAGNUSSON, D. (1967) *Test Theory* — Addison-Wesley Publishing Company. Reading, Massachusetts.
- SAX, G. (1980) *Principles of educational and psychological measurement and evaluation*. 2 ed. Wadsworth Publishing Company. Belmont, California.