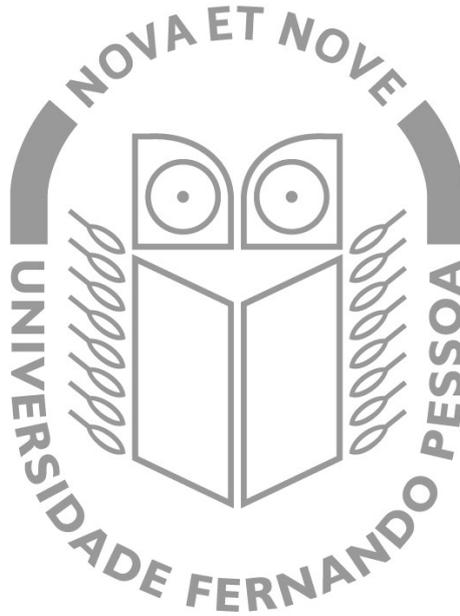


Universidade Fernando Pessoa



Provas de Agregação

Lição

A fiabilidade dos instrumentos de medição

Alexandre Miguel Fernandes Gomes da Silva

Lição no âmbito do ramo do conhecimento de Ciências da Informação, especialidade Sistemas e Tecnologias da Informação, nos termos da alínea c) do número 2 do artigo 5 do Decreto-Lei nº 239/2007

2022

Índice

1.	Introdução	3
2.	Estrutura da lição	3
3.	Tema da lição	3
4.	Enquadramento do tema na unidade curricular de estudos de mercado	4
4.1	Questionários	6
4.2.	Escalas	8
5.	Objetivos	10
6.	Conteúdos	10
7.	Metodologia	11
8.	Desenvolvimento da aula	11
9.	A fiabilidade dos instrumentos de medição	13

1. Introdução

A presente lição foi desenvolvida para a candidatura a provas de Agregação de acordo com a alínea c) do número 2 do artigo 5 do Decreto-Lei nº 239/2007 de 19 de junho, onde é requerido ao candidato uma lição sobre um tema no âmbito do ramo do conhecimento em que são prestadas as provas.

Assim, sendo o ramo do conhecimento de Ciências da Informação, especialidade Sistemas e Tecnologias da Informação, para a lição o candidato optou por desenvolver o tema “fiabilidade dos instrumentos de medição” inserido na unidade curricular de Estudos de Mercado, lecionada na licenciatura em Marketing e Negócios Internacionais, do Instituto Superior de Contabilidade e Administração do Instituto Politécnico de Coimbra. Esta unidade Curricular foi alvo de desenvolvimento em relatório próprio produzido em documento independente apresentado para as provas de Agregação.

2. Estrutura da lição

A lição está estruturada em seis partes distintas:

1. Tema da aula;
2. Enquadramento na unidade curricular de Estudos de Mercado;
3. Objetivos da aula;
4. Conteúdos da aula;
5. Metodologia utilizada na concretização da lição;
6. Desenvolvimento da lição sustentada num referencial científico-pedagógico.

3. Tema da lição

O tema da lição versa a avaliação da fiabilidade de um instrumento de medição a recolha de informação, no caso por fontes primárias, o caso do questionário, sendo tratadas as seguintes linhas de ação:

1. Introdução
2. Análise fatorial exploratória
3. Equações Estruturais
4. Análise Fatorial Confirmatória
5. Índices de Ajustamento
6. Fiabilidade
7. Estabilidade temporal
 - 7.1. Teste-reteste
 - 7.2. Formas equivalentes
8. Consistência interna
9. Metade-Metade
10. Análise de sensibilidade dos itens da escala
11. Considerações finais

4. Enquadramento do tema na unidade curricular de estudos de mercado

A forma de medir é uma questão central em muitas áreas não só na unidade curricular de estudos de mercado, mas também em outras disciplinas e cursos, como auditoria, recursos humanos, secretariado, comércio e finanças, no âmbito do Instituto Superior de Contabilidade e Administração de Coimbra. O uso de questionários como fonte de informação primária é um recurso habitual até como tema autónomo de formações ou cursos breves, como se destacam:

2011 Curso de Limesurvey Aperfeiçoamento

2019/11/29 XII Workshop Grudis – A análise de dados de questionários na investigação em contabilidade – ISCAL (IPL),– Um balanço

Instituto Politécnico de Lisboa Instituto Superior de Contabilidade e Administração de Lisboa, Portugal

2019/07/05 XI Workshop Grudis – A utilização de questionários na investigação em contabilidade – ESCE IP Setúbal,

Instituto Politécnico de Setúbal Escola Superior de Ciências Empresariais, Portugal

Ora constitui uma preocupação a relação e consistência do que é perguntado com o que é efetivamente medido. O grau de concordância (ou não) entre a conceção do instrumento, no pressuposto dos seus objetivos e hipóteses colocadas, e o entendimento dos respondentes traduzido nas respostas dadas. A própria dimensão do questionário e o conjunto de perguntas associado deve ser aferido, determinando a sua adequação.

Assim, o tema apresentado tem dois objetivos, o primeiro a aprendizagem dos conceitos relativos à avaliação da fiabilidade de um instrumento de medição, o segundo é pedagógico, mostrando o impacto do desenho experimental na qualidade do resultado.

No âmbito da unidade curricular de estudos de mercado contribui para a concretização do seu ponto 6, subsecção 6.2 fontes primárias.

Conteúdos programáticos da unidade curricular

- 1 Definição do problema**
 - Objectivos**
 - Estrutura**
- 2 Contexto interno**
 - Dados da organização (organização/cultura/I&D)**
 - Capacidade financeira**
 - Preço**
 - Produto**
 - Logística/stocks**
- 3 Contexto externo**
 - Contexto país/mundo (balança de pagamentos etc)**
 - Distribuição**
 - Mercado**

- Cientes**
- Oferta**
- Procura**
- 4 Operacionalização**
 - Mercado - grupo alvo**
 - Variáveis – instrumentos**
 - Técnicas**
- 5 Recolha informação**
 - Fontes primárias**
 - Fontes secundárias**
- 6 Análise de dados**
 - Procedimentos analíticos (quantitativos ou qualitativos)**
 - Agregação de informação**
 - Inferência**
- 7 Interpretação**
 - Organização do estudo, cúmulo que combina todas as etapas anteriores e permite concretizar o objectivo**
- 8 Conclusão**
 - Report e decisão**

Na unidade curricular de estudos de mercado os questionários sobressaem por corresponderem a uma técnica importante e frequente na recolha de informação. Assim, é valorizada a sua elaboração e entendimento da sua estrutura com o correspondente numero de aulas teórico-práticas disponibilizadas.

A lição aqui apresentada corresponde à 22ª aula, 11ª semana de aulas, do primeiro semestre do terceiro ano.

É precedida por um conjunto de aulas onde se introduz globalmente o tema de estudos de mercado, o problema, a definição das variáveis analisando o contexto interno à organização e o contexto externo onde a organização opera. Insistindo nos fatores e determinantes para o problema em questão e sua possível mensuração. Também é lecionada uma aula (nona aula) de escalas, inserida no tópico variáveis, e outras de construção e implementação dum questionário (10ª – 12ª aula). Nesta ultima aula o modelo do questionário online (comum) é colocado numa *nuvem* partilhada, solicitando-se que cada aluno promova a resposta pelo publico alvo, obtendo-se um ficheiro de dados reais de resposta para posterior análise. Pretende-se uma ilustração em ambiente de vida real onde os alunos trabalham com os seus próprios dados.

Como anteriormente referido a mensuração é uma questão com particular relevância pela sua subjetividade nas áreas do marketing, comportamento, recursos humanos, auditoria, sustentabilidade e responsabilidade social (referindo apenas as prementes no contesto do Instituto Superior de Contabilidade e Administração de Coimbra), por exemplo a medição da produtividade, motivação, liderança, comportamento do consumidor, responsabilidade social ou fiscal ou ambiental etc. Por contraste com a contabilidade e finanças onde, em regra, as variáveis são quantitativas reais e determinísticas.

Apresenta-se de seguida uma síntese de conteúdos lecionados anteriormente ao tema da presente lição reforçando a ideia da necessidade da construção, com rigor, de um instrumento de medida:

4.1. Questionários

São uma técnica estruturada para recolha de dados, que consiste numa série de questões escritas ou orais a que os inquiridos devem responder.

Devem traduzir a informação necessária através de um número de questões específicas, motivando e encorajando o inquirido a cooperar e a responder ao questionário até ao final. Deve minimizar o número de respostas com erro (respostas imprecisas).

Note-se que os questionários são um instrumento de medida que, como balanças ou metros, têm que ser “aféridos” e “calibrados”, bem como as suas unidades devem ser traduzidas em unidades “SI”. Quer isto dizer que a elaboração de um questionário deve ser cuidadosamente ponderada para traduzir medidas corretas e comparáveis quer temporalmente quer contextualmente.

Os questionários podem ser elaborados com os seguintes objetivos básicos:

- Estimar grandezas “absolutas”, tais como: quantidades/tempo, percentagens;
- Estimar grandezas “relativas”, tais como: proporções de tipos numa população;
- Descrever uma população ou subpopulação, tais como: características de consumidores, formadores de opinião;
- Verificar hipóteses, tais como: variações de natureza ou de frequência de comportamento com a idade, entre outras;

Assim a forma de conceção e redação de um questionário é pré-determinada pelos objetivos a serem alcançados em termos estatísticos, ou seja, antes da elaboração propriamente dita das questões há que definir muito bem os objetivos principais da investigação e fazer um planeamento minucioso de todas as técnicas e modelos estatísticos que irão ser utilizados para testar as hipóteses já definidas. É com base nos parâmetros mencionados anteriormente que se vão elaborar as perguntas do questionário.

