

6 Modelo de Equações Estruturais

6.1. Introdução

No capítulo 5 utilizamos a TRIN para construir as escalas dos construtos capital social, capital cultural e capital econômico. Neste capítulo apresentamos uma outra metodologia para a construção das escalas, a Análise Fatorial Confirmatória (AFC). Esta metodologia produz uma escala contínua e permite testar a hipótese de que, de fato, existem três variáveis ou fatores latentes (capital social, capital cultural e capital econômico) para aquele conjunto de itens. Estes itens ou variáveis indicadoras são os mesmos utilizados no capítulo 5.

Nas próximas seções serão desenvolvidos os conceitos teóricos da análise de fatores confirmatória e as diversas fases de modelagem. No final apresentamos os resultados.

6.2. Modelo Estrutural Generalizado

O modelo de equações estruturais (SEM, do inglês *Structural Equation Modeling*) engloba uma família de modelos conhecida por diversos nomes, entre eles a análise de caminhos, a análise de estrutura de covariância, a análise de variáveis latentes e **Análise Fatorial Confirmatória**. É uma técnica de amplo uso e com muitas possibilidades. Aqui estamos interessados somente na análise fatorial. As técnicas SEM se distinguem por duas características: i) estimação de múltiplas e inter-relacionadas relações de dependência, ou seja, a estimação de uma série de equações de regressão múltipla interdependentes de modo simultâneo, pela especificação do modelo estrutural, e, ii) a possibilidade de considerar variáveis observadas com erro de medição e variáveis não observadas,

ou variáveis latentes: variáveis teóricas que só podem ser estimadas indiretamente através de variáveis observadas¹¹ (Hair J. et al., 2000).

O modelo de equações estruturais está composto por dois sub modelos: o modelo de medição e o modelo estrutural¹².

O modelo estrutural define a relação entre as variáveis latentes exógenas e endógenas. Conseqüentemente, este modelo especifica quais variáveis latentes (exógenas) influenciam direta ou indiretamente mudanças nos valores da outra variável latente (endógena ou dependente).

O modelo estrutural está representado assim:

$$\eta_{mx1} = B_{mxm} \eta_{nx1} + \Gamma_{mxm} \xi_{nx1} + \zeta_{mx1} \quad (6.1)$$

Onde:

- O vetor η representa as variáveis latentes endógenas e é de ordem $mx1$.
- Γ é uma matriz de coeficientes de mxn que relaciona os n fatores exógenos com os m fatores endógenos.
- B é uma matriz de coeficientes de mxm que relaciona os m fatores endógenos um com outro.
- O vetor ξ é de ordem $nx1$ representa as n variáveis exógenas latentes.
- ζ é um vetor de resíduos de $mx1$, representa os erros na equação que relaciona η e ξ . Assume-se que ζ não está correlacionado com ξ . O modelo assume que a matriz $(I-B)$ é não singular.

Por sua parte, o modelo de medição descreve as relações entre as variáveis latentes e seus indicadores observados. Para cada construto que aparece no modelo é necessário determinar quais são suas variáveis indicadoras. O objetivo fundamental deste modelo é corroborar a idoneidade dos indicadores selecionados na medição dos construtos de interesse. O modelo de medição pode ser especificado em termos das variáveis exógenas (variáveis X) ou em termos das variáveis endógenas (variáveis Y).

¹¹ Também conhecidos como variáveis indicadoras ou variáveis manifestas.

$$X_{qx1} = \Lambda_{xqxn} \xi_{nx1} + \delta_{qx1} \quad (6.2)$$

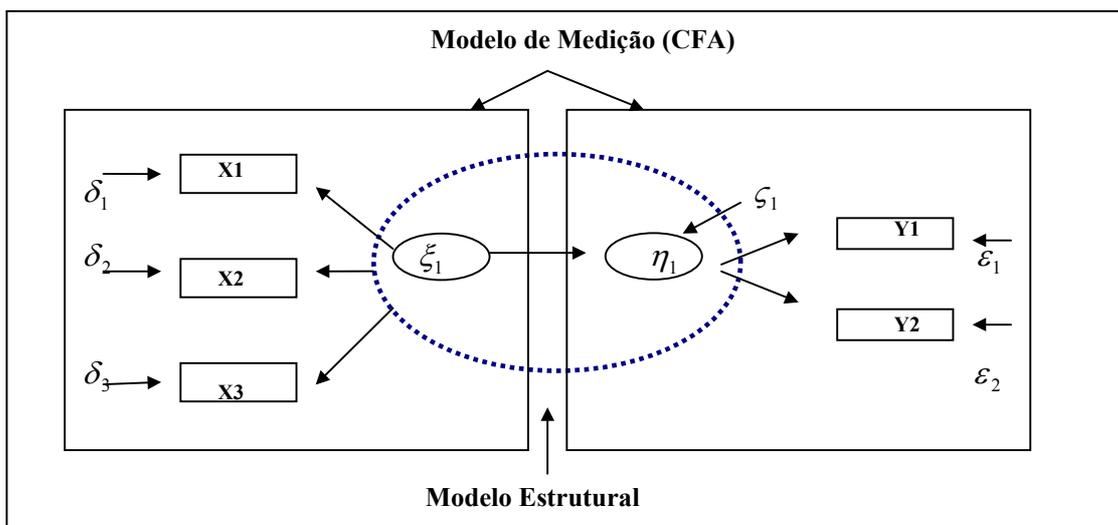
$$Y_{px1} = \Lambda_{ypxm} \eta_{mx1} + \varepsilon_{px1} \quad (6.3)$$

