

## ESTUDOS CORRELACIONAIS EM EDUCAÇÃO: POTENCIALIDADES E LIMITAÇÕES

Clara Pereira Coutinho<sup>(1)</sup>  
Instituto de Educação e Psicologia Universidade do Minho

### Resumo

Este artigo descreve os métodos de investigação correlacional e a forma como estes podem ser usados na pesquisa educativa em estudos descritivos que envolvem a análise de relações entre variáveis. Nesse sentido, começaremos por diferenciar a investigação correlacional de outros métodos de pesquisa educativa, equacionando tanto as suas vantagens como as suas limitações. Numa fase seguinte, procederemos à caracterização dos planos que integram esta família metodológica, apresentando exemplos de contextos de investigação em que podem ser aplicados. Por último, apresentaremos as técnicas estatísticas que permitem aferir da magnitude da relação entre variáveis, do poder de previsão das mesmas bem como técnicas de análise multivariada que tanto potencial podem ter no estudo e análise estatística da complexidade dos fenómenos educativos mas que, no nosso país, são ainda muito pouco utilizadas na investigação realizada no âmbito das Ciências da Educação.

PALAVRAS-CHAVE: *Investigação, educação, métodos, correlação, análise multivariada.*

### 1. Introdução

Sob a designação genérica de "correlacionais" englobamos todos os planos de investigação educativa em que, de uma forma ou de outra, o objectivo central da pesquisa é encontrar e avaliar a intensidade de relações entre variáveis, sem manipulação e sem pretensões de causalidade (Almeida e Freire, 1997; Anderson e Arsenault, 1999; Black, 1999; Meltzoff, 1998; Moore, 1983; Stern e Kalof, 1996). A busca de relações ou de associações entre as variáveis processa-se através de procedimentos estatísticos que nos fornecem uma medida quantificada do grau de relação entre as variáveis, a que se costuma chamar de *coeficiente de correlação*.

Trata-se, contudo, de uma família metodológica tão ampla, tanto ao nível dos procedimentos metodológicos como dos objectivos científicos que visa al-

Morada (address): Instituto de Educação e Psicologia, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Telemóvel: 9381 898. Email: ccoutinho@iep.uminho.pt

<sup>(1)</sup> Professora Auxiliar Departamento de Currículo e Tecnologia Educativa do Instituto de Educação e Psicologia, Universidade do Minho

cancar, que continua a persistir em torno destes modelos metodológicos uma clara ambiguidade conceptual patente nas múltiplas definições e propostas de classificação que, na literatura, surgem associadas a este métodos. A título de exemplo, e tal como referem Bravo e Eisman (1998), enquanto para uns os métodos correlacionais se podem incluir no grupo mais amplo dos estudos chamados de *comparativos* (cf. Pereda citado em Bravo e Eisman, 1998), para outros autores os métodos correlacionais e os comparativos são ambos modalidades da chamada investigação *ex post facto* (Bisqueria, 1989); a investigação *ex post facto*, por seu lado, seria, na perspectiva de Campbell e Stanley (1963) um caso especial dos planos ditos *quase-experimentais*.

Pela parte que nos toca e na ausência de unanimidade, vamos adoptar a posição dos que consideram os estudos correlacionais como uma variante dos desenhos descritivos caracterizados por fazerem a ponte entre os métodos compreensivos da realidade (estudos qualitativos) e os explicativos (estudos experimentais) (Almeida e Freire, 2000). Em relação aos primeiros, os métodos correlacionais conseguem ir para além da mera descrição dos fenómenos uma vez que possibilitam que o investigador estabeleça relações entre as variáveis, quantificando, inclusive, tais relações; relativamente aos segundos (estudos experimentais), estes modelos não conseguem fundamentar o significado da *causalidade* ao nível das relações encontradas (Lukas e Santiago, 2004).

De um maneira geral, podemos apontar como características individualizadoras destes modelos metodológicos os seguintes aspectos: a) o investigador trabalha em contextos naturais, b) sem exercer qualquer tipo de influência directa sobre o curso dos acontecimentos (não se manipulam variáveis), e c) serem estudos destinados a testar hipóteses que permitam corroborar teorias ou confrontá-las (Bravo e Eisman, 1998; Lukas e Santiago, 2004). Relativamente a este último aspecto convém ter presente que a qualidade de um estudo correlacional depende muito, como veremos, da fundamentação e profundidade dos constructos teóricos de que derivam as hipóteses do investigador. Sem fundamentação teórica prévia, a busca de relações e, sobretudo, a interpretação dos coeficientes de correlação encontrados pode ser um exercício académico estéril ou mesmo inútil (Borg e Gall, 1989).

## 2. O âmago da investigação correlacional

Para desenvolver um estudo de análise correlacional temos obrigatoriamente de ter medições/pontuações em pelo menos duas variáveis distintas; a primeira variável da qual obtemos dados empíricos costuma ser designada

pela notação X e a segunda por Y. Na terminologia da investigação correlacional a variável X é denominada variável antecedente ou *preditora* e a variável Y costuma ser designada por *variável crítico* ou *predita* (Cohen, Manion e Morrison, 2001). Na prática, o objectivo último de um estudo deste tipo é, pois, o de obter pares de pontuações X e Y para cada sujeito da amostra, e essa é uma das características identificadoras deste modelo metodológico que estabelece a fronteira relativamente aos planos de investigação que designamos por experimentais (Borg e Gall, 1989).

Nesse sentido, quando dizemos que há uma relação entre duas variáveis X e Y, queremos referir que as pontuações nas duas variáveis covariam, ou seja, mudam ao mesmo tempo, uma com a outra. Por exemplo, acredita-se hoje que os hábitos de leitura influenciam o rendimento académico; assim sendo, se existir de facto uma relação entre as duas variáveis o número de horas de leitura e rendimento académico devem *covar* (cf. Heiman, 1996, p. 182 e ss.).

Dizer que duas variáveis estão relacionadas (ou correlacionadas) não significa que se trate de uma relação causal entre elas. De facto, para se demonstrar uma relação de causalidade entre as duas variáveis que nos serviram de exemplo, haveria que se planificar um estudo experimental e manipular o número de horas de leitura para aferir de uma eventual relação causal; teríamos de formar, por exemplo, três grupos equivalentes, em que um grupo de sujeitos fosse classificado como leitores assíduos, outro grupo como leitores médios e um terceiro de sujeitos sem hábitos de leitura. No final do ano lectivo, todos os sujeitos seriam alvo de uma avaliação num mesmo teste geral de avaliação de conhecimentos e, só depois, se poderia inferir se os hábitos de leitura, resultam, de facto, num melhor desempenho académico dos sujeitos.

Contudo, é possível que esta mesma questão de investigação possa ser alvo de uma abordagem de tipo correlacional. Neste caso, não haveria manipulação ou selecção de pessoas para condições experimentais específicas como no caso anterior. Pelo contrário, apenas se mediriam os mesmos sujeitos nas duas variáveis para averiguar se a relação estava (ou não) presente. Poderíamos, por exemplo, perguntar aos alunos quantos livros tinham lido no último mês e avaliar o desempenho de cada um pelo cálculo da média das avaliações finais do ano anterior. Por um procedimento idêntico poderíamos saber, por exemplo, da relação entre as pontuações dos alunos num teste de personalidade e as suas pontuações numa escala de atitudes, ou se haveria relação entre um bom desempenho em Matemática e Português correlacionando as notas finais dos mesmos sujeitos nessas disciplinas nos exames finais do 12º ano, etc, etc.