Uma boa questão nunca deve proporcionar tendência a uma resposta específica, não deve expressar nenhuma expectativa e as opções de resposta, se existirem, não devem excluir nenhuma possibilidade plausível de resposta do entrevistado.

Aspetos a ter em conta na construção de um questionário

Aquando da construção de um questionário, devem ser definidos os seguintes aspetos:

- Definição do tipo de resposta mais adequada a cada pergunta (ter em conta que cada pergunta corresponde uma variável e a natureza dessa variável determina a sua análise);
- Definição do tipo de escala de medida a associar às respostas;
- Definição da metodologia para análise dos dados;

Para alcançar os aspetos mencionados anteriormente, há que:

- Listar todas as variáveis de investigação, incluindo as características dos casos;
- Especificar o número de perguntas para medir cada uma das variáveis;

- Escrever uma versão inicial para cada pergunta;
- Pensar nas hipóteses em estudo, bem como nas variáveis e perguntas a ela associadas;
- Consoante as hipóteses do estudo (converter em hipóteses estatísticas), decidir as técnicas e estatísticas adequadas para testar a hipótese, tendo em atenção os pressupostos destas técnicas;
- Face às mudanças que possam ter sido introduzidas, verificar se as versões finais das perguntas e das respostas ainda estão adequadas para testar as hipóteses;
- Perceber se o respondente (público-alvo) sabe como responder;
- Planear as secções do questionário;
- Recolher apenas as características dos casos estritamente relevantes para a investigação;
- Estabelecer o modo como as perguntas devem ser escritas (construção, extensão e clareza), em função do tipo de pergunta.

Estrutura do questionário

As secções dos questionários

É muito importante antes de se começar a elaborar as questões definir um plano para as secções. Este plano tem dois objetivos, dar uma estrutura ao questionário e ajudar a justificar a inclusão das perguntas no questionário.

Quando não existe uma planificação adequada das secções do questionário, é muito simples incluir perguntas desnecessárias ou omitir perguntas essenciais, em termos práticos isto poderá por em causa o trabalho de investigação. Para fazer um plano das secções será necessário mencionar todas as hipóteses, bem como a literatura que deu origem a essas mesmas hipóteses.

Alguns cuidados a ter na elaboração dos questionários

Na elaboração de um questionário é importante, antes de mais, ter em conta as habilitações do público-alvo. É de destacar que o conjunto de questões deve ser muito bem organizadas e conter uma forma lógica para quem responde, evitando as irrelevantes, insensíveis, intrusivas, desinteressantes, com uma estrutura (ou formato) demasiado confusos e complexos, ou ainda questões demasiado longas.

Deve, o investigador, ter o cuidado de não utilizar questões ambíguas que possam, por isso, ter mais do que um significado, que por sua vez, levem a diferentes interpretações.

Não deve incluir duas questões numa só, pois pode levar a respostas induzidas ou nem sempre relevantes, além de não ser possível determinar qual das “questões” foi respondida, aquando o tratamento da informação.

O investigador deve ainda evitar questões baseadas em pressuposições, pois parte-se do princípio que o inquirido encaixa numa determinada categoria e procura informação baseada nesse pressuposto.

É também necessário redobrar a atenção ao formular questões de natureza pessoal, ou que abordem assuntos delicados ou incómodos para o inquirido.

As questões devem ser reduzidas e adequadas à pesquisa em questão. Assim, elas devem ser desenvolvidas tendo em conta três princípios básicos:

- Princípios da clareza - devem ser claras, concisas e unívocas;
- Princípios da coerência - devem corresponder à intenção da própria pergunta;
- Princípio da neutralidade - não devem induzir uma dada resposta mas sim libertar o inquirido do referencial de juízos de valor ou do preconceito do próprio autor.

A escolha das questões é extremamente importante, algumas regras para a escolha e elaboração das mesmas:

- Incorporar apenas questões relacionadas com o problema investigado;
- Não incorporar questões cujas respostas podem ser obtidas através de meios mais exatos;
- Evitar perguntas que invadam a intimidade e privacidade dos inquiridos;
- As questões devem ser formuladas de uma maneira clara, concreta e precisa;
- Dever-se-á ter em consideração o sistema de referência e de informação do inquirido;
- A questão deve possibilitar uma única interpretação por parte do inquirido;
- A questão não deve sugerir respostas;
- A questão deve tratar de uma única ideia.

A elaboração de um questionário deriva de um processo de aperfeiçoamento, resultado de várias revisões e aplicação do pré – teste, que forem necessários.

Cada questão deve ser analisada individualmente para garantir que é importante, que não é ambígua ou de difícil compreensão.

4.2. Escalas

Quando o questionário contém perguntas fechadas é preciso escolher um conjunto de respostas alternativas para cada uma destas perguntas. É possível associar números às respostas para analisá-las por meio de técnicas estatísticas, os números associados com um conjunto de respostas apresentam uma escala de medida.

Há vários tipos destas escalas. Os dois tipos de escala frequentemente usados em questionários que são as escalas nominais, ordinais, intervalares e de razão.

A natureza da escala condiciona o nível de informação e a análise estatística.

As escalas nominais consistem num conjunto de categorias de resposta qualitativamente diferentes e mutuamente exclusivas. É possível atribuir um número a cada categoria para codificar as respostas, mas estes números não implicam diferenças em quantidade sendo o seu objetivo identificar a categoria qualitativa. As escalas nominais fornecem dados na forma de frequências. Isto significa que o que se tem são o número de respostas em cada categoria da escala.

As escalas ordinais mantêm as características da escala nominal, mas tem a capacidade de ordenar os dados.

Nas escalas intervalares conhecemos as distâncias entre quaisquer dois números, ou pontos dessa escala. O investigador tem a possibilidade de perceber as diferenças que separam dois objetos. A escala possui propriedades de ordem e diferença, mas o zero da escala não corresponde à anulação da característica em estudo.

Sobre as escalas de razão (rácio) podemos referir que têm todas as características das escalas de intervalares mas têm uma característica adicional. O valor “zero” não é arbitrário - é absoluto ou “real”. Por esta característica adicional, numa escala de razão é possível fazer inferências sobre um rácio de valores na escala.

Para terminar, podemos classificar as escalas nas que são comparativas, com a possibilidade de comparação direta de dois ou mais objetos e as não comparativas em que cada objeto é avaliado independentemente da existência de outros.

Na escolha de uma escala devemos ter em conta aspetos importantes como o nível de informação, nominal, ordinal, intervalar, escala de razão, não podendo esquecer a capacidade dos que estão a ser inquiridos, as características do objeto que é medido e aspetos relacionados com o método, custo e contexto do inquérito. É eficaz que o investigador escolha a escala que conduz a uma maior quantidade de informação fiável, para se poder usar o maior número de técnicas estatísticas possíveis na análise dos dados.

Em conclusão

Nas investigações em que o questionário é utilizado como uma das técnicas do processo de recolha de dados, sendo através dos dados obtidos que se vão testar as hipóteses formuladas e responder às questões da investigação, assim, é fundamental que este (as variáveis criadas) estejam de facto adequadas aos conceitos teóricos que visam medir.

Bibliografia usada nestas aulas

Chaudhuri A., Stenger H. (2005). Sampling: theory and methods. 2ª Ed Chapman & Hall

Hill, M. & Hill, A. (2005). Investigação por Questionário. Lisboa: Edições Sílabo

Ghiglione, Rodolphe; Matalon, Benjamin (2001, [1977]). O Inquérito: Teoria e Prática. 4ª Ed. Oeiras: Celta Editora.

Vicente, P., Reis, E. (2011). Internet surveys: Opportunities and challenges. 10.4018/978-1-60960-042-6.ch050.

Vieira, A. (2011) Escalas Em Marketing - Métricas de Respostas do Consumidor e de Desempenho Empresarial. Atlas

É criado, assim, juntamente com o plano amostral, um contexto onde a análise estatística e de fiabilidade são aplicados no sentido de mostrar a adequabilidade dos instrumentos criados para medir dando assim sentido à própria disciplina de estudos de mercado – isto é – estamos de facto a responder à questão que suscitou o estudo?

5. Objetivos

Os objetivos de aprendizagem, expressos em conhecimentos e competências a desenvolver pelos estudantes na sequência da lição, são os seguintes:

- Perceber os instrumentos de medida devem ser aferidos
- Conhecer técnicas de validação/aferição da qualidade de informação dum instrumento de medida
- Relacionar a informação pretendida com a forma de a obter e as questões a fazer
- Relacionar o instrumento de medida com o tipo de variáveis e com a subsequente análise quantitativa ou qualitativa
- Compreender o problema em estudo e as hipóteses colocadas em relação ao desenho experimental
- Aferir a qualidade e rigor de informação obtida, compreender os efeitos de manipulação de variáveis, estrutura ou linguagem no resultado

6. Conteúdos

Os conteúdos programáticos a analisar na lição são os seguintes:

A fiabilidade dos instrumentos de medição

12. Introdução
13. Análise fatorial exploratória
14. Equações Estruturais
15. Análise Fatorial Confirmatória
16. Índices de Ajustamento
17. Fiabilidade
18. Estabilidade temporal
 - 18.1. Teste-reteste
 - 18.2. Formas equivalentes
19. Consistência interna
20. Metade-Metade
21. Análise de sensibilidade dos itens da escala
22. Considerações finais

7. Metodologia

A lição corresponde a uma aula teórico-prática de 90 minutos (aqui adaptada a 60 minutos). Transmitem-se os conteúdos fundamentais do tema usando um método expositivo apoiado em recursos multimédia.