Onde:

- X é um vetor de $qx1$ de variáveis exógenas observáveis
- Y é um vetor de $px1$ de variáveis endógenas observáveis.
- ξ é um vetor de $nx1$ de variáveis latentes exógenas
- η é um vetor de $mx1$ de variáveis latentes endógenas
- δ é um vetor de $qx1$ de erros de medição em x .
- ε é um vetor de $px1$ de erros de medição em y
- Λ_x é uma matriz qxn dos coeficientes de regressão que relaciona as n fatores exógenos com cada um das q variáveis observáveis designadas para medi-las.
- Λ_y é uma matriz pxm dos coeficientes de regressão que relaciona as m fatores endógenos com cada um das p variáveis observáveis designadas para medi-las.

No quadro 6.1 a seguir mostramos um diagrama para a melhor compreensão do SEM.

Quadro 6.1- Representação esquemática do Modelo de Equações Estruturais (SEM).



¹² As definições teóricas a serem desenvolvidas estão baseados nos livros de Kenneth (1989) e

6.3. Análise de caminhos

O diagrama de caminhos é uma representação gráfica de um modelo de equações estruturais (SEM). Uma das principais vantagens desta representação é que as relações entre as variáveis podem ser visualizadas graficamente. Para entender melhor o diagrama de caminhos é necessário definir os símbolos utilizados, que são apresentados no quadro 6.2 .

Quadro 6.2- Símbolos utilizados no diagrama de caminhos.

Símbolo	Representação
	<ul style="list-style-type: none"> • Variável latente
	<ul style="list-style-type: none"> • Variável observável
	<ul style="list-style-type: none"> • Coeficiente da regressão da variável observável sobre a variável latente.
	<ul style="list-style-type: none"> • Coeficiente da regressão da variável latente sobre outra variável latente.
	<ul style="list-style-type: none"> • Distúrbio na predição da variável latente
	<ul style="list-style-type: none"> • Erro de medição associado com a variável observável
	<ul style="list-style-type: none"> • Relação direcional
	<ul style="list-style-type: none"> • Relação não direcional
	<ul style="list-style-type: none"> • Associação não analisada entre duas variáveis.

6.4. Análise Fatorial Confirmatória

Spearman (1904) tem o crédito de ter criado a técnica conhecida como análise de fatores. No seu artigo pioneiro ele usou a análise de fatores para determinar o fator de inteligência geral de um indivíduo. Este fator único (o fator G de Spearman) está fundamentado no resultado do indivíduo em questões de um teste.

A Análise Fatorial Exploratória (AFE) e a Análise Fatorial Confirmatória (AFC) são dois métodos distintos de análise de fatores. As diferenças entre os dois são que na análise exploratória não se assume um modelo a priori, não se impõe o número de fatores latentes, e tipicamente todas as variáveis latentes influenciam todas as variáveis observadas. Também, não são permitidos que os erros de medição estejam correlacionados e é comum a subidentificação dos parâmetros. Em contraste, a Análise Confirmatória é construída a priori, sendo o número de variáveis latentes definido pelo pesquisador; se uma variável latente influencia alguma variável observada, esta relação é especificada. Alguns efeitos diretos da variável latente sobre a variável observada são fixados em zero ou, às vezes, em uma constante. Os erros de medição podem ser correlacionados. As covariâncias das variáveis latentes podem ser estimadas e, para isso, é preciso que os parâmetros sejam identificáveis.

6.4.1. Variância implicada pelo modelo

A hipótese fundamental de um modelo de equações estruturais é que a matriz de covariância das variáveis observadas é função dos parâmetros do modelo. Se o modelo fosse correto e se conhecêssemos os parâmetros, a matriz de covariância populacional poderia ser reproduzida exatamente. A eq. (6.4) mostra a formalização desta hipótese:

$$\Sigma = \Sigma(\theta), \quad (6.4)$$

onde Σ é a matriz de covariância populacional das variáveis observadas (X 's), θ é o vetor que contém os parâmetros do modelo, e $\Sigma(\theta)$ é a matriz implicada pelo modelo.

O modelo geral da Análise Confirmatória é dado por uma das equações que definem o modelo de medição (eq. 6.2 e eq. 6.3), por exemplo:

$$X_{qx1} = \Lambda_{xqxn} \xi_{nx1} + \delta_{qx1}. \quad (6.5)$$

As variáveis observadas dependem da variável latente e do erro de medição. Os erros de medição são não correlacionados com as variáveis latentes, $E(\xi\delta') = 0$, e $E(\delta) = 0$. Os coeficientes (Λ_y, Λ_x) descrevem o efeito da variável latente sobre as variáveis observadas.

Na AFC a matriz implicada pelo modelo é função da matriz de covariâncias entre os fatores latentes, Φ ; da matriz de coeficientes, Λ_x e da matriz de erros de medição, Θ_δ . Assim, a variância implicada pelo modelo é:

$$\Sigma(\theta) = \Lambda_x \Phi \Lambda_x' + \Theta_\delta \quad (6.6)$$

A estimação dos parâmetros se baseia na seleção dos valores dos parâmetros estruturais que reproduzem a matriz de covariância. Assim, é importante entender a relação entre os elementos da covariância e os parâmetros.

Por exemplo, para simplificar, considere-se um modelo com um fator e duas variáveis indicadoras ou observadas :

$$x_1 = \lambda_{11}\xi_1 + \delta_1 \quad (6.7)$$

$$x_2 = \lambda_{21}\xi_1 + \delta_2 \quad (6.8)$$

Seja a matriz de covariância das variáveis observadas:

$$\Sigma = \begin{bmatrix} \text{var}(x_1) & \text{cov}(x_1, x_2) \\ \text{cov}(x_2, x_1) & \text{var}(x_2) \end{bmatrix} \quad (6.9)$$

A matriz de covariâncias implicada pelo modelo pode ser expressa como:

$$\begin{aligned} \Sigma(\theta) &= E(XX') = E[(\Lambda_x\xi + \delta)(\xi'\Lambda_x + \delta')] \\ &= \begin{bmatrix} \lambda_{11}^2 \text{var}(\xi_1) + \text{var}(\delta_1) & \lambda_{21}\lambda_{11} \text{var}(\xi_1) \\ \lambda_{21}\lambda_{11} \text{var}(\xi_1) & \lambda_{21}^2 \text{var}(\xi_1) + \text{var}(\delta_2) \end{bmatrix} \end{aligned} \quad (6.10)$$

Se, para cada elemento da eq. (6.9) e eq. (6.10) se verifica a identidade, então as variâncias e covariâncias observadas são decompostas em seus parâmetros estruturais. Por exemplo a $Var(x_1)$ é dada por:

$$Var(x_1) = \lambda_{11}^2 \phi_{11} + var(\delta_1) . \quad (6.11)$$

A eq. (6.11) nos diz que a variância de x_1 mudará se λ_{11} , ϕ_{11} ou $var(\delta_1)$ tomam valores diferentes.

Por outra parte, existem vários métodos para estimar a função de ajuste, $F(S, \Sigma(\theta))$. Entre os mais comuns estão: máxima verossimilhança, mínimos quadrados não ponderados e mínimos quadrados generalizados. Dentre as inúmeras funções de ajuste $F(S, \Sigma(\theta))$, deve-se escolher aquelas que produzem estimadores de $\Sigma(\theta)$ que possuam propriedades estatísticas desejáveis (consistência, ausência de viés e eficiência).

6.4.2. Avaliação do ajuste do modelo

A avaliação do ajuste do modelo pode ser feita em duas partes: o exame do ajuste local e o exame do ajuste global. A avaliação do ajuste local é feita através da estatística de teste “t”. Deste modo pode ser verificado se os parâmetros são significativamente diferentes de zero. Informação sobre o ajuste local também pode ser obtida, analisando o coeficiente de determinação¹³ R^2 e os resíduos. A avaliação do ajuste global do modelo, pode ser feita pelo teste de ajuste exato, usando a estatística de teste qui-quadrado, e pelos teste de ajuste aproximado usando a estatística $RMSEA$, a raiz do erro quadrático médio. Outros critérios de ajuste aproximado estão disponíveis como o índice de qualidade de ajuste (GFI), o índice de ajuste comparativo (CFI), e o índice de ajuste normalizado (NFI).

¹³ O R^2 mede a porcentagem de variação de certa variável observada explicada por certo fator latente.

A estatística qui-quadrado é a razão de verossimilhança que se utiliza para avaliar a significância estatística de que todos os elementos da matriz de resíduos $(S - \Sigma(\theta))$ ¹⁴ sejam nulos. Esta estatística é igual a $(N - 1)F_{\min}$ ¹⁵.

A raiz do erro quadrático médio (RMSEA) leva em consideração o erro de aproximação na população, e faz a seguinte pergunta “ *quão bem o modelo se ajustaria à matriz de covariâncias populacional se esta estivesse disponível (com parâmetros desconhecidos, mas com a escolha ótima dos parâmetros)?*”. Esta discrepância ou erro é medida pelo RMSEA. Browne & Cudeck (1993) sugerem que o valor de RMSEA de 0,05 ou menor indicam um ajuste muito bom, e valores de RMSEA de 0,08 ou inferiores indicam um bom ajuste do modelo.

O índice de qualidade de ajuste (GFI) também é outra medida da discrepância entre a matriz de covariância amostral (S) e a matriz de covariância obtida do modelo ($\Sigma(\theta)$). O AGFI se obtém ajustando o GFI pelos graus de liberdade. Este indicador toma um valor máximo de 1, valores maiores ou iguais a 0,80 são, geralmente, considerados aceitáveis.

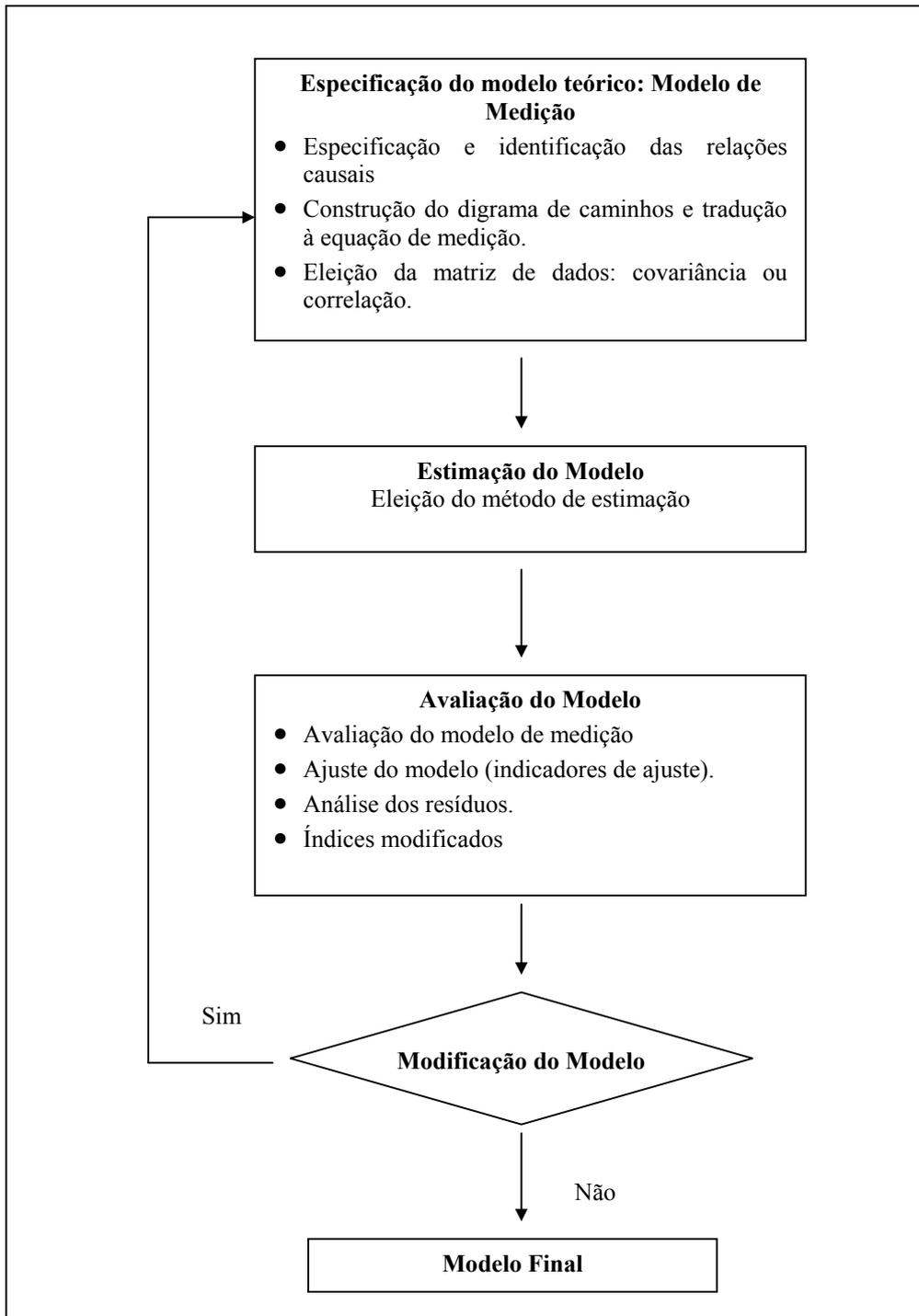
Adicionalmente, tem-se ainda duas medidas de ajuste aproximado: i) o índice de ajuste normalizado (NFI) e ii) o índice de ajuste comparativo (CFI), que avaliam o ajuste do modelo comparado com um modelo nulo. Nesse caso o modelo nulo é teorizado como um modelo com apenas um fator e sem erro de medição. Os valores em que oscilam NFI e CFI são de zero a um, valores acima de 0,90 indicam um ajuste aceitável.

6.5. Modelagem

O processo de estimação do modelo consiste de quatro etapas: i) especificação, ii) estimação, iii) avaliação e, às vezes, uma quarta, iv) modificação do modelo. No quadro 5.6, apresentamos as quatro etapas deste procedimento.

¹⁴ Ho: $S - \Sigma(\theta) = 0$

Quadro 6.3- Algoritmo do procedimento de estimação.



¹⁵ F_{\min} é a função de ajuste.

6.6. Resultados

Estimaram-se dois modelos de medição para os três construtos: capital social, capital cultural e capital econômico. Para representar estas três variáveis latentes, utilizaram-se 28 itens no primeiro modelo e 18 no segundo modelo. A escolha das variáveis indicadoras para cada modelo foi baseada nas escalas construídas (com os dois procedimentos) com o modelo não paramétrico da TRI, desenvolvido no capítulo 5. No primeiro modelo se excluíram itens com covariância negativa e H_i muito baixo ($< 0,20$) e no segundo modelo, que seria um sub-modelo deste, somente se considerou itens com poder de discriminação razoável ou alto ($H_i > 0,36$)¹⁶. Na tabela 6.1 mostramos os indicadores das variáveis latentes.

Tabela 6.1- Construtos e as variáveis indicadoras.