Tal como o nome indica, na investigação correlacional usam-se sempre indicadores estatísticos da relação para sumariar a magnitude da correlação entre as variáveis. No entanto, não basta que se usem coeficientes de correlação para que um estudo seja correlacional: é a ausência de manipulação de variáveis que transforma automaticamente um estudo em correlacional. Esse é o âmbito da investigação correlacional.

Em suma, podemos concluir que a investigação correlacional se caracteriza por colocar hipóteses e objectivos de investigação sobre factos e fenómenos que não são intrinsecamente manipuláveis, com o objectivo de decifrar e analisar relações subjacentes aos mesmos numa tentativa de exploração de possíveis relações entre elas (Lukas e Santiago, 2004; Borg e Gall, 1989).

### 3. Correlação e causalidade

Quando dizemos que há uma relação entre duas variáveis X e Y, temos tendência para inferir de uma possível causalidade, ou seja, que X causa Y. Assim, se num estudo verificarmos que os sujeitos menos ansiosos também têm tendência para ter melhor desempenho académico, temos quase sempre a tentação de querer concluir que mais ansiedade é sinónimo de pior desempenho.

Esse é um dos problemas subjacente aos estudos correlacionais e de que devemos estar bem conscientes: a investigação correlacional apenas nos diz como duas variáveis X e Y se associam ou relacionam, e o facto de haver uma relação entre as duas variáveis não significa de forma alguma que a relação é causal, ou seja, que alterações numa variável causem alterações na outra (Heiman, 1996).

Haver associação ou relação não significa que se trate de uma relação de causalidade. Na perspectiva de Heiman (1996), há dois requisitos ou exigências para se poder inferir de relações causais entre variáveis. A primeira é que a variável X tem de ocorrer temporalmente sempre antes de Y. Ora na investigação correlacional podemos não conseguir sabê-lo de forma inequívoca. Voltando ao exemplo atrás referido, nós apenas avaliamos as horas de leitura dos sujeitos e o seu desempenho académico. Nada nos garante que tenham sido os hábitos de leitura a aumentar o rendimento académico; não seria antes o rendimento académico a motivar o interesse pela leitura? Ou seja, em qualquer correlação tanto pode ser X a causar Y, como Y a causar X.

Em segundo lugar, a variável X teria de ser a única a causar Y. Acontece que na investigação correlacional nunca somos capazes de controlar ou eliminar outras variáveis que, potencialmente, podem causar as mudanças ocorri-

das. Por exemplo, pode acontecer que os sujeitos que dizem gostar mais de ler sejam também mais inteligentes, e nesse caso nunca poderemos saber com certeza se são os hábitos de leitura ou a inteligência os responsáveis pelo melhor aproveitamento dos sujeitos.

Em suma, uma correlação, por si, não permite aferir de causalidade. Só em estudos experimentais em que nos é possível manipular a variável independente (VI) e manter sob controlo todas as outras potenciais variáveis estranhas capazes de contaminarem a variável dependente (VD), é que temos base de sustentação para aferir de uma relação causal entre variáveis (Charles, 1998; Moore, 1983; Punch, 1998; Schutt, 1999; Stern e Kalof, 1996; Vogt, 1999).

### 4. Investigação correlacional versus experimental

Se a manipulação de variáveis é o ponto de clivagem crucial no que toca a separar a investigação correlacional da experimental, outros aspectos podem ainda ajudar a diferenciar estas duas grandes famílias metodológicas da investigação educativa (cf. Heiman, 1996):

1. A investigação experimental normalmente utiliza dois (ou mais) grupos de sujeitos e a análise estatística baseia-se na comparação de dois (ou mais) valores grupais (normalmente a média das pontuações de cada um dos grupos); na investigação correlacional usamos apenas um grupo e sumariamos o valor da relação entre as variáveis em análise num valor único calculado a partir de todos os pares de pontuações X-Y dos sujeitos que a integram (N representa o número de pares de pontuações/observações);
2. Ao contrário da metodologia experimental em que a VI (variável independente) tem de ser sempre anterior no tempo à VD (variável dependente), num estudo correlacional nunca podemos identificar qual é a variável anterior e a posterior e por isso não podemos falar de variáveis independentes ou dependentes. Então como decidimos qual das variáveis é X e qual é a Y? Tudo depende da forma como formulamos as nossas hipóteses de investigação. Se, relativamente ao problema que nos tem servido de exemplo a questão de investigação for: "Será que os alunos que lêem mais têm melhor desempenho académico?" então a leitura será a variável X, também denominada variável antecedente ou preditora e o desempenho académico a variável critério ou predita. De forma inversa, se a questão de in-

investigação for: "Será que um melhor desempenho académico fomenta o gosto pela leitura?" então X será o desempenho e Y o gosto pela leitura.

3. Por último, na investigação correlacional a representação gráfica dos pares de pontuações X-Y dos sujeitos em cada uma das variáveis faz-se com recurso aos chamados gráficos de pontos ou *scatterplots* o que não acontece nos estudos experimentais em que se usam gráficos de barras, histogramas, etc.

## 5. Classificação dos estudos correlacionais

De acordo com Cohen, Manion e Morrison (2001, p. 199) os estudos correlacionais podem ser divididos em duas grandes famílias: os estudos de relação e os estudos de previsão. Nos primeiros, o objectivo central da investigação é tentar compreender melhor a complexidade do fenómeno educativo estudando as relações entre as variáveis sobre as quais o investigador coloca hipótese sobre possíveis associações. Nesse sentido, trata-se de uma metodologia de investigação particularmente importante em áreas do conhecimento em que ainda há pouca investigação realizada e em que há necessidade de encontrar pistas que possibilitem a compreensão da complexidade do fenómeno educativo pela descoberta de eventuais relações entre os constructos e conceitos que o integram e explicam (Anderson, 1998; Charles, 1998; Lukas e Santiago, 2004).

Talvez por isso a investigação correlacional de tipo relacional seja muitas vezes alvo de intensas críticas por parte de alguns autores que subvalorizam a qualidade científica de tais abordagens metodológicas (cf. Cohen, Manion e Morrison, 2001). É o que acontece quando os estudos correlacionais são conduzidos sob o que alguns designam de *shotgun approaches* (Borg e Gall, 1989, p. 581) ou ainda *looking approaches* (Anderson, 1998, p. 113) e que se caracterizam pelo facto do investigador partir à procura de relações entre as variáveis sem ter em conta um referencial teórico que norteie e justifique a sua busca de eventuais relações/associações (uma lógica de *looking for* a que alude Anderson, *id*, *ibid*).

Tomemos como exemplo o conceito complexo e ambíguo de "eficácia docente" de que nos fala também Cohen, Manion e Morrison (2001). Trata-se de um constructo amplo que depende de outras variáveis de menor complexidade que operam em conjunto ou separadamente sobre ele. Factores como a inteligência, a motivação, a percepção de auto-eficácia, as

destrezas verbais, a empatia surgem logo na nossa mente como podendo afectar o desempenho docente. Uma revisão de literatura pode confirmar ou rejeitar tal possibilidade. Assim que um conjunto significativo de factores sejam identificados e justificados pela literatura, faz então sentido que sejam desenvolvidos instrumentos de medida para a sua avaliação. Estes instrumentos são então aplicados a uma amostra representativa de sujeitos e dferidas pontuações nas referidas variáveis; estas são depois correlacionadas com a medida da variável mais complexa que está a ser investigada, neste caso concreto, o conceito de "eficácia docente". Nesta fase exploratória a análise estatística consiste muitas vezes no cálculo simples dos coeficientes de correlação bivariados (porque aplicados a duas variáveis) que nos dão logo um esclarecimento adicional sobre a variável que queremos investigar. Neste exemplo, o estudo correlacional permitiu obter dados que podem ser usados e desenvolvidos por outros investigadores como hipóteses adicionais em estudos futuros, e é essa a mais valia que pode advir de um estudo correlacional de tipo exploratório bem conduzido, porque sustentado numa correcta revisão da literatura (cf. Cohen, Manion e Morrison, 2001, p. 199 e ss.).

Em contraste com o estudo exploratório de relação entre variáveis, os estudos correlacionais de previsão são desenvolvidos em áreas do saber em que o conhecimento está mais desenvolvido e alicerçado em investigação anterior (Heiman, 1996). A previsão a partir de técnicas de correlação baseia-se no pressuposto de que, pelo menos, um dos factores de que se tem uma avaliação/medição está presente e conduz/justifica o comportamento da variável que se quer prever (Punch, 1998). Se, por exemplo, quisermos prever a probabilidade de sucesso de um grupo de vendedores num curso de formação intensivo, devemos começar por encontrar as variáveis que, em estudos prévios, se verificou estarem relacionadas com o sucesso nas vendas. Essas características poderiam ser o empenhedorismo, a capacidade de comunicação verbal, a motivação para o sucesso, a maturidade emocional, a sociabilidade, etc. O grau de correlação com que estes preditores se relacionam com o comportamento que queremos prever, neste caso, o sucesso a vender, determinará a acurácia/justeza da nossa previsão. As variáveis cruciais para o sucesso não podem ser previstas se não estiverem presentes no momento em que se faz a previsão. Em termos de potencial preditivo, o grau de associação entre as duas variáveis tem de ser muito forte, e quanto mais forte maior o poder preditivo. Em teoria, só uma correlação perfeita evitaria erros na previsão de uma variável conhecida a outra. Também é importante que a previsão se reporte a um futuro próximo; previ-

sões de longo prazo terão poucas possibilidades de sucesso porque as variáveis mudam com o evoluir do tempo.

Não devemos esquecer nunca que a correlação é um conceito grupal, uma medida genérica importante para se prever o desempenho de todo um grupo no seu conjunto ou seja, válida para todo o grupo que é alvo de uma avaliação/medição; se é possível prever que um grupo de crianças sobredotadas como um todo terá sucesso académico, já não é possível prever com exactidão que uma dada criança sobredotada venha a ser excelente. Por fim, baixos coeficientes de correlação têm fraco poder preditivo e só correlações altas podem ser tidas em conta em estudos de previsão (Cohen, Manion e Morrison, 2001, p.200 e ss).

## 6. O coeficiente de correlação

As diversas técnicas estatísticas utilizadas em estudos de relação têm como objectivo responder a três questões acerca da associação entre variáveis: 1<sup>ª</sup>: há relação entre as variáveis? Se a resposta for "Sim", então 2<sup>ª</sup>: Qual o sentido da relação e 3<sup>ª</sup>: Qual a sua magnitude?

O coeficiente de correlação dá-nos a medida quantificada do grau de associação entre duas variáveis e assume valores que podem variar entre -1 e +1. Se os dois conjuntos de dados variam no mesmo sentido (os dois crescem ou decrescem) como acontece, por exemplo, com as variáveis inteligência e desempenho académico, temos o que se designa por uma *correlação positiva*. Na *correlação negativa* uma variável cresce e a outra decresce, como acontece, por exemplo, entre o número de erros que uma criança dá num ditado e a nota que obtém.

### 6.1 Tipos de coeficientes de correlação

Existem coeficientes específicos para o cálculo da correlação entre variáveis dependendo da escolha do investigador da natureza e do tipo das variáveis que pretende correlacionar. Assim sendo,

1. no caso da variável ser nominal, usa-se o *coeficiente de contingência* que parte de uma tabela de contingências 2x2 (variáveis dicotómicas), ou o *coeficiente Cramer C* quando as variáveis apresentam mais do que duas categorias (cf. Siegel e Castellan, 1988, p.225);
2. no caso das variáveis apresentarem uma ordenação (ou seja, serem ordinais) podem usar-se o *coeficiente de correlação de Spearman rank*

order ( $r_s$ ) ou o coeficiente de correlação de Kendall, entre outros (cf. Siegel e Castellan, 1988, p.245);

3. se uma variável é contínua (por exemplo "aproveitamento em Matemática") e a outra é dicotómica (por exemplo "ter/não ter computador pessoal") pode usar-se o *coeficiente de correlação biserial* ( $r_{pb}$ ) (cf. Charles, 1998, p.267; Heiman, 1006, p. 200);
4. quando ambas as variáveis são intervalares, a medida da correlação é muito mais potente e rigorosa, e obtém-se pelo *coeficiente de correlação de Pearson* ou *produto-momento* ( $r$ ), que nos mostra tanto a direcção como a força (o como e o quanto) da relação entre as duas variáveis (cf. Punch, 1998)

A dispersão dos pontos também nos dá indicação da força da correlação: quanto mais os pontos se concentram mais acentuada é a elipse e maior a correlação, quanto mais a elipse tende para a forma circular menor a magnitude da correlação (Wiersma, 1995). A figura 4, adaptada de Shavelson (1996, p.153) apresenta-nos alguns exemplos de correlações lineares, com os correspondentes valores da magnitude do coeficiente de correlação  $r$ , e onde podemos ver como há uma associação entre a forma do scatterplot e a magnitude da relação entre as variáveis.

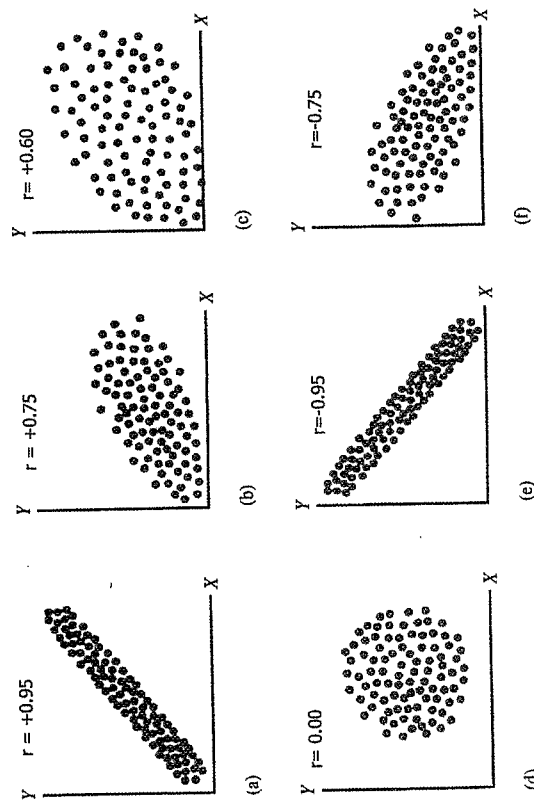


Figura 4 – Representação gráfica de alguns coeficientes de correlação

## 6.2 Interpretação do coeficiente de correlação

Calculado o valor do coeficiente de correlação  $r$ , há que se proceder à sua interpretação o que nem sempre é tarefa simples. Para Cohen, Manion e Morrison (2000), a interpretação de um coeficiente de correlação deve fazer-se segundo três perspectivas:

- examinando a pontuação obtida, ou seja o valor numérico do coeficiente;
- analisando a significância estatística da correlação;
- mediante a análise do quadrado do coeficiente de correlação (o coeficiente de determinação).

O *valor numérico da correlação* dá-nos logo uma ideia da força da relação entre as variáveis como referimos atrás: valores perto de zero indicam correlações fracas e próximos de +1 ou -1 correlações fortes. As dificuldades prendem-se precisamente na interpretação dos valores intermédios e é, por isso, que se revela importante aferir da *significância estatística da correlação*. A questão pode colocar-se do seguinte modo: *quão forte é que deve ser uma correlação para podermos aceitá-la como válida, ou seja, como sendo estatisticamente significativa?*

Vamos imaginar que calculámos o coeficiente de correlação entre as variáveis "altura" e "peso" de 30 crianças e obtivemos um coeficiente de correlação com o valor de 0,3879. Será que posso concluir que as duas variáveis estão correlacionadas? Ou seja, a correlação terá significância estatística? Para o sabermos teremos de consultar uma tabela de valores de significância estatística (em qualquer manual de estatística); neste caso concreto, verificamos que, para um nível de significância de menos de 0,05 (nível mais divulgado na investigação educativa), o valor do coeficiente de correlação para 30 sujeitos terá de ser igual ou superior a 0,36. No nosso caso, o valor encontrado é superior pelo que podíamos concluir que a relação entre as duas variáveis, para o grupo de 30 sujeitos avaliados nas duas variáveis tinha significância estatística. O nível de significância estatística varia em função do número de elementos da amostra por isso é fundamental que, para além do valor numérico do coeficiente de correlação, o investigador tenha em conta a dimensão da amostra quando está a interpretar. Quanto maior for o tamanho da amostra menor é o valor numérico necessário para que a correlação seja significativa mas a inversa é igualmente verdadeira (no nosso exemplo, se em vez de 30 tivéssemos apenas 10 sujeitos o valor do coeficiente de correlação teria de ser muito superior para ter significância estatística, neste caso, ser superior a

0,65). A questão da significância estatística de uma correlação é particularmente importante em estudos correlacionais exploratórios, porque, em estudos de previsão, o que importa mesmo é a magnitude da relação.

Por último, outro indicador que nos pode ajudar a interpretar melhor esta questão da intensidade de uma correlação é a análise do *quadrado da correlação*, o chamado *coeficiente de determinação*  $r^2$ . Este valor indica-nos a proporção de variância de uma primeira variável X que pode ser atribuída a uma relação linear com a segunda variável Y. Dito de outra forma, indica-nos a proporção de variância comum às duas variáveis X e Y.

Por exemplo: se duas variáveis X e Y têm um coeficiente de correlação de 0,65, então  $42,5\%$  (quase metade) da variação de X pode ser fruto da tendência que X tem de covariar com Y (ou de variar em função de Y). Na figura 5 está representada graficamente uma correlação de 0,65 entre duas variáveis X e Y:  $0,65^2 = 42,25$ , o que significa que as duas variáveis partilham  $42,25\%$  da variância comum.

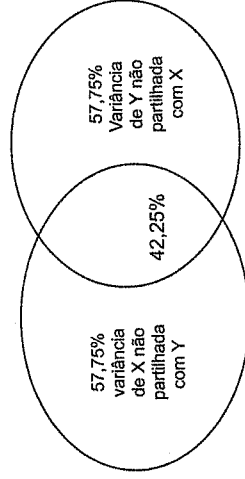


Figura 5 – Representação gráfica de uma correlação de 0,65 entre X e Y

Embora os três procedimentos indicados nos ajudem a interpretar um dado coeficiente de correlação é importante termos sempre presente que há muitos outros factores que podem influenciar o valor de uma correlação e que, por isso mesmo, há sempre que ser muito cauteloso no que toca a extrapolar resultados para além dos dados obtidos na amostra a que o coeficiente de correlação se refere.

No entanto e sempre numa lógica de ajudar os investigadores a analisarem os coeficientes de correlação obtidos num estudo concreto surgem muitas vezes, na literatura, informações que podem ajudar na tarefa de interpretar resultados; tal é o caso da proposta de Borg (1963), que apresentamos no quadro 1 abaixo representado, e que refere índices aproximativos para estimar a magnitude de coeficientes de correlação relativos a amostras de 100 (ou mais) sujeitos.

Quadro 1.: Sugestões para interpretar a magnitude de coeficientes de correlação

Coeficientes de correlação	Interpretação
Entre 0,35 e 0,65	Revelam relações muito débeis entre variáveis mesmo que possam ser significativas. Uma correlação de 0,20 mostra que apenas 4% da variância é comum às duas variáveis
Entre 0,35 e 0,65	Neste nível as correlações têm significância estatística para o nível de <math>\alpha=0,1</math>. A nível de estudos de previsão, podem ser importantes analisadas em conjunto com outras numa análise de regressão. Usadas isoladamente são de muito pouca utilidade para previsão individualizada.
Entre 0,65 e 0,85	A partir deste nível é possível fazer previsão para grupos com alguma certeza.
Superiores a 0,85	Este nível revela uma relação muito forte entre as variáveis correlacionadas. Uma correlação de 0,85 indica que a medida usada para prever tem cerca de 72% de variância comum com a variável prevista. Estudos de previsão em educação raramente obtêm correlações deste nível mas, caso se obtentham, são muito úteis seja para previsão individual seja de grupo.

### 6.3 Curvilinearidade

Até agora assumimos sempre as correlações serem lineares, ou seja que as variáveis covariavam no mesmo sentido, o que leva a que, no gráfico do *scatter-plot*, a correlação se materialize sob a forma de uma linha recta que é possível traçar entre os pontos. No entanto, na investigação educativa, nem sempre é possível assumir a linearidade na relação entre duas variáveis, como é o caso, por exemplo da variável stress: um nível baixo de stress pode aumentar positivamente a performance dos sujeitos mas se for em demasia pode levar a um retrocesso no rendimento. O resultado não é uma relação linear mas uma linha curva indicadora da existência de uma correlação curvilinear. Da mesma forma acontece, por exemplo, com a temperatura da casa e o grau de conforto: aquecer a casa aumenta o conforto - correlação positiva, mas aquecer demais pode ser factor de desconforto - correlação negativa. Outros exemplos de correlação curvilineas entre variáveis sócio-educativas são os que a seguir se indicam:

- Pressão do director da escola sobre os professores
- Pressão do professor sobre o rendimento do aluno
- Grau de desafio e desempenho académico
- Idade e concentração
- Idade e sociabilidade
- Idade e aptidão cognitiva.

O que se pretende dizer é que, na prática da pesquisa, antes de assumir o pressuposto da linearidade o investigador deve estar sensível à circunstância de isso não acontecer: nesse caso, ou seja, se a curvilinearidade de for uma possibilidade a admitir, terá de utilizar técnicas estatísticas apropriadas a correlações não lineares entre as variáveis como é o caso do coeficiente  $\eta$ (eta) (cf. Cohen, Manion e Morrison, 2001, p. 198), ou então, em alternativa, proceder a uma transformação da correlação não linear em linear.

## 7. Uso da correlação para previsão: estudo da regressão

A regressão linear consiste no processo que nos permite determinar o valor a obter numa variável, conhecida a outra variável. A variável que permite a previsão é, como vimos, chamada de *preditora* e a variável a prever costuma chamar-se *variável critério*. A previsão em educação pode tomar formas muito variadas, por exemplo prever o aproveitamento/sucesso (critério) a partir de testes de inteligência (preditora) como se faz em certas universidades americanas para seleccionar os alunos que aí querem ingressar. A precisão da previsão é tanto maior quanto maior for a correlação entre as duas variáveis (preditora e critério) (Heiman, 1996).

