

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/233735561>

O Uso da Análise Fatorial: Algumas Diretrizes para Pesquisadores

Chapter · January 2012

CITATIONS

81

READS

11,125

1 author:



Jacob Arie Laros

University of Brasília

89 PUBLICATIONS 568 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Normatization and validation of the SON-R nonverbal intelligence tests [View project](#)



Acompanhamento de sistemas de avaliação educacional em larga escala. [View project](#)

Capítulo 7

O Uso da Análise Fatorial: Algumas Diretrizes para Pesquisadores

Jacob A. Laros, PhD, UnB

Introdução

Análise fatorial é um dos procedimentos psicométricos mais freqüentemente utilizados tanto na construção, quanto na revisão e avaliação de instrumentos psicológicos, como no desenvolvimento de teorias psicológicas. A análise fatorial é particularmente útil quando aplicada a escalas que consistem de uma grande quantidade de itens utilizados para medir personalidade, estilos de comportamento ou atitudes. Uma outra utilidade de análise fatorial é na verificação da unidimensionalidade, um pressuposto central da Teoria de Resposta ao Item (TRI). O pressuposto da unidimensionalidade implica que todos os itens de um instrumento estejam medindo um único construto. Se o instrumento está de fato medindo mais do que um fator, um escore total individual deve ser calculado para cada fator e todas as análises estatísticas subseqüentes devem ser feitas independentemente para cada fator.

O ponto central de partida na análise fatorial é o princípio da parcimônia: um grande número de variáveis observadas pode ser explicado por um número menor de variáveis hipotéticas, não-observadas. Estas variáveis hipotéticas, também chamadas de fatores, são responsáveis pelo fato de as variáveis observadas correlacionarem entre si. Em relação ao princípio de parcimônia, Kerlinger (1986) tem caracterizado a análise fatorial como um dos métodos psicométricos mais poderosos para reduzir a complexidade de uma grande quantidade de variáveis a uma estrutura relativamente simples, consistindo de um número menor de fatores. Os fatores são combinações lineares de variáveis observadas.

Um outro uso fundamental da análise fatorial ocorre no processo da validação de instrumentos psicológicos. Neste processo, a técnica de análise fatorial é imprescindível. A grande maioria dos pesquisadores em ciências comportamental e social reconhece o relacionamento próximo entre análise fatorial e questões de validade. Análise fatorial e validade de construto têm sido associadas durante muito tempo. A este respeito, Nunnally (1978) fez a seguinte observação: *“Análise fatorial é intimamente relacionada com as questões de validade de instrumentos psicológicos. Análise fatorial está no coração da mensuração dos construtos psicológicos”*. De fato, validade fatorial é o nome histórico para indicar o que hoje chamamos de validade de construto.

Apesar da popularidade da análise fatorial, tanto a complexidade matemática dos procedimentos analíticos quanto a flexibilidade da técnica contribuíram para que esta, hoje em dia, seja uma das técnicas menos entendidas por psicólogos e estudantes de psicologia. Existe uma ambigüidade considerável sobre práticas de pesquisa apropriadas em relação à análise fatorial. O presente capítulo tem como objetivo diminuir esta ambigüidade e oferecer dicas sobre práticas de pesquisa apropriadas.

Sempre que um pesquisador estiver realizando uma análise fatorial, ele terá que tomar várias decisões importantes sobre os seguintes tópicos: (1) a natureza e o tamanho da amostra que formará a base de dados da análise fatorial; (2) a adequação das distribuições das variáveis que entram numa análise fatorial; (3) a seleção de variáveis a serem submetidas à análise fatorial; (4) o uso de correlações corrigidas por atenuação ou o uso de correlações não corrigidas; (5) o tipo de correlações a ser utilizado; (6) a inclusão de variáveis marcadoras (*marker variables*) entre as variáveis a serem analisadas; (7) o número de fatores a serem extraídos; (8) a utilização da análise fatorial exploratória ou da análise fatorial confirmatória; (9) o tipo de análise fatorial a ser utilizado para extrair os fatores; (10) o procedimento de rotação a ser usado a fim de direcionar os fatores; (11) a interpretação dos resultados da análise fatorial; (12) a investigação de uma solução hierárquica; (13) a necessidade e a forma de calcular os escores fatoriais; (14) a investigação de validade cruzada da solução fatorial encontrada; (15) a utilização de grupos no estudo de invariância da estrutura fatorial encontrada; e (16) a seleção dos resultados mais relevantes da análise fatorial para a publicação da pesquisa.

Este capítulo tem por objetivo esclarecer as decisões a serem tomadas ao executar uma análise fatorial, de modo que o pesquisador possa tomá-las de forma mais adequada e melhor fundamentada. Várias questões relacionadas ao uso de análise fatorial serão discutidas a fim de que os pesquisadores possam tornar-se mais informados sobre as opções de escolha. Além disso, espera-se demonstrar que o uso de técnicas analíticas particulares requer justificativas; que existem padrões norteadores para a escolha dessas técnicas; e que as técnicas utilizadas em qualquer estudo merecem ser relatadas de forma suficientemente detalhada, de modo que os leitores interessados possam fazer interpretações mais completas dos resultados. Antes disso, entretanto, serão discutidos alguns usos e métodos de análise fatorial para ajudar o leitor a entender melhor os seus objetivos.

Usos básicos e métodos de análise fatorial

Discutindo-se análise fatorial, faz-se necessário distinguir o seu uso do método empregado. A questão do uso da análise fatorial diz respeito ao objetivo da análise, enquanto a questão da técnica ou método refere-se aos procedimentos matemáticos e estatísticos particulares utilizados para alcançar esses objetivos. Dois tipos de uso da análise fatorial podem ser identificados: o exploratório e o confirmatório. De maneira semelhante, dois tipos de técnicas de análise fatorial podem ser distinguidos: procedimentos exploratórios e procedimentos confirmatórios.

A fim de avaliar construtos psicológicos, a análise fatorial utiliza dois tipos de uso exploratório que estão alinhados com os objetivos de explicação e redução dos dados. O primeiro uso, relacionado à explicação, consiste em identificar as dimensões subjacentes de um determinado domínio (por exemplo: personalidade, criatividade, inteligência) que o instrumento em questão está medindo. Neste caso, análise fatorial é levada a efeito, a fim de identificar as dimensões subjacentes que representam os construtos teóricos do instrumento. Este procedimento é exploratório porque, presumivelmente, o investigador não tem qualquer expectativa firme *a priori* baseada em teoria ou pesquisa anterior acerca da composição de subescalas. Assim, a análise fatorial é utilizada para descobrir as variáveis latentes que estão subjacentes à escala. Para alcançar este fim, a análise fatorial usa a matriz de correlações ou a de covariâncias entre as variáveis mensuradas. Na teoria, essas variáveis latentes são as causas subjacentes das variáveis mensuradas. A análise fatorial produz cargas fatoriais, as quais podem ser consideradas pesos de regressão das variáveis mensuradas para prever o construto subjacente. Nos casos onde existe mais de um fator subjacente aos dados, a análise fatorial também produz correlações entre os fatores. Além disso, divide a variância de cada variável mensurada em duas

partes: (a) variância comum, ou variância associada às variáveis latentes, que é estimada com base na variância partilhada com outras variáveis mensuradas na análise, e (b) variância única, que é a combinação da variância fidedigna específica para a variável mensurada e a variância do erro randômico na mensuração da variável. Assim, *análise fatorial exploratória* (AFE), também chamada de *análise fatorial comum* (AFC), pode proporcionar valiosas informações sobre a estrutura multivariada de um instrumento de mensuração, identificando os construtos teóricos.