Considerando-se a natureza teórica e pratica da lição os conceitos teóricos são intercalados com exemplos de natureza prática com o intuito de aplicabilidade autónoma em contexto real.

Assim, os conteúdos são apresentados de forma oral, e sequencial, iniciando com a explicitação dos objetivos da aula e o conteúdo a analisar seguido do desenvolvimento individual de cada tópico do conteúdo com referência ao enquadramento teórico e a exemplos práticos.

Ao longo da lição é proporcionado um ambiente de dialogo com os alunos promovendo questões relativas aos conteúdos lecionados e usando as suas respostas para debate e troca de ideias, focando-os nas principais noções.

Durante a exposição são invocados os conhecimentos adquiridos anteriormente na unidade curricular para os relacionar com o desenvolvimento do tema da lição.

Na parte final é feita uma síntese do tema, usando um exemplo prático, para ilustrar todas as noções chave e esclarecer eventuais dúvidas. É, ainda, enquadrado o tema seguinte e a aula de aplicação pratica usando software próprio.

Palavras chave:

Fiabilidade – escalas – variáveis – questionário – variáveis latentes – dimensões – constructos – ajustamento – variância – correlação – significância -

8. Desenvolvimento da aula

O desenvolvimento do tema da lição corresponde ao enquadramento dos conteúdos que servem de suporte à aula.

As razões que presidiram à escolha deste tema prendem-se com a sua atualidade e transversalidade no uso nas mais variadas ciências e atividades. Resulta também, de uma experiencia de 20 anos como docente nas áreas de marketing, gestão, análise de dados, auditoria e controlo de gestão e gestão da qualidade onde esta temática constitui, na grande maioria das vezes, a forma de análise de instrumentos de recolha de informação, seja em projetos

(vide Auditoria à Rede Nacional Apoio à Vitima 01CP-SGPCM/2020),

seja em teses/dissertações

(a titulo de exemplo no Instituto Superior de Contabilidade e Administração de Coimbra: Engenheiro, Susana Raquel Jesus. 2021 Impacto da Inteligência artificial nas empresas: estudo

de mercado em empresas da região de Coimbra. Mestrado em Gestão Empresarial; Oliveira, Cátia Vanessa Gonçalves de. 2021 Motivação e Satisfação nos Trabalhadores Temporários. Mestrado em Gestão de Recursos Humanos; Roberto, Andreia Raquel Ferreira. 2021 A curadoria de dados na gestão empresarial com estudo sobre a perspetiva empresarial portuguesa. Mestrado em Gestão Empresarial),

seja em serviços à comunidade

(exemplo como coordenador do PolLab: elaboração do “Questionário da indústria cerâmica” para a APICER pelo Pollab 9/5/2016; elaboração de relatórios de caracterização de empresas dos concelhos de Castanheira de Pêra, Figueiró Dos Vinhos, Góis, Lousã, Miranda do Corvo, Pampilhosa da Serra, Penela, Vila Nova de Poiares. Para a AESL – Associação Empresarial da Serra da Lousã, 26/10/2015).

Adicionalmente o tema apresenta uma estreita relação com outros conteúdos abordados na unidade curricular, aproveitando-se a aula como meio de convergência de conhecimentos e associado a uma forte aplicação prática constitui também uma motivação para o interesse dos alunos.

O texto apresentado representa um compromisso entre o discurso usado na sala de aula, com a sua fluidez própria, assente essencialmente no entendimento e motivação do aluno do ensino superior, e a revisão sistemática do tema, resultando num texto com uma sequência fundamentada, mas não exaustiva.

9. A fiabilidade dos instrumentos de medição

Introdução

A aula deve iniciar com o seguinte exemplo

Sigerson, L.; Cheng, C (2018) Scales for measuring user engagement with social network sites: A systematic review of psychometric properties. *Computers in Human Behavior* **83** 87e105

Mantym, M; Islam, A (2016) The Janus face of Facebook: Positive and negative sides of social networking site use. *Computers in Human Behavior* 61

<http://dx.doi.org/10.1016/j.chb.2016.02.078>

Medir significa estabelecer uma “concordância” entre a forma de medir e o que está a ser medido. Este processo é relativamente fácil quando estamos em presença de variáveis diretamente observáveis. Contudo, quando as variáveis são latentes (não observadas nem medidas diretamente), esse processo torna-se mais difícil, como é o caso da medição da ansiedade, da satisfação ou da qualidade de vida quando medidas através de escalas com vários itens.

Atualmente é grande a variedade de instrumentos de medição, que podem assumir a designação de testes, escalas ou inventários “de acordo com a relação que se assume existir entre os itens, ou questões, que compõem o questionário” (Ribeiro, 1999, p. 76). Em cada ano cresce o número de instrumentos de medida elaborados em diferentes países, o que obriga à sua validação e à realização de testes de fiabilidade. Neste artigo focar-nos-emos na fiabilidade.

Até há relativamente pouco tempo, a inspeção da fiabilidade era um processo moroso e difícil de realizar se tivermos por referência os curtos prazos estabelecidos para as investigações académicas contemporâneas. Com o incremento da informática e devido à maior facilidade de acesso a programas estatísticos, estes processos tornaram-se cada vez mais fáceis de realizar, pois é agora possível realizar cálculos complexos em frações de tempo, impensáveis há umas décadas atrás.

Um instrumento de medição deve ser inequivocamente fiável. A fiabilidade dos instrumentos de medição depende, em última instância, das boas características dos seus itens. Para se obter um instrumento fiável poder-se-á ter que equacionar eliminar, substituir ou rever os itens.

O termo “fiabilidade” pretende significar a qualidade daquilo ou de quem é fiável. No contexto da inquirição o termo apresenta especificidades que importa refletir. Fiabilidade reporta-se à consistência de resultados obtidos pelos mesmos indivíduos, quando inquiridos em diferentes momentos ou num determinado momento, sendo neste último caso a fiabilidade determinada a partir de itens equivalentes (Anastasi, 1977; Freeman, 1980). Na linguagem científica encontramos uma panóplia terminológica para nomear o substantivo, que aqui designamos por fiabilidade (do inglês, reliability). Alguns clássicos nesta área apresentam nas suas traduções para língua portuguesa os termos garantia (Freeman, 1980), precisão (Anastasi, 1977) ou mesmo fidelidade (Bryman & Cramer, 1993). Em autores que escrevem em língua portuguesa os termos fiabilidade (Hill & Hill, 2012; Maroco & Garcia-Marques, 2006) e fidedignidade (Vaz-Serra, Ponciano, & Freitas, 1980) são os utilizados.

Estimar a fiabilidade de um teste é um procedimento fundamental quando recorremos a instrumentos de medição que utilizam escalas de vários itens. Na década de noventa do século passado e no dealbar do século XXI, estimar a fiabilidade de um teste ou de escala ou de um inventário tornou-se um procedimento rápido, resultado do incremento de pacotes estatísticos que, de forma célere em associação com plataformas cada vez mais interativas e amigáveis, permitiram a utilização massiva destes procedimentos.

A fiabilidade pressupõe também a reprodutibilidade de resultados. Anne Anastasi, cientista cujo brilhantismo é plasmado numa vasta obra onde a psicometria é desconstruída, afirma que a precisão se refere “à consistência de resultados obtidos pelos mesmos indivíduos em diferentes oportunidades ou com diferentes conjuntos de itens equivalentes” (Anastasi, 1977, p. 84). No mesmo sentido, Freeman assevera que a garantia se reflete na consistência dos resultados efetuados em várias avaliações, significando “o grau em que os resultados obtidos estão isentos daqueles defeitos internos suscetíveis de provocar erros de medição inerentes aos próprios itens e à standardização” (Freeman, 1980, pp. 73-74).

Análise fatorial exploratória

A análise fatorial exploratória é um de uma família de métodos estatísticos multivariados que tenta identificar o menor número de construtos hipotéticos (também conhecidos como fatores, dimensões, variáveis latentes, variáveis sintéticas ou atributos) que podem explicar com parcimonia a covariância observada entre um conjunto de variáveis medidas (também chamadas de variáveis observadas, variáveis manifestas, efeitos). Ou seja, identificar os fatores comuns que explicam a ordem e a estrutura entre as variáveis medidas. Nas ciências sociais e comportamentais, os fatores são assumidos como características não observáveis dos indivíduos, que se manifestam em diferenças nas pontuações alcançadas nas variáveis medidas (Tucker & MacCallum, 1997, Brown 2015).

Um fator é uma variável não observável que assume a variância entre variáveis observadas. Noutras palavras, as medidas observadas estão inter-relacionadas porque compartilham uma causa (dimensão) comum. Fundada em princípios filosóficos e estatísticos (Mulaik, 1987), a AFE foi aplicada pela primeira vez por Spearman (1904) e rapidamente se tornou uma ferramenta fundamental na avaliação de teorias e validação de instrumentos de medida (Haig, 2014; Henson & Roberts, 2006; Izquierdo, Olea e Abad, 2014). Conforme observado por Edwards e Bagozzi (2000), as relações entre construtos e suas variáveis indicadoras são importantes porque esse conhecimento permite um “mapeamento de construtos teóricos em fenómenos empíricos”.

O termo AFE é frequentemente usado de forma bastante imprecisa para se referir a dois modelos que diferem em propósito e computação: especificamente, análise de componentes principais (ACP) e análise fatorial comum (Fabrigar et al., 1999). A ACP analisa toda a matriz de correlação (incluindo as autocorrelações de 1,00 encontradas na diagonal) e “pretende reduzir os dados preservando o máximo possível de informações do conjunto de dados original” (Norris & Lecavalier, 2010). Para fazer isso, a ACP calcula combinações lineares das variáveis originais medidas que explicam o máximo de informações possível sobre essas variáveis originais. Chamadas de componentes, essas novas variáveis medidas são representações parcimoniosas das variáveis medidas originais, mas não são construções latentes (Cudeck, 2000). Em vez disso,

as variáveis medidas influenciam os componentes. Assim, é usual referir essas combinações lineares como componentes, não como fatores.

Assumindo que as variáveis medidas estão correlacionadas porque são influenciadas pelo mesmo construto latente subjacente, a análise fatorial comum tenta separar a variância total das variáveis medidas na variância que é comum às variáveis medidas.

A AFE é usada para descobrir – explorar um modelo sem à partida um conhecimento prévio das relações (e da sua força) entre variáveis observadas.

Conceitos importantes a reter na implementação desta temática (assumindo o uso de SPSS e AMOS) ver Maroco 2021 a,b.

Teste de esfericidade de Bartlett

Calcula o determinante da matriz das somas dos produtos e produtos cruzados (S) a partir da qual a matriz de correlação é obtida. O determinante da matriz S é convertido numa estatística da distribuição do qui-quadrado χ^2 e testado quanto à significância. A hipótese nula é que a matriz de inter-correlação vem de uma população na qual as variáveis são não colineares (ou seja, uma matriz identidade). e que as correlações não nulas na matriz amostral são devidas ao erro amostral.

Medida Kaiser-Meyer-Olkin de Adequação da Amostragem (KMO)

Se duas variáveis compartilham um fator comum com outras variáveis, a sua correlação parcial (a_{ij}) será pequena, indicando a variância única que elas compartilham.

$$a_{ij} = (r_{ij} \bullet 1, 2, 3, \dots k)$$

$$KMO = (\sum \sum r_{ij}^2) / (\sum \sum r_{ij}^2 + \sum \sum a_{ij}^2)$$

Se $a_{ij} = 0,0$ as variáveis estão a medir um fator comum e $KMO = 1,0$

Comunalidade

Comunalidade, h^2 , é o quadrado da correlação múltipla para uma dada variável usando os fatores como preditores. A comunalidade mede a percentagem de variância numa determinada variável explicada por todos os fatores em conjunto e pode ser interpretada como a confiabilidade do indicador.

Quando uma variável indicadora tem baixa comunalidade, o modelo fatorial não está a funcionar bem para aquele indicador e possivelmente deve ser removido do modelo. No entanto, as comunalidades devem ser percebidas em relação à interpretabilidade dos fatores. Uma comunalidade de 0,75 parece alta, mas não tem sentido, a menos que o fator no qual a

variável está “carregada” seja interpretável. Uma comunalidade de 0,25 parece baixa, mas pode ser significativa se o item estiver contribuindo para um fator bem definido. Ou seja, o que é crítico não é o coeficiente de comunalidade em si, mas sim a extensão em que o item desempenha um papel na interpretação do fator, embora muitas vezes esse papel seja maior quando a comunalidade é alta.

Se a comunalidade exceder 1,0, há uma solução espúria, que pode refletir uma amostra muito pequena ou o investigador tem muitos ou poucos fatores.

A comunalidade de uma variável é calculada como a soma das cargas fatoriais ao quadrado para essa variável. Lembrar de que r -quadrado é a porcentagem de variância explicada e, como os fatores não são correlacionados, as cargas quadradas podem ser adicionadas para obter a porcentagem total explicada, que é a comunalidade. Para ACP ortogonal completa, a comunalidade inicial será de 1,0 para todas as variáveis e toda a variância nas variáveis será explicada por todos os fatores, que serão tantos quantas forem as variáveis. A comunalidade "extraída" é a porcentagem de variância numa determinada variável explicada pelos fatores que são extraídos, que geralmente será menor que todos os fatores possíveis, resultando em coeficientes menores que 1,0.

Valores próprios

O valor próprio para um dado fator mede a variância em todas as variáveis que são contabilizadas por aquele fator. A razão de valores próprios é a razão de importância explicativa dos fatores em relação às variáveis. Se um fator tem um valor próprio baixo, então está contribuindo pouco para a explicação das variâncias nas variáveis e pode ser ignorado como redundante com fatores mais importantes.

Assim, os valores próprios medem a quantidade de variação na amostra total contabilizada por cada fator. Observar que o valor próprio não é a porcentagem de variância explicada, mas sim uma medida de "quantidade", usada para comparação com outros valores próprios. O valor próprio de um fator pode ser calculado como a soma de suas cargas fatoriais ao quadrado para todas as variáveis.

Equações Estruturais

A modelação através de equações estruturais (MEE) é uma técnica muito geral, engloba variados métodos, e é muito usada em ciências sociais/comportamentais.

A aplicação de MEE destina-se à análise de construtos teóricos representados por variáveis latentes (fatores). A relação entre os construtos é representada por coeficientes de regressão. Esta modelação implica uma estrutura para as covariâncias entre as variáveis observadas.

Jöreskog (1989) refere-se a estes modelos como *linear structural relations* ou modelos *Lisrel* contudo a metodologia evoluiu permitindo relações não lineares. Este modelo pode ser representado matematicamente da seguinte forma:

$$\eta = B\eta + \Gamma\xi + \zeta, \quad y = \Lambda_y\eta + e, \quad x = \Lambda_x\xi + \delta$$

onde η e ξ são vetores de variáveis latentes (fatores), e x e y são vetores de variáveis observadas. As componentes de η correspondem a variáveis latentes endógenas; as componentes de ξ correspondem a variáveis latentes exógenas. Os fatores exógenos e endógenos estão ligados por um sistema de equações lineares (o modelo estrutural) com matrizes de coeficientes B e Γ e um vetor de erros ζ . A matriz $I - B$ é não singular. Os vetores aleatórios x e y correspondem a variáveis observadas que estão relacionadas com as variáveis latentes por dois sistemas de equações com os coeficientes Λ_x e Λ_y e com os respetivos erros δ e e .

A estrutura de covariância:

$$C = J(I - A)^{-1} P \left((I - A)^{-1} \right)' J' \quad \text{onde}$$

$$A = \begin{pmatrix} 0 & 0 & \Lambda_y & 0 \\ 0 & 0 & 0 & \Lambda_x \\ 0 & 0 & B & \Gamma \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad e \quad P = \begin{pmatrix} \Theta_\varepsilon & & & \\ & \Theta_\delta & & \\ & & \Psi & \\ & & & \Phi \end{pmatrix} \quad \text{com}$$

$$\Phi = \varepsilon \{ \xi\xi' \}, \quad \Psi = \varepsilon \{ \zeta\zeta' \}, \quad \Theta_\delta = \varepsilon \{ \delta\delta' \} \quad e \quad \Theta_\varepsilon = \varepsilon \{ ee' \}.$$

Assim a MEE proporciona uma moldura para a análise estatística que permite enquadrar vários métodos como por exemplo análise fatorial, regressão, análise discriminante e correlação.

É também possível tratar esta técnica com uma representação gráfica (um grafo). Os grafos mostram as variáveis interligadas indicando o fluxo e o peso de cada arco (relação) e nó (variável latente/observada). Podemos, assim, visualizar, de forma simples, quais as variáveis que *causam* outras. É de notar que o grafo orientado proposto para representar o fluxo causal do sistema de variáveis, mesmo que bem ajustado aos dados, não implica a veracidade da causalidade.

Análise Fatorial Confirmatória

A análise fatorial assume que as covariâncias entre um conjunto de variáveis observadas pode ser explicada por um menor número de fatores subjacentes. Na análise fatorial exploratória não presumimos nenhum conhecimento sobre o número de fatores nem das suas relações com as variáveis observadas. Em contraste na AFC o modelo é imposto aos dados, isto é, o modelo é previamente especificado de forma exata. Assim, a modelação através de equações estruturais neste caso tem dois objetivos, o primeiro prevê a obtenção de estimativas dos parâmetros do modelo (dos fatores, das suas variâncias, covariâncias e o erro das variáveis observadas); o segundo é verificar o ajustamento do modelo aos dados.

Para cada fator devemos fixar o parâmetro em 1, permitindo, assim, uma interpretação, dado que em caso contrário desconhecemos a escala do fator. Alternativamente podemos fixar as variâncias dos fatores e determinar os parâmetros (standardização da escala).

Hoje em dia há vários softwares que tratam esta temática sem necessidade de especificar todas as matrizes do sistema de equações. Programas como o AMOS, no qual foi realizado o trabalho aqui apresentado, basta apenas desenhar o grafo.

Índices de Ajustamento

Em AFC temos, como já foi referido, que especificar o modelo antes de efetuar a análise. A especificação resulta da combinação da teoria com resultados práticos já obtidos em análises anteriores.

Após a obtenção do modelo podemos efetuar um teste do χ^2 para verificar o ajustamento do modelo aos dados. Obviamente que a aplicação do teste e a sua potência dependem da normalidade dos dados e do tamanho da amostra.

O teste de Wald ou rácio crítico (CR), que resulta da estimativa dividida pelo seu erro padrão, e o critério de informação de Akaike (AIC), que resulta da diferença entre o dobro da estatística do χ^2 e os graus de liberdade, são duas medidas do grau de ajustamento ao modelo ideal. Outra medida é a RMSEA (root mean square error approximation) que trata como o actual modelo aproxima o modelo verdadeiro. É desejável ter uma RMSEA inferior a 0.05 (ver Byrne 2001, Kline 2005)

Fiabilidade

Existem várias estimativas da fiabilidade de um teste e “pode haver tantas variedades de precisão do teste quantas as condições que influem nos seus resultados, pois qualquer uma dessas condições pode não ter significado para determinado objetivo e, assim, ser classificada como variância do erro identificadas por cada uma” (Anastasi, 1977, p. 85). Dependendo da forma de estimar a fiabilidade utilizam-se vários procedimentos. O pressuposto subjacente ao estudo da fiabilidade é calcular o tamanho do erro. Assim, no caso da inexistência de variabilidade nos resultados, não haveria erro e a fiabilidade seria igual a 1. De notar que a definição de erro é complexa porque está associada ao conceito mensurado. Se o teste mede a depressão, a diferença encontrada nos resultados entre as duas administrações (variância do erro) pode ser atribuída tanto a flutuações aleatórias (e neste caso falamos de erro aleatório), como podem igualmente ser devidas a um diferente estado emocional (e neste caso não podemos afirmar que se trata de erro não aleatório). Se objetivarmos medir a estabilidade temporal de um teste temos que nos assegurar que não existem diferenças no estado emocional entre as duas aplicações. Este procedimento pode ser controlado através da inclusão de perguntas sobre experiências intermediárias significativas entre as administrações.

Estabilidade temporal

Diz-se que um instrumento apresenta estabilidade temporal se os seus resultados se mantiverem constantes ao longo do tempo, ou seja, “realizando-se duas vezes o mesmo teste,

a correlação entre os resultados dá-nos uma indicação acerca da estabilidade dos resultados no tempo” (Laveault & Grégoire, 2002, p. 150).

Os procedimentos para avaliar a estabilidade temporal ou fiabilidade externa implicam a administração de duas versões do mesmo teste (teste-reteste) ou de duas versões de testes equivalentes.

Teste-reteste

Uma forma de estimarmos a estabilidade temporal de um conjunto de perguntas ou itens equivalentes pressupõe a utilização de um teste-reteste. Assim, para estimar esta estabilidade temporal é necessário administrar o conjunto de perguntas ou itens equivalentes a um grupo e correlacionar administração os valores obtidos com os de uma outra efetuada em um outro momento. Se o teste for fiável é expectável que, apesar do tempo decorrido entre as duas administrações e se nenhuma alteração significativa ocorrer entretanto, as respostas registadas para cada indivíduo se mantenham inalteradas, ou quase.

As diferenças encontradas nos testes podem assim ser atribuídas a diferentes fatores: (i) administrador, (ii) administrado, ou (iii) cenário.

- Relativamente ao administrador imaginemos que este se apresentou, numa das aplicações do teste, com um comportamento percecionado pelo respondente como inapropriado. Será que este comportamento não se refletirá negativamente nos resultados? Não poderá contribuir para um preenchimento pouco rigoroso, pouco profissional ou mesmo desonesto do teste, da escala ou do inventário, logo interferindo nos resultados? Por outro lado, imaginemos que o administrador demonstrou excessiva simpatia a par de uma narrativa longa e inapropriada sobre o instrumento de medição. Não poderão estes comportamentos ativar no inquirido respostas de acordo com as normas sociais? É consabido que “o desejo de aceitação social pode enviesar parâmetros avaliados em investigações científicas, constituindo uma ameaça à sua validade, pelo que deve ser controlado” (Poínhos et al., 2008, p. 223).
- No tocante ao administrado imagine uma de múltiplas possibilidades: no espaço que mediou as duas administrações alguns dos inquiridos ficaram desempregados. Esta nova situação de inatividade forçada é vivida por alguns dos respondentes com elevados níveis de ansiedade. Se o conceito a medir for a ansiedade-estado (estado emocional transitório ou condição do organismo humano caracterizado por sentimentos desagradáveis de tensão e apreensão conscientemente percebidos consentâneos com o aumento na atividade do sistema nervoso autónomo), as diferenças encontradas não deverão ser atribuídas a uma baixa fiabilidade teste-reteste mas sim a alterações nos níveis de ansiedade-estado em resultado da nova situação vivida. Neste caso estaríamos em presença de um teste que apesar de apresentar uma aparente baixa fiabilidade teste-reteste é fiável. Parece-nos aqui pertinente mencionar que a fiabilidade dos instrumentos de medição que avaliam estados de humor podem demonstrar elevada oscilação, isto é, é provável verificarem-se diferenças significativas entre as pontuações. Contrariamente os instrumentos de medição que avaliam características de personalidade devem apresentar pontuações estáveis ao longo do tempo.

- Por último, relativamente ao cenário, imagine que a sala onde está a ser administrado o conjunto de perguntas ou itens equivalentes está mal isolada e o frio, ou mesmo o barulho do exterior, provocam mal-estar. É expectável que também nesta situação os resultados reflitam o cenário ambiental vivido.

Importa aqui referir que quando utilizamos um instrumento de medição devemos sempre recorrer ao seu manual ou artigo original ou, caso não exista, aos respetivos autores para obtermos indicação sobre o intervalo que deve mediar as duas administrações. Esta informação é preciosa porque a duração temporal que medeia as administrações poderá criar diferentes efeitos relacionados com o conceito a ser medido. Intervalos curtos podem provocar efeitos relacionados com a memória; pelo contrário, longos intervalos podem possibilitar a aquisição de novos conhecimentos. Anastasi sobre este assunto refere que “podem ser citados facilmente exemplos de testes que apresentem alta precisão em períodos de poucos dias ou semana, mas cujos resultados revelam uma perda quase completa de correspondência quando o intervalo se estende a até dez ou quinze anos. Por exemplo, muitos testes de inteligência para crianças em idades pré-escolares fornecem medidas moderadamente estáveis no período pré-escolar, mas são virtualmente inúteis para predizer o QI na idade escolar ou adulta” (Anastasi, 1977, p. 32).

A idade dos sujeitos tem igualmente importância crucial. Anastasi sugere que o período que medeia um teste-reteste deve ser menor em crianças, porque as mudanças desenvolvimentais nas crianças são discerníveis em períodos curtos (um mês ou menos), independentemente da idade o intervalo teste-reteste não dever exceder seis meses (Anastasi, 1977). O mesmo pode suceder quando estamos perante testes de fiabilidade em condições de saúde em relação às quais se espera progressões muito diferentes.

Além disto, convém aqui também reportar que a interpretação do resultado da fiabilidade teste-reteste deve ter em consideração o facto de um teste se seguir ao outro. A experiência da primeira administração pode afetar o desempenho do sujeito na segunda administração.

Nenhum teste, mesmo que meça um traço da personalidade estável, se deve esperar que demonstre uma fiabilidade teste-reteste perfeita com correlação igual a 1, pois existem inúmeros fatores que influenciam as pontuações. Em síntese, a fadiga, os níveis diferentes de concentração e de motivação, as diferentes condições do meio ambiente (a temperatura, os barulhos, as distrações ambientais), os efeitos da prática e da aprendizagem, o intervalo entre as aplicações, os acontecimentos pessoais inesperados no decorrer deste intervalo, e os erros inerentes à administração, são alguns das possíveis variações que interferem com a variância do erro.

Quando um instrumento de medição é operacionalizado numa escala compósita de itens múltiplos (e.g., através de soma ou médias dos seus itens) pode utilizar-se o coeficiente de correlação linear de Pearson para obter uma estimativa da fiabilidade entre as formas. Em baixo poderá observar uma das fórmulas utilizadas no cálculo deste coeficiente de correlação r :

$$r = \frac{n \times \sum XiYi - \sum Xi \times \sum Yi}{\sqrt{[n \times \sum Xi^2 - (\sum Xi)^2] \times [n \times \sum Yi^2 - (\sum Yi)^2]}}$$

onde x e y representam as variáveis.

Por exemplo considerar administrar um instrumento de medição antes e depois da implementação de medidas de desempenho, da introdução de novos equipamentos ou procedimentos.

Formas equivalentes

Segundo Anastasi o coeficiente de fiabilidade “é uma medida tanto da estabilidade temporal como da coerência de resposta a diferentes formas de itens (ou formas de teste)” (Anastasi, 1977, p. 95).

O método denominado por formas equivalentes, formas alternativas ou paralelas é semelhante ao método teste-reteste na medida em que são feitas duas administrações do mesmo teste. Existem, contudo, diferenças assinaláveis entre estes dois métodos. Enquanto no método teste-reteste são administradas as mesmas versões nas duas sessões, nas formas equivalentes as versões equivalem-se, ou seja, as versões devem ser iguais nas instruções, na forma e em todas as demais características e similares no conteúdo. Pretende-se, com este método, eliminar dois tipos de vieses encontrados no método teste-reteste: o facto de os indivíduos poderem recordar-se do teste anterior bem como os possíveis efeitos da prática. Assim, este método pretende eliminar os efeitos da prática e da memória ao testar indivíduos através da utilização de versões, comparáveis, mas não iguais nas duas sessões.

Alta fiabilidade em formas equivalentes sugere que os itens das duas versões do teste são representativos de uma mesma população de itens que hipoteticamente representam o conceito que está a ser medido. Baixa fiabilidade nas formas equivalentes sugere que os dois formatos do teste não estão a medir a mesma coisa. Convém realçar que a fiabilidade através de um teste de formato equivalente contém muitas das mesmas limitações da fiabilidade do teste-reteste.

Aquando da mensuração da fiabilidade teste-reteste e das formas equivalentes é obrigatório mencionar o intervalo de tempo que mediou entre as duas administrações. Se as duas formas são aplicadas uma a seguir à outra, em sucessão imediata, a correlação mostra apenas a fiabilidade entre as formas (conteúdo) e não entre as ocasiões (estabilidade temporal). Pelas características acima aduzidas percebe-se que esta forma não seja muito utilizada em investigação.

Consistência interna

No caso da consistência interna, designada também por fiabilidade interna por Bryman e Cramer (1993), apenas se aplica uma versão e uma única vez. Imagine-se que se quer avaliar uma variável latente a partir de múltiplos itens. Como é expectável, o grupo de itens deve operacionalizar a variável latente e não uma outra variável. Será que os itens têm consistência interna? Só é possível afirmar que um instrumento de medida tem consistência interna se todos os seus itens contribuírem para a medição da mesma característica. O procedimento da consistência interna para estimar a fiabilidade é hoje um dos métodos mais utilizados na investigação, na medida em que é um método que, além de ser económico (Polit & Hungler, 1992) pois requer apenas só uma prova, é também o melhor método para avaliar uma das fontes mais importantes de erros de medição que é a seleção dos itens do teste.

O coeficiente alfa de Cronbach é comumente utilizado para estimar a fiabilidade de instrumentos nos quais os itens apresentam múltiplas respostas. A regra básica, aqui também, é de que os valores se devem situar entre 0,8 e 1,0 (Bryman & Cramer, 1993). Quando um conceito e a sua medição compreendem várias dimensões, é habitual calcularem-se os coeficientes de fiabilidade para cada uma das dimensões subjacentes em vez de calcular um só para a medida no seu todo.

$$\alpha = \frac{k}{k-1} \left[1 - \sum_{i=1}^k \frac{S_i^2}{S_t^2} \right]$$

A fórmula para calcular o coeficiente alfa é a seguinte: onde K é número de itens da escala, S_i^2 é variância dos resultados do teste no item I e S_t^2 é variância do teste.

O processo de medição da consistência interna através do coeficiente alfa de Cronbach produz baixas estimativas da fiabilidade do teste, mas sobrestima a fiabilidade de testes de velocidade. Consequentemente, os procedimentos de consistência interna são considerados inapropriados para determinar a fiabilidade de testes de velocidade (Ary, Jacobs, & Razavieh, 1990).

Apesar de ser consensual que uma escala deva ser fidedigna os valores a partir dos quais se infere da fiabilidade da escala parecem não o ser. Senão vejamos a tabela produzida por Peterson (1994).

Autor	Situação	Níveis recomendados
Davis (1994)	Preditor individual	Acima de 0,75
	Previsão para grupos de 25-50	0,5
	Previsão para grupos acima de 50	Abaixo de 0,5
Kaplan e Sacuzzo (1982)	Investigação fundamental	0,7-0,8
	Investigação aplicada	0,95
Murphy e Davidsholder (1988)	Nível inaceitável	Abaixo de 0,6
	Nível baixo	0,7
	Nível moderado a elevado	0,8-0,9
	Nível Elevado	0,9
Nunnally (1978)	Investigação preliminar	0,7
	Investigação fundamental	0,8
	Investigação aplicada	0,9-0,95

Adaptado (Peterson, 1994)

Quadro 1 – Valores propostos por vários autores para o α de Cronbach

Este coeficiente, por apresentar algumas fraquezas, tem levantado algumas críticas tendo autores como Maroco e Garcia-Marques (2006) apresentado alternativas. A fiabilidade compósita definida por Fornell e Larcker (1981) para um fator j com k itens é dada por $fc = \frac{(\sum_{i=1}^k p_i)^2}{(\sum_{i=1}^k p_i)^2 + \sum_{i=1}^k e_i}$, onde p representam os pesos fatoriais de cada item e os erros de cada item.

Enquanto o coeficiente alfa de Cronbach é utilizado em testes com itens com múltiplas respostas, a fórmula KR 20 destina-se a testes em que os itens oferecem apenas duas hipóteses alternativas, como é o caso de verdadeiro/falso, sim/não, ou certo/errado.

$$R_{KR-} = \frac{K}{K-1} \left[1 - \frac{\sum_{i=1}^K p_i q_i}{\sigma_X^2} \right]$$

em que K é o número de itens, p a proporção de respostas corretas, q a de incorretas e σ_X^2 a variância.

Metade-Metade

Um outro procedimento utilizado para medir se um conjunto de questões ou itens equivalentes apresentam fiabilidade designa-se como metade-metade, do inglês *split-half*, também designada como fiabilidade das metades ou da bipartição. Pode chegar-se a um valor de fiabilidade a partir de uma única administração. Neste caso correlacionam-se as duas metades de um teste. A divisão das metades deve ser baseada em critérios que terão em conta tanto o conceito como a forma como este se encontra operacionalizado. Uma variante deste processo é conhecida como o método par-ímpar, que é talvez o método mais antigo para estimar a consistência interna (Polit & Hungler, 1994).

A fiabilidade metade-metade é semelhante à forma alternativa, dividindo uma única escala em duas. Este método avalia o grau de consistência entre os itens, determinando a consistência interna da escala. Não mede a estabilidade temporal, mas oferece a vantagem de permitir obter uma medida de fiabilidade a partir de uma única administração e assume que todos os itens contribuem de igual forma para a mensuração de um conceito central (Anastasi, 1977; Freeman, 1980).

Imagine-se em presença de uma escala composta por dez itens. Uma das formas de dividirmos a escala seria decompô-la segundo a numeração que lhe foi atribuída inicialmente, ou seja, uma das metades corresponde aos itens pares e a outra é constituída pelos itens ímpares. Outra forma é dividir os itens em duas grandes metades, ou seja, os cinco primeiros itens com os da segunda metade. Importa referir que a opção pela divisão deve estar relacionada com o tipo de teste. A divisão por duas grandes metades, ou seja, a primeira metade com a segunda metade, pode provocar muitos inconvenientes especialmente em testes organizados com um grau crescente de dificuldade. As metades podem ainda ser criadas através de uma seleção aleatória.

Qualquer coeficiente de correlação obtido através da técnica metade-metade tende a gerar uma estimação sistematicamente inferior ao da escala na sua totalidade, ou seja, a correlação aqui obtida é relativa a metade do teste. As escalas com maiores números de itens geram maiores valores de fiabilidade (Polit & Hungler, 1992). Com o objetivo de superar esta dificuldade foi criada uma fórmula para ajustar o coeficiente de correlação para toda a escala. Obtêm-se, assim, um coeficiente que pode ser interpretado da mesma forma que o coeficiente de correlação de

Pearson, na medida que varia de 0 a 1. Idealmente, ele deve ser maior ou igual a 0,8 (Bryman & Cramer, 1993) e obtido através da fórmula seguinte:

$$r = \frac{\frac{1}{2}(s_p^2 - s_{p1}^2 - s_{p2}^2)}{s_{p1}s_{p2}}$$

onde s_p , s_{p1} e s_{p2} representam respetivamente os valores dos desvios padrão para todos os itens e para cada uma das metades.

Por outro lado, conhecendo o valor de correlação entre as duas metades, é possível estimar o coeficiente de correlação para toda a escala. A equação de correção é denominada de Spearman-Brown e traduz-se no seguinte algoritmo:

$$r_{CSB} = \frac{2r_{xy}}{1 + r_{xy}}$$

Nesta fórmula, r_{CSB} representa a correção introduzida por Spearman-Brown e r_{xy} a correlação entre as duas metades.

O facto de existirem muitas formas de dividir os itens em dois grupos faz com que se possa também obter múltiplas estimativas de fidedignidade. Por esse facto normalmente utilizam-se apenas o coeficiente alfa de Cronbach e a fórmula de Kunder-Richardson pois traduzem a média de todos as bipartições possíveis.

Análise de sensibilidade dos itens da escala

Existem vários procedimentos que poderão ser realizados para avaliar a sensibilidade dos itens de uma escala. Referimo-nos à análise da matriz de correlação inter-item, ao valor do alfa de Cronbach se cada um dos itens fosse eliminado e à correlação item-total.

A matriz de correlações apresenta a correlação de cada item com todos os outros. Na diagonal da matriz deve encontrar o valor 1, já que a correlação de um item com ele próprio será sempre igual à unidade. Os valores das restantes correlações devem ser elevados e positivos, indicando que os itens medem a mesma variável latente. Se existirem valores negativos deve verificar se os itens em causa estavam na mesma direção conceptual dos outros itens e, se for esse o caso, deve proceder-se à sua recodificação.

É sempre possível verificar, como análise de sensibilidade, o impacto de cada item no modelo global com a medição do alfa sem esse item ou com a correlação desse item com a soma dos restantes. Caso não haja impacto na escala, é preferível a mais simples. Duas medidas que mostram isso são o alfa item-total e a correlação item-total como se pode ver nas fórmulas seguintes.

A análise do valor do alfa de Cronbach quando um determinado item é eliminado permite-nos analisar o impacto da exclusão de um determinado item. É dado pela fórmula

$$\bar{A}_i = \frac{k-1}{k-2} \left(1 - \sum_{\substack{l=1 \\ l \neq i}}^k S_l^2 / \bar{S}_i^2 \right)$$

em que k é o número de itens, S_i^2 é variância dos resultados do teste no item i e \tilde{S}_i^2 é variância do teste sem o item

Se da análise da eliminação de determinados itens da escala se verificarem valores de alfa de Cronbach superiores ao valor de alfa da escala total deve considerar a eliminação desses itens da escala. Importa mencionar que valores do alfa de Cronbach inferiores a 0,7 não são considerados favoráveis daí que a possibilidade de eliminação do item deva ser equacionada (Pallant, 2007). Note que só é aconselhável a eliminação de itens de uma escala quando esta está a ser construída ou quando se pretende reduzir o número dos seus itens. Caso contrário, o procedimento de eliminação inibe, em rigor, qualquer comparação dos nossos resultados com a escala original.

Uma outra forma de obtermos informações acerca da consistência interna de uma escala é através da correlação item-total que nos indica o grau segundo o qual cada item se correlaciona com a pontuação total. Esta correlação é dada por $R_i = \frac{cov(X_i, P) - S_i^2}{S_i \tilde{S}_i}$. Caso seja necessário eliminar itens, a correlação item-total permite-nos eliminar os itens que apresentam correlações mais baixas com o resto da escala. Teremos assim uma escala com itens com maior consistência interna, apesar de a escala no seu conjunto, caso se tenham eliminado muitos itens, diminuir em termos de fiabilidade medida através da consistência interna. Segundo Pallant (2007) valores baixos (menores que 0,3) indicam que o item não está a medir a escala como um todo, podendo existir alternativa à escala apresentada. Se o alfa de Cronbach da escala global for baixo, por exemplo, inferior a 0,7; podem existir itens incorretos, podendo haver necessidade da sua remoção de acordo com baixas correlações item-total.

Considerações finais

Todos os instrumentos de medição apresentam vantagens e desvantagens pois dependem de vários fatores como, por exemplo, o tipo de dados (escala), a forma como foi operacionalizado o constructo (unifatorial ou multifatorial), o objetivo do instrumento (avaliação de um traço de personalidade ou conhecimentos) e os respetivos itens. Estes fatores ditarão qual ou quais as medidas para cada caso. Estamos perante uma situação de medida relativa, dado que a fiabilidade ou a consistência está sempre diretamente relacionada com o que se pretende medir ou comparar.

Todos os estimadores apresentarão resultados diferentes para a mesma situação. Geralmente o teste-reteste apresenta valores mais baixos pois depende de mais do que uma avaliação. Por outro lado algum cuidado deve ser tido no desenho experimental, dado que a análise é feita normalmente em estudos quasi-experimentais não havendo muitas vezes aleatoriedade e equilíbrio nas medições.

Assim, não podemos postular uma “receita” padrão para o uso dos referidos estimadores, antes propomos uma ponderação cuidada na escolha do mais adequado ao estudo em causa.

Referências

Anastasi, A. (1977). *Testes Psicológicos* (2.a Edição., p. 798). São Paulo: Editora Pedagógica e Universitária Ltda.

Ary, D., Jacobs, L. C., & Razavieh, A. (1990). *Introduction to research in education*. (I. Harcourt Brace Jovanovich, Ed.) (4th ed., p. 536). Fort WORTH: Ted Buchholz.

Brown, TA (2015). *Análise fatorial confirmatória para pesquisa aplicada* (2ª ed.). Nova York, NY : Guilford Press .

Bryman, A., & Cramer, D. (1993). *Análise de Dados em Ciências Sociais: Introdução às Técnicas Utilizando o SPSS* (2.a Edição., p. 346). Oeiras: Celta Editora.

Byrne, B.M. (2001) *Structural equation modelling with AMOS, basic concepts, applications and programming*. Lawrence Erlbaum Associates, Publishers

Cudeck, R. (2000). *Análise fatorial exploratória*. Em Tinsley, HEA, Brown, SD (Eds.), *Manual de estatística multivariada aplicada e modelagem matemática* (pp. 265 - 296). Nova York, NY : Academic Press .rs, London

Edwards, JR, Bagozzi, RP (2000). *Sobre a natureza e direção das relações entre construtos e medidas* . *Métodos Psicológicos*, 5, 155 - 174.

Fabrigar, LR, Wegener, DT (2012). *Análise fatorial exploratória*. Nova York, NY: Oxford University Press .

Fabrigar, LR, Wegener, DT, MacCallum, RC, Strahan, EJ (1999). *Avaliar o uso da análise fatorial exploratória na pesquisa psicológica* . *Psychological Methods*, 4, 272 - 299. doi: 10.1037/1082-989X.4.3.272

Freeman, F. (1980). *Teoria e Prática dos Testes Psicológicos* (2.a Edição., p. 780). Lisboa: Fundação Calouste Gulbenkian.

Haig, BD (2014). *Investigando o mundo psicológico: método científico nas ciências comportamentais*. Cambridge : MIT Press .

Henson, RK, Roberts, JK (2006). *Uso da análise fatorial exploratória em pesquisas publicadas* . *Medição Educacional e Psicológica*, 66, 393 - 416. doi: 10.1177/0013164405282485

Hill, M. M., & Hill, A. (2012). *Investigação por questionário*. (M. Robalo, Ed.) (2.a Edição., p. 377). Lisboa: Edições Sílabo.

Hox, J.J., Bechger, T.M. (2004) *An introduction to structural equation modeling*. *Family Science Review*, 11:354-373

IBM SPSS. (2011). *IBM SPSS Statistics 20 Algorithms*. Retrieved from ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/en/client/Manuals/IBM_SPSS_Statistics_Algorithms.pdf

Izquierdo, I., Olea, J., Abad, FJ (2014). *Análise fatorial exploratória em estudos de validação: Usos e recomendações*. *Psicothema*, 26, 395 - 400 . doi: 10.7334/psicothema2013.349.

Jöreskog, K.G., Sörbom, D. (1989) *Lisrel 7: A guide to de program and applications*. Chicago: SPSS

Laveault, D., & Grégoire, J. (2002). *Introdução às Teorias dos Testes em Ciências Humanas* (p. 335). Porto: Porto Editora, Lda.

Kline, R.B. (2005, 2016) Principles and practice of structural equation modelling. 2ª Ed. Guilford Press

Maroco, J. (2021)a Análise Estatística com o SPSS Statistics. ReportNumber

Maroco, J. (2021)b Análise de Equações Estruturais Fundamentos teóricos, software & aplicações. ReportNumber

Maroco, J. & Garcia-Marques, T. (2006). Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas? *Laboratório de Psicologia*, 4, 65-90.

Mulaik, SA (1987). Uma breve história dos fundamentos filosóficos da análise fatorial exploratória. *Pesquisa comportamental multivariada*, 22, 267 - 305 . doi: 10.1207/s15327906mbr2203_3

Norris, M., Lecavalier, L. (2010). Avaliando o uso da análise fatorial exploratória na pesquisa psicológica de deficiência do desenvolvimento. *Journal of Autism and Developmental Disorders*, 40, 8 - 20 . doi: 10.1007/s10803-009-0816-2

Pallant, J. (2007). *SPSS Survival Manual: a step by step guide to data analysis using SPSS for Windows (3th ed.)*. New York: Mc Graw Hill.

Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research*, 21(2), 381–391.

Póinhos, R., Correia, F., Faneca, M., Ferreira, J., Gonçalves, C., Pinhão, S., & Medina, J. L. (2008). Desejabilidade Social e Barreiras ao cumprimento da terapêutica dietética em mulheres com excesso de peso. *Acta Médica Portuguesa*, 21, 221–228.

Polit, D., & Hungler, B. (1992). *Investigacion Cientifica en Ciencias de la Salud (3.a Edició., p. 563)*. México: Nueva Editorial Interamericana, S.A.

Ribeiro, J. L. P. (1999). *Investigação e avaliação em psicologia da saúde. (C. Editores, Ed.) (1.a Edição., p. 147)*. Lisboa.

Spearman, CE (1904). "Inteligência geral", objetivamente determinada e medida . *American Journal of Psychiatry*, 15, 201 - 293 .

Tucker, LR, MacCallum, RC (1997). Análise fatorial exploratória. Recuperado de <https://www.unc.edu/~rcm/book/factor.pdf>

Vaz-Serra, A., Ponciano, E., & Freitas, F. (1980). Resultado da aplicação do Eysenk personality inventory a uma amostra da população portuguesa. *Psiquiatria Clínica*, (21), 127–132.

Watkins MW. Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*. 2018;44(3):219-246. doi:10.1177/0095798418771807

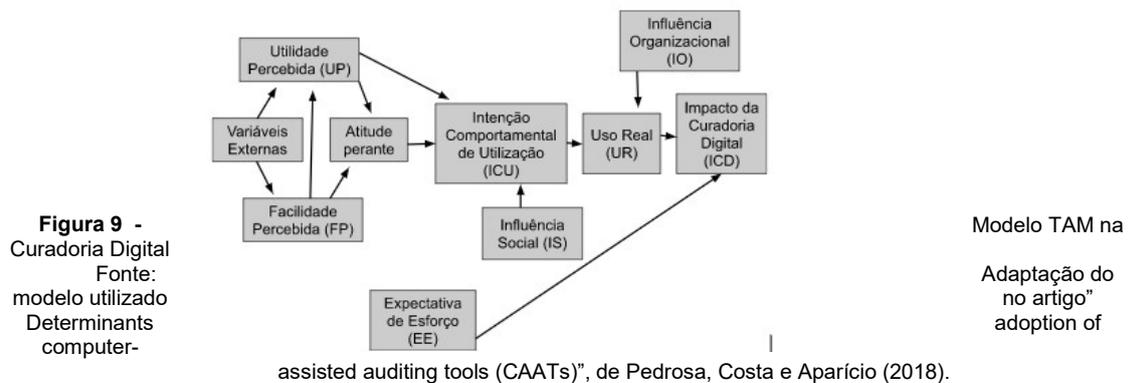
Complementar

Daniel, F; Silva, A; Ferreira, P. "Contributions to the Discussion on the Assessment of the Reliability of a Measurement Instrument". *Referência IV 7 (2015): 129-137*. <https://doi.org/10.12707%2Ffriv15003>.

Exemplo

Roberto, Andreia Raquel Ferreira. 2021 A curadoria de dados na gestão empresarial com estudo sobre a perspectiva empresarial portuguesa. Tese de Mestrado em Gestão Empresarial

O modelo construído foi inspirado no modelo TAM já mencionado anteriormente, tendo sido essencialmente uma adaptação do modelo utilizado no artigo “Determinants adoption of computer-assisted auditing tools (CAATs)”, de Pedrosa, Costa e Aparício (2018), conforme a figura abaixo.



Como podemos constatar na figura 9, foram oito as variáveis avaliadas neste modelo, sendo elas de forma mais detalhada:

1. **Facilidade:** esta variável está associada à dificuldade de utilização da tecnologia (neste caso, a curadoria digital), percebida pelo participante; para a avaliar as questões utilizadas incidiram sobre se a empresa possuía recursos e conhecimento necessários para a usar, assim como suporte por parte da equipa;
2. **Utilidade:** já nesta segunda o foco é se efetivamente a curadoria digital é útil no dia-a-dia da empresa, trazendo benefícios concretos para o trabalho dos colaboradores; neste sentido as questões focam-se nos tópicos da eficácia que traz aos processos, assim como se facilita a tomada de decisão;
3. **Esforço:** aqui estamos perante a perspectiva de esforço para utilização da curadoria digital, que irá diferir consoante vários fatores como cultura e estratégia da empresa, assim como receptividade a novas técnicas; desta forma, é questionado aos participantes se consideram que esta está integrada nas mesmas, e se a sua implementação é fácil;
4. **Influência externa:** em quarto lugar temos então os fatores externos à empresa que irão influenciar a sua atitude, nomeadamente se o mercado envolvente e os seus concorrentes apoiam a utilização da curadoria digital; neste sentido, as questões focam-se nestes pontos;
5. **Influência interna:** por outro lado, aqui o objetivo é perceber se os gestores/CEO’s suportam e apoiam a utilização da curadoria digital; as questões suportam-se então nestes elementos, assim como se a empresa como um todo apoia essa utilização;
6. **Intenção:** já nesta variável o foco é perceber se os participantes têm verdadeiramente o objetivo de utilizar a curadoria digital no seu dia a dia; desta forma, as questões incidem essencialmente sobre se na sua empresa existe o hábito de organizar e preservar documentos para possível posterior consulta;

7. **Uso real:** em penúltimo lugar, esta variável tem como objetivo perceber se a empresa onde o participante trabalha utiliza de facto a curadoria digital no seu dia-a-dia; neste sentido, as questões incidem sobre questões práticas como se organizam os documentos, possuem algum colaborador especificamente para a curadoria digital, entre outros;
8. **Impacto:** por fim, a última variável, como o nome indica, pretende confirmar se efetivamente os participantes são da opinião de que a curadoria digital tem um impacto positivo na empresa, permitindo um melhor desempenho, aumentando a produtividade da empresa pelo facto de acederem em tempo real aos dados, entre outros, sendo as questões direcionadas para estes tópicos.

Capítulo III - As Hipóteses

Para aplicação foi feita uma adaptação à Curadoria Digital, e tendo em conta que Taylor e Todd (1995b) e Temesgen (2005) estudaram o efeito da influência dos pares na influência social e na intenção comportamental, as hipóteses foram as seguintes:

1. H1: Os pares têm um impacto positivo na Influência Social;
2. H2: O número de gestores na empresa tem um impacto positivo na Influência Social;
3. H3: O número de colaboradores na empresa tem um impacto positivo na Influência Social;
4. H4: A Influência Social tem um impacto positivo na Intenção Comportamental de Utilização da Curadoria Digital (ICU);
5. H5: A utilidade percebida tem um impacto positivo na ICU;
6. H6: A Expectativa de Esforço tem um impacto positivo no Impacto Individual;
7. H7: A Facilidade percebida tem um impacto positivo no Uso Real do Sistema (UR);
8. H8: A Intenção Comportamental de utilização da Curadoria Digital tem um impacto positivo no UR;
9. H9: O UR influencia de forma positiva o Impacto da Curadoria Digital.

Estrutura do Inquérito

No que toca à construção da estrutura do inquérito e respetivas questões foi tido em conta, como referido anteriormente, o estudo de Pedrosa, Costa e Aparício (2018), tendo este sido adaptado ao caso específico da Curadoria Digital, como podemos constatar no **Quadro 3**.

Tópico	Código	Indicadores	Fonte
Facilidade Percebida	FP1	A minha empresa possui os recursos necessários para utilizar a Curadoria Digital.	Venkatesh et al. (2003)
	FP2	A minha empresa possui o conhecimento necessário para utilizar a Curadoria Digital.	
	FP3	A minha empresa possui suporte da equipa para utilizar a Curadoria Digital.	
Utilidade Percebida	UP1	A Curadoria Digital trás eficácia aos processos.	Davis (1989)
	UP2	A Curadoria Digital auxilia na tomada de decisão.	
	UP3	A Curadoria Digital é útil no trabalho da empresa.	

Expectativa do esforço	EE1	A Curadoria Digital está integrada na estratégia, cultura e nos vários departamentos da empresa, sendo a sua utilização clara e compreensível.	Venkatesh et al. (2003)
	EE2	É fácil para a empresa tornar-se competente no uso da Curadoria Digital.	
	EE3	A implementação da Curadoria Digital é fácil para a empresa.	
Influência Social	IS1	O mercado/ concorrentes que influenciam o comportamento da empresa acham que esta deve utilizar a Curadoria Digital.	Venkatesh et al. (2003)
	IS2	As referências no nosso setor entendem que a empresa deve utilizar a Curadoria Digital.	
Influência da Empresa	IE1	Os gestores/CEO's responsáveis da empresa facilitam a utilização da Curadoria Digital.	Venkatesh et al. (2003)
	IE2	Em geral, a nossa empresa apoiou o uso da Curadoria Digital.	
Impacto da Curadoria Digital (CD)	ICD1	A CD permite que a empresa realize tarefas mais rapidamente.	Urbach et al. (2010a)
	ICD2	AA CD aumenta a produtividade da empresa.	
	ICD3	A CD permite um melhor desempenho.	
	ICD4	A CD fornece informação sobre clientes e fornecedores em tempo real.	
	ICD5	A CD permite aceder a todas as dimensões da empresa em tempo real.	
	ICD6	A CD permite atuar em tempo real.	
	ICD7	A CD permite acompanhar todos os processos em tempo real.	
Intenção Comportamental de Utilização	ICU1	Na nossa empresa digitalizamos documentos que tenhamos apenas em formato físico, para posterior consulta.	Andreia Roberto (2021) Construção própria pela autora deste projeto.
	ICU2	Na nossa empresa achamos que os processos dos clientes devem ser guardados durante alguns anos, pois podem ser-nos úteis a longo prazo.	
	ICU3	Na nossa empresa gostaríamos de ter acesso a todos os documentos (faturas, notas de encomenda, contratos, fichas de funcionários), de forma intuitiva e rápida.	
	ICU4	Se a nossa empresa praticasse a curadoria digital poderíamos obter melhores resultados.	
Uso Real do Sistema	UR1	Na nossa empresa organizamos todos os documentos em formato digital por temas/pastas.	Andreia Roberto (2021)
	UR2	Na nossa empresa temos alguém responsável pelos dados/documentos da empresa, que os recolhe e organiza.	

	UR3	Na nossa empresa fazemos um filtro por todos os dados que possuímos, de forma a avaliar quais deles devem ser armazenados e quais devem ser eliminados.	Construção própria pela autora deste projeto.
	UR4	Na nossa empresa colocamos ao dispor da comunidade de potenciais utilizadores os dados que possuímos no nosso repositório.	
	UR5	Na nossa empresa reutilizamos dados que foram armazenados anteriormente, sendo estes úteis na tomada de decisões.	
	UR6	Na nossa empresa utilizamos ferramentas de gestão e organização digitais.	

Quadro 3 - Questões Curadoria Digital Inquérito

Fonte: Adaptação do modelo utilizado no artigo "Determinants adoption of computer-assisted auditing tools (CAATs)", de Pedrosa, Costa e Aparício (2018).

Ao observar a tabela abaixo conseguimos perceber que a correlação entre a Utilidade e a Facilidade é a mais forte, apresentando um valor de 0,556.

Podemos constatar que são estas as correlações mais fortes, sendo que a Facilidade só atua sobre a Utilidade; esta só influencia a Intenção; que o Uso real é a variável influenciada por maior número de variáveis (Influencia, Intenção e Esforço); esta última atua sobre o impacto.

	Estimate
Utilidade_1 <--- Facilidade_1	0,556
Intenção_1 <--- Utilidade_1	0,434
Uso_real_1 <--- Influencia_1	0,213
Uso_real_1 <--- Intenção_1	0,253
Uso_real_1 <--- Esforço_1	0,271
Impacto_1 <--- Uso_real_1	0,287

Tabela 3 - Valores de correlação de variáveis

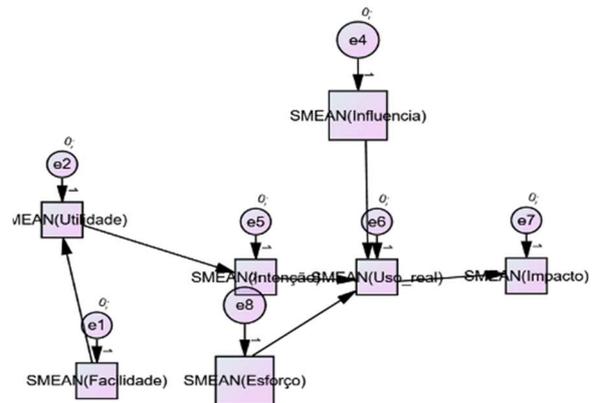


Figura 10 - Grafo de correlação de variáveis