Vejamos então como a correlação fundamenta o modelo de regressão linear que vínhamos delineando: se duas variáveis estão correlacionadas (seja positiva ou negativamente) tendem a agrupar-se ao longo de uma linha diagonal que cruza o eixo do *scatter* respectivo. Podemos então tentar traçar a linha para onde tendem os pontos do gráfico. Essa linha - *line of best fit* - "...que melhor se ajusta aos pontos" (Calder, 1996: 254), obtém-se calculando a média dos desvios dos pares de pontuações, e substitui como que "...uma linha de previsão das duas variáveis" (Punch, 1998:123): a recta de regressão linear simples.

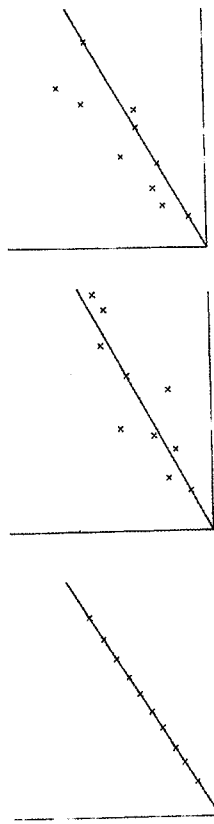


Figura 7 - Rectas de regressão perfeitamente ajustada, ajustada e pouco ajustada aos dados



Com o auxílio do computador é possível ajustar a recta de regressão linear aos dados com rigor e precisão (Punch, 1998), e obter uma *linha perfeitamente ajustada, ajustada ou pouco ajustada* aos dados (ver figura 7, adaptada de Calder 1996, p.256). A medida do quanto *ajustada* é a recta de regressão aos dados, é-nos fornecida pelo valor do coeficiente de correlação de Pearson  $r$  respectivo: se for alto, significa que os pontos de agrupam junto à recta e o poder preditivo é grande, se baixo, os pontos afastam-se dela e o poder preditivo é reduzido (cf. Heiman, 1996).

### 8. Fases de um estudo correlacional

A planificação de um estudo correlacional é em tudo semelhante ao de um qualquer plano descritivo de cariz quantitativo (cf. Lukas e Santiago, 2004 e ainda Charles, 1998). Nesse sentido, regra geral, os passos que pressupõem o desenvolvimento de um plano de investigação deste tipo são os seguintes:

1. Formulação do problema de investigação;
2. Formulação de hipóteses acerca de relações/associações entre variáveis;
3. Procedimentos metodológicos:
  - a. Seleção e definição operacional das variáveis (fundamentadas na revisão de literatura);
  - b. Seleção ou elaboração do instrumento de recolha de dados;
  - c. Seleção da amostra (não aleatória mas que deve ser o mais homogénea possível, caso contrário as correlações podem ser obscurecidas pela presença de sujeitos que diferem muito uns dos outros);
4. Recolha de dados (por instrumentos tão variados como sejam testes estandarizados, questionários, entrevistas ou mesmo técnicas de observação);
5. Análise estatística dos dados e discussão dos resultados;
6. Redacção do relatório de investigação.

### 9. Tipologia dos desenhos correlacionais

Tal como seria de prever a tipologia dos estudos correlacionais está muito associada ao tipo de técnicas estatísticas utilizadas para o cálculo dos indicadores seja da relação seja da previsão; a proposta que passamos a apresentar baseia-se na proposta andóloga sugerida por Bravo e Buendía (1998), mas

conta ainda com o contributo de mais alguns dos autores que consultámos para a elaboração deste artigo (Calder e Sapsford, 1996; Heiman, 1996; Lukas e Santiago, 2004).

#### 9.1 Estudos de relação

Trata-se do modelo clássico dentro da família dos estudos correlacionais em que, como já referimos anteriormente, o objectivo do investigador é o de avaliar a magnitude da relação entre variáveis pelo cálculo e interpretação dos coeficientes de correlação (Almeida e Freire, 2000).

A análise de tipo relacional pode envolver a relação de duas ou mais variáveis. No primeiro caso, como vimos atrás, a análise diz-se bivariada e é o modelo mais simples dentro da análise correlacional; mesmo que a análise envolva várias variáveis, o investigador calcula os coeficientes de correlação entre duas das variáveis em estudo de cada vez e interpreta os resultados obtidos normalmente centrando a atenção nos indicadores mais altos e/ou mais baixos.

No segundo caso, ou seja, se a análise relacional envolver acção simultânea de mais do que duas variáveis a análise correlacional diz-se *multivariada*. A *regressão múltipla*, a *análise discriminante*, a *análise factorial*, a *análise de clusters* e a *análise de sendas* são os exemplos mais relevantes dentro desta família de modelos de análise correlacional multivariada (Borg e Gall, 1989; Calder e Sapsford, 1996).

Antes de passarmos a apresentar cada uma dessas técnicas não queremos deixar de referir a chamada *correlação parcial* que, de certa forma, faz a ponte entre os modelos de análise bi e multivariadas que vimos apresentando e que consta de todos os programas de análise estatística como é caso do SPSS 14.0. De facto, o cálculo da correlação parcial envolve a presença de, pelos menos, três variáveis mas não as analisa em simultâneo como acontece nos modelos de tipo multivariado; na correlação parcial o que o investigador faz é calcular o coeficiente de correlação entre duas variáveis mantendo uma terceira variável sob controlo ou neutralizada (cf. Cohen, Manion e Morrison, 2000). Um exemplo é a alta correlação que se verifica entre os valores do desempenho em testes de leitura e de vocabulário já que ambas as variáveis são também altamente correlacionadas com o factor inteligência. Tentar explorar relações entre as duas primeiras variáveis *per si*, implica neutralizar a influência da inteligência, o que é conseguido por meio do recurso ao cálculo da *correlação parcial* cujo coeficiente é designado por "Partial  $r$ " (Charles, 1998).



## 9.2 Regressão múltipla

A regressão múltipla é uma técnica de análise multivariada que permite determinar a relação entre uma variável critério e uma combinação de duas ou mais variáveis preditoras. Trata-se de uma das técnicas de análise multivariada mais popularizadas na investigação educativa pela sua grande versatilidade e pela qualidade da informação que nos pode fornecer acerca da relação entre múltiplas variáveis, já que permite estimar tanto a magnitude como a significância estatística da relação entre elas (Borg e Gall, 1989). Para além da investigação correlacional esta técnica pode ser utilizada para a análise de dados de estudos causais comparativos e até de estudos experimentais, podendo ser usada com variáveis de vários níveis ou escalas de medida (intervalar, ordinal e nominal).

O cálculo da equação da regressão múltipla envolve operações matemáticas bastante complexas que possibilitam a obtenção de coeficientes de previsão (os coeficientes  $\beta$ ) da variável critério tomando como base os valores obtidos nas diversas variáveis preditoras; quanto maior for o coeficiente  $\beta$  de uma variável preditora maior o seu peso relativo na equação da recta de regressão múltipla para a qual participam (contribuem) todas as variáveis preditoras na proporção da sua influência relativa (Borg e Gall, 1989; Cohen *et al.*, 2003).

Um exemplo concreto é-nos descrito por Borg e Gall (1989) tomando como base o estudo desenvolvido por Paula Jorde-Bloom e Martin Ford (1988): os dois investigadores queriam identificar factores que distinguíssem directores de escolas primárias que fomentavam o uso do computador na escola para fins administrativos e para usos pedagógicos (as duas variáveis critério) dos que não o faziam. A revisão de literatura fundamentou cinco ordens de factores possíveis: a) a experiência prévia com o computador, b) as expectativas face ao potencial das TIC, c) as percepções de auto-eficácia, d) propensão para inovar e e) o apoio externo. O objectivo da análise de regressão múltipla foi o de verificar, de entre os vários factores possíveis, quais os que se revelavam mais importantes na tomada de decisões dos directores de escola para os dois fins em vista (usos administrativos e pedagógicos). Com base na regressão múltipla foi possível verificar que, para usos administrativos o factor que mais pesava na decisão do director de escola era a "experiência prévia com o computador", mas, para usos pedagógicos a "percepção de auto-eficácia" era o factor determinante.

## 9.3 Análise discriminante

Permite diferenciar grupos de indivíduos segundo os seus perfis em função de um conjunto de variáveis. Usa-se muito em pedagogia diferencial para estabelecer grupos diferenciados de sujeitos como base, por exemplo, para conceber programas ou modelos educativos.

A análise discriminante é muito semelhante à da regressão múltipla na medida em que ambas as técnicas envolvem duas ou mais variáveis preditoras e uma variável critério. A análise discriminante, contudo, está limitada ao caso especial da variável critério ser dicotómica, como seja, por exemplo, o caso de uma pessoa pertencer ou não a um determinado grupo: ser homem ou mulher; ter bom/mau rendimento escolar; engenheiros versus médicos, etc, etc. A equação da análise discriminante usa as pontuações de uma pessoa na variável preditora para tentar prever a que grupo essa pessoa pode pertencer (critério) (Bravo, 1998; Anderson, 1998; Cohen *et al.*, 2003).

## 9.4 Estudos baseados na análise factorial

Estes estudos pretendem conhecer a estrutura interna de determinados factores capazes de explicarem um número mais amplo de variáveis, ou seja, trata-se de um método que pretende trazer à luz "padrões" de variáveis: um conjunto de variáveis estão correlacionadas e a presença de inter-correlações altas entre elas pode ser indiciadora da presença de um "factor" mais geral que as explica conceptualmente. Por exemplo, podemos medir vários aspectos ligados ao sucesso escolar, como seja as notas em disciplinas como Português e Matemática, as percepções de auto-eficácia, os hábitos de estudo, e concluir que altas correlações entre elas poderia revelar a presença de um factor comum capaz de as explicar a todas. Ou seja, se as variáveis não estão correlacionadas não poderá haver nenhum factor que as explique; no entanto, se um grupo de variáveis estão fortemente relacionadas pode existir acima delas uma variável mais genérica a que chamamos "factor" e que as explica a todas (Kline, 1986). A análise factorial constituiu-se como um método que fornece ao investigador uma base empírica para reduzir um grande número de variáveis que estão moderada a fortemente correlacionadas a um (ou mais) factores que as integra. Um factor representa as variáveis mais correlacionadas. Quando a busca dos factores é mais empírica que teórica utiliza-se o chamado método das componentes principais que se tem revelado muito útil para a análise e validação de instrumentos de medida de variáveis, em especial escalas de atitudes e de valores numa fase inicial de desenvolvimento (Anderson, 1998).

### 9.5 Análise de conglomerados ou de clusters

Consiste em classificar os sujeitos em grupos o mais semelhantes possível entre si e o mais diferentes possível em relação a outros grupos. Para além da sua utilização para a elaboração e refinamento de instrumentos de medida, também tem sido usada para detectar constructos pessoais e percepções de grupos sobre temáticas educativas. Em conjunto com os estudos de análise discriminante são também chamados de estudos de classificação (Stevens, 2002; Cohen et al., 2003).

### 9.6 Path analysis (Análise de sendas)

Servem para contrastar a plausibilidade ou a coincidência empírica de modelos teóricos causais. O seu principal objectivo é o de estudar os efeitos directos e indirectos de variáveis, consideradas "causas", sobre variáveis consideradas "efeito", sem manipular as mesmas. Podem considerar-se uma alternativa aos estudos experimentais puros que são muito difíceis de empreender, uma vez que, embora não sejam desenhados para o estabelecimento de relações causais entre variáveis, são no sim para comprovar o ajuste ou a coincidência com modelos teóricos causais previamente planeados (Borg e Gall, 1989; Bravo, 1998).

A questão central numa *path analysis* tem a ver com a interpretação que o investigador faz do modelo causal que quer testar e nunca deve ser conduzida com o fim de descobrir causas mas de as testar. Um estudo em que um modelo causal foi testado com base na análise de sendas foi o desenvolvido por Bárbara Munro (1981) sobre uma amostra nacional de alunos que tinham concluído o ensino secundário uns anos antes. Nesse estudo, a investigadora pretendia testar um modelo teórico que apontava como principal causa do abandono da universidade a incapacidade do estudante de fazer a transição para o ensino superior e de se integrar na vida académica. Através da *path analysis* foi possível obter um modelo causal mais completo e detalhado. De facto, as características dos alunos finalistas foram capazes de prever a integração académica mas não foram capazes de justificar de forma directa o abandono escolar. O desejo do aluno completar o curso foi o factor que teve o efeito mais positivo na decisão de permanecer na instituição e concluir os estudos superiores.

### 9.7 Estudos de previsão ou predictivos

Procuram prever comportamentos/attitudes (variável critério) a partir do conhecimento que se tem do valor de determinadas variáveis ditas predictoras. Baseiam-se na correlação e na regressão (linear, múltipla, polinómica, não li-

near). Através destas técnicas define-se um grupo de variáveis critério com o objectivo de ver os níveis de contribuição de cada variável predictor para o prognóstico da variável critério. Para além da validação de instrumentos de medida, as maiores aplicações são no campo da orientação escolar profissional e diagnóstico pedagógico (Lukas e Santiago, 2004)

O quadro que segue, traduzido de Bravo e Eisman (1998) sintetiza os principais métodos correlacionais apresentados com os objectivos que permitem alcançar.

Quadro2: Síntese dos principais métodos correlacionais

Métodos correlacionais	Características
Estudos correlacionais	Conhecer a magnitude da relação entre variáveis
Estudos preditivos	Prever determinados comportamentos/attitudes conhecidos os valores de variáveis predictoras
Regressão múltipla	Permite determinar a relação entre uma variável critério e uma combinação de variáveis predictoras
Análise factorial	Simplificar e organizar em estruturas mais gerais um grande número de variáveis
Análise discriminante	Diferenciar grupos em função do seu perfil característico num certo grupo de variáveis
Análise de conglomerados ou clusters	Classificar os sujeitos em conglomerados o mais semelhantes possível entre si ou o mais diferentes entre si relativamente a outros conglomerados
Análise de sendas ou path analysis	Contrastar a plausibilidade de modelos teóricos causais

## 10. Aplicação dos métodos correlacionais ao campo educativo

Os métodos correlacionais são muito úteis para nos ajudar a compreender acontecimentos, condições ou comportamentos que estão relacionados entre si (Charles, 1998).

Para Juste (1985), os métodos correlacionais podem ser desenvolvidos com os estudos preparatórios de estudos experimentais, já que a existência de

a correlação não estabelece relações de causalidade e não permite um controlo rigoroso das variáveis estranhas como acontece na investigação experimental, pelo que os resultados destes estudos terão sempre um valor precário.

2. a selecção das variáveis para análise não se fundamentar numa sólida revisão de literatura que inclua as teorias mais representativas e a investigação prévia como acontece nas abordagens de tipo *shotgun*;
3. utilizarem-se coeficientes bivariados que não são apropriados ao tipo de variáveis em estudo;
4. a imprecisão da informação fornecida pelos coeficientes de correlação que está muito condicionada pela fiabilidade do processo de recolha de dados;
5. as variáveis podem apresentar correlações débeis, comprometendo a relevância científica do estudo que se limita a sugerir outras formas mais rigorosas de pesquisa;
6. ter presente que os coeficientes de correlação não podem ser interpretados como valores absolutos, ou seja, que são relativos a uma dada população/grupo e que, muito provavelmente, não se repetem noutra;
7. falhas no desenvolvimento de variáveis critério capazes de servirem para estudos de previsão futuros;
8. em estudos que incorporam muitas variáveis, não se deve limitar a análise a coeficientes de correlação bivariados deve-se avançar com técnicas de análise multivariada que permitam clarificar como todas as variáveis se relacionam e interagem entre si;
9. falhas na interpretação prática da significância estatística dos coeficientes de correlação uma vez que, perante um grande número de casos, não são necessários coeficientes de correlação elevados para que esta se produza; apesar de poderem ser mais do que fruto do "mero acaso" (ou seja, apesar de terem significação estatística), tais correlações terão sempre pouca ou nenhuma validade prática;
10. Para o caso da análise de sendas (*path analysis*), acrescem ainda:
  - a. falhas na especificação de uma variável causal importante;
  - b. não verificar se os dados satisfazem os pressupostos teóricos que estão na base da condução da análise.

uma correlação entre variáveis não tem grande valor em si mesma se não tiver continuidade na procura de justificações possíveis para as relações encontradas, o que só é possível, como vimos, com a manipulação de variáveis num estudo de tipo experimental.

Numa mesma lógica mas analisados desde um outro ponto de vista, os estudos correlacionais também podem ser muito úteis ao permitirem identificar possíveis variáveis estranhas que importa controlar num estudo experimental que, no futuro, possa vir a ser implementado.

Por último, podem também constituir uma alternativa válida em situações em que, por razões organizacionais ou éticas, não é possível realizar uma investigação com as exigências dos desenhos experimentais ou quase-experimentais (Bravo e Eisman, 1996)

## 10. Vantagens e limitações dos desenhos correlacionais

Relativamente a vantagens Borg e Gall (1989) e também Cohen, Manion e Morrison (2001) destacam as seguintes:

1. aproximações exploratórias a áreas de investigação em que pouco se conhece sobre o objecto de estudo;
2. não exigirem amostras grandes e permitirem trabalhar com número elevado de variáveis estudando as relações entre elas simultaneamente;
3. ao contrário da investigação experimental que só permite a manipulação de uma variável única e que, por isso mesmo, introduz alguma "irrealidade" na pesquisa, a investigação correlacional, embora menos rigorosa, permite o estudo das variáveis em situações muito mais próximas da realidade educativa;
4. por último, o facto de nos fornecer indicadores quantificados do grau de relação entre as variáveis - o coeficiente de correlação dá-nos a medida numérica da relação para todo um grupo de sujeitos - ajuda-nos muito a encontrar "sentidos" na complexidade dos fenómenos educativos e a procurar novas respostas.

Entre os inconvenientes (e os erros) mais frequentemente praticados nos estudos que recorrem a este modelo metodológico Borg e Gall (1989) e também Anderson e Arsenault, (1999) destacam:

1. a tentação do investigador em considerar os resultados da investigação correlacional como prova de uma relação causa-efeito entre variáveis:

## 11. Investigação correlacional em Educação

Embora os métodos de análise correlacional estejam extremamente divulgados na investigação em diversas áreas do conhecimento como é o caso da Psicologia ou mais ainda, na pesquisa em Ciências da Saúde, relativamente à pesquisa no domínio das Ciências da Educação são ainda de utilização relativamente restrita. Numa pesquisa realizada no motor de busca *scholar.google*, na base de dados internacional no domínio das Ciências da Educação e no *repositorium* da Universidade do Minho ([www.sdum.uminho.pt](http://www.sdum.uminho.pt)) constatámos que, nas publicações, escasseiam as referências a estudos que, de uma forma explícita, reportem os modelos metodológicos de que vimos falando.

No entanto, a análise das relações entre variáveis é a base de muitos dos modelos teóricos da análise estatística inferencial, e desde esta perspectiva, podemos dizer a análise correlacional é parte integrante da análise estatística de todos os estudos de tipo experimental! O que se passa é que, nesses estudos, os métodos correlacionais não são o centro da questão metodológica desde o ponto de vista do investigador. Dito de outro modo, as técnicas correlacionais são muito usadas na prática da pesquisa educativa mas não são o cerne do projecto de investigação a que se reportam e daí não constarem do título respectivo. Isto pode ser comprovado se, em vez do título fizermos uma pesquisa por “descritores” ou “palavras-chave” e veremos como a situação se modifica consideravelmente: no ERIC, por exemplo, se fizermos uma busca a publicações em revistas e dissertações (Mestrado e Doutoramento) pelo descritor “correlation” obtemos como resultado da pesquisa mais de 4000 ocorrências!

Em Portugal a situação não é muito diferente. Tomando como exemplo o caso da investigação realizada no domínio mais restrito da Tecnologia Educativa em que desenvolvemos a nossa actividade de investigação, verificamos uma clara supremacia dos planos de investigação de cariz quantitativo com particular destaque para as modalidades survey e quase experimental (Coutinho, 2005b); os estudos classificados como “correlacionais” eram muito raros (cf. Coutinho, 2006; Coutinho e Gomes, 2006a; Coutinho e Gomes, 2006b), escasseando também os estudos em que as técnicas de análise correlacional constituíam parte substantiva da análise estatística realizada (cf. Coutinho e Chaves, 2001; Coutinho, 2005a).

## 12. Considerações finais

Neste artigo apresentámos a ampla diversidade de estudos correlacionais com o objectivo de sensibilizar os investigadores portugueses para as suas potencialidades bem como para as limitações que decorrem da sua utilização na investigação educativa.

Trata-se de modelos metodológicos muito versáteis por permitirem a análise conjunta de um grande número de variáveis educativas, analisando a forma como estas se relacionam entre si, em separado ou em simultâneo, sem as enormes exigências que implica a planificação de verdadeiros estudos experimentais *strictus sensu* (Borg e Gall, 1989; Charles, 1998).

Nesse sentido, e tal como refere Anderson (1998), trata-se de abordagens muito úteis numa fase inicial da exploração de relações entre variáveis quando estas são em grande número. São também muito eficazes no desenvolvimento de instrumentos para a medida de variáveis educativas já que possibilitam a análise de volumes enormes de dados que importa simplificar para compreender. Uma outra vantagem é a possibilidade que dão ao investigador de trabalhar em ambientes naturais sem os inconvenientes que pressupõe a preparação de ambientes “assépticos” como acontece na investigação experimental. Por último, se bem utilizados, podem servir para testar a teoria como é o caso da análise de sendas (*path analysis*) (cf. Anderson, 1998).

Face a todas estas potencialidades podemos concluir que os inconvenientes são proporcionalmente reduzidos; de facto, as principais desvantagens reportadas na literatura têm mais a ver com questões de ética do investigador – não possibilitam aferição de relações de causalidade, não possibilitam a extrapolção de resultados para além da amostra a que se reportam – do que com os métodos em si. Na nossa opinião, a pouca divulgação destes modelos na pesquisa educativa no nosso país tem sobretudo a ver com dificuldades que se prendem com a complexidade dos modelos matemáticos que os sustentam, bem como com o desconhecimento das suas virtualidades metodológicas.

Esperamos que este nosso contributo possa servir como meio de divulgação e como incentivo a uma maior utilização destes métodos na investigação educativa no nosso país.

## Referências

- Almeida, L. e Freire, T. (1997). *Metodologia da Investigação em Psicologia e Educação*. Coimbra: APPORT.
- Anderson, G. e Arsenault, N. (1999). *Fundamentals of Educational Research*. London: Falmer Press Teachers Library.
- Anderson, G. (1999). Correlational Research. In G. Anderson & N. Arsenault (Eds.), *Fundamentals of Educational Research*. (pp.111-118). London: Falmer Press.
- Bisquerra, R. (1989). *Métodos de Investigación Educativa: Guía Práctica*. Barcelona: Ediciones CEAC.
- Black, T. (1993). *Evaluating Social Science Research: An Introduction*. London: SAGE Publications.
- Borg, W. (1963). *Educational Research: An Introduction*, London: Longman.
- Borg, W. e Gall, M. (1989) *Educational research: An Introduction*. New York: Longman, 5th Ed.
- Bravo, M. P. C. e Eisman, L. B. (1998). *Investigación Educativa*. Sevilla: Ediciones Alfar. 3ª Ed.
- Bravo, M. P. C. (1998). Los métodos correlacionales. In M. P. Bravo & L. B. Eisman (Eds.), *Investigación Educativa*. (155-175). Sevilla: Ediciones Alfar.
- Calder, J. (1996). Statistical Techniques. In R. Sapsford & V. Jupp (Eds.) *Data Collection and Analysis*. London: SAGE Publications. pp. 195-261.
- Calder, J. e Sapsford, R. (1996) *Multivariate Analysis*. In R. Sapsford & V. Jupp (Eds.) *Data Collection and Analysis*. (262-281). London: SAGE Publications.
- Campbell, D. e Stanley, J. (1963). *Experimental and Quasi Experimental Designs*. Boston: Houghton Mifflin Company.
- Charlès, C. M. (1998). *Introduction to Educational Research*. New York: Longman. 3rd Ed.
- Cohen, J.; West, S. G.; Aijien, L. & Cohen, P. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Mahwah, NJ: Lawrence Erlbaum.
- Cohen, L., Manion, L. e Morrison, K. (2001). *Research Methods in Education*. New York: Routledge Falmer. 5th Ed.
- Coutinho, C. P. (2005a). *Percursos da Investigação em Tecnologia Educativa em Portugal: uma abordagem temática e metodológica a publicações científicas (1985-2000)*. Braga: CIED-UM.
- Coutinho, C. P. (2005b). ICT in Education in Portugal: a review of 15 Years of Research. In F. Welsh & A. Tremante F. Malpica (Ed.), *Proceedings of the 3rd International Conference on Education and Information Systems: Technologies and Applications*. (pp. 40-44). Orlando. FL.
- Coutinho, C. P. (2006). Aspectos metodológicos da investigação em Tecnologia Educativa em Portugal (1985-2000), Comunicação apresentada no XIV Colóquio AFIRSE, *Para um balanço da Investigação em Educação de 1960 a 2005 – Teorias e Práticas*, realizado de 16-18 Fevereiro de 2006, FPCE-UL.
- Coutinho, C. P. e Chaves, J. H. (2001). Investigação em Tecnologia Educativa na Universidade do Minho: uma abordagem temática e metodológica às dissertações de mestrado concluídas nos cursos de Mestrado em Educação. In A. Estrela & J. Ferreira (Eds.), *Tecnologias em Educação: estudos e investigações – X Colóquio AFIRSE*. (pp. 289-302). Lisboa: FPCE.
- Coutinho, C. P. e Gomes, M. J. (2006a). Critical review of research in Educational Technology in Portugal (2000-2005). In E. Pearson & P. Bohman (Eds.), *ED-MEDIA 2006* (pp. 2679-2686). Orlando, FL: AACE- Association for the Advancement of Computing in Education.
- Coutinho, C. e Gomes, M. J. (2006b). Uma análise da investigação realizada no âmbito do Mestrado em Tecnologia Educativa da Universidade do Minho. In *Actas do XIV Colóquio Afirse/Aipelf-FPCE - Universidade de Lisboa* .
- Heiman, G. (1996). *Basic Statistics for the Behavioral Sciences*. Boston: Houghton Mifflin Company. 2nd Ed.
- Jorde-Bloom, P., Ford, M. (1988). Factors influencing Early Childhood Administrators' Decisions regarding the Adoption of Computer Technology. *Journal of Educational Computing Research*, 4, 31-47.
- Juste, P. (1985). Estudios Correlacionales. In A. De La Orden (Ed.), *Diccionario de Investigación Educativa*. Madrid: Anaya.
- Kline, P. (1986). *A handbook of test construction: introduction to psychometric design*. New York: Methuen & Co.
- Lukas, J. F. & Santiago, K. (2004). *Evaluacion Educativa*. Madrid: Alianza Editorial.
- Meltzoff, J. (1998). *Critical Thinking About Research: Psychology and Related Fields*. Washington DC: American Psychology Association.
- Moore, G. W. (1983). *Developing and Evaluating Educational Research*. New York: HarperCollins Publishers.
- Munro, B. (1981). Dropouts from Higher Education: Path Analysis of a National Sample. *American Educational Research Journal*, 18 (2), pp. 133-141.
- Punch, K. (1998). *Introduction to Social Research: quantitative & qualitative approaches*. London: SAGE Publications.
- Schutt, R. K. (1999). *Investigating the Social World: The process and Practice of Research*. (2ª Ed.). Thousand Oaks, CA: Pine Forge Press.
- Shavelson, R. (1996). *Statistical Reasoning for the Behavioral Sciences*. (3ª Ed.). Needham Heights, MA: Allyn & Bacon.

- Siegel, S. e Castellan, N. (1988). *Non Parametric Statistics for the Behavioral Sciences*. New York: MacGraw-Hill International Editions.
- Stern, P.e Kalof, (1996). *Evaluating Social Science Research*. Oxford: Oxford University Press.
- Stevens, J. (2002). *Applied Multivariate Statistics for the Social Sciences*. Mahwah, NJ: Lawrence Erlbaum.
- Vogt, W. P. (1999) *Dictionary of Statistics and Methodology: a non-ethical guide for the Social Sciences*. (2ª Ed) Newbury Park, CA: SAGE Publications.
- Wiersma, W. (1995). *Research Methods in Education: An Introduction*. (6ª Ed). Boston: Allyn and Bacon.

## CORRELATIONAL METHODS IN EDUCATION: STRENGTHS AND LIMITATIONS

Clara Coutinho  
Minho University, Braga, Portugal  
ccoutinho@ep.uminho.pt

**Abstract:** This paper describes correlational research methods in educational inquiry. In fact, correlational research makes a bridge between descriptive research methods and true experiments, as the researcher can study all kinds of relations between variables but not make inferences from samples to population. We begin by presenting main features of correlational research methods and present real examples of problems and contexts where these research techniques can be successfully used. We will also present the statistical methods to evaluate and analyse the magnitude of relations between variables. We finish presenting multivariate analysis, that although being so useful to investigate in educational settings are scarcely used in educational research in Portugal.

**KEY-WORDS:** *Research, education, methods, correlations, multivariate analysis.*