O segundo uso de análise fatorial exploratória é relacionado à redução de dados. O objetivo da redução de dados é descobrir ponderações ótimas para as variáveis mensuradas, de forma que um grande conjunto de variáveis possa ser reduzido a um conjunto menor de índices sumários que tenham máxima variabilidade e fidedignidade. A meta da redução de dados é tipicamente atingida pelo uso da análise dos componentes principais (ACP) e não pelo uso da análise fatorial comum (AFC). Uma diferença fundamental entre os dois métodos é que a ACP trabalha com a variância total observada, enquanto a AFC trabalha somente com a variância partilhada dos itens (variância erro e variância única são excluídas). Em outras palavras, a ACP tem como ponto de partida a variância observada das variáveis e a AFC a covariância observada entre as variáveis. Na AFC, os fatores são estimados para explicar as covariâncias entre as variáveis observadas, portanto os fatores são considerados como as causas das variáveis observadas. Em contraste, na ACP, os componentes são estimados para representar a variância das variáveis observadas de uma maneira tão econômica quanto possível. Os componentes principais são somas otimamente ponderadas das variáveis observadas, neste sentido, as variáveis observadas são consideradas as causas dos componentes principais.

Em contraste à pura exploração, a análise fatorial é freqüentemente utilizada para confirmar hipóteses de vários tipos. Pesquisadores usualmente

geram hipóteses com relação a fatores que devem ser representados em um dado domínio de inquérito. Essas hipóteses são baseadas na teoria ou nos resultados de estudos empíricos prévios. A análise fatorial resultante tipicamente revela alguns dos construtos esperados e, outras vezes, revela fatores adicionais não esperados. Se esses fatores adicionais puderem ser significativamente interpretados e demonstrarem ser fenômenos fidedignos e replicáveis, então a sua identificação poderá contribuir substancialmente para o entendimento do domínio da pesquisa e interpretação dos escores da escala. Como um procedimento confirmatório, a análise fatorial é primariamente um método utilizado para avaliar a validade de construto das medidas e não para a redução de dados. A validade de construto é reforçada se a estrutura fatorial da escala é consistente com os construtos que o instrumento propõe medir. Se a análise fatorial falha na detecção dos construtos subjacentes ou se os construtos detectados são inconsistentes com as expectativas, a validade de construto da escala é comprometida.

A confirmação de uma estrutura fatorial hipotetizada é mais adequadamente estabelecida com as técnicas de análise fatorial confirmatória, que constitui um caso especial da técnica estatística denominada Modelagem por Equação Estrutural (Bentler, 1989; Jöreskog & Sörbom, 1989). Na análise fatorial confirmatória é testado se a estrutura fatorial teórica se adequa aos dados observados. De maneira adicional, a análise fatorial confirmatória permite testar o ajuste relativo a modelos fatoriais concorrentes. Este tipo de análise é de grande valor no processo de revisão e refinamento de instrumentos psicológicos e suas estruturas fatoriais. Algumas de suas extensões permitem a investigação da invariância da estrutura fatorial encontrada em grupos ou amostras diferentes (Reise, Widaman & Pugh, 1993). A análise fatorial confirmatória pode ser utilizada também para testar a validade convergente e discriminante dos fatores utilizando o método *multitraço-multimétodo* (Kenny & Kashy, 1992; Widaman, 1985). Os procedimentos analíticos implicados atualmente pelo termo *análise*

fatorial confirmatória foram desenvolvidos, em grande parte, nos últimos 20 anos (Jöreskog, 1969, 1971). Até a década de 70, os pesquisadores utilizaram técnicas da análise fatorial exploratória para alcançar tanto fins exploratórios como objetivos confirmatórios. Atualmente, as novas técnicas de análise fatorial confirmatória podem atingir os mesmos objetivos ao reproduzir a estrutura fatorial e confirmar uma teoria. Ela possibilita hoje a testagem de hipóteses cruciais, o que não era possível com o uso das técnicas analíticas exploratórias. Uma vez tendo apresentado, brevemente, alguns usos básicos e métodos da análise fatorial, a seguir serão discutidas algumas relevantes decisões que um pesquisador, executando uma análise fatorial, precisa tomar.

1 – Decisão sobre a natureza e o tamanho de amostra

Quanto mais heterogênea uma amostra, em relação às variáveis que estão sendo mensuradas, mais altas as correlações encontradas entre os escores de teste (a variância tem uma relação positiva com a correlação). Uma possível consequência disto é que, na análise fatorial, o contraste entre correlações relativamente altas e baixas aumentará, tornando, com isso, a estrutura fatorial mais evidente. Portanto, ao se utilizar a técnica de análise fatorial, melhores resultados são obtidos com amostras heterogêneas do que com amostras homogêneas. O pesquisador precisa estar consciente deste fato. Assim, recomenda-se aos pesquisadores, que pretendem utilizar análise fatorial, o uso de amostras apropriadamente heterogêneas que representem toda a extensão da população alvo (Clark & Watson, 1995). Quando uma escala é designada para uso clínico, por exemplo, é de grande importância utilizar, desde o início da pesquisa, amostras clínicas, em vez de basear-se somente em amostras de estudantes universitários.

Sobre o tamanho mínimo de uma amostra na análise fatorial, Crocker e Algina (1986) indicam a regra geral de usar 10 sujeitos por variável, com um mínimo de 100 sujeitos na amostra total. Uma outra regra é de Gorsuch (1983) que declarou que na análise fatorial a amostra deve conter pelo menos 5 participantes por variável e uma amostra total de pelo menos 200 sujeitos. Guadagnoli e Velicer (1988), porém, desafiaram o critério de Gorsuch e argumentaram que nenhuma base teórica ou empírica existe para recomendações de relação entre o número de participantes e o número de variáveis. A partir de um estudo tipo *Monte Carlo*, os autores sugerem que o tamanho desejado de uma amostra depende do tamanho das cargas fatoriais obtidas. Com cargas fatoriais em torno de 0,80, obtêm-se soluções fatoriais altamente estáveis em amostras de 50 pessoas. Quando as cargas fatoriais estão ao redor de 0,40, porém, amostras de 300 a 400 sujeitos são necessárias para atingir soluções estáveis.

Um outro autor, Wolins (1995), que desafia o critério de Gorsuch e o de Crocker e Algina, afirma que não existe um tamanho de amostra mínimo para efetuar uma análise fatorial com um determinado número de variáveis. Segundo ele, é incorreto supor que estudos fatoriais que envolvem um grande número de variáveis requerem amostras maiores do que estudos com menos variáveis.

Em relação ao tamanho da amostra requerido para a análise fatorial, Pasquali (1999) relata, como regra geral, um mínimo de 100 sujeitos por fator medido. Comrey e Lee (1992) classificam, em relação à análise fatorial, amostras de 50 como muito inferiores, de 100 como inferiores, de 200 como razoáveis, de 300 como boas, de 500 como muito boas e de 1.000 ou mais como excelentes.

2 – Decisão sobre a adequação das distribuições das variáveis observadas

Construtores de testes ou escalas deveriam examinar as distribuições dos itens individuais, antes levar a efeito uma análise fatorial ou qualquer outra análise estrutural complexa. Ao inspecionar estas distribuições, duas considerações são fundamentais.

A primeira consideração refere-se à importância de identificar e, possivelmente, excluir itens que possuem distribuições extremamente assimétricas e desequilibradas. Em escalas com um formato *verdadeiro-falso*, estes são os itens que quase todo mundo (por exemplo, 95% ou mais) endossa ou nega; em escalas com um formato tipo *Likert*, estes são os itens para os quais a grande maioria dos respondentes escolhe a mesma opção (por exemplo, "concordo completamente"). Itens altamente desequilibrados são indesejáveis por várias razões. Em primeiro lugar, quando a maioria dos sujeitos responde de maneira semelhante, os itens contêm pouca informação. Em segundo lugar, devido à variabilidade limitada, estes itens provavelmente mostram uma correlação fraca com os outros itens do conjunto e, portanto, terão um pobre desempenho nas análises fatoriais subseqüentes. Em terceiro lugar, itens com distribuições extremamente desequilibradas podem produzir correlações altamente instáveis. Porém, antes de excluir definitivamente um item com base em uma distribuição desequilibrada, faz-se necessário examinar os dados de diversas amostras da população alvo. Uma primeira ponderação que o pesquisador deve fazer refere-se à desejabilidade da retenção de itens que contêm informação importante para avaliar o construto sob investigação, mesmo se estes itens possuem distribuições extremamente desequilibradas (e propriedades psicométricas relativamente pobres). A segunda consideração refere-se à desejabilidade de se reter itens que mostram uma grande amplitude

de distribuições em grupos diferentes. É importante reter itens que discriminam em diferentes pontos ao longo de toda a escala. Uma das vantagens essenciais oferecidas pela Teoria de Resposta ao Item (TRI) é que as estimativas dos parâmetros relatam em que ponto da escala um determinado item é maximamente informativo. Estas estimativas podem ser usadas como base para a escolha de um conjunto eficiente de itens que tornam o instrumento preciso em toda a escala, isso quer dizer, para todos os possíveis valores da escala, sejam baixos, médios ou altos.

A segunda consideração fundamental em relação ao exame das distribuições dos itens individuais, refere-se à normalidade das distribuições. Idealmente, a análise fatorial é aplicada aos dados que são distribuídos com normalidade multivariada. É importante ressaltar que a normalidade multivariada é uma suposição razoável somente nos casos em que cada variável apresenta normalidade univariada. Normalidade multivariada é uma suposição rígida, necessária somente para a execução de certos tipos de métodos de estimação de parâmetros, como, por exemplo, o método de máxima verossimilhança (*maximum likelihood*). A estipulação de que os dados utilizados numa análise fatorial apresentem normalidade multivariada pode parecer muito rigorosa para a maioria dos investigadores, uma vez que a análise fatorial dos eixos principais (*Principal Axis Factoring*), procedimento mais comumente utilizado na análise fatorial exploratória, não requer a suposição de normalidade multivariada. Não obstante, todos os métodos de análise fatorial provavelmente produziriam estruturas fatoriais mais claras, mais replicáveis, se os dados apresentassem normalidade multivariada. Ainda assim, cabe destacar que, na prática, tanto a análise fatorial exploratória como confirmatória parecem ser relativamente robustas contra violações de normalidade (Gorsuch, 1983). Porém, um estudo de tipo *Monte Carlo* de análise fatorial confirmatória realizado por Hu, Bentler e Kano (1992) mostrou a dificuldade de se obter soluções fatoriais confirmatórias aceitáveis quando distribuições não-normais

ocorrem juntamente com outras violações de pressupostos como, por exemplo, a falta de independência entre as variáveis e os erros de mensuração e um número insuficiente de pessoas na amostra.

3 – Decisão a respeito da seleção das variáveis a serem submetidas à análise fatorial

É importante ressaltar que na análise fatorial, como, aliás, em qualquer outra análise quantitativa, nada que não tenha sido incluído anteriormente no instrumento pode aparecer nos resultados. Com a seleção das variáveis que são submetidas à análise fatorial, o pesquisador está limitando as possibilidades de explicar os fatores. Os críticos de análise fatorial expressaram esta constatação como *"Lixo entra, lixo sai"*. A exigência mais básica para uso otimizado de análise fatorial é que os dados sejam mensurados em nível de mensuração intervalar ou *quase-intervalar*. O modo mais direto para assegurar a qualidade dos dados é levar a efeito uma seleção de itens e uma cuidadosa análise exploratória de itens (Haynes, Richard & Kubany, 1995). As propriedades psicométricas das variáveis que são submetidas a uma análise fatorial inspiram grande preocupação, considerando-se o fato de que as variáveis individuais tendem a ser menos fidedignas que escalas que consistem de múltiplas variáveis. A fase inicial, na qual se determinam quais itens serão excluídos e quais serão retidos no conjunto de itens, é um momento crucial da análise.

Uma análise prévia ou exploratória de itens precisa ser executada de modo a assegurar que os itens construídos para medir um determinado construto comum tenham correlações moderadas entre eles. Se um item não mostra, pelo menos, uma correlação moderada (por exemplo, $r = 0,20$) com os outros itens do construto, o item provavelmente terá um desempenho pobre numa análise fatorial. Porém, este processo de verificação das correlações individuais entre os

itens de um instrumento, torna-se muito trabalhoso conforme o número de itens da escala aumenta. Por exemplo, um conjunto de 30 itens gera 435 correlações a serem inspecionadas, e um conjunto de 40 itens, 780. No caso de um elevado número de itens a ser analisado, os psicometristas recomendam fortemente o uso de análise fatorial na análise prévia de itens (Comrey, 1988; Cortina, 1993; Floyd & Widaman, 1995). Neste caso, a análise fatorial é utilizada para identificar os itens que mostram uma relação muito fraca com o construto alvo; estes itens serão excluídos nas análises fatoriais subseqüentes.

Outro assunto importante relacionado à qualidade dos dados diz respeito ao formato de resposta a ser utilizado. Claramente, as duas formas de resposta dominantes em testes contemporâneos de personalidade e na área clínica são o formato dicotômico, com duas opções (por exemplo, verdadeiro-falso) e o formato politômico (como, por exemplo a escala *Likert*), com três ou mais opções. Problemas com análise fatorial de variáveis dicotômicas já foram reconhecidos nos anos quarenta, com a identificação de "fatores de dificuldade" que só eram devidos à variação na dificuldade dos itens e que não mostravam uma relação com o(s) construto(s) subjacente(s) investigado(s). Embora itens dicotômicos possam ser submetidos a uma análise fatorial usando técnicas padrão, os resultados podem ser enviesados. Soluções analíticas adequadas foram desenvolvidas para estes problemas de viés nos anos setenta, mas devem ser utilizados programas de *software* especiais que aplicam estas técnicas. Os leitores interessados podem verificar programas como, por exemplo, *TESTFACT* (Wilson, Wood, & Gibbons, 1991) e *NOVAX* (Waller, 1994), com os quais é possível se extrair fatores sem viés usando dados dicotômicos. Um segundo método para lidar com itens dicotômicos é calcular escores que são as somas de dois ou mais itens semelhantes. Tais escores, chamados de parcelas de itens (*item parcels*, veja Kishton & Widaman, 1994), não são dicotômicos e, assim, são mais imunes ao problema de isolar fatores de dificuldade.

