

Benefícios da modelação de equações estruturais: Ilustração com variáveis ordinais no *LISREL*

ANTÓNIO AUGUSTO PINTO MOREIRA DINIZ



ciep|ue

CENTRO DE INVESTIGAÇÃO EM EDUCAÇÃO E PSICOLOGIA
DA UNIVERSIDADE DE ÉVORA

**ANTÓNIO AUGUSTO PINTO
MOREIRA DINIZ**

**Benefícios da modelação de
equações estruturais:
Ilustração com variáveis ordinais no
*LISREL***

Apresentado à Universidade do Minho para admissão a Provas de Agregação
dentro do âmbito do ramo do conhecimento em Ciências da Educação, na
especialidade de Psicologia da Educação.

ÉVORA

ABRIL de 2022

FICHA TÉCNICA

Título:

Benefícios da modelação de equações estruturais: Ilustração com variáveis ordinais no LISREL

Autor:

António Augusto Pinto Moreira Diniz

Edição:

© Centro de Investigação em Educação e Psicologia da Universidade de Évora
(CIEP | UE), 1.ª Edição, Évora. Abril de 2022 www.ciep.uevora.pt

Morada:

Colégio Pedro da Fonseca, Rua da Barba Rala, n.º 1, Parque Industrial e Tecnológico de Évora, 7005-345 Évora, Portugal

Revisão e paginação:

Mafalda Pequeno

Capa e Design gráfico:

©mr-creative.net

Impressão e acabamento

VASP Digital Printing Services – www.vasp.pt

ISBN

978-972-778-241-3

Depósito legal

507641/22

É expressamente proibido reproduzir esta obra, na totalidade ou em parte, sob qualquer forma ou meio, exceto para fins de ensino e investigação. Autorizações especiais podem ser requeridas para ciep@uevora.pt

«Este trabalho é financiado por fundos nacionais através da FCT – Fundação para a Ciência e a Tecnologia, I.P., no âmbito do projeto UIDB/04312/2020»

ÍNDICE

1. Objetivo	7
2. O Primado da Teoria na Análise de Dados Multivariados (ADM).....	9
3. Benefícios da Utilização da Instrumentação de ADM de Segunda Geração.....	15
3.1. Benefício 1: Acomodação do Dinamismo entre Teoria, Modelos e Dados...	15
3.1.1. Ilustração: Análise Fatorial Exploratória (com o SPSS) versus Confirmatória (com o AMOS e o LISREL).....	17
3.2. Benefício 2: Possibilidade de Perseguir o Ideal de Unificação das Teorias...	32
3.2.1. Ilustração: Estimação e Teste de um Modelo Fatorial Hierárquico da Integração Social no Ensino Superior.....	33
3.3. Benefício 3: Incremento nos Processos de Generalização de Validade.....	41
3.4. Benefício 4: Possibilidade de Teste de Relações Não-Recursivas.....	41
3.5. Benefício 5: Possibilidade de Teste de Relações Diretas e Indiretas.....	42
4. Nota Final.....	45
5. Referências.....	47

ÍNDICE DE FIGURAS

Figura 1. Modelo uni-fatorial com variáveis observadas ordinais: diagrama conceptual.....	20
Figura 2. Análise fatorial exploratória da EAPN no SPSS pelo método do eixo principal sobre uma matriz de covariâncias de Pearson: Teste gráfico de Cattell. Valores próprios no eixo das ordenadas e fatores no das abcissas.....	27
Figura 3. Modelos alternativos do Questionário de Integração Social no Ensino Superior (QISES).....	36
Figura 4. Modelo pentafatorial oblíquo do Questionário de Integração Social no Ensino Superior (QISES).....	37
Figura 5. Componente estrutural do modelo bifatorial hierárquico de primeira ordem derivado do (QISES).....	39
Figura 6. Componente estrutural do modelo unifatorial hierárquico de terceira ordem derivado do (QISES).....	40
Figura 7. Modelo não-recursivo: Diagrama conceptual.....	42
Figura 8. Modelo de mediação: Diagrama conceptual.....	42
Figura 9. Modelos de mediação alternativos: Diagramas conceptuais.....	43

ÍNDICE DE TABELAS

Tabela 1. Análise fatorial exploratória comum vs. Análise fatorial confirmatória.....	17
Tabela 2. Análise Fatorial Exploratória da EAPN: Critério de Keiser para a Retenção.....	24
Tabela 3. EAPN: Resultados Derivados a Partir da Análise Fatorial Exploratória.....	26
Tabela 4. Estatísticas de Ajustamento no AMOS dos Modelos Tri e Bifatorial Oblíquo da EPAN.....	28
Tabela 5. Estatísticas de Ajustamento do Modelo Bifatorial Oblíquo da EPAN no AMOS e no LISREL.....	29
Tabela 6. Modelo Bifatorial Oblíquo da EAPN: Resultados da Análise Fatorial com o SPSS, o AMOS e o LISREL.....	30
Tabela 7. Modelo Pentafatorial Oblíquo do QISES: Resultados Psicométricos.....	38

1. Objetivo

Com esta lição pretende-se evidenciar o potencial para a reflexão teórica e para a investigação empírica em ciências sociais, nomeadamente em psicologia, da instrumentação estatística de análise de dados multivariados (IADM) de segunda geração, em especial aquelas que respeitam à modelação de equações estruturais (MEE: neste caso o M da abreviatura significa modelação; adiante deve ser lido como modelação ou modelo consoante o contexto frásico). Este desígnio é perseguido através de uma arqueologia que contrasta a utilização feita da IADM de primeira geração com a de segunda geração e da apresentação de um conjunto estudos que respeitam aos benefícios resultantes da MEE, contemplando variáveis latentes e variáveis observadas ordinais. Para demonstração de dois desses benefícios, apresentam-se dois estudos originais que envolvem a análise fatorial.

Entretanto, realça-se a subordinação da utilização da IADM de segunda geração à teoria, particularmente a MEE, legitimando-se os dados enquanto parte integrante da teoria substantiva sobre o domínio em apreciação, através de um enfoque renovado na teoria dos dados e da estimação estatística.

2. O Primado da Teoria na Análise de Dados Multivariados (ADM)

Num esquema heurístico de categorização que articula propósitos de investigação e procedimentos de síntese estatística de dados, a IADM pode ser classificada como exploratória ou confirmatória. Na investigação que visa a formalização de hipóteses (ou teorias) sobre fenómenos mal conhecidos, ou empiricamente pouco estudados, utilizam-se técnicas estatísticas exploratórias, tais como a análise fatorial exploratória (AFE), ainda que com *nuances* em função do método de estimação utilizado: componentes principais *versus* análise fatorial comum, nas suas vertentes de estimação pelo método de máxima verosimilhança ou pelo método do eixo principal (Costello & Osborne, 2005; Fabrigar, Wegener, MacCallum, & Strahan, 1999). Na investigação que visa a refutação de hipóteses (ou teorias) formalizadas *a priori* sobre fenómenos já conhecidos e empiricamente estudados, utilizam-se técnicas estatísticas confirmatórias, como a análise fatorial confirmatória (AFC: Bollen, 1989; Brown, 2015). Ora, a instrumentação de análise de dados multivariados (IADM) de segunda geração, seguindo a terminologia de Fornell (1982), tal como o LISREL (*Linear Structural Relationships*; Jöreskog & Sörbom, 1993, 1996), serve estas últimas.

Fazendo uma breve arqueologia da IADM, ancorada em Fornell (1982), pode dizer-se que foi na década de 1960, com a rápida disseminação e desenvolvimento dos computadores, que apareceram programas estatísticos de primeira geração, como o SPSS, os quais suscitaram o início de uma mudança na orientação da investigação em ciências sociais, particularmente em psicologia. Das teorias abstratas, muitas vezes vazias de factos empíricos, passou-se para um crescente empirismo, muitas vezes sem interesse pela justificação teórica. Nesta senda, assistimos a mudanças dramáticas durante a década de

1970 na maneira como os dados passaram a ser analisados. Uma vez que a utilização desta IADM de primeira geração requeria poucas suposições estatísticas e menos conhecimento *a priori*, ela entrou facilmente nas ciências sociais, trazendo doses massivas de empirismo, muitas vezes ingênuo, onde os dados precediam a conceptualização: fenómenos complexos podiam ser estruturados de forma interpretável; vastos conjuntos de informação podiam ser reunidos em simultâneo; a multidimensionalidade podia ser investigada; e, os investigadores puderam reduzir a sua invocação do critério de suficiência a um nível de significância estatística.

O fascínio pelo potencial mediador da IADM de primeira geração imperou durante a década de 1970, suscitando uma inércia nas práticas de ADM em ciências sociais que, conforme atrás referido, aconteceu num contexto maioritariamente acrítico de *ciência normal* (Kuhn, 1970) assente numa utilização ritualista da estatística (Gigerenzer, 2018; Gigerenzer, Krauss, & Vitouch, 2004). Este foi o preço a pagar pela sua utilização. Porém, hoje em dia podemos dizer que disso houve uma tomada de consciência, atendendo à evolução e sofisticação entretanto ocorrida nessa instrumentação (ver, e.g., Field, 2013; Marôco, 2011; Tabachnick & Fidell, 2013).

Contudo, na década de 1980, com a IADM de segunda geração, a ênfase passa a ser posta no aspeto cumulativo do desenvolvimento teórico, através do qual o conhecimento adquirido é posto à prova da realidade dos dados e incorporado na análise dos mesmos. Esse conhecimento *a priori* pode derivar da teoria, do prévio conhecimento sobre modelos, métodos e técnicas de investigação, ou de prévia evidência empírica. Os conceitos assumem um duplo papel: como factos adquiridos, combinando a teoria com os dados, ou como hipóteses *a priori* a considerar, confrontando a teoria com os dados.

Sabemos que as variáveis conceptuais são essenciais para o progresso científico. Não há maneiras racionalmente otimizadas para extrair conhecimento da investigação empírica sem previamente fazer certas assunções quanto à matéria substantiva em análise, tal como quanto à natureza métrica dos dados empíricos coletados associada às técnicas para os tratar de forma a assegurar a validade das conclusões estatísticas, entre outros aspetos da validade interna dos estudos (Shadish, Cook, & Campbell, 2002).

E isto representou uma nova maneira de pensar a investigação em ciências sociais, oferecendo um potencial para o conhecimento científico que ultrapassava a mera descrição e associação empírica. Agora, com a IADM de segunda geração, o cerne do desafio é o de conjugar a teoria e os dados, reconhecendo que essa correspondência não é perfeita ou, por outras palavras, que essa correspondência é probabilística, sendo isso feito pelo teste empírico de modelos formalizados *a priori* que representam as teorias.

De um ponto de vista teórico, existem vantagens em tomar os modelos como conjuntos de hipóteses probabilísticas através da sua associação a um erro de medição. Eles não representam processos determinísticos, mas sim estocásticos, tratando explicitamente o erro de medição como uma observação imperfeita das variáveis empíricas ou como uma correspondência imperfeita entre variáveis abstratas e empíricas.

A IADM de primeira geração apresentou durante décadas limitações para avaliar a qualidade das medidas ou para corrigir o erro de medição. Ignorando-os, incorremos no risco de estimar parâmetros estatísticos enviesados na medida em que as relações estabelecidas entre variáveis observadas possuem uma componente de erro (resíduos ou ruído) de distribuição que atenua a sua magnitude e incrementa o seu erro-padrão. Isto tem influência direta e negativa nas estimativas, inflacionando o erro de tipo II. Nestas condições, relações

que seriam estatisticamente significativas poderão não ser observadas por razões meramente artefactuais.

Em síntese, enquanto a IADM de primeira geração pouco requeria em termos de especificação *a priori*, tanto teórica quanto estatística, a IADM de segunda geração, tal como o *LISREL*, exige-a. Depois do seu aparecimento, situado em 1973 quando Jöreskog apresentou a primeira versão do *modelo de equações lineares estruturais para variáveis latentes (LISREL)*, um programa que permitia aplicar um modelo geral a uma grande variedade de modelos substantivos, ocorre a sua progressiva disseminação na comunidade científica das ciências sociais. Inicialmente lenta, dado que ainda no final da década de 1980 a utilização do *LISREL* exigia o domínio de uma sintaxe complexa de símbolos estatísticos gregos e de noções de cálculo matricial (Jöreskog & Sörbom, 1988), a disseminação desta IADM de segunda geração alarga-se com o aparecimento e desenvolvimento de três outros programas, também concebidos para a análise de estruturas, ou de matrizes, de variâncias-covariâncias (adiante, simplesmente covariâncias): o *EQS* (Bentler, 1985), para utilização com dados intervalares, e o *LISCOMP* (Muthén, 1987) e o *PRELIS* (Jöreskog & Sörbom, 1986) que possibilitavam a utilização de dados categoriais.

Entretanto, a vulgarização da utilização da IADM de segunda geração é acelerada com a publicação de livros complementares aos manuais do *LISREL* e do *EQS* (e.g., Bollen, 1989; Byrne, 1989; Loehlin, 1987); com a introdução da linguagem de programação *SIMPLIS* (*SIMPLE LISREL*; Jöreskog & Sörbom, 1993); com o aparecimento, em 1994, da revista *Structural Equation Modeling: A Multidisciplinary Journal*; e, com uma cada vez maior divulgação através da sua utilização em artigos técnicos e substantivos noutras publicações (Hershberger, 2003). Neste contexto, uns autores centraram-se mais no desenvolvimento e

demonstração algébrica relativas à ligação da teoria dos dados com a da estimação estatística (e.g., Browne & Cudeck, 1993) e outros penderam mais para a demonstração dessa ligação via simulação computacional (e.g., Hu & Bentler, 1998).

Hoje em dia, a IADM de segunda geração permite pôr em marcha um alargado conjunto de técnicas comumente designadas por MEE que são similares no racional e relacionadas no procedimento (ver, e.g., Kline, 2016; Schumacker & Lomax, 2010).

Na base da MEE está a vantagem de com ela se poder incorporar uma componente estrutural com variáveis hipotéticas expurgadas de erro. Mais precisamente, os construtos, variáveis abstratas não observadas (adiante fatores ou variáveis latentes) que respeitam ao modelo estrutural, são qualidades representadas por uma outra componente, o modelo de medição, que integra as respetivas variáveis quantitativas observadas (adiante itens ou indicadores) com o erro que lhes é inerente. Personalidade ou inteligência são conceitos sem ligação direta a medidas quantitativas observáveis, mas a sua operacionalização enquanto construtos num MEE permite estimar de forma purificada as suas relações quer com outros construtos, quer com variáveis observadas (e.g., modelos de múltiplas causas e múltiplos indicadores; Jöreskog & Goldberger, 1975): o erro está do lado do modelo de medição e não no do modelo estrutural (para uma ilustração empírica, ver Amado & Diniz, 2017).

Nesta virtuosa possibilidade de analisar relações com, ou entre, variáveis latentes, radica um conjunto de benefícios da IADM de segunda geração que a seguir se apresentam.

3. Benefícios da Utilização da Instrumentação de ADM de Segunda Geração

3.1. Benefício 1: Acomodação do Dinamismo entre Teoria, Modelos e Dados

As teorias (relações entre conceitos), enquanto abstrações relativas a certos segmentos do mundo real, são testadas através de modelos que as representam (relações entre construtos e de construtos com os seus indicadores), apreciando-se a adequação do seu ajustamento ao segmento do mundo real que ele representa (plausibilidade do modelo, envolvendo significância estatística e parcimónia). E o resultado obtido pode exigir a sua reespecificação, a redefinição dos construtos e/ou das relações entre construtos, revendo-se, assim, a teoria. Ora, a IADM de segunda geração, nomeadamente a MEE, permite acomodar este dinamismo entre teoria, modelos e dados.

Neste contexto, considera-se que não há modelos verdadeiros. Eles só “providenciam uma aproximação ao mundo real que tem algum significado substantivo e alguma utilidade” (MacCallum, 2003, p. 115) e, porventura, uns são mais adequados do que outros. É, pois, prudente considerar que todos os modelos são “imperfeitos tanto no sentido fraco de *incompletos*, quanto no sentido forte de *falsos*” (Meehl, 1990, p. 110, sublinhados do autor). Quanto à sua *incompletude* temos a impossibilidade de representação de toda a complexidade do mundo real, traduzida na omissão de variáveis importantes, e quanto à sua *falsidade* temos as contradições entre o modelo e o mundo real, por exemplo traduzida no recurso a modelos lineares para representar relações não-lineares (MacCallum, 2003).

Focando-nos neste último caso, tomemos o exemplo de MEE com variáveis ordinais: médias, variâncias e covariâncias não fazem sentido uma vez que variáveis ordinais não possuem origem ou unidades de medição (Jöreskog, 2005), potenciando distribuições não-normais e gerando relações não-lineares com outras variáveis (Bentler & Dudgeon, 1996).

Acresce que um modelo pode ser plausível, quando apresenta resultados estatisticamente significativos e é suficientemente parcimonioso, mas não ser verosímil, no sentido de revelar uma verdade aparente, mas não de facto, atendendo ao seu valor heurístico ou interpretativo. Por exemplo, um modelo bifatorial pode estar bem ajustado aos dados empíricos mas apresentar colinearidade entre os fatores. Neste caso é adequado utilizar o argumento Ockhamiano da parcimónia e colapsar os seus indicadores num só fator: não devemos considerar mais entidades do que as necessárias para descrever um dado fenómeno.

A IADM de primeira geração prometeu orientar o investigador no seio das hipóteses com uma preocupação mitigada quanto à teoria dos dados e à teoria substantiva, fazendo com que as hipóteses *a priori* fossem frequentemente substituídas por interpretações *a posteriori*, dando-se pouca atenção à relação entre esquemas conceptuais abstratos e dados empíricos. Parafraseando a metáfora de Fornell (1982, p. 2), os dados conduziam a teoria, que ia no banco detrás.

Ora, a derivação empírica acompanhada da falta de sustentação teórica gera consequências indesejáveis. Por exemplo, na AFE, a falta de conhecimento *a priori* não produz, geralmente, soluções otimizadas para os fatores, ou interpretações únicas para os mesmos, o que torna difícil justificar os resultados. Contrariamente, a AFC, implementável na IADM de segunda geração, produz soluções teoricamente derivadas e estatisticamente otimizadas para os fatores, clarificando a sua eventual ambiguidade interpretativa gerada pela AFE (Bollen, 1989).

Este benefício é elucidado através de um estudo em que, com o mesmo ficheiro de dados, se confrontam os resultados da AFE com os da AFC, realçando as virtudes desta

última. Assim, com o mesmo ficheiro de dados ordinais, pretende-se comparar os resultados da AFE no *IBM SPSS (version 24)* com os da AFC no *IBM SPSS AMOS (version 24;* Arbuckle, 2016) e no *LISREL 8.80* (Jöreskog & Sörbom, 2006).

3.1.1. Ilustração: Análise Fatorial Exploratória (com o SPSS) versus Confirmatória (com o AMOS e o LISREL).

Com este estudo pretende-se examinar as principais diferenças entre a AFE comum (quer pelo método do eixo principal, quer pelo de máxima verosimilhança) e a AFC, as quais são apresentadas nos quatro primeiros tópicos da Tabela 1.

Tabela 1. Análise fatorial exploratória comum vs. Análise fatorial confirmatória	
Análise fatorial exploratória comum	Análise fatorial confirmatória
Objetivo: gerar uma teoria sobre a medida; explorar a estrutura fatorial (EF).	Objetivo: testar uma teoria sobre a medida; corroborar ou infirmar a EF.
1. EF determinada <i>a posteriori</i> a partir das inter-relações entre os itens.	1. EF especificada <i>a priori</i> a partir da teoria e/ou de investigação empírica forte.
2. Número de fatores e respetivos itens empiricamente derivado.	2. Número de fatores e respetivos itens teoricamente derivado.
3. Todos os itens estão relacionados com todos os fatores (com pesos fatoriais desatenuados).	3. Cada item está relacionado com o seu fator específico (com peso fatorial desatenuado).
4. Os erros de medição não podem ser correlacionados.	4. Os erros de medição podem ser correlacionados.
5. EF ortogonais ou oblíquas de 1ª ordem.	5. EF ortogonais ou oblíquas de 1ª ordem e EF de 2ª (fatores de 1ª ordem correlacionados) e 3ª ordem (fatores de 2ª ordem correlacionados).

Método

Participantes

Participaram no estudo 327 idosos comunitários voluntários (leque etário = 65-97 anos, *Mdn* = 74) sem compromisso cognitivo (rastreo com o *MMSE*: Folstein, Folstein, & McHugh,1975; Morgado, Rocha, Maruta, Guerreiro, & Martins, 2009) e sem depressão (rastreo com a *GDS-15*: Sheikh & Yesavage, 1986; Tumas, Rodrigues, Farias, & Cripa

(2008), com menos do que o 1.º ciclo (26.6%) e com o 1.º ciclo (42.8%). A amostra era maioritariamente composta por mulheres (63.3%).

Instrumento e Procedimento

Utilizou-se a Escala de Afetos Positivos e Negativos (EAPN: Amado, 2008; Diniz & Amado, 2005), baseada na *Positive and Negative Affect Schedule* (PANAS: Watson, Clark, & Tellegen, 1988), a qual possui dois fatores, afetos positivos (AP) e afetos negativos (AN), cada um com 10 itens de formato de resposta de tipo-Likert com cinco níveis (desde “Muito pouco ou nada” até “Muitíssimo”).

A amostragem foi não-probabilística e, após obtenção do consentimento informado dos participantes, os dados relativos à frequência com que sentiam cada afeto “normalmente” foram recolhidos mediante entrevista estruturada, conduzida de forma diretiva, atendendo a aspetos de ordem instrumental e ético-deontológica para otimizar o processo (Diniz & Amado, 2014).

Análise estatística. Com o SPSS, recorreu-se à AFE comum (porque contrariamente à AFE pelo método dos componentes principais, é expurgada da análise a componente de erro das variáveis) sobre a matriz de covariâncias de Pearson das variáveis ordinais da EAPN, através do método do eixo principal (MEP), que tem a vantagem de não implicar pressupostos distribucionais das variáveis, e do método de máxima verosimilhança (MMV) (note-se que estes métodos podem apresentar soluções muito similares), com rotação oblíqua (*promax*) dos fatores retidos e com leitura dos resultados da matriz padrão dos fatores/itens (Costello & Osborne, 2005; Fabrigar et al., 1999). No contexto da decisão sobre o número de fatores a reter na AFE da EAPN fez-se, ainda, uma comparação entre o critério

de Keiser (1960), a análise do teste gráfico de Cattell (1966; Cattell & Vogelman, 1977), e a análise paralela (*a priori*) baseada no MEP com permutação de dados (O'Connor, 2000).

Entretanto, o modelo fatorial empiricamente derivado foi testado com o *AMOS*, através da AFC, pelo método de máxima verosimilhança (MMV) sobre uma matriz de covariâncias de Pearson. Este método de estimação é o comumente utilizado no *AMOS*, mesmo para escalas ordinais, ainda que seja possível utilizar no programa uma abordagem Bayesiana (procedimentalmente bastante mais complexa) com uma amostra de grande dimensão (> 1000) para se obter estimativas muito próximas das do MMV (Arbuckle, 2016).

Por último, o modelo fatorial resultante da AFC com o *AMOS* foi testado com o *LISREL*, pelo MMV, mas mediante a abordagem bivariada latente normal (ABLN: Jöreskog, 2005) e com o robusto χ^2 de Satorra e Bentler ($SB\chi^2$; 1994), que corrige as estimativas, os seus erros-padrão e as estatísticas de ajustamento dos modelos aos dados empíricos sob condições de não-normalidade (Curran, West, & Finch, 1996). Acresce que tanto no *AMOS* quanto no *LISREL*, cada fator do modelo foi identificado fixando a um (1.00) o peso fatorial de um dos seus itens (variável de referência) e que o modelo fatorial foi especificado com a independência do erro de medição entre os itens. Note-se que no caso da existência de modificações de modelos estatisticamente derivadas (e.g., covariâncias de erro livremente estimadas entre pares de itens), devem procurar-se razões substantivas que as justifiquem (MacCallum, Roznowski, & Necowitz, 1992).

Entretanto, no *LISREL* as variáveis ordinais requerem um tipo particular de parametrização (Jöreskog, 2005). Num primeiro passo, procede-se à ABLN no *PRELIS 2* (Jöreskog & Sörbom, 1996). Na Figura 1 ilustra-se a relação entre fatores e variáveis ordinais de acordo com essa abordagem. Nela, as distribuições latentes normais (y^*_{ij}) das variáveis

ordinais observadas (y_{ij}), seccionadas por $m - 1$ limites distribucionais (m = número de opções de resposta), servem para computar a matriz de correlações (ou de covariâncias) policóricas dessas respostas latentes, conjuntamente com a sua matriz de covariâncias assintóticas para ajudar à estimação e teste do modelo de interesse. Por analogia, trabalha-se com o genótipo correspondente ao fenótipo.

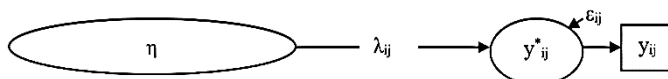


Figura 1. Modelo uni-fatorial com variáveis observadas ordinais: diagrama conceptual . η = fator; λ_{ij} = pesos fatoriais; y_{ij} = variáveis observadas categoriais; y^*_{ij} = respostas latentes (ou “traços” latentes) das y_{ij} ; ϵ_{ij} = termos de erro (ou resíduos ou ruído), ou quantidade da variância das y^*_{ij} não capturada por η (erro aleatório mais fator específico).

Justifica-se o recurso a esta abordagem porque Babakus, Ferguson e Jöreskog (1987) constataram que na AFC com variáveis ordinais derivadas da discretização de variáveis contínuas (estimação pelo MMV) as correlações policóricas são preferíveis a outras estatísticas de associação bivariadas (r , rho de Spearman ou tau-b de Kendall). As outras correlações têm um efeito de atenuação sobre os pesos fatoriais estimados. Só as correlações policóricas não produziram parâmetros subestimados e erros-padrão sobrestimados.

Bollen (1989) acrescenta que essa “atenuação é inversamente relacionada com o número de categorias, sendo maior com 2 ou 3 e muito menor com 7 ou mais” e que a estimativa do ajustamento de um MEE (o χ^2) é mais inflacionada pela curtose e pela assimetria do que pelo número de categorias: claro que quanto menor o seu número maior a não-normalidade. Em todo o caso, fica o alerta: “se as correlações momento-produto forem

utilizadas com dados ordinais, uma variável observada pode ser descartada ... [quando] de facto é um bom indicador” (Babakus et al., 1987, p. 227).

Também Flora e Curran (2004) observaram que correlações policóricas estimadas a partir de distribuições latentes contínuas de variáveis ordinais geram estimativas multivariadas robustas sob condições de não-normalidade moderada (assimetria = |1.25| e curtose = |3.75|), reduzindo a probabilidade de cometer o erro de Tipo I no teste de MEE (i.e., rejeitar um modelo quando ele está corretamente especificado). E isto relaciona-se com o segundo passo.

No segundo passo, as matrizes calculadas no *PRELIS* sevem de *input* para estimar o modelo no *LISREL 8* (Jöreskog & Sörbom, 1993) com a já referida correção de Satorra e Bentler (1994), a qual, segundo Curran et al. (1996), apresenta um bom desempenho na estimação de modelos por MV baseada em correlações policóricas com amostras de pequena dimensão ($N = 200$), mesmo com distribuições severamente não-normais (assimetria = |3.0|; curtose = |21.0|). Flora e Curran (2004) acrescentam que essa correção também apresenta um bom desempenho em amostras de pequena dimensão ($N = 100$) com um método alternativo ao de MV (também disponível no *LISREL*) que combina correlações ou covariâncias policóricas e estimação pelos mínimos quadrados diagonalmente ponderados (*DWLS*).

Por outro lado, sabe-se do problema conceptual e prático envolvido na utilização do qui-quadrado na análise do ajustamento (ou identidade) de um modelo aos dados empíricos (Bollen, 1989). Em termos conceptuais, o facto de os modelos serem tidos como aproximações da realidade implica que um teste de ajustamento como o do qui-quadrado, que implica a presunção que um modelo é uma representação exata da realidade, não é

verossímil. Em termos práticos, o facto de o qui-quadrado depender da dimensão da amostra, da não-normalidade das variáveis observadas e da complexidade do modelo pode gerar resultados artefactuais porque quanto maior a dimensão da amostra, a não-normalidade e a complexidade mais inflacionado será o qui-quadrado, conduzindo, erroneamente, à rejeição do modelo.

Assim, foram criados outros índices de ajustamento alternativos ao qui-quadrado, designados de práticos (também nomeados heurísticos, subjetivos ou psicológicos) para avaliar a precisão preditiva interpolativa (PPI) de um modelo, isto é, dentro do domínio de origem dos dados observados (Forster, 2002). Segundo a simulação de Hu e Bentler (1998), o *root mean square error of approximation (RMSEA)*, o *standardized root mean square residual (SRMR)* e o *comparative fit index (CFI)* são adequados para esse efeito. Quanto aos seus valores de referência, um *RMSEA* [observar IC 90%] próximo ou menor do que .06 indica um bom ajustamento (Hu & Bentler, 1998), igual a .08 indica um ajustamento aceitável e igual ou maior do que .10 indica um ajustamento não aceitável (Browne & Cudeck, 1993); um *SRMR* próximo ou menor do que .08 indica um bom ajustamento (Hu & Bentler, 1998) e igual a .10 indica um ajustamento aceitável (por convenção); por último, um *CFI* próximo ou maior do que .95 indica um bom ajustamento (Hu & Bentler, 1998) e igual a .90 indica um ajustamento aceitável (por convenção). Contudo, os valores de referência para os índices de PPI não devem ser tidos na prática como regras de ouro na decisão de aceitar, ou não, o ajustamento de um modelo (Marsh, Hau, & Wen, 2004). Entretanto, um outro índice, o *expected cross-validation index (ECVI)* (Browne & Cudeck, 1993) serve à avaliação da precisão preditiva extrapolativa (PPE) de um modelo (para amostras da mesma dimensão), isto é, fora do domínio interpolativo (Forster, 2002). Com ele visa-se a comparação entre

modelos alternativos ou concorrentes, selecionando-se o com menor *ECVI*, que corresponde ao mais parcimonioso.

Por último, numa abordagem em dois passos (Anderson & Gerbing, 1988; Jöreskog & Sörbom, 1993), estando o MME ajustado aos dados empíricos deve ser feito o exame das propriedades psicométricas dos fatores que o integram (Fornell & Larcker, 1981) antes de os relacionar com outras variáveis ou com outros fatores. Examina-se a validade convergente (VC), através da utilização da fórmula que permite o cálculo da variância média extraída (VME) dos itens pelo respetivo fator [$VME = \Sigma \beta_{ij}^2 / (\Sigma \beta_{ij}^2 + \Sigma \varepsilon_{ij})$], que deve ser igual ou maior do que .50 (na fórmula, β_{ij} = pesos fatoriais estandardizados e ε_{ij} = resíduos, ou variâncias do erro, estandardizados). Também se utiliza uma variação dessa fórmula para calcular a fiabilidade compósita (FC) dos fatores [$FC = (\Sigma \beta_{ij})^2 / (\Sigma \beta_{ij})^2 + \Sigma \varepsilon_{ij}$], que deve ser igual ou maior do que .70 para estados iniciais de construção de um instrumento, .80 para comparação entre grupos e .90 para diagnóstico individual (Nunnally & Bernstein, 1994). Acresce, ainda, o exame da validade discriminante (VD) entre fatores. Para que ela seja assegurada, a variância partilhada entre fatores (ϕ^2 ; ϕ = correlação desatenuada) deve ser menor do que a VME dos fatores.

Resultados e Discussão

Iniciando a análise da estrutura fatorial da EAPN através da maneira mais convencional de determinar o número de fatores a reter através da AFE (valores próprios > 1; Keiser, 1960), constatou-se que a matriz de covariâncias era fatorizável ($KMO = .86$; χ^2 de Bartlett = 2531.14, $gl = 190$, $p < .001$) e a existência de quatro fatores, conforme pode observar-se na Tabela 2.

Tabela 2. Análise Fatorial Exploratória da EAPN: Critério de Keiser para a Retenção

Fator	Valor próprio	% Variância	% Acumulada
1	8.05	29.64	29.64
2	4.49	16.51	46.15
3	2.00	7.37	53.52
4	1.41	5.19	58.71

Nota. Análise no SPSS pelo método do eixo principal sobre uma matriz de covariâncias de Pearson.

Contudo, a solução com quatro fatores gerou um deles com somente dois itens, pelo que foi rejeitada. Na Tabela 3 apresenta-se os resultados dos parâmetros estimados através da AFE, pelo MEP e pelo MMV, para o modelo oblíquo com três fatores empiricamente derivado da EAPN que extraíram mais do que 50% da variância dos itens (Tabela 2). Nela constam também os resultados da AFC com o AMOS para comparação com os da AFE, bem como os resultados da AFC com dois fatores para comparação com os da AFC com três fatores.

Quanto à AFE, na tabela observa-se que, conforme o expectável, a generalidade dos itens (salvo os itens Receoso e Amedrontado) apresentaram pesos fatoriais similares em ambos os métodos de estimação (MEP e MMV), apesar dos resultados do AMOS apontarem para a existência de não-normalidade multivariada: curtose = 86.83, RC (rácio crítico = parâmetro/EP) = 26.47, $p < .001$.

Entretanto, o item Receoso apresentou-se como ambíguo, com pesos fatoriais semelhantes para os fatores AN2 e AN1 em ambos os métodos de estimação, o conduziria à sua exclusão.

Também os itens Chateado e Amedrontado apresentaram ambiguidade na estimação por MMV. Por último, na estimação pelo MMV a VME e a FC foram notoriamente mais baixas

do que pelo MEP por causa da diferença dos pesos fatoriais dos itens Amedrontado e Receoso.

Quanto à comparação de resultados entre a AFE e AFC (Tabela 3), para além da diferença, para bastante melhor, no que respeita à VC e FC dos dois fatores relativos aos afetos negativos (AN1 e AN2), constatou-se que a solução da AFE com três fatores não era verosímil uma vez que na solução da AFC com os mesmos três fatores, os fatores AN1 e AN2 não apresentavam validade discriminante ($\varphi^2 > \text{VME}$ de cada um). Faz-se notar que este resultado estatístico vai ao encontro da falta de justificação substantiva para uma bipartição dos itens dos afetos negativos (Watson et al., 1988; Zevon & Tellegen, 1982), sendo a solução da AFC com dois fatores (Tabela 3) a que melhor traduz a estrutura latente dos dados.

Entretanto, os resultados do teste gráfico do cascalho (*scree plot test*; Cattell, 1966) também apontavam para um modelo de dois fatores. Existem dois valores próprios ≥ 1 antes do ponto de inflexão da curva assintótica que define os fatores (cf. seta na Figura 2): segundo Cattell e Vogelmann (1977, p. 308), “o último fator real é o anterior ao início do cascalho (*scree*)”.

Tabela 3. EAPN: Resultados Derivados a Partir da Análise Fatorial Exploratória

Fatores e itens (<i>N</i> = 327)	AFE		AFC	
	3 Fs MEP	3 Fs MMV	3 Fs MMV	2 Fs MMV
	β	β	β	β
Afetos Positivos (AP)				
Ativo	.76	.77	.73	.73
Forte	.68	.69	.66	.66
Entusiasmado	.67	.66	.69	.69
Decidido	.66	.66	.63	.63
Interessado	.66	.66	.69	.69
Orgulhoso	.65	.63	.60	.60
Animado	.59	.59		.65
Inspirado	.55	.54	.54	.54
Atencioso	.49	.49	.48	.48
Atento	.43	.44	.43	.43
VC (VME) / FC	.39 / .86	.39 / .86	.40 / .87	.38 / .86
Afetos Negativos 1 (AN1)				
Nervoso	.95	.96	.66	.57
Agitado	.64	.65	.62	.56
Angustiado	.62	.64	.73	.68
Irritável	.43	.45	.69	.66
Chateado	.41	.44(AN2; .31)	.69	.66
VC (VME) / FC	.41 / .76	.42 / .77	.47 / .81	---
Afetos Negativos 2 (AN2)				
Culpado	.82	.79	.66	.60
Hostil	.81	.83	.64	.59
Envergonhado	.57	.50	.68	.63
Amedrontado	.51	.43(AN1; .34)	.75	.69
Receoso	.40(AN1; .32)	.31(AN1; .39)	.70	.67
VC (VME) / FC	.41 / .77	.37 / .72	.47 / .82	.39 / .87
ϕ AP-AN1	-.31	-.31	-.33	
ϕ AP-AN2	-.27	-.27	-.28	
ϕ AN1-AN2	.55	.55	.75	
ϕ^2 AN1-AN2	.30	.30	.56	
ϕ AP-AN	---	---	---	-.34

Nota. Fs = fatores; AFE = análise fatorial exploratória no SPSS pelo método do eixo principal (MEP) e pelo método de máxima verisimilhança (MMV) sobre uma matriz de covariâncias de Pearson (rotação *promax*; resultados da matriz padrão); AFC = análise fatorial confirmatória no AMOS pelo MMV sobre uma matriz de covariâncias de Pearson.

β = peso fatorial estandardizado; VME = variância média extraída; FC = fiabilidade compósita; ϕ = correlação desatenuada; ϕ^2 = variância partilhada.

Destques a negrito: resultados notoriamente diferentes para o modelo trifatorial (viz., FC e ϕ); (.

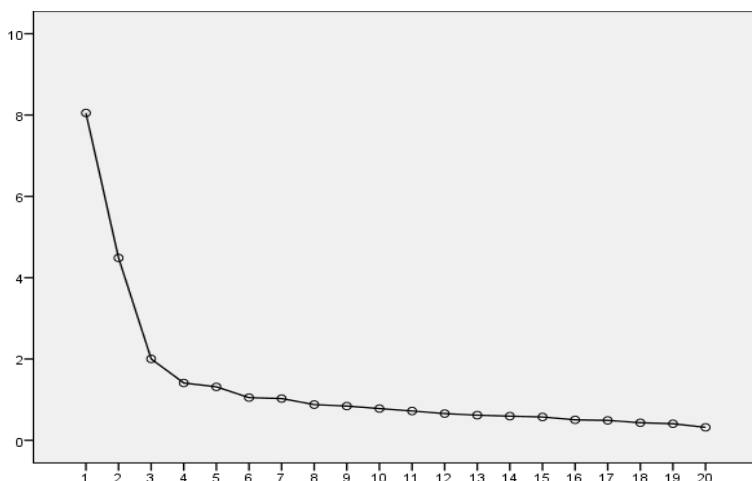


Figura 2. Análise fatorial exploratória da EAPN no SPSS pelo método do eixo principal sobre uma matriz de covariâncias de Pearson: Teste gráfico de Cattell. Valores próprios no eixo das ordenadas e fatores no das abcissas.

O mesmo indicou a análise paralela (O'Connor, 2000): sintaxe SPSS retirada de <https://people.ok.ubc.ca/briocconn/nfactors/rawpar.sps>. A simulação foi baseada no MEP com permutação de dados (indicada para a não-normalidade das variáveis): número de conjuntos de casos = 100; Nível de confiança = 95%, com resultados para o valor próprio ≥ 1 do Fator 1 = 5.27 e do Fator 2 = 2.75, mas não do Fator 3 = .86.

Em suma, na AFE devem usar-se ambas as técnicas, a análise paralela e o teste gráfico, para decidir quantos os fatores a reter.

Comparando o ajustamento aos dados empíricos dos modelos oblíquos testados através da AFC no AMOS, o trifatorial apresentou melhor ajustamento (PPI e PPE) do que o bifatorial (Tabela 4). Contudo, apesar de ser mais plausível, ele não era, como atrás foi referido, substantivamente verosímil, pelo que o modelo a selecionar deve ser o bifatorial. Para além desta justificação estatística acresce que o modelo bifatorial corresponde ao da

PANAS (Watson et al., 1988), o que reforça essa opção. Porém, ele apresentava-se no limiar de rejeição (Tabela 4; limite superior do IC 90% do *RMSEA* e *CFI*).

Tabela 4. Estatísticas de Ajustamento no AMOS dos Modelos Tri e Bifatorial Oblíquo da EPAN

Modelo	MMV χ^2 (gl)	RMSEA [IC 90%]	SRMR	CFI	ECVI
3 Fatores	507.55 (167)	.08 [.07-.09]	.07	.86	1.82
2 Fatores	605.72 (169)	.09 [.08-.10]	.07	.82	2.11
2 Fatores ^a	527.52 (167)	.08 [.07-.09]	.07	.85	1.88

Nota. MMV = método de máxima verosimilhança; *RMSEA* = root mean square error of approximation; *SRMR* = standardized root mean square residual; *CFI* = comparative fit index; *ECVI* = expected cross-validation index.

^aModelo com duas covariâncias de erro (“Receoso” e “Amedrontado”; “Nervoso” e “Agitado”).

Se um modelo não se ajustar aos dados empíricos ele deve ser rejeitado mas, dentro de certos limites, ele pode ser reespecificado e melhorado (e.g., adição de novos parâmetros relevantes ou subtração de parâmetros irrelevantes). A AFC exige uma detalhada especificação do modelo inicial mas, no caso de ele não ser adequado, podemos, numa estratégia exploratória, alterá-lo. Todavia, como já foi referido, essas modificações exigem a devida justificação substantiva (MacCallum et al., 1992).

Para o modelo bifatorial apresentar um ajustamento mais aceitável foi necessário correlacionar os resíduos de dois pares de itens (Tabela 4). Trata-se de um processo de acomodação do modelo aos dados observados que envolve um incremento da sua precisão preditiva. Note-se que este processo é aplicado ao domínio abarcado pelo modelo (PPI). Noutro domínio o modelo pode não ajustar (Forster, 2002). Deve então perguntar-se: estas indesejáveis modificações (MacCallum et al., 1992) ficaram a dever-se à componente

aleatória (idiossincrasias da amostra) ou sistemática (fator específico) dos resíduos dos itens?

Contudo, para estes dois pares de itens existe uma justificação substantiva (fator específico dos itens). Eles representam as categorias de conteúdo de Zevon e Tellegen (1982; Watson et al., 1988), Temeroso (“Receoso” e “Amedrontado”) e Agitado (“Nervoso” e “Agitado”), muito relacionáveis com o lado negativo da experiência de envelhecer. Desta maneira, temos uma condição necessária, apesar de não suficiente, para admitir as duas modificações realizadas. Para o esclarecer cabalmente deve observar-se, num processo de generalização de validade, se elas também são necessárias em amostras independentes da agora utilizada.

Entretanto, comparando os resultados no *AMOS* e no *LISREL* do teste do ajustamento aos dados empíricos do modelo bifatorial com as duas covariâncias de erro, pode observar-se na Tabela 5 que o *LISREL* apresentou melhor PPI (salvo para o *SRMR*) e melhor PPE.

Mas o modelo sem essas covariâncias de erro já apresentava um ajustamento aceitável no *LISREL*, e tinha melhor PPI e PPE do que no *AMOS* sem as covariâncias de erro.

Tabela 5. Estatísticas de Ajustamento do Modelo Bifatorial Oblíquo da EPAN no AMOS e no LISREL

Modelo	$\chi^2 (gl)$	RMSEA [IC 90%]	SRMR	CFI	ECVI
AMOS ^a (MMV)	527.52 (167)	.08 [.07-.09]	.07	.85	1.88
LISREL ^a (MMV _S)	456.53 (167)	.07 [.06-.08]	.08	.96	1.66
LISREL (MMV _{SE})	520.50 (169)	.08[.07-.09]	.08	.95	1.85

Nota. SB = correção de Satorra-Bentler. Ver Tabela 4 para outras abreviaturas.

^aModelo com duas covariâncias de erro (“Receoso” e “Amedrontado”; “Nervoso” e “Agitado”).

Ora, a razão para a diferença encontrada entre o *LISREL* e o *AMOS* é a inadequação da estimação de MEE pelo MMV sobre uma matriz de covariâncias de Pearson com dados ordinais.

Pode então afirmar-se que os resultados obtidos para a EPAN com o *AMOS* eram *falsos* (Meehl, 1990; MacCallum, 2003), enquanto os do *LISREL* pelo MMV sobre uma matriz de covariância assintótica de covariâncias policóricas (i.e., com o robusto χ^2 de Satorra & Bentler, 1994) não, porque este método é o adequado à não-normalidade distribucional das variáveis observadas (Jöreskog, 2005).

Por fim, acresce que os resultados psicométricos da EPAN produzidos pelo *LISREL*, comparativamente com os do *SPSS* e com os do *AMOS*, foram os melhores (Tabela 6).

Assim, observou-se neste processo de exame da estrutura latente da EPAN que com o *LISREL* foi possível acomodar o dinamismo entre teoria, modelos e natureza dos dados. Entretanto, este benefício está na base daqueles outros que a seguir se apresentam.

Tabela 6. Modelo Bifatorial Oblíquo da EAPN: Resultados da Análise Fatorial com o SPSS, o AMOS e o LISREL

Fatores e itens (<i>N</i> = 327)	AFE _{SPSS}		AFC(1) _{AMOS}		AFC(2) _{LISREL}	
	β	R^2	β	R^2	β	R^2
Afetos Positivos (AP) Ativo	.76	.58	.73	.53	.77	.59
Forte	.68	.46	.66	.44	.70	.49
Entusiasmado	.67	.45	.69	.48	.73	.53
Decidido	.66	.44	.63	.40	.66	.44
Interessado	.66	.44	.69	.48	.72	.52
Orgulhoso	.64	.41	.60	.36	.63	.40

Animado	.58	.34	.65	.42	.67	.45
Inspirado	.54	.29	.54	.29	.58	.34
Atencioso	.48	.23	.48	.23	.52	.27
Atento	.43	.18	.43	.18	.45	.20
VC (VME) / FC	.38 / .86		.38 / .86		.42 / .88	
Afetos Negativos (AN) Nervos	.61	.37	.57	.32	.59	.35
Agitado	.60	.36	.56	.31	.59	.35
Angustiado	.65	.42	.68	.46	.71	.50
Irritável	.66	.44	.66	.44	.71	.50
Chateado	.61	.37	.66	.44	.69	.48
Culpado	.61	.37	.60	.36	.69	.48
Hostil	.58	.34	.59	.35	.68	.46
Envergonhado	.65	.42	.63	.40	.68	.46
Amedrontado	.70	.49	.69	.48	.75	.56
Receoso	.65	.42	.67	.45	.72	.52
VC (VME) / FC	.40 / .87		.39 / .87		.46 / .89	
ϕ AP-AN	-.29		-.34		-.34	

Nota. AFE = análise fatorial exploratória no SPSS pelo método de máxima verossimilhança (MMV) sobre uma matriz de covariâncias de Pearson (MCP) e rotação *promax* (resultados da matriz padrão); AFC(1) = análise fatorial confirmatória no AMOS pelo MMV sobre uma MCP; AFC(2) = análise fatorial confirmatória no LISREL pelo MMV sobre uma matriz de covariância assintótica de covariâncias policóricas (i.e., com o robusto χ^2 de Satorra e Bentler, 1994).

Curtose multivariada calculada pelo AMOS = 86.83, RC = 26.47, $p < .001$; RC = rácio crítico (parâmetro/EP) = valores z ($z \geq |1.96|$, $p < .05$, indicando não-normalidade); β = peso fatorial estandardizado; R^2 (comunalidade) = $1 - \varepsilon$ (variância do resíduo estandardizada). VME = variância média extraída; FC = fiabilidade compósita; ϕ = correlação desatenuada.

AMOS vs. LISREL: resultados notoriamente diferentes ($R^2 \geq .05$ e VME de .30 para .40) a negrito.

3.2. Benefício 2: Possibilidade de Perseguir o Ideal de Unificação das Teorias

Implementáveis na IADM de segunda geração, os MEE que contemplam fatores e variáveis observadas que os representam podem acomodar teorias de elevada complexidade. Por exemplo, a teoria hierárquica de três estratos da inteligência (fatores de primeira, segunda e terceira ordem) de Cattell-Horn-Carroll (CHC; Almeida, 2002; Schnelder & McGrew, 2012).

Esta teoria tem a virtude de aproximar as tradições psicométrica e cognitivista em psicologia, para além de conciliar as duas teorias clássicas da inteligência: a de Spearman, cujo modelo contempla o fator *g* a capturar correlações entre medições de fatores específicos, e a de Thurstone, cujo modelo contempla fatores específicos (aptidões mentais primárias; AMPs) independentes de um fator de ordem superior, *g*. Entretanto, esta última resultou na teoria de Cattell-Horn (CH), cujo modelo expandido (Horn & Blankson, 2012) contempla fatores de segunda ordem independentes (e.g., inteligência fluida e cristalizada) a capturar correlações entre fatores de primeira ordem específicos (AMPs).

Ora, comparando na AFC o ajustamento aos dados empíricos de modelos hierárquicos, tais como o modelo de CHC, com o de outros modelos neles aninhados (*nested models*), tais como o modelo de CH, podemos perseguir o ideal de parcimónia e de integração das teorias que caracteriza a ciência (Forster & Sober, 1994).

Em termos práticos, através da fixação *a priori* de certos parâmetros e do constrangimento e libertação de outros desconhecidos (estimação), obtemos formalizações de MEE fatoriais hierárquicos, testando, simultaneamente, fatores de ordem superior (tal como *g*) a subsumirem outros de ordem menos abstrata (fatores mais específicos oblíquos) que, por sua vez, subsumem os respetivos indicadores observados. E isto não era possível

com a instrumentação de primeira geração. Assim, a evolução das teorias fatoriais da inteligência, desde a de Spearman até à de CHC, demonstra que elas nunca são perfeitamente corretas, dependendo, entre outros, da IADM que, numa dada época, permite a sua demonstração.

Este benefício é elucidado através de um estudo original, agora só com o *LISREL 8.80*, em que, com o mesmo ficheiro de dados ordinais, é apresentado um procedimento que culmina com o teste de um modelo fatorial hierárquico de terceira ordem (último tópico da Tabela 1), perseguindo, como já foi referido, o ideal de unificação e parcimónia das teorias que caracteriza a ciência (Forster & Sober, 1994).

3.2.1. Ilustração: Estimação e Teste de um Modelo Fatorial Hierárquico da Integração Social no Ensino Superior.

Com este estudo pretende-se testar um modelo unificador das relações intrauniversitárias (Tinto, 1975) e extrauniversitárias (Bean, 1985), comparando, com o mesmo ficheiro de dados ordinais, três estruturas fatoriais alternativas (Figura 3) do Questionário de Integração Social no Ensino Superior (QISES; Diniz, 2017): oblíqua, hierárquica (H) com dois fatores oblíquos de segunda ordem; e, H com um fator de terceira ordem (subsumindo os de 2ª ordem e estes os de 1ª ordem com os respetivos itens).

Importa referir que em casos de seleção de modelos concorrentes, a escolha do melhor modelo deve observar a ponderação de critérios ligados à plausibilidade de um modelo (i.e., validade das conclusões estatísticas do seu teste), bem como critérios ligados à sua verosimilidade (uma vez garantida a plausibilidade), tomada, como já foi referido, no sentido de valor heurístico, ou interpretativo de um modelo.

Método

Participantes

Foi constituída uma amostra de 739 estudantes universitários voluntários do 1º ano (leque etário = 17-34 anos, $Mdn = 18$), na sua maioria mulheres (65.8%), que frequentavam as Escolas de Artes ($n = 123$), de Enfermagem ($n = 38$), de Ciências Sociais ($n = 348$) e de Ciências e Tecnologias ($n = 230$) da Universidade de Évora.

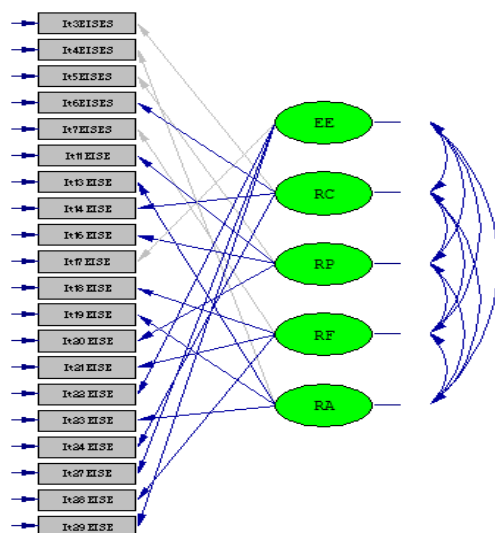
Instrumento e Procedimento

Utilizou-se o QISES (Diniz, 2017), que é composto por 20 itens com forma de resposta de Likert, igualmente distribuídos por cinco fatores: Equilíbrio Emocional, Relações com os Colegas e Relações com os Professores, representantes da experiência intrauniversitária (Tinto, 1975); e, Relações com a Família e Relações com os Amigos, representantes da experiência extrauniversitária (Bean, 1985).

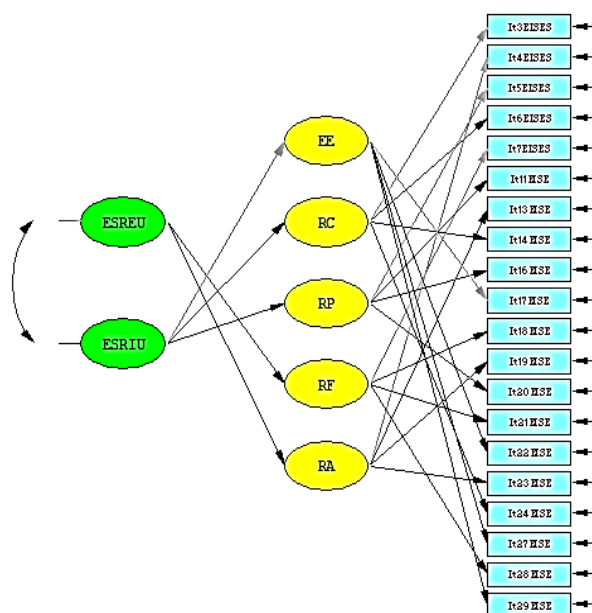
A amostragem foi não-probabilística e os dados foram recolhidos em contextos informais entre o final de outubro e o final de novembro, após obtenção do consentimento informado dos estudantes.

Análise estatística. Salvo os detalhes que a seguir se explicitam, a análise dos dados deste estudo foi a mesma da do Estudo 1. No teste dos modelos hierárquicos dos painéis B e C da Figura 3, a estimação das relações entre os fatores de segunda ordem foi realizada com a sua identificação através da fixação a um (1.00) de uma das suas trajetórias para os seus fatores de primeira ordem. No caso do modelo do painel C, como o fator de ordem superior tinha relações com somente os dois fatores de ordem inferior, a sua identificação exigiu a fixação a um (1.00) das suas trajetórias para ambos os fatores.

A



B



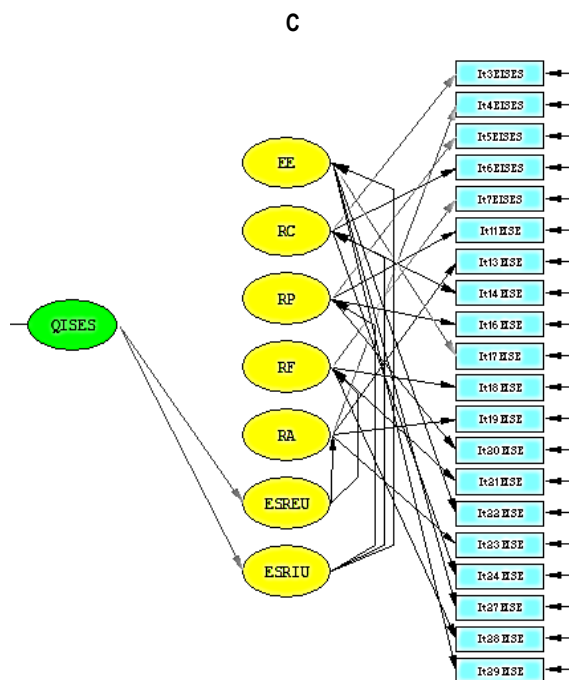


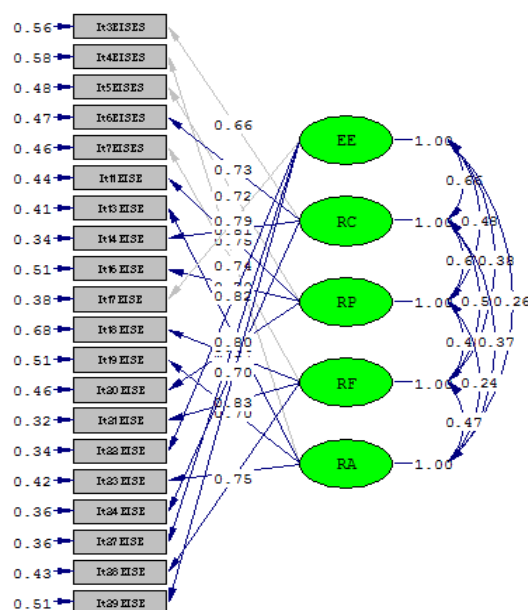
Figura 3. Modelos alternativos do Questionário de Integração Social no Ensino Superior (QISES): Diagramas conceituais. Painel A= modelo oblíquo; painel B = modelo hierárquico (H) com dois fatores oblíquos de segunda ordem; painel C = modelo H com um fator de terceira ordem. ESREU = escala de suporte relacional extrauniversitário; ESRIU = escala de suporte relacional intrauniversitário; RC = relações com colegas; RP = relações com professores; RF = relações com família; RA = relações com amigos; EE = equilíbrio emocional.

Como já foi referido no Benefício 1, o *ECVI* serve à comparação de modelos alternativos. Um dos possíveis modelos alternativos do modelo estimado é o modelo saturado (i.e., modelo em que todas as variáveis estão correlacionadas), cujo resultado, $ECVI_{Msat}$, deve ser superior ao do *ECVI* obtido para o modelo em apreciação. Existe, ainda, a possibilidade de apreciar a precisão da estimativa pontual do *ECVI* estimado, considerando um IC de 90%. Esta análise fornece-nos indicações acerca da parcimónia do modelo e da sua qualidade de aproximação na população. Se o limite superior do IC for superior ao $ECVI_{Msat}$ e se o limite inferior do IC lhe for inferior, isso indica uma razoável aproximação na

matriz de covariâncias que seria obtida numa outra amostra da mesma dimensão. Se o limite superior do IC for inferior ao $ECVI_{Msat}$, isso indica uma boa aproximação nessa matriz. Se o limite inferior do IC for superior ao $ECVI_{Msat}$, isso indica problemas de aproximação nessa matriz.

Resultados e Discussão

O modelo de um estrato fatorial com cinco fatores oblíquos (Figura 4) apresentou uma boa PPI ($RMSEA < .06$, $SRMR < .08$ e $CFI > .95$) e uma PPE aceitável para amostras com a mesma dimensão (limite inferior do IC de 90% do $ECVI < ECVI_{Msat}$).



$MVSB\chi^2 = 356.78$, $g/ = 160$; $RMSEA$ [IC 90%] = .041 [.035-.047]; $SRMR = .056$; $CFI = .987$; $ECVI$ [IC 90%] = .619 [.550-.698]; $ECVI_{Msat} = .569$

Figura 4. Modelo pentafatorial oblíquo do Questionário de Integração Social no Ensino Superior (QISES): Resultados da solução completamente estandarizada. Ver Figura 3 para abreviaturas.

Quanto às propriedades psicométricas dos fatores (Tabela 7), obteve-se uma boa VC (VME > .50) e uma boa FC para comparação entre grupos (> .80). Também se observou uma boa VD entre fatores [variância partilhada (ϕ^2) entre fatores < VME].

Tabela 7. Modelo Pentafatorial Oblíquo do QISES: Resultados Psicométricos

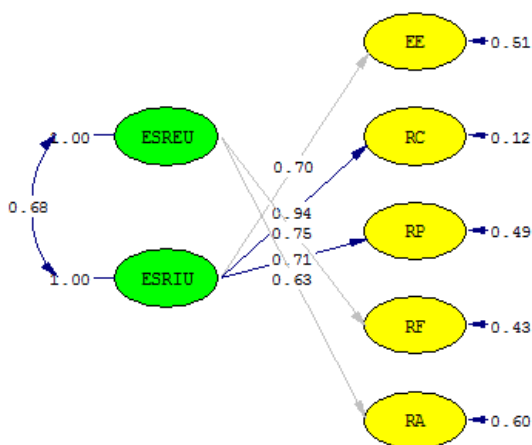
Fator	VME	FC	ϕ
Equilíbrio emocional (EE)	.61	.86	
Relações com colegas (RC)	.57	.84	
Relações com professores (RI)	.53	.82	
Relações com família (RF)	.53	.82	
Relações com amigos (RA)	.52	.81	
EE-RC			.67
EE-RP			.48
EE-RF			.38
EE-RA			.26
RC-RP			.67
RC-RF			.50
RC-RA			.37
RP-RF			.48
RP-RA			.25
RF-RA			.47

Nota. Resultados baseados na solução completamente estandardizada. VME = variância média extraída; FC = fiabilidade compósita; ϕ = correlação desatenuada.

O modelo de dois estratos com dois fatores oblíquos de segunda ordem (Figura 5) apresentou pior PPI (> *SRMR*; < *CFI*) e pior PPE (> *ECVI*) do que o pentafatorial oblíquo (Figura 4). Contudo, este seria o modelo selecionado porque tem um maior valor heurístico, acrescentando informação sobre a importância relativa dos fatores de primeira ordem para

representar as relações intrauniversitárias (ESRIU) e extrauniversitárias (ESREU) e a sua inter-relação.

Acresce ainda que, em termos absolutos, este modelo era estatisticamente significativo (PPI) e suficientemente parcimonioso (limite inferior do IC 90% do $ECVI < ECVI_{Msat}$).



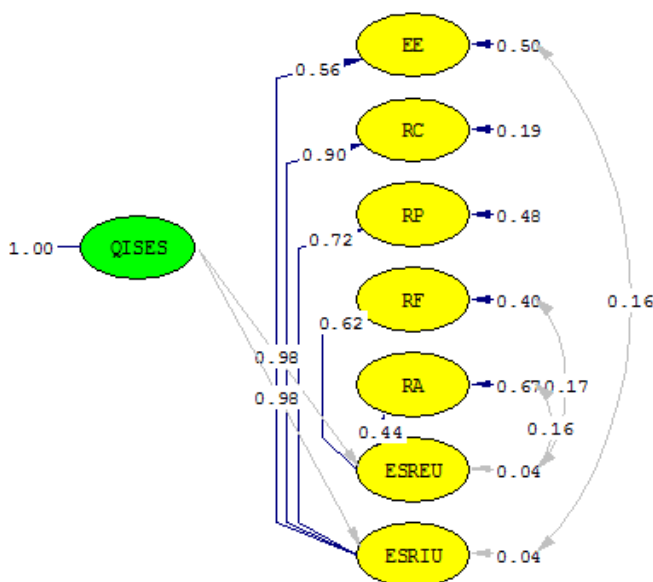
$MVSB\chi^2 = 380.36$, $gl = 165$; $RMSEA$ [IC 90%] = .041 [.036-.047]; $SRMR = .067$; $CFI = .986$; $ECVI$ [IC 90%] = .627 [.557-.709]; $ECVI_{Msat} = .569$

Figura 5. Componente estrutural do modelo bifatorial hierárquico de primeira ordem derivado do QISES: Resultados da solução completamente estandardizada. Ver Figura 3 para abreviaturas.

Quanto ao modelo de três estratos com dois fatores de segunda e um de terceira ordem que unifica as teorias de Bean (1985) e de Tinto (1975) respeitantes à integração social no ensino superior (Figura 6), ele era estatisticamente significativo (PPI) e suficientemente parcimonioso (limite inferior do IC 90% do $ECVI < ECVI_{Msat}$). Note-se que no Modelo o fator de terceira ordem subsume dois fatores de segunda ordem (e não três), pelo que a sua identificação obriga igualização dessas trajetórias.

Acresce ainda que, em termos relativos, este modelo apresentou melhor PPI ($< SRMR$) e melhor PPE ($< IC\ 90\%$ do $ECVI$) do que o modelo de dois estratos com dois fatores oblíquos de segunda ordem (Figura 5).

Conclui-se, assim, neste processo de testagem e de seleção de modelos em competição que o modelo de três estratos com dois fatores de segunda e um de terceira ordem é o melhor modelo operacionalizado pelo QISES e que através da MEE foi possível integrar as duas teorias em apreciação, perseguindo o ideal de unificação das teorias que caracteriza a ciência.



$MVSB\chi^2 = 370.72$, $gl = 164$; $RMSEA [IC\ 90\%] = .041 [.036-.047]$; $SRMR = .058$; $CFI = .986$; $ECVI [IC\ 90\%] = .627 [.556-.708]$; $ECVI_{Msat} = .569$

Figura 6. Componente estrutural do modelo unifatorial hierárquico de terceira ordem derivado do QISES: Resultados da solução completamente estandarizada. Ver Figura 3 para abreviaturas.

3.3. Benefício 3: Incremento nos Processos de Generalização de Validade

Com a IADM de segunda geração, e através do teste da invariância de estruturas de médias e covariâncias (Cheung & Rensvold, 2002; Jöreskog, 2005; Little, 1997; Little, Preacher, Selig, & Card, 2007; Meredith, 1993; Vandenberg & Lance, 2000), é possível comparar os resultados de MEE em múltiplos grupos (para uma ilustração empírica, ver Diniz, Pocinho, & Almeida, 2011; Diniz et al., 2018), e em múltiplas ocasiões (para uma ilustração empírica, ainda que somente aproximada em termos técnicos, ver Diniz & Almeida, 2006), analisando eventuais variações pontuais na sua estrutura (invariância parcial; Byrne, Shavelson, & Muthén, 1989) quer na componente de medição (funcionamento diferencial de indicadores ao nível dos pesos fatoriais, das médias ajustadas, ou *intercepts*, ou dos resíduos), quer na componente estrutural (diferenças nas variâncias, covariâncias e médias fatoriais). Assim, é possível testar um MEE quer em estudos transversais com amostras independentes, quer em longitudinais, na tentativa de assegurar a sua *validade externa* (Shadish et al., 2002), perseguindo o objetivo de generalização da teoria que ele representa (*generalização de validade*; Bagozzi & Yi, 1988).

3.4. Benefício 4: Possibilidade de Teste de Relações Não-Recursivas

Também contrariamente à IADM de primeira geração, com a de segunda geração pode-se não só testar modelos recursivos, ou seja modelos de relações entre variáveis (observadas e/ou latentes) do tipo A prediz B, mas também modelos não-recursivos onde A prediz B e, simultaneamente, B prediz A.

Para uma ilustração empírica num estudo longitudinal, ver Diniz & Almeida (2006; relações não recursivas entre B-E e EE no diagrama apresentado na Figura 7).

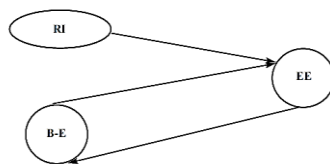


Figura 7. Modelo não-recursivo: Diagrama conceitual

3.5. Benefício 5: Possibilidade de Teste de Relações Diretas e Indiretas

Acresce, ainda, a vantagem da possibilidade oferecida pela IADM de segunda geração do teste simultâneo de relações diretas e indiretas em modelos de mediação (Iacobucci, Saldanha, & Deng, 2007), permitindo saber se um efeito indireto contribui significativamente para o efeito total. E isso pode ser feito examinando a diferença (Δ) entre o efeito total e o efeito direto (não-estandardizados) através da seguinte fórmula (Diniz, Pocinho, & Almeida, 2011): $\Delta z = \text{efeito total} - \text{efeito direto} / \text{raiz quadrada } [(EP^2_{\text{efeito total}} + EP^2_{\text{efeito direto}})/2]$. Um valor da estatística Δz superior a 1.96, $p < .05$, significa que o efeito indireto contribuiu significativamente para o efeito total.

Para uma ilustração empírica com múltiplos grupos e variáveis latentes, ver Diniz et al. (2011; diagrama conceitual apresentado na Figura 8).

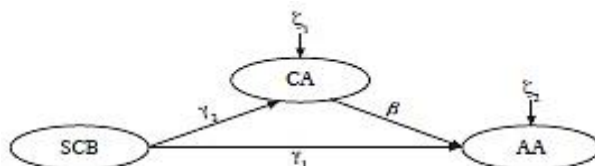


Figure 1. Path Model with Latent Variables for Sociocultural Background Influence on Academic Achievement: Conceptual diagram. SCB= sociocultural background; CA= cognitive abilities; AA= academic achievement. γ_1 = direct effects of the latent endogenous variable SCB, or latent predictor, on the latent exogenous variables, or latent criteria, AA and CH (higher SCB, higher AA and higher CA); β = direct effect between the two latent criteria (higher CA, higher AA); $\gamma_2 \times \beta$ = indirect effect of SCB on AA; ζ (random disturbance or structural residual)= amount of latent criteria's variance not accounted by latent predictor(s)

Figura 8. Modelo de mediação: Diagrama conceitual.

Para uma ilustração empírica com teste de modelos alternativos num estudo transversal, ver Alfonso et al. (2020; diagramas conceptuais apresentados na Figura 9).

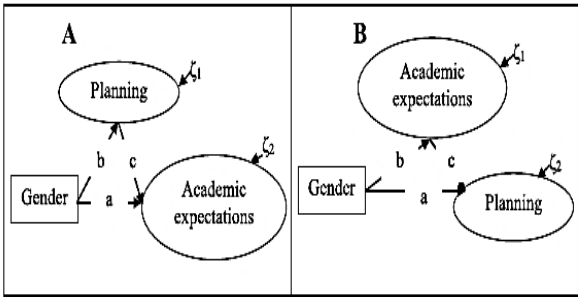


Fig. 1 Alternative mediation models. Conceptual diagrams for planning as mediator in **A** and for academic expectations as mediators in **B**. *a* = direct effect of the observed predictor (OP) on the latent criterion (LC); *b* = direct effect of the OP on the latent mediator (LM); *c* = direct effect of the LM on the LC; *b* × *c* = indirect effect of the OP on the LC; *a* + (*b* × *c*) = total effect; ζ_1 = structural residuals, random disturbances, or amount of mediators (ζ_1) and criteria's (ζ_2) variance not accounted by predictor(s)

Figura 9. Modelos de mediação alternativos: Diagramas conceptuais.

4. Nota Final

A década de 1980 trouxe uma virtuosa mudança de *paradigma* (Kuhn, 1970; conceito aqui tomado não no sentido alargado de matriz disciplinar mas, mais estritamente, de um conjunto de práticas discursivas mediadas pela instrumentação) na análise de dados multivariados em ciências sociais, nomeadamente em psicologia. O aparecimento e disseminação da IADM de segunda geração nessa década suscitaram dois benefícios primordiais: o da primazia atribuída à teoria substantiva, enquanto norteadora da investigação empírica, e o da primazia atribuída à teoria dos dados e da estimação estatística, pela importância atribuída ao erro de distribuição e de medição – aleatório ou sistemático – na interpretação dos resultados dessa investigação.

Estes dois benefícios gerais puderam ser desdobrados nos cinco casos particulares que atrás foram apresentados, dois deles ilustrados com estudos originais relativos à estimação e teste de MEE com variáveis observadas ordinais. Nestes dois estudos o foco metodológico foi colocado na operacionalização dos construtos e na correspondente adequação das variáveis observadas que os representam, sempre atendendo ao erro de medição derivado, entre outros fatores, à sua natureza métrica.

Sabemos que muitos dos testes de desempenho e de questionários possuem itens com escalas ordinais. Ora, este tipo de escalas exige uma parametrização particular para garantir a verosimilidade das conclusões estatísticas derivadas da estimação e teste de modelos que as contemplam. E aqui evoca-se o Benefício 1 e a sua ilustração, na qual se procedeu à comparação entre a AFE com o *SPSS* e a AFC com o *AMOS* e com o *LISREL*. Demonstrou-se que este último programa, através da ABLN (Jöreskog, 2005) com o robusto

χ^2 de Satorra e Bentler (1994), foi o que evidenciou o melhor desempenho, acomodando o dinamismo entre teoria, modelos e natureza dos dados.

Nesta incursão pelas virtudes da utilização da IADM de segunda geração, foi também possível demonstrar que com os fatores devidamente especificados e identificados é possível estimar a sua relação, desatenuada para o erro de medição, em modelos fatoriais hierárquicos de elevada complexidade, perseguindo o ideal de unificação das teorias, como se constatou através do Benefício 2 e respetiva ilustração.

Por último, também se destacou o potencial IADM de segunda geração para incrementar processos de generalização de validade através de estudos com grupos múltiplos e de estudos diacrónicos (ver Benefício 3), para testar simultaneamente relações de ação e retroação em modelos não-recursivos (ver Benefício 4) e para testar simultaneamente relações diretas e indiretas em modelos de mediação (ver Benefício 5).

Deve destacar-se, ainda, que os benefícios apresentados resultam do aperfeiçoamento técnico-teórico-empírico dos MEE. Melhorou-se a relação da estrutura técnica (matemática-estatística) com a estrutura conceptual (teoria substantiva) através da evidência empírica decorrente da prática de recolha e legitimação dos dados enquanto parte integrante das teorias (teste e valor heurístico).

5. Referências

- Alfonso, S., Diniz, A. M., Deaño, M., Tellado, F., García-Señorán, M., Conde, Á., & Iglesias-Sarmiento, V. (2020). Gender, planning and academic expectations in first-year higher education students: Testing two alternative mediation models. *Psicologia: Reflexão e Crítica*, 33, 1-9. doi:10.1186/s41155-020-00142-z
- Almeida, L. S. (2002). As aptidões na definição e avaliação da inteligência: O concurso da análise fatorial. *Paidéia*, 12(23), 5-17. Retirado de <http://www.scielo.br/pdf/paideia/v12n23/02.pdf>
- Amado, N. (2008). Sucesso no envelhecimento e histórias de vida em idosos sócio-culturalmente muito e pouco diferenciados (Dissertação de doutoramento não publicada). Universidade Nova de Lisboa e Instituto Superior de Psicologia Aplicada, Lisboa. Retirado de <http://hdl.handle.net/10400.12/39>
- Amado, N., & Diniz, A. M. (2017). Strength of religious faith in the Portuguese Catholic elderly. *Archive for the Psychology of Religion*, 39, 82-98. doi:10.1163/15736121-12341333
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modelling in practice: A review and recommended two step approach. *Psychological Bulletin*, 103, 411-423. doi:10.1037//00332909.103.3.411
- Babakus, E., Ferguson, C. E., & Jöreskog, K. G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumption. *Journal of Marketing Research*, 24, 222-228. doi:10.2307/3151512

- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16, 74-94. doi:10.1177/009207038801600107
- Bean, J. P. (1985). Interaction effects based on class level in an exploratory model of college student dropout syndrome. *American Educational Research Journal*, 22, 35-64. doi:10.3102/00028312022001035
- Bentler, P. M. (1985). *Theory and implementation of EQS: A structural equations program*. Los Angeles, CA: BMDP Statistical Software.
- Bentler, P. M., & Dudgeon, P. (1996). Covariance structure analysis: Statistical practice, theory, and directions. *Annual Review of Psychology*, 47, 563-592. doi:10.1146/annurev.psych.47.1.563
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Press.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). London, England: SAGE.
- Byrne, B. M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York, NY: Springer-Verlag.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456-466. doi:10.1037/00332909.105.3.456

- Cattell, R. B. (1966). The scree plot test for the number of factors. *Multivariate Behavioral Research*, 1, 140-161. doi:10.1207/s15327906mbr0102_10
- Cattell, R. B., & Vogelmann, S. (1977). A comprehensive trial of the scree and Kg criteria for determining the number of factors. *Multivariate Behavioral Research*, 12, 289-325. doi:10.1207/s15327906mbr1203_2
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233-255. doi:10.1207/S15328007SEM0902_5
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7), 1-8. Retrieved from <http://pareonline.net/pdf/v10n7.pdf>
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1, 16-29. doi:10.1037/1082-989X.1.1.16
- Diniz, A. M. (2017). Questionário de Integração Social no Ensino Superior (QISES). In L. Almeida, M. R. Simões, & M. Gonçalves (Coords.), *Adaptação, desenvolvimento e sucesso académico dos estudantes do ensino superior: Instrumentos de avaliação* (pp. 21-32). Retirado de <http://www.adipsieduc.pt/wpcontent/uploads/2014/12/9789899951716.pdf>

- Diniz, A. M., Alfonso, S., Araújo, A. M., Deaño, M., Costa, A. R., Conde, Á., & Almeida, L. S. (2018). Gender differences in first-year college students' academic expectations. *Studies in Higher Education*, 43, 689-701. doi:10.1080/03075079.2016.1196350
- Diniz, A. M., & Almeida, L. S. (2006). Adaptação à Universidade em estudantes do primeiro ano: Estudo diacrónico da interacção entre o relacionamento com pares, o bem-estar pessoal e o equilíbrio emocional. *Análise Psicológica*, 24(1), 29-28.
- Diniz, A. M., & Amado, N. (2005). Escala de Afetos Positivos e Negativos (EAPN). Instrumento não publicado, derivado da Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988).
- Diniz, A. M., & Amado, N. (2014). Procedures for successful data collection through psychological tests in the elderly. *Psicologia: Reflexão e Crítica*, 27, 491-497. doi:10.1590/1678-7153.201427309
- Diniz, A. M., Pocinho, M. D., & Almeida, L. S. (2011). Cognitive abilities, sociocultural background and academic achievement. *Psicothema*, 23(4), 695-700. Retrieved from <http://www.psicothema.com/pdf/3943.pdf>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272-299. doi:10.1037/1082-989X.4.3.272
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.) [Adobe Digital Editions version]. London, UK: Sage.

- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9, 466-491. doi:10.1037/1082-989X.9.4.466
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini- Mental State Examination: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198. doi:10.1016/0022-3956(75)90026-6
- Fornell, C. (1982). A second generation of multivariate analysis: An overview. In C. Fornell (Ed.), *A second generation of multivariate analysis: An overview* (Vol. 1, pp. 1-21). New York: Praeger Publishers.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39-50. doi:10.2307/3151312
- Forster, M. R. (2002). Predictive accuracy as an achievable goal of science. *Philosophy of Science*, 69, S124-S134. doi:10.1086/341840
- Forster, M., & Sober, E. (1994). How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *British Journal for the Philosophy of Science* 45(1), 1-35. Retrieved from <http://www.jstor.org/stable/687960>
- Gigerenzer, G. (2018). Statistical rituals: The replication delusion and how we got there. *Advances in Methods and Practices in Psychological Science*, 1, 198-218. doi:10.1177/2515245918771329

- Gigerenzer, G., Krauss, S., & Vitouch, O. (2004). The null ritual: What you always wanted to know about null hypothesis testing but were afraid to ask. In D. Kaplan (Ed.), *Handbook on quantitative methods in the social sciences* (pp. 391-408). Thousand Oaks, CA: Sage.
- Hershberger, S. L. (2003). The growth of structural equation modeling: 1994 to 2001. *Structural Equation Modeling: A Multidisciplinary Journal*, 10, 35-46. doi:10.1207/S15328007SEM1001_2
- Horn, J. L., & Blankson, A. N. (2012). Foundations for better understanding of cognitive abilities. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 73-98). New York, NY: Guilford Press.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparametrized model misspecification. *Psychological Methods*, 3, 424-453. doi:10.1037/1082-989X.3.4.424
- Iacobucci, D., Saldanha, N., & Deng, X. (2007). A meditation on mediation: Evidence that structural equations models perform better than regressions. *Journal of Consumer Psychology*, 17, 139-153. doi:10.1016/S1057-7408(07)70020-7
- Jöreskog, K. G. (1973). A general method for estimating a linear structural equation system. In A. Goldberger, & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 85-112). New York, NY: Academic Press.
- Jöreskog, K. G. (2005). *Structural equation modeling with ordinal variables using LISREL*. Retrieved from: [http:// www.ssicentral.com/lisrel/techdocs](http://www.ssicentral.com/lisrel/techdocs)

- Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, 70, 631-639. doi:10.1080/01621459.1975.10482485
- Jöreskog, K. G., & Sörbom, D. (1986). PRELIS: A program for multivariate data screening and data summarization. Mooresville, IN: Scientific Software.
- Jöreskog, K. G., & Sörbom, D. (1988). LISREL 7 [Computer software]. Chicago, IL: Scientific Software.
- Jöreskog, K. G., & Sörbom, D. (1993). LISREL 8: Structural equation modeling with the SIMPLIS command language. Chicago, IL: Scientific Software.
- Jöreskog, K. G., & Sörbom, D. (1996). PRELIS 2: User's reference guide. Chicago, IL: Scientific Software.
- Jöreskog, K. G., & Sörbom, D. (2006). LISREL 8.80 for Windows [Computer software]. Chicago, IL: Scientific Software International.
- Keiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141-151. doi:10.1177/001316446002000116
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4rd ed.). New York, NY: Guilford Press.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed., enlarged). Chicago, IL: University of Chicago Press.

- Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research*, 32, 53-76. doi:10.1207/s15327906mbr3201_3
- Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development*, 31, 357-365. doi:10.1177/0165025407077757
- Loehlin, J. C. (1987). *Latent variable models: An introduction to factor, path, and structural analysis*. Hillsdale, NJ: Erlbaum.
- MacCallum, R. C. (2003). Working with imperfect models. *Multivariate Behavioral Research*, 38, 113-139. doi:10.1207/S15327906MBR3801_5
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modification in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111, 490-504. doi:10.1037/0033-2909.111.3.490
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11, 320-341. doi: 10.1207/s15328007sem1103_2
- Marôco, J. (2011). *Análise estatística com o SPSS Statistics (5a ed.)*. Pero Pinheiro: ReportNumber.

- Meehl, P. E. (1990). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry*, 1, 108-141. doi:10.1207/s15327965pli0102_1
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525-543. doi:10.1007/BF02294825
- Morgado, J., Rocha, C. S., Maruta, C., Guerreiro, M., & Martins, I. P. (2009). Novos valores normativos do Mini-Mental State Examination. *Sinapse*, 9, 10-16. doi:10.1111/j.1468-1331.2009.02907.x
- Muthén, B. (1987). LISCOMP: Analysis of linear structural equations with a comprehensive measurement model. Theoretical integration and user's guide. Mooresville, IN: Scientific Software.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32, 396-402. doi:10.3758/BF03200807
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye, & C.C. Clogg (Eds.), *Latent variable analysis* (pp. 399-419). Thousand Oaks, CA: SAGE.

- Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist*, 5, 165-173. doi:10.1300/J018v05n01_09
- Schumacker, R. E., & Lomax, R. G. (2010). A beginner's guide to structural equation modelling (3rd ed.). New York, NY: Routledge.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, MA: Houghton Mifflin.
- Schnelder, W. J., & McGrew, K. S. (2012). The Cattell–Horn–Carroll model of intelligence. In D. Flanagan, & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99-144). New York, NY: Guilford Press.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.
- Tinto, V. (1975). Dropouts from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45, 89-125. doi:10.3102/00346543045001089
- Tumas, V., Rodrigues, G. G. R., Farias, T. L. A., & Cripa, J. A. S. (2008). The accuracy of diagnosis of major depression in patients with parkinson's disease: A comparative study among the UPDRS, the Geriatric Depression Scale and the Beck Depression Inventory. *Arquivos de Neuro-psiquiatria*, 66, 152-156. doi:10.1590/S0004-282X2008000200002
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for

organizational research. *Organizational Research Methods*, 3, 4-69.
doi:10.1177/109442810031002

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology*, 54, 1063-1070. doi:10.1037/0022-3514.54.6.1063

Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic nomothetic analysis. *Journal of Personality and Social Psychology*, 43, 111-112.
doi:10.1037/0022-3514.43.1.111



cieplue

CENTRO DE INVESTIGAÇÃO EM EDUCAÇÃO E PSICOLOGIA
DA UNIVERSIDADE DE ÉVORA



UNIVERSIDADE DE ÉVORA
INSTITUTO DE INVESTIGAÇÃO
E FORMAÇÃO AVANÇADA



UNIVERSIDADE
DE ÉVORA

FCT

Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR