

CONSTRUÇÃO DE UMA ESCALA DE AVALIAÇÃO DE PRODUTOS EM SUPORTE SCRIPTO COM APLICAÇÕES EM CONTEXTOS DE FORMAÇÃO PROFISSIONAL¹

Vito Carioca

Jorge Matos

Escola Superior de Educação de Beja

1. Introdução

Ao reflectir um quadro de questões que se prendem com a validação de escalas tipo Likert, de forma a ser possível efectuar a indexação aos níveis (Níveis 1, 2, 3, 4 e 5) da Formação Profissional (Anexo ao Despacho Normativo nº42/B/2000, de 20 de Setembro) de produtos resultantes de projectos comunitários e concebidos em suporte *scripto* (e.g. manuais), assumimos a existência de um quadro limitativo no campo das fontes de pesquisa nesta vertente de análise.

As condicionantes referidas situam o estudo a dois níveis: i) a elaboração de itens a incluir na escala de atitudes², de forma a ser possível efectuar a indexação; ii) a identificação das características gerais do produto, de forma a permitir uma maior optimização do(s) mesmo(s) em contextos de Formação Profissional.

O primeiro aspecto implica, objectivamente, efectuar a avaliação para indexação do produto. O segundo aspecto implica a conceptualização, *a priori*, de um conjunto de categorias³ que estruturam a escala de atitudes.

Este estudo assume o quadro de pressupostos referenciados, sendo o seu objecto concreto a elaboração de uma escala que permita a avaliação de produtos. Este tipo de avaliação, embora directamente relacionada com os aspectos técnicos de construção (design e realização), pode, no entender de Costa (1999), ser extremamente enriquecedora ao permitir a reflexão sobre o potencial de um determinado produto para exploração pedagógica e, no estudo em questão, em contextos de Formação Profissional.

2. Desenho do Estudo

2.1. Introdução

A investigação que nos propúnhamos realizar, no quadro dos pressupostos referidos, deveria ter como suporte um conhecimento profundo do seu âmbito,

critério assumido com as pesquisas efectuadas relativamente a este suporte mediatizado (e.g. Cabero, 1993; Gérard, e Roegiers, 1998; <http://www.deb.min-edu.pt/1ciclo/manuaiscolares.htm>).

Por outro lado, refiramos, também, que o mesmo não teria sido possível sem as orientações metodológicas dos estudos de Raposo (1981), Estrela (1986), Moreira (1994), Trindade (1991), Carioca (1998) e Tuckman (2000), obras que estruturam toda a fundamentação das técnicas e dos instrumentos de recolha de dados adoptados.

A escala deveria obedecer a uma lógica organizacional que permitisse apreciações valorativas, resultantes do somatório dos seus itens, que identificam, no seu todo, “um perfil de cada produto” e, com alguma razoabilidade, permitam uma avaliação global, em termos do seu potencial pedagógico e didáctico-educativo, enquanto ferramenta de apoio às aprendizagens em Formação Profissional (indexação e avaliação global das características do produto).

Esse instrumento, que se apresenta na sua versão definitiva no ANEXO II, foi elaborado no quadro dum processo metodológico faseado e de acordo com o modelo adaptado de Collis e Ollila (cit. por Helena Pereira, 1990), que resumidamente se apresenta:

- reflexão sobre o objecto da pesquisa através de leituras situadas no âmbito temático do estudo. Esta reflexão e o conhecimento empírico-especulativo que possuíamos neste campo permitiu-nos elaborar um “conjunto total de itens” que, após análise minuciosa, originou o conjunto de itens do anteprojecto de escala;
- foi então feita a aplicação-piloto do pré-questionário a especialistas no sentido de validar o mesmo, os quais constituíam um grupo integrado com valências nas áreas da educação ambiental e na técnica de construção de escalas tipo Likert;
- Com o mesmo objectivo, foi feita a aplicação a um número determinado de sujeitos do universo de formadores considerado;
- As recomendações dos especialistas e as sugestões dos formadores permitiram estruturar o pré-questionário na sua versão definitiva;
- Aplicação do pré-questionário, na sua versão definitiva a um grupo de 10 formadores;
- Aplicação do coeficiente de correlação r de Pearson (Raposo, 1981; Toucinho da Silva, 1994; Selwyn, 1997; Carioca, 1998) aos dados obtidos através da aplicação do pré-questionário aos formadores, o que permitiu a elaboração da versão definitiva do questionário (ANEXO II).

Por outro lado, o conteúdo funcional da escala seria constituído por um conjunto de itens que, no seu todo, procuravam identificar a especificidade do produto, de forma a ser possível realizar a sua indexação aos níveis de formação profissional (Níveis 1, 2, 3, 4 e 5) e efectuar uma apreciação das suas características gerais. Relativamente a esta questão, considerou-se o seguinte:

- elaborar um conjunto de itens, formulados de acordo com a análise de conteúdo ao Despacho Normativo nº42/B/2000, de 20 de Setembro, e ao Quadro de Referência das Tipologias de Projecto do Programa Operacional Emprego, Formação e Desenvolvimento Social (POEFDS), que permitiriam realizar a indexação e que foram agrupados segundo duas vertentes: i) *conhecimento técnico*; ii) *competência intrínseca à execução de tarefas*. Os itens considerados seriam agrupados na constelação semântica da categoria Conteúdos;
- os itens elaborados nestas vertentes seriam os seguintes, a saber:

i. Conhecimento técnico:

- O nível de profundidade do conteúdo é manifestamente muito limitado (item negativo);
- O produto/aplicação abrange uma grande diversidade de assuntos (item positivo);
- Os conteúdos do produto/aplicação explicitam, em rigor, os fundamentos científicos da área temática em questão (item positivo).

ii. Competência intrínseca à execução de tarefas

- O roteiro de conteúdos não integra actividades que preparem para o desempenho de funções de coordenação (item negativo);
- O produto/aplicação contempla situações que permitem desenvolver competências de autonomia (item positivo);
- O produto/aplicação não contempla situações que permitem desenvolver a autonomia ao nível da profissão, mas unicamente ao nível técnico (item negativo).