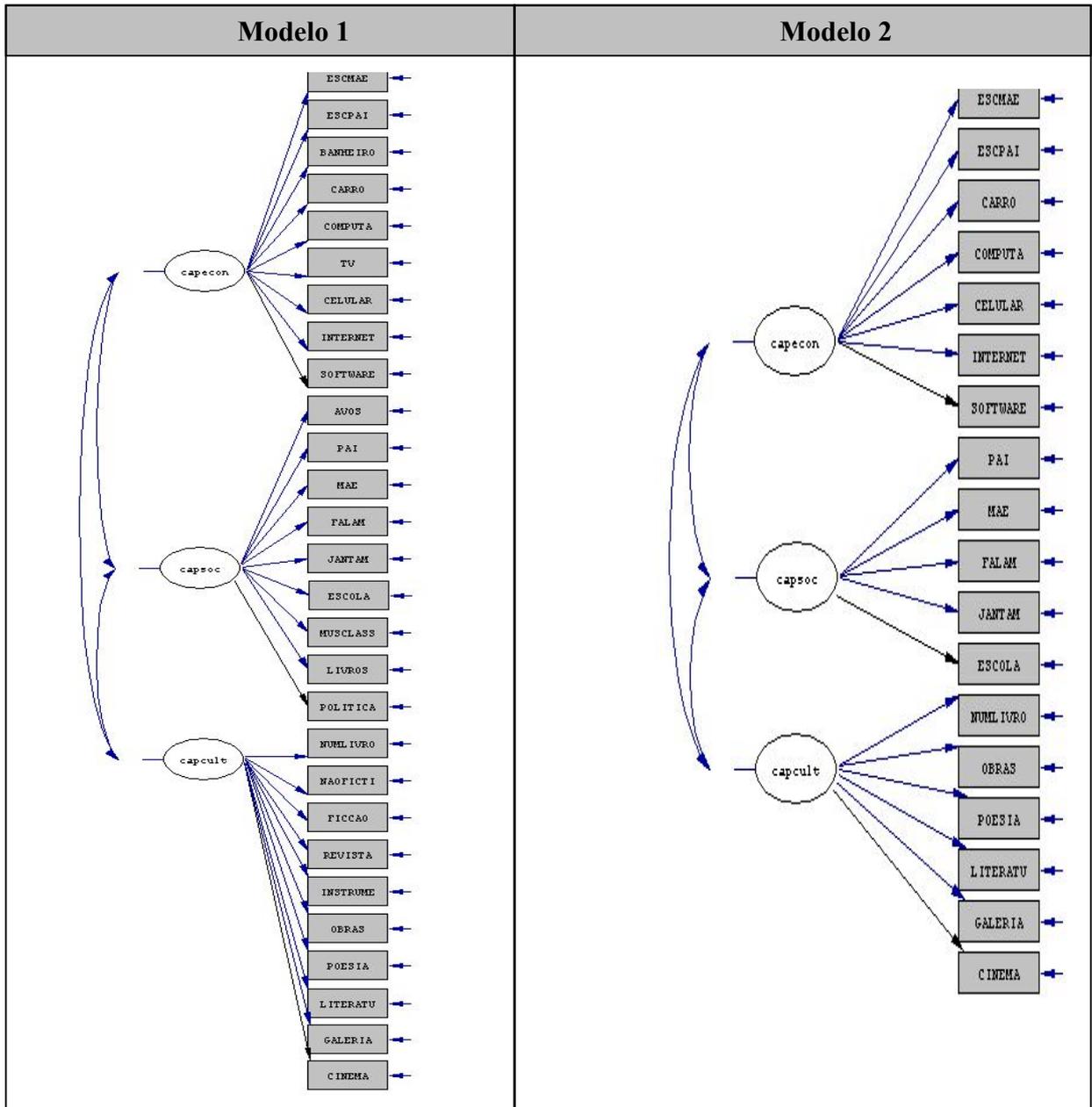
Modelo 1 (28 variáveis)	Modelo 2 (18 variáveis)
<p>Itens: capital cultural (10 variáveis)</p> <ol style="list-style-type: none"> 1. Foi para o cinema (cinema) 2. Vai para museus, galerias de arte (galeria) 3. Livros de literatura clássica (literatu) 4. Livro de poesia (poesia) 5. Obras de arte (obras) 6. Instrumento musical (instrume) 7. Lê revistas (revista) 8. Lê livros de ficção (ficcao) 9. Lê livros que não são de ficção (naofic) 10. Número de livros (numlivros) <p>Itens: capital social (9 variáveis)</p> <ol style="list-style-type: none"> 1. Falam de política (politica) 2. Falam de livros (livros) 3. Ouvem música clássica (musclass) 4. Falam sobre como esta indo à escola (escola) 5. Jantam juntos (jantam) 6. Falam de qualquer tema (falam) 7. Sua mãe ajuda a fazer o dever de casa (mae) 8. Seu pai ajuda a fazer o dever de casa (pai) 9. Seus avôs ajudam a fazer o dever de casa (avos) <p>Itens: capital econômico (9 variáveis)</p> <ol style="list-style-type: none"> 1. <i>Software</i> educativo (software) 2. Conexão a internet (internet) 3. Celular (celular) 4. Televisor (tv) 5. Computador (computa) 6. Carro (carro) 7. Banheiro (banheiro) 8. Escolaridade do pai (escpai) 9. Escolaridade da mãe (escmae) 	<p>Itens: capital cultural (6 variáveis)</p> <ol style="list-style-type: none"> 1. Vai para o cinema (cinema) 2. Vai para museus, galerias de arte (galeria) 3. Livros de literatura clássica (literatu) 4. Livro de poesia (poesia) 5. Obras de arte (obras) 6. Número de livros (numlivros) <p>Itens: capital social (5 variáveis)</p> <ol style="list-style-type: none"> 1. Falam como vai na escola (escola) 2. Jantam juntos (jantam) 3. Falam de qualquer tema (falam) 4. Sua mãe ajuda a fazer o dever de casa (mae) 5. Seu pai ajuda a fazer o dever de casa (pai) <p>Itens: capital econômico (7 variáveis)</p> <ol style="list-style-type: none"> 1. Software educativo (software) 2. Conexão a internet (internet) 3. Celulares (celular) 4. Computador (computa) 5. Carro (carro) 6. Escolaridade do pai (escpai) 7. Escolaridade da mãe (escmae)

¹⁶ Com exceção dos itens do capital cultural, os quais têm coeficientes H_i entre 0.27 e 0.34.

- **Especificação do modelo**

Em ambos os modelos as variáveis indicadoras estão associadas unicamente a um fator. Os fatores latentes estão correlacionadas entre si, e assume-se, também que os erros de medição não estão correlacionados. Estas relações são expressas graficamente através do diagrama de caminhos mostrado no figura 6.2.

Figura 6.1- Modelo 1 e modelo 2 hipotéticos.



- **Estimação do modelo**

Os dois modelos foram estimados no programa LISREL 8.53 de Jöreskog & Sörbom (2002) através do método mínimos quadrados ponderados¹⁷. Este método usa a matriz de correlação tetracórica e/ou policórica¹⁸ e adicionalmente requer a estimação da matriz de covariância assintótica das correlações amostrais.

- **Modificação e avaliação do modelo**

Os dois modelos propostos inicialmente foram reespecificados, o que permitiu a obtenção de um melhor ajuste. Estas modificações se deram nas correlações entre os erros de algumas variáveis. No modelo 1 estabeleceu-se as correlações entre os erros de medição das variáveis:

- “ficção” e “revista”,
- “revista” e “livros que não são de ficção”,
- “ficção” e “livros que não são de ficção”,
- “escolaridade do pai” e “escmae”,
- “pai” e “mãe”, “escolaridade da mãe” e “mãe”,
- “escolaridade do pai” e “pai”.

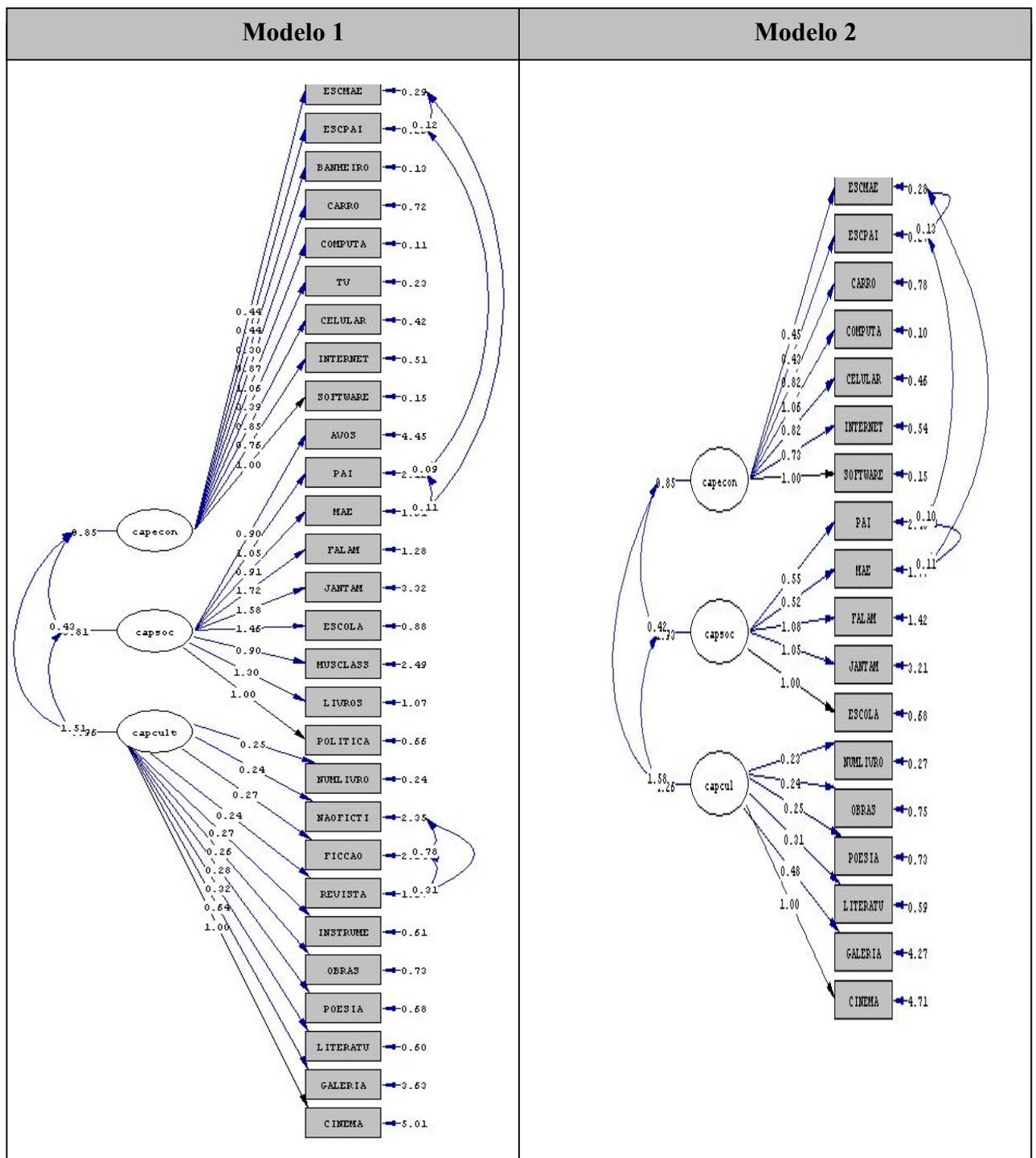
No modelo 2 estabeleceu-se as correlações entre os erros de medição das variáveis:

- “escolaridade do pai” e “escolaridade da mãe”,
- “pai” e “mãe”,
- “escolaridade da mãe” e “mãe”,
- “escolaridade do pai” e “pai”.

Estas relações podem ser observadas na figura 6.2.

¹⁷ Jöreskog & Sörbom (2002) recomendam que, se os dados são ordinais ou categóricos, o método de estimação deve ser mínimos quadrados ponderados e a matriz de correlação utilizada deve ser a tetracórica ou policórica.

Figura 6.2- Modelo 1 e modelo 2 modificados.



A avaliação do ajuste global dos modelos pode ser determinada a partir dos diversos índices da tabela 6.1. Nos dois modelos, o teste qui-quadrado rejeita a

¹⁸ A correlação policórica é uma medida de associação empregada como uma substituição para a correlação produto-momento quando ambas as variáveis são medidas ordinais com três ou mais categorias. Já a correlação tetracórica é usada para relacionar duas medidas binárias.

hipótese nula a um nível de significância de 0,05 (p-valor =0,00). No entanto, os índices aproximados de ajuste apresentaram melhor desempenho. No caso do modelo 1, a raiz do erro quadrático médio (RMSEA) é menor que 0,05, o que sugere um ajuste bom. O índice de qualidade de ajuste (AGFI) é de 0,97, o que é considerado aceitável. Os índices de ajuste normalizado e comparativo (NFI e CFI) são 0,88 e 0,90, respectivamente, e estão ligeiramente abaixo do valor aceitável.

No modelo 2, os resultados foram melhores, e, de forma geral, o ajuste deste modelo foi muito bom. A raiz do erro quadrático médio (RMSEA) foi 0,035, menor que no modelo 1. O índice de qualidade de ajuste (AGFI) foi 0,99, valor perto de um. Os índices de ajuste aproximados também tiveram ótimo desempenho, com valores acima de 0,95.

Tabela 6.2- Medidas de ajuste.

Modelo 1	Modelo 2
<i>Com 340 graus de liberdade</i>	<i>Com 128 graus de liberdade</i>
Qui-quadrado = 1767.6 (p = 0,00)	Qui-quadrado = 451.6 (p = 0,00)
Raiz do erro quadrático médio (RMSEA) = 0,045	Raiz do erro quadrático médio (RMSEA) = 0,035
Índice de qualidade de ajuste (AGFI) = 0,97	Índice de qualidade de ajuste (AGFI) = 0,99
Índice de ajuste normado (NFI) = 0,88	Índice de ajuste normado (NFI) = 0,96
Índice de ajuste comparativo (CFI) = 0,90	Índice de ajuste comparativo (CFI) = 0,97

Na avaliação local, para ambos os modelos, os “t” valores foram todos significativamente diferente de zero, o que sugere relações importantes entre as variáveis observadas e os fatores latentes. Analisando o R^2 no modelo 1, observa-se que só 4 das 28 variáveis observadas possuem valores abaixo de 15%; por sua parte as variáveis “**software**” e “**computador**” apresentam os valores mais altos, 85% e 90% respectivamente. No modelo 2, cinco variáveis têm valores entre 19% e 27%, já as variáveis restantes estão acima de 37%.

Tabela 6.3- Coeficiente de determinação.

Modelo 1						Modelo 2					
cinema	galeria	literatu	poesia	obras	instrume	cinema	galeria	literatu	poesia	obras	numlivro
0,44	0,31	0,40	0,32	0,28	0,31	0,47	0,19	0,41	0,27	0,25	0,46
revista	ficcao	naoficti	numlivro								
0,149	0,11	0,09	0,51								
politica	livros	musclass	escola	jantam	falam	escola	jantam	falam	mae	pai	software
0,55	0,56	0,21	0,66	0,38	0,65	0,74	0,40	0,62	0,22	0,19	0,85
mae	pai	avos									
0,29	0,30	0,13									
software	internet	celular	tv	computa		internet	celular	computa	carro	escpai	escmae
0,85	0,49	0,59	0,36	0,90		0,46	0,56	0,91	0,43	0,37	0,39
carro	banheiro	escpai	escmae								
0,47	0,38	0,38	0,36								

Em geral, as medidas de ajuste do modelo 2 foram melhores do que as medidas de ajuste do modelo 1. Mas existem indicadores específicos que servem para comparar e por conseguinte determinar qual modelo é o melhor: i) o índice de validação cruzada (ECVI), ii) a medida de Akaike-AIC e iii) a medida de Bozdogan-CAIC. O melhor modelo será aquele que apresente menores valores nos três indicadores. Da tabela 6.4 pode se verificar que o modelo 2 possui menores valores, por tanto é o que apresenta melhor ajuste.

Tabela 6.4- Medidas para comparar os modelos.

Modelo 1	Modelo 2
Índice de validação cruzada (ECVI)= 0,89	Índice de validação cruzada (ECVI) = 0,25
AIC = 1899.60	AIC = 537.75
CAIC = 2339.17	CAIC = 824.13

6.7.

TRI não-paramétrica versus a Análise Fatorial Confirmatória

Nesta seção comparamos e descrevemos os resultados obtidos através das duas metodologias: TRIN e AFC.

Primeiro, analisamos o grau de associação entre os fatores latentes de diferentes metodologias. Em segundo lugar, analisamos as correlações entre fatores latentes de uma mesma metodologia. E finalmente mostramos as variáveis usadas em cada um dos métodos.

Tabela 6.5-Coeficientes de correlação entre os fatores dos modelos.

Capital social		
	ACF: escala 1	ACF: escala 2
TRIN:escala 1	0,93	0,83
TRIN: escala 2	0,84	0,89
Capital cultural		
	ACF: escala 1	ACF: escala 2
TRIN:escala 1	0,75	0,68
TRIN: escala 2	0,76	0,77
Capital econômico		
	ACF: escala 1	ACF: escala 2
TRIN:escala 1	0,91	0,87
TRIN: escala 2	0,90	0,90

Na tabela 6.5 mostram-se os coeficientes de correlação de cada fator latente. No construto capital social, os coeficientes de correlação entre os escores obtidos via TRIN e AFC estão acima de 0,82. No construto capital econômico, os coeficientes de correlação estão entre 0,87 e 0,91. No caso do construto capital cultural, os coeficientes de correlação estão na faixa de 0,68 e 0,75, e são valores menores comparados com os outros dois.

As correlações elevadas entre os escores dos construtos de diferentes metodologias confirmam e validam a medição dos três construtos através das variáveis indicadoras propostas. De certa forma, a alta correlação observada era de se esperar, posto que na análise de fatores confirmatória não se consideraram os itens “ruins”¹⁹ identificados na análise previa via TRIN.

Na tabela 6.6 mostramos os coeficientes de correlação entre os escores dos fatores latentes. Podemos verificar que os coeficientes de correlação, em geral, são maiores na AFC. Ao mesmo tempo, a correlação entre capital cultural e econômico é a mais forte e a correlação entre capital social e econômico é a mais fraca.

¹⁹ São itens que apresentaram covariância negativa com um ou mais itens e possuem baixo poder de discriminação ($H_i < 0,20$).

Um coeficiente de correlação alto entre o capital cultural e econômico nos indicaria que existe uma forte relação entre estes dois construtos. Assim, para possuir bens culturais, ao menos na sua forma objetivada, é necessário ter simplesmente capital econômico, por exemplo a compra de livros, de pinturas. Todavia, para apropriar-se simbolicamente destes bens é necessário possuir os instrumentos desta apropriação e os códigos necessários para decifrá-los, ou seja, é necessário possuir capital cultural no estado incorporado. Por sua parte, a correlação entre o capital cultural e social é intermediária. Isto nos mostra que a transmissão de capital cultural de pais para filhos e a aquisição de capital cultural (filhos) é possível, se existe uma relação estreita e fluída entre eles. Finalmente, a correlação entre capital social e capital econômico é baixa. Isto indicaria que não necessariamente a maior posse de capital econômico assegura maior posse de capital social, ou vice-versa. Por exemplo, segundo Coleman, é o caso das famílias asiáticas, quem tinham nível educativo baixo (pouco capital econômico), sem embargo o capital social disponível na família para a educação dos filhos era extremamente alto.

Tabela 6.6-Correlação entre os fatores latentes

Análise de fatores confirmatória (AFC)		
	Escala 1	Escala 2
Corr (cap.cultural,cap. social)	0,55	0,33
Corr(cap.cultural,cap. econômico)	0,82	0,83
Corr (cap. social, cap. econômico)	0,52	0,32
TRI não-paramétrica		
	Escala 1	Escala 2
Corr (cap.cultural,cap. social)	0,36	0,22
Corr(cap.cultural,cap. econômico)	0,41	0,43
Corr (cap. social, cap. econômico)	0,22	0,16

Tabela 6.7- Variáveis utilizadas nas duas metodologias: TRIN e AFC.

	TRI não-paramétrica		Análise de fatores confirmatória	
	Escala 1	Escala 2	Escala 1	Escala 2
Capital cultural				
1.	cinema	cinema	cinema	cinema
2.	galeria	galeria	galeria
3.	literatura	literatura	literatura	literatura
4.	poesia	poesia	poesia
5.	obras	obras	obras	obras
6.	instrumento	...	instrumento	...
7.	revista	...
8.	ficcao	...
9.	naoficcão	...
10.	numlivros	numlivros	numlivros	numlivros
Capital social				
1.	política	...	política	...
2.	livros	livros	...
3.	musica clássica	...	musica clássica	...
4.	escola	escola	escola	escola
5.	jantam	jantam	jantam	jantam
6.	falam	falam	falam	falam
7.	mãe	mãe	mãe	mãe
8.	pai	pai	pai	pai
9.	avos	...	avos	...
Capital econômico				
1.	internet	internet	internet	internet
2.	computador	computador	computador	computador
3.	software	software	software	software
4.	carro	...	carro	carro
5.	banheiro	...	banheiro	...
6.	celular	celular	celular	celular
7.	tv	...	tv	...
8.	escmai	escmai	escpai	escpai
9.	escpai	escpai	escmae	escmae

No capítulo 5 constatou-se que a escala 2 é que possui as melhores propriedades psicométricas, respeito à escala 1, e, a sua vez está formada por um número menor de itens. De igual maneira, no caso da AFC, a melhor escala é a escala 2, que possui também menor número de variáveis. Esses resultados sugerem que podemos obter boas medidas dos construtos sem a necessidade de um questionário grande com muitos itens, que muitas vezes é cansativo e implica

maiores custos. A modo de comparação, mostramos na tabela 6.6 todas as variáveis utilizadas nas duas metodologias.

No próximo capítulo estima-se um modelo de regressão multinível, onde as variáveis explicativas são os três construtos mensurados neste capítulo e no capítulo 5 e a variável dependente é a proficiência do aluno no teste de Matemática.