4 – Decisão sobre a inclusão das variáveis marcadoras na análise fatorial

Na fase inicial de coleta de dados, é prática comum administrar o conjunto de itens preliminar sem qualquer item ou escala adicional. Esta prática é lamentável porque não permite o exame dos limites do construto alvo. Explorar estes limites é absolutamente crítico para entender o construto do ponto de vista teórico e empírico. Portanto, deveriam ser incluídos escalas ou itens marcadores na coleta inicial de dados (Clark & Watson, 1995). Muito freqüentemente construtores de testes descobrem muito tarde, no processo de validação, que a escala desenvolvida tem uma correlação de 0,85 com uma outra medida existente. Uma análise fatorial bem planejada pode desempenhar um papel decisivo na melhoria da validade discriminante de uma nova medida. Por exemplo, muitas escalas novas na área da psicologia clínica têm uma validade discriminante fraca, porque elas não se diferenciam claramente do amplo construto de neuroticismo. O modo mais fácil de evitar a criação de mais uma medida de neuroticismo consiste em sujeitar os itens da nova escala – junto com um número aproximadamente igual de itens de neuroticismo – a uma análise fatorial comum. Nesta instância, o pesquisador extrairia dois fatores e os rotaria a fim de obter uma estrutura simples (por exemplo, usando *Varimax* ou *Promax*). Idealmente, todos os itens da nova escala (mais freqüentemente somente um subconjunto da escala) mostrarão cargas altas em um fator, e os itens de neuroticismo carregarão altamente no outro fator. Se isto não ocorre, a nova escala aparentemente está medindo o construto neuroticismo e não o construto proposto. Caso contrário, os itens que carregam fortemente no fator da escala em desenvolvimento, mas bastante debilmente no fator de neuroticismo, são candidatos excelentes para retenção. Contudo, os itens com cargas relativamente altas no fator de neuroticismo têm validade discriminante pobre e provavelmente deveriam ser excluídos. Este procedimento pode ser seguido no desenvolvimento de qualquer escala de avaliação de um construto

que pretende ser diferente de uma escala já existente. Para isso, itens marcadores que avaliam o construto, do qual deve se diferenciar, precisam ser incluídos na coleta inicial de dados.

5 – Decisão sobre o número de fatores a extrair

A seleção do correto número de componentes para retenção na análise fatorial é um passo crucial na construção de instrumentos psicométricos e na elaboração de teorias psicológicas. Em relação a esta questão, Zwick e Velicer (1986) afirmaram que *"A determinação do número de componentes ou fatores a extrair é provavelmente a decisão mais importante que um investigador, executando análise fatorial, tomará"*. A ocorrência de erros nesta fase afetará a interpretação de todos os resultados subseqüentes. A este respeito, Velicer e Jackson (1990) mostraram que a *superextração* (a extração de um número maior de fatores dos que realmente existem) é um erro tão sério como a *subextração* (a extração de um número menor de fatores dos que realmente existem); ambos levarão a resultados e conclusões distorcidas. Os autores afirmaram que tanto a superestimação quanto a subestimação do número de fatores são provavelmente as principais causas das discrepâncias entre ACP (Análise de Componentes Principais) e AFC (Análise Fatorial Comum). A extração do correto número de fatores é essencial e, somente nesta instância, os diferentes métodos de análise fatorial levam às mesmas conclusões.

Existem vários critérios para determinar o número correto de fatores a serem extraídos. Os critérios mais comumente considerados são: (1) o critério de autovalor maior do que 1,0 de Guttman-Kaiser (Guttman, 1954; Kaiser, 1960); (2) o critério baseado no teste de qui-quadrado de Bartlett (Bartlett, 1950); (3) o teste *scree* de Cattell (Cattell, 1966); (4) o critério da média mínima de correlações parciais de Velicer (Velicer, 1976); e (5) o critério de análise

paralela de Horn (Horn, 1965). Estes 5 critérios mais utilizados serão brevemente explicados a seguir.

1) O critério de Guttman-Kaiser (GK – denominado de K-1 no SPSS).

O critério de autovalor (também chamado de *eigenvalue*) maior do que 1,0 (Guttman, 1954; Kaiser, 1960) é o mais comumente utilizado, dado que é a opção padrão nos pacotes estatísticos como SPSS, SAS e BMDP. O autovalor corresponde à quantidade da variância explicada por um componente, sendo que um autovalor igual a 1,0 representa a totalidade de porcentagem da variância explicada por uma única variável. A soma da quantidade de autovalores corresponde ao número de variáveis analisadas. O critério GK é baseado na consideração de que um fator precisa explicar pelo menos a quantidade de variância que é explicada por uma única variável. Gorsuch (1983) relatou que o número de componentes retidos pelo critério GK normalmente se encontra entre 1/3 e 1/6 do número de variáveis incluídas na matriz de correlação.

2) O critério baseado no teste de qui-quadrado de Bartlett

Este critério de Bartlett (1950, 1951) consiste na testagem estatística da hipótese nula segundo a qual os autovalores de todos os componentes são iguais. Começando com o autovalor do último componente, cada autovalor é excluído sequencialmente até o teste de qui-quadrado da hipótese nula ser rejeitado. O teste parece ser sensível ao tamanho da amostra, levando à retenção de mais componentes em amostras grandes. Gorsuch (1973) demonstra que os valores aumentados dos autovalores em amostras grandes podem levar à retenção de mais componentes do que existem. Horn e Engstrom (1979) sugeriram mudar o nível do α de acordo com o tamanho da amostra. O critério de qui-quadrado de Bartlett não está disponível nos pacotes estatísticos mais utilizados, como SPSS, SAS e BMDP. Este critério de Bartlett não deve ser confundido com o teste de Bartlett de

Esfericidade, que está disponível na maioria dos pacotes estatísticos e que oferece uma indicação da fatorabilidade da matriz de correlações. A hipótese nula, no teste de Esfericidade é que a matriz de correlação é uma matriz de identidade, isso quer dizer que todas as variáveis têm uma correlação de zero. A hipótese nula tem que ser rejeitada para a análise fatorial fazer sentido.

3) O teste *scree* de Cattell.

O teste de *scree* foi desenvolvido por Cattell em 1966. O critério é baseado em um gráfico dos autovalores. O eixo Y representa os valores dos autovalores e o eixo X mostra o número sequencial dos componentes. O teste de *scree* separa o '*scree*' de fatores triviais do início de fatores não triviais por intermédio de uma inspeção visual do gráfico. O julgamento subjetivo está então baseado no uso de uma linha reta colocada ao longo da parte do fundo do gráfico onde os pontos formam uma linha aproximadamente reta. Os pontos acima da linha reta são associados com fatores não triviais, enquanto os pontos restantes representam os fatores triviais. Uma outra maneira para descrever o teste de *scree* consiste na busca de um “cotovelo” no gráfico”. Este cotovelo separa os fatores não triviais dos fatores triviais. Com o teste *scree* diversas complicações podem ocorrer, tais como: (a) os valores de autovalores podem mostrar um declive gradual com nenhum ponto de quebra claramente marcado e (b) os valores de autovalores podem demonstrar mais do que um ponto de quebra.

Horn e Engstrom (1979) notaram semelhanças subjacentes às lógicas do teste de *scree* e o teste do qui-quadrado de Bartlett: ambos são uma análise da igualdade dos autovalores dos componentes, sendo que a primeira é uma análise visual, a outra é uma análise estatística. Zwick e Velicer (1982) acharam que o teste *scree* é mais acurado com amostras grandes e componentes fortes.

4) O critério da média mínima de correlações parciais (MMCP) de Velicer.

Velicer (1976) sugeriu um critério baseado na matriz das correlações parciais. A matriz de correlações parciais é calculada após a extração de cada um dos componentes sequencialmente. A média das correlações parciais levadas ao quadrado é calculada baseada nesta matriz. Quando o valor mínimo das correlações parciais levadas ao quadrado é encontrado, nenhum componente mais é extraído. A correlação encontra um mínimo quando a matriz residual se assemelha a uma matriz de identidade. Velicer (1976) aponta para a exatidão do critério e afirma que este método pode ser aplicado com qualquer matriz de correlações ou de covariância.

5) O critério de análise paralela de Horn

Horn (1965) propôs comparar os autovalores obtidos empiricamente com os autovalores obtidos através de matrizes que contêm variáveis randômicas não-correlacionadas, tendo tamanhos de amostra iguais aos da matriz de correlação empiricamente obtida. O autor sugeriu gerar aproximadamente 50 matrizes randômicas de correlação, efetuar análises fatoriais destas matrizes e calcular a média dos 50 autovalores obtidos. Os autovalores da matriz de correlação com os dados empíricos seriam então comparados com os valores médios dos autovalores das matrizes com valores randômicos. Calculando os autovalores baseados nos dados randômicos, são criados critérios para julgar o grau de "*randomicidade*" dos autovalores empiricamente determinados. Somente fatores que correspondem a autovalores empíricos, que excedem os valores médios dos autovalores obtidos randomicamente, seriam extraídos. Autovalores empíricos menores ou iguais aos autovalores randômicos seriam considerados como sendo devidos à variância amostral randômica.

Várias pesquisas, principalmente estudos utilizando dados simulados, investigaram os méritos relativos de diferentes critérios, inclusive o critério de

análise paralela de Horn. Na apresentação original do procedimento de Análise Paralela (AP) em 1965, foi demonstrado que o critério de Guttman-Kaiser (GK) resulta em severa *superextração* de fatores em comparação com o critério de AP. Linn (1968) executou um estudo tipo *Monte Carlo* em relação ao critério de GK baseado em sete fatores predeterminados, de 20 e 40 variáveis, e amostras de 100 e 500 sujeitos. Os resultados deste estudo mostraram que o grau de superestimação do número de fatores era extremo com o uso do critério de GK. O critério de GK superestimou o número correto de fatores em 66% de casos. Este resultado pobre levou a uma recomendação contra uso continuado do critério de GK. Humphreys e Montanelli (1975) realizaram um estudo de simulação que empregou matrizes de correlação com três e sete fatores, 20 e 40 variáveis, amostras de 1.000 e 500 sujeitos e entre 40 e 50 replicações por condição. O critério de AP de Horn foi comparado com o critério de verossimilhança máxima de Jöreskog. O critério de verossimilhança máxima mostrou uma tendência consistente para superestimar o verdadeiro número de fatores, enquanto o critério de AP de Horn estava quase sempre correto.

No estudo de simulação mais abrangente realizado até hoje, Zwick e Velicer (1986) compararam o critério de GK, o teste de Bartlett, o teste scree de Cattell, o critério da média mínima de correlações parciais (MMCP) de Velicer e o critério de AP de Horn. Os dados simulados foram baseados em dois níveis fatoriais com 6 e 12 fatores conhecidos e dois níveis de saturação nos fatores com cargas fatoriais de 0,50 e 0,80. Os autores construíram matrizes de correlação, de 36 e 72 variáveis, tamanhos de amostra diferentes (com 2 e 5 vezes mais pessoas do que variáveis) e seis níveis de complexidade fatorial diferente, que variavam de uma estrutura fatorial simples até estruturas muito complexas. O critério de GK mostrou resultados mais pobres, indicando o número correto de fatores em apenas 22% das vezes, ou seja, errou em 78% dos casos. O critério de GK superestimou o verdadeiro número de fatores todas as vezes que indicou o número de fatores incorretamente. Surpreendentemente, ele

continua sendo amplamente utilizado por pesquisadores, embora tenha sido demonstrado de maneira clara que é um procedimento altamente inexato.

O teste de Bartlett estava correto em aproximadamente 30% das vezes, tendendo a superestimar o número de fatores com o tamanho da amostra crescente. O teste de *scree* mostrou 57% de precisão e superestimou o número correto de fatores em 90% dos casos que estava incorreto. O critério MMCP de Velicer era preciso em 84% do tempo e mostrou uma tendência para subestimação do número de fatores em 90% dos casos quando estava errado. O critério AP de Horn foi o procedimento mais preciso, indicando o número correto de fatores em 92% dos casos e mostrou ser o melhor método em todas as condições de decisão. Quando o critério AP de Horn estava errado, mostrou uma leve tendência a superestimar o verdadeiro número de fatores em aproximadamente 66% das vezes.

Zwick e Velicer (1968) notaram a tendência do critério AP de Horn para superestimar o número de fatores e sugeriram sua utilização em conjunto com o critério de MMCP de Velicer que mostrou a tendência de subestimar o número de fatores.

Hubbard e Allan (1987) demonstraram por meio da investigação de 30 bancos de dados de vários estudos empíricos que o critério AP de Horn tem uma vantagem considerável na determinação do número correto de fatores em comparação com o critério de GK, o teste de Bartlett e o teste *scree* de Cattell. Outros estudos, como os de Crawford e Koopman (1973) e Humphreys e Ilgen (1969), forneceram amplo apoio para a utilidade do critério AP de Horn. Existe pouca fundamentação para escolher quaisquer um dos outros critérios menos precisos, uma vez que existe uma grande quantidade de evidência que indica que o critério de AP é relativamente preciso e capaz de determinar o número correto de fatores a extrair.

Infelizmente, a falta de familiaridade com o critério de AP de Horn, assim como a falta de conhecimento das sérias negligências dos outros critérios e o fato que o procedimento não está disponível nos pacotes estatísticos principais existentes no mercado, limitou seu uso difundido por pesquisadores.

6 – Decisão de utilizar análise de componentes principais ou análise fatorial comum

Um dos procedimentos multivariados mais utilizado por pesquisadores das ciências sociais e do comportamento é a análise fatorial comum (AFC). Apesar de ser frequentemente confundida com a análise de componentes principais (ACP), a FC e a ACP possuem objetivos diferentes. Embora o propósito de ACP seja a redução de dados, baseada numa transformação matemática de uma matriz de correlação ou covariância, o objetivo primário de AFC é de uma ordem mais elevada. O propósito de AFC é descobrir e identificar variáveis subjacentes não-observáveis que são manifestas por um conjunto maior de variáveis diretamente observáveis. O modelo matemático conceitual que forma a base de AFC requer a estimação do número correto de fatores a reter. Este modelo apresenta uma complexidade consideravelmente maior do que a ACP. Dentro da análise fatorial comum (AFC) existem varias técnicas, a saber (a) *Principal Axis Factoring*; (b) *Unweighted Least Squares*; (c) *Generalized Least Squares*; (d) *Maximum Likelihood*; (e) *Alpha Factoring* e (f) *Image Factoring*. Uma percentagem considerável dos pesquisadores da área aplicada, que usam a técnica de análise fatorial, está confusa a respeito de quando utilizar a ACP e quando a AFC. Uma das razões principais para esta confusão é que a ACP também pode ser utilizada para a extração de fatores, embora existam muitos métodos de AFC especializados nisso. Num recente debate relativo ao uso comum de ACP como método de AFC, Velicer e Jackson (1990) enfatizaram uma observação muito importante: "*Se o número correto de fatores é extraído, então*

a interpretação substantiva dos resultados da análise será a mesma indistintamente se é utilizada a ACP ou a AFC". O ponto principal do Velicer e Jackson é que a decisão mais importante a ser tomada na análise fatorial diz respeito aos meios utilizados para determinar o número correto de fatores a reter e não qual método específico de análise fatorial é empregado.

Os especialistas na área de análise fatorial diferem muito em suas opiniões em relação à utilidade da ACP em comparação à análise fatorial comum (AFC). A diferença entre os dois métodos envolve as entradas utilizadas na diagonal da matriz de correlações que é analisada. Quando uma matriz de correlação é analisada, a ACP usa os valores 1,0 na diagonal, enquanto a AFC usa estimativas de comunalidade, normalmente calculadas por um processo iterativo. Quanto mais confiáveis as variáveis a serem analisadas, mais equivalentes são os resultados dos dois métodos de análise fatorial. A mesma lógica aplica-se em relação ao número de variáveis que entra numa análise fatorial; quanto mais variáveis entram, mais equivalentes os resultados da ACP e AFC (Daniel, 1990). Porém, se só um número pequeno de variáveis é analisado e as variáveis mostram comunalidades baixas, então as diferenças entre os resultados obtidos com a ACP e a AFC podem ser consideráveis. Nestes casos, a AFC conduz a estimativas precisas de cargas fatoriais e correlações entre fatores; enquanto a ACP tende a superestimar as cargas fatoriais e subestimar as correlações entre as dimensões. As diferenças nas estimativas obtidas pelas duas técnicas não são triviais. O mais importante, talvez, é que as estimativas baseadas na AFC são comparáveis com as estimativas obtidas usando técnicas de análise confirmatória, isso em contraste com as estimativas inexatas baseadas na ACP. Levando este fato em conta, as técnicas de AFC deveriam ser preferidas sobre as técnicas da ACP (Floyd & Widaman, 1995).

Pesquisadores podem usar tanto a ACP como a AFC na extração dos fatores ou outras alternativas razoáveis; mas eles precisam citar explicitamente o

método de extração e oferecer uma justificativa para uma determinada escolha. A análise dos componentes principais é uma alternativa razoável em alguns casos, especialmente dado que o método produz escores que têm os mesmos coeficientes de correlação que os fatores rotados. Além disso, a ACP não capitaliza indevidamente no erro amostral como o custo a pagar pela estimação do erro de mensuração. Nas seguintes situações, outros métodos de extração do que ACP parecem ser mais adequados: (1) quando o posto da matriz fatorizada é pequeno; (2) quando o erro de mensuração é considerável; (3) quando há muita variação nos erros de mensuração entre as variáveis, e (4) quando o erro amostral é pequeno devido ao tamanho da amostra.

7 – Decisão sobre o procedimento de rotação

Depois da extração, os fatores retidos geralmente são rotados para tornar a solução fatorial mais interpretável, mantendo as propriedades matemáticas da solução iguais. O objetivo do processo de rotação é conseguir uma estrutura fatorial simples. Uma estrutura simples é alcançada quando cada variável, preferencialmente, tem uma única carga alta em um único fator. Em outras palavras, uma estrutura fatorial simples existe quando cada variável tem uma carga principal em um único fator. A interpretação dos fatores, para a maioria dos pesquisadores, é simplificada quando todas as cargas fatoriais são positivas. Geralmente, a obtenção de cargas positivas na análise fatorial é relativamente fácil com instrumentos psicológicos porque na maioria das escalas que medem personalidade, estilos de comportamento, atitudes e outros processos psicológicos, a escala de resposta pode ser invertida. Por exemplo, numa escala Likert de 5 alternativas, os valores 1 e 2 podem ser convertidos para os valores 4 e 5 e vice versa; o valor 3 que representa o ponto neutro da escala, permanece como está.

O procedimento de rotação pode ser ortogonal ou oblíquo. Na rotação ortogonal, os fatores são mantidos não-correlacionados e na rotação oblíqua, em contraste, os fatores podem se correlacionar. Os procedimentos de rotação ortogonal mais comuns, na análise fatorial exploratória, são *Varimax*, *Quartimax*, *Equamax*, *Orthomax* e *Parsimax*. Destes procedimentos, o *Varimax* é o mais utilizado. O objetivo de *Varimax* é maximizar a variância das cargas fatoriais para cada fator por meio do aumento das cargas altas e a diminuição das cargas baixas. Este procedimento é a opção padrão em quase todos os pacotes estatísticos e produz uma estrutura fatorial razoavelmente simples na maioria das situações. Contudo, os pesquisadores precisam ser encorajados a verificarem também as soluções oblíquas para os seus dados. Os seguintes procedimentos de rotação oblíqua são os mais utilizados na análise fatorial exploratória: *Direct Oblimin*, *Quartimin*, *Procrustes* e *Promax*. Procedimentos de rotação oblíqua podem conduzir a uma estrutura simples mais convincente e melhor interpretável do que uma solução ortogonal. Os pesquisadores podem permitir vários graus de correlação entre os fatores nos procedimentos de rotação oblíqua, levando em conta o fato de que a permissão de correlações muito altas entre fatores pode ir contra os propósitos de análise fatorial de identificar construtos *distintos*. Quando a correlação entre os fatores é muito baixa ($r < 0,30$), o uso de um procedimento de rotação ortogonal é geralmente mais adequado.

8 – Decisão sobre a interpretação dos resultados

Análise fatorial exploratória (AFE) produz, para cada variável, uma carga fatorial em cada fator. A carga fatorial indica, em porcentagem, quanta covariância existe entre o fator e o item. O valor da carga fatorial varia entre -1,00 e +1,00, sendo que um valor de 0 indica a total ausência de covariância entre a variável e o fator. Depois da rotação ortogonal, a carga fatorial pode ser

interpretada como a correlação entre uma variável e o fator. Depois de uma rotação oblíqua, entretanto, a interpretação da carga fatorial não é mais tão simples: neste caso, ela deve ser interpretada como uma medida da relação única entre a variável e o fator. Uma vez que a rotação oblíqua produz fatores correlacionados, as cargas fatoriais superestimam a correlação entre a variável e o fator. Uma variável pode ter uma carga importante num determinado fator, não por causa de uma relação direta, mas por causa de uma correlação do fator com um outro fator. A rotação oblíqua produz duas matrizes, a matriz de padrões (*pattern matrix*) e a matriz de estrutura (*structure matrix*). Normalmente, as cargas da matriz de padrões são analisadas, porque as diferenças entre as cargas altas e baixas são mais evidentes nesta do que na matriz de estrutura.

Geralmente, as cargas fatoriais são consideradas significativas em análises exploratórias, quando elas excedem o valor absoluto 0,30. Este valor é considerado uma carga mínima necessária para a variável ser um representante útil do fator. Tabachnick e Fidell (1996) adotaram o valor 0,32 como valor mínimo para qualificar o item como um representante útil da variável, uma vez que este valor corresponde a 10% da variância explicada ($0,32^2 \approx 0,10$). Quanto mais alto o valor da carga fatorial, melhor a variável representa o fator. Desta maneira, a carga fatorial dá uma indicação sobre a qualidade da variável. Comrey e Lee (1992) sugerem que as cargas maiores que 0,71 são excelentes, maiores que 0,63 são muito boas, maiores que 0,55 boas, maiores que 0,45 razoáveis e maiores que 0,32 pobres. Relatar somente as cargas fatoriais acima de 0,30 ou 0,40 pode melhorar a aparência da estrutura fatorial obtida, entretanto, todas as cargas fatoriais deveriam ser informadas para assegurar suficiente informação para uma plena avaliação dos resultados. Quando uma variável não tem uma carga fatorial substancialmente alta em nenhum dos fatores, esta variável pode ser excluída e a análise fatorial deveria ser refeita com o subconjunto restante de itens.

Outro critério pragmático para soluções fatoriais satisfatórias é que os fatores deveriam explicar uma porcentagem significativa da variância total das variáveis medidas. Streiner (1994) sugeriu que os fatores deveriam explicar pelo menos 50% da variância total. Porém, uma exigência mais razoável é que o conjunto de fatores explique pelo menos 80% da variância comum (isto é, a soma das comunalidades calculadas inicialmente). Se menos variância é explicada, as comunalidades das variáveis são baixas e dever-se-ia considerar a possibilidade de excluir as variáveis com cargas fatoriais relativamente baixas em todos os fatores, para melhorar a solução fatorial geral. A razão da exclusão destes itens é para aumentar a porcentagem da variância comum explicada. A explicação de pouca variância comum desafia a importância relativa de fatores comuns vis-à-vis a variância de fatores específicos associados com variáveis individuais.

9 – Decisão sobre a investigação de uma solução hierárquica

Freqüentemente, espera-se que um fator geral possa explicar a maior parte da variância dos escores das variáveis medidas. Uma solução fatorial hierárquica pode ser o modelo mais apropriado para os dados nos quais fatores de primeira ordem, que são relativamente específicos, podem ser agrupados em fatores mais gerais, de um nível mais alto. Às vezes é desejável avançar e analisar as correlações entre os fatores de primeira ordem para obter fatores de segunda ordem. O número de fatores de segunda ordem depende do número de fatores de primeira ordem. Com dois fatores de primeira ordem no máximo um fator de segunda ordem pode ser obtido.

Quando um fator de uma ordem mais alta é esperado, a matriz fatorial de estrutura inicial deve ser submetida à rotação oblíqua para permitir correlações entre os fatores de primeira ordem. A matriz de correlações entre os fatores de

primeira ordem pode ser submetida a uma análise fatorial para identificar o(s) fator(es) de segunda ordem.

Análise fatorial hierárquica é um procedimento muito pouco utilizado no desenvolvimento e na avaliação dos instrumentos psicológicos. Soluções hierárquicas provavelmente são apropriadas para muitos instrumentos psicológicos porque muitos construtos psicológicos são compostos de múltiplas facetas correlacionadas.

10 – Decisão sobre a utilização da análise fatorial exploratória ou confirmatória

Na obra-prima do Carroll (1983) "*Human Cognitive Abilities: A survey of factor-analytic abilities*", o autor usa, em relação à decisão de utilizar análise fatorial exploratória (AFE) ou análise fatorial confirmatória (AFCF), o critério a seguir: quando as hipóteses a ser testadas têm apoio forte de uma teoria psicológica ou de análises empíricas anteriores de um banco de dados diferente do banco atual, a AFCF deveria ser utilizada. Em todos os demais casos, a AFE deveria ser utilizada. Aplicação dos métodos de AFCF gera informação acerca da probabilidade de que os dados se conformam com o modelo hipotetizado, ou melhor expressado, acerca da probabilidade de que os dados poderiam ser gerados se o modelo hipotetizado for correto. Pode acontecer que os dados possam ser gerados igualmente bem supondo outros modelos alternativos, caso em que é difícil escolher entre os modelos. Portanto, métodos de AFCF têm certas limitações. Métodos de AFE, por outro lado, são projetados para deixar os dados falarem por si mesmos, isso é, deixar a estrutura dos dados sugerir o modelo fatorial mais provável. Eles não requerem a testagem das hipóteses relativamente a modelos; nenhuma hipótese precisa ser montada com antecedência. Deste ponto de vista, métodos de AFE parecem ser mais flexíveis.

11 – Decisão sobre se e como calcular escores fatoriais

Escores fatoriais são mais fidedignos e acurados na amostra na qual as análises fatoriais foram baseadas, quando são calculados utilizando os pesos fatoriais derivados do padrão fatorial comum (Harman, 1976). Porém, ponderar com o valor 1,0 todos os itens com cargas principais nos fatores, produz escores fatoriais que são virtualmente muito precisos na amostra original; e, ademais, os pesos com o valor 1,0 funcionam melhor que os pesos exatos em qualquer amostra nova (Gorsuch, 1983). Assim, é desnecessário utilizar escores fatoriais exatos que marcam coeficientes para diferenciar itens de peso diferente quando se calcula escores de fator. Ao invés, os escores fatoriais podem ser calculados somando os escores de todos os itens ou variáveis com cargas significantes no fator principal e devem ser excluídos os itens que dividem cargas por mais de um fator. É importante se lembrar que todos os itens devem estar na mesma escala (por exemplo, escala Likert de 1 a 7 pontos) para assegurar pesagem igual. Se as medidas que carregam em um fator particular têm escalas muito diferentes, todas as variáveis deveriam ser unificadas antes de serem somadas em escores fatoriais.

12 – Decisão sobre a realização de um estudo de validade cruzada

O processo de validade cruzada implica, em relação à análise fatorial, que a solução fatorial obtida na base de dados de uma primeira amostra é verificada através dos dados de uma segunda amostra independente da primeira. A investigação da validade cruzada da solução fatorial é desejável tanto em relação à análise fatorial exploratória como à análise confirmatória. Idealmente, o tamanho da amostra deve ser suficiente para ser dividida em, pelo menos, dois

grupos, usando um grupo como a amostra de derivação e o outro como a amostra para a realização de validação cruzada. É importante ressaltar aqui que a amostra total precisa ser dividida de tal forma que não existam diferenças nas características dos grupos que poderiam afetar sua estrutura fatorial. Através da comparação de soluções fatoriais para grupos de mulheres e homens ou para grupos de diferentes etnias é possível oferecer uma resposta sobre a questão da *generalizabilidade* da estrutura fatorial, mas não serve como validação cruzada da solução fatorial. Não obstante, diferenças entre grupos em relação à solução fatorial são causa de preocupação. Em alguns casos, as soluções fatoriais não têm soluções reproduzidas em grupos com diferenças conhecidas, como amostras clínicas e não-clínicas, por causa da restrição de amplitude (*restriction of range*) das variáveis em um ou ambos os grupos. Nesta situação, a solução fatorial para a amostra total pode ser a mais apropriada para determinar a composição de subescalas. Porém, o fracasso para reproduzir a estrutura fatorial para grupos com variâncias semelhantes nas variáveis, como grupos de homens e mulheres, pode indicar que a estrutura fatorial da escala (e por dedução o construto avaliado) é realmente diferente para os dois grupos. Neste momento, o pesquisador precisa decidir entre desenvolver *escalas* separadas para os grupos ou só focalizar nas variáveis e características do construto que mostram uma estrutura consistente em ambos os grupos. O desenvolvimento de escalas separadas pode aumentar a sensibilidade das medidas para variância dentro grupos, mas complica ou elimina totalmente a possibilidade de comparações entre grupos.

Uma prática potencialmente útil para a validação cruzada é realizar uma análise fatorial exploratória com a metade da amostra e efetuar uma análise confirmatória, para verificar a estrutura fatorial, com a outra metade da amostra. Contudo, é necessário entender que a confirmação da estrutura fatorial não acontecerá se a solução exploratória não explica a maioria da variância nos dados. Neste caso, é aconselhável efetuar uma validação cruzada da solução

levando a efeito uma segunda análise fatorial exploratória com uma segunda amostra independente e comparar as soluções fatoriais utilizando o coeficiente de congruência que avalia a semelhança de cargas fatoriais das duas amostras.

13 – Decisão de investigar a invariância da estrutura fatorial em amostras diferentes

Uma pergunta que surge freqüentemente em relação aos instrumentos psicológicos é se a sua estrutura fatorial também é pertinente em amostras com características diferentes, em relação a variáveis como gênero, idade, nível socioeconômico ou alguma outra característica que poderia afetar as respostas do instrumento. A questão básica não é se as duas amostras têm um nível médio diferente em relação ao construto mensurado pelo instrumento, mas se o instrumento mede o mesmo construto em amostras diferentes. Se a estrutura fatorial não mostra invariância entre grupos, não podem ser levadas a efeito comparações significativas dos escores fatoriais entre esses grupos. Porém, se a invariância da estrutura fatorial é estabelecida, diferenças entre grupos em relação ao escore total do instrumento refletem com precisão as diferenças em relação às características subjacentes avaliadas pelo fator. O reconhecimento crescente das influências culturais, desenvolvimentais e contextuais, nos construtos psicológicos aumentou o interesse em demonstrar, empiricamente, a invariância da estrutura fatorial de instrumentos psicológicos em vez de simplesmente assumir que as medidas são equivalentes entre amostras diferentes.

Jöreskog e Sörbom (1989) descreveram os procedimentos para utilizar uma extensão de análise fatorial confirmatória com múltiplas amostras para avaliar a invariância da estrutura fatorial entre grupos. No caso menos restritivo, o investigador examina se as variáveis mostram o mesmo padrão de cargas

fatoriais significativas entre grupos, verificando a adequação de ajuste (*goodness-of-fit*) de um modelo de múltiplas amostras com padrões semelhantes de cargas fatoriais fixas e livres em cada grupo. Para um teste mais rigoroso e para a avaliação correta da invariância de mensuração, é preciso restringir todas as cargas fatoriais comuns como sendo invariantes entre os grupos.

14 – Decisão sobre a seleção dos resultados na publicação

Com a finalidade de interpretação e replicação dos resultados das análises fatoriais exploratórias, é importante oferecer informação sobre os seguintes aspectos: (a) o modelo de análise fatorial utilizado; (b) o método utilizado para o cálculo das comunalidades (somente para AFC); (c) o método escolhido para a extração de fatores; (d) o critério adotado de retenção fatores e (e) a justificativa da escolha para o método de rotação fatorial. Relatórios sobre análises exploratórias também deveriam incluir uma tabela que informe as cargas fatoriais para todas as variáveis em todos os fatores. Nesta tabela, as variáveis são agrupadas normalmente pelos fatores nos quais eles têm cargas principais e os fatores são colocados na ordem na qual foram identificados, com as cargas principais em tipo de letra negrito. Adicionalmente, a tabela ou o texto deveria informar os autovalores e a porcentagem de variância explicada por cada fator.

Na análise fatorial confirmatória, a natureza da estrutura fatorial proposta deve ser especificada anteriormente e deve ser incluída se é considerado que os fatores são ortogonais ou são correlacionados e se são esperadas cargas secundárias para quaisquer uns dos itens. Os resultados deveriam incluir os erros-padrão para cada carga, as correlações entre os fatores e indicações do significado destas correlações, assim como as estimativas da porcentagem de discrepância consideradas por cada fator e estatísticas de ajuste de modelo globais. A solução fatorial pode ser apresentada em uma tabela ou figura que

ilustra o modelo estrutural que lista todas as cargas de fator, intercorrelações de fatores e coeficientes de perturbação.

Uma vez que há decisões de caráter subjetivo, tanto na execução da análise fatorial exploratória quanto na confirmatória, os investigadores devem disponibilizar ao leitor o maior número possível de informações. Sempre que possível, a matriz de correlação completa entre as variáveis deveria ser dada em uma tabela ou em apêndice, juntamente com as médias dos itens e seus desvios-padrão, de forma que análises exploratórias e confirmatórias dos dados possam ser executadas por leitores interessados. Se o espaço de artigo não permite o informar a tabela de correlações, os autores deveriam estar dispostos a prover esta informação sob pedido.

Discussão Final

Ao longo de anos foi reconhecido cada vez mais que nenhum método analítico, inclusive análise fatorial, nos dita como nossos construtos ou nossas teorias deveriam ser (veja Mulaik, 1994). Como Mulaik enfatizou: “somos nós que criamos significados para coisas decidindo como eles deveriam ser usados. Assim, nós deveríamos ver a loucura de supor que a análise fatorial nos ensinará o que é que é inteligência, ou o que é personalidade”. Tais estudiosos discutem que nossos dados podem informar nosso julgamento, mas que nós somos responsáveis pelo julgamento que temos que exercitar inevitavelmente, e nossas definições de construto devem ser fundamentadas na teoria e não nos dados. Thompson e Daniel (1996), porém, assinalam que os resultados analíticos podem informar as definições que nós desejamos criar, embora nós permaneçamos responsáveis por nossas elaborações e podemos até mesmo desejar reter definições que não têm sido apoiadas empiricamente.