- estes 6 itens, cuja variação de pontuação se situa entre 6 e 30 (=24), serão valorados e somados de forma idêntica aos outros da escala, e a integração nos Níveis de Formação será feita de acordo com a seguinte escala de valores:

I		II		III		IV		V	
6	10	11	15	16	20	21	25	26	30

- elaborar um conjunto de itens, agrupados segundo um conjunto de categorias, que permitiriam identificar as características do produto, e que adiante se explicitam;
- para efeitos de optimização dos produtos avaliados e integrados, admite-se a possibilidade de produtos classificados como “adequados para um determinado nível de formação”, poderem ser utilizados em situações de Formação Profissional em níveis inferiores, mas nunca o inverso.

2.2. As Técnicas e os Instrumentos de Pesquisa Utilizados

2.2.1. Nota inicial

A metodologia assumida por Raposo (1981), ao referenciar numa primeira fase o recurso ao método especulativo e ao método empírico como fundamentais na elaboração da escala, a conceptualização do modelo adaptado de Mc Kennell já referido e os estudos de Trindade (1991), Castaño (1992) e Carioca (1998), permitiram-nos situar a primeira fase do estudo, desde a reflexão inicial sobre o objecto da pesquisa, passando pela elaboração do anteprojecto de escala de atitudes até ao questionário definitivo.

A reflexão inicial incidiu em leituras no âmbito da temática que nos propúnhamos tratar, as quais nos ajudaram a circunscrever e a delimitar o problema. Essas fontes e o conhecimento que possuíamos na matéria permitiram a elaboração de um conjunto total de itens.

A maior parte destes itens foram testados e têm validade e fidelidade asseguradas nos estudos referidos, e os restantes foram validados no próprio processo de investigação. Estes itens foram agrupados segundo um conjunto de categorias que pretendíamos focar, facto que o quadro nº1 a seguir apresentado, detalhadamente nos apresenta.

Refiramos, igualmente, que os itens considerados no conjunto total não esgotam, logicamente, o universo de situações que certamente se poderão colocar relativamente à temática em questão. No entanto, pretendem traduzir as expectativas de uma determinada vertente de investigação relativamente à mesma, bem como a nossa própria expectativa.

2.2.2. A estrutura do pré-questionário e sua fundamentação

O estudo criterioso da técnica de Likert nos seus pontos essenciais levou-nos a realizar um estudo-piloto a aplicar a um painel de especialistas nesta área, no sentido de validar o pré-questionário. Com idêntico objectivo foi distribuído idêntica versão do instrumento a uma amostra de 30 formadores com experiência na Formação Profissional, quer no âmbito da Formação Inicial de Formadores, quer na Formação de Base.

Nesta fase da investigação, e relativamente ao grupo de formadores, assumimos a metodologia de Trindade (1991), fazendo a aplicação do instrumento em grupo, o que permitiu uma análise e comentários críticos ao seu teor e conteúdo funcional. Este duplo procedimento permitiu-nos clarificar o sentido de alguns dos itens e eliminar a sua ambiguidade.

A análise dos resultados obtidos com a aplicação destes dois instrumentos permitiu-nos elaborar a versão final do pré-questionário. Para a sua formulação na versão definitiva tivemos em conta as seguintes sugestões dos especialistas :

- Apresentar, com a razoabilidade possível, um número idêntico de itens negativos e itens positivos em cada categoria⁴ da escala;
- procurar - sempre que possível - que as categorias tivessem o mesmo número de itens, aspecto em que sentimos algumas limitações, procurando, no entanto, que existisse alguma homogeneidade numérica⁵;
- corrigir alguns dos itens da(s) escala(s), que apresentavam uma classificação incorrecta em termos de atitude favorável ou desfavorável;
- alterar o conteúdo formal e funcional de alguns itens de forma a uma melhor clarificação do seu sentido.

O quadro referenciado de dúvidas e sugestões permitiu estruturar a versão definitiva do pré-questionário que seria aplicado a um grupo de 10 formadores, no sentido de completar o processo de validação do mesmo e, simultaneamente, estruturar a versão definitiva do questionário.

Considerando o processo anterior, no que respeita à validação do pré-questionário, concretamente a sua aplicação a um Painel de Juizes (Estrela, 1986) e a um conjunto de formadores na área da Educação Ambiental, que nos permitiram

identificar algumas limitações na construção da escala e (re)estruturar a mesma, pareceu-nos que seria suficiente um número de 10⁶ para complementar o processo de validação⁷ do pré-questionário, mediante o recurso ao coeficiente de correlação *r* de Pearson.

A aplicação dos procedimentos metodológicos específicos da técnica de construção de uma escala tipo Likert permitiu a distribuição seguinte:

Quadro n.º 1
Distribuição dos itens do Pré-Questionário por Categorias
(Discurso scripto)

Categorias	Itens
CONTEÚDOS ©	1, 9, 10, 16, 17, 21, 22, 25, 26, 31, 36, 37, 38, 44, 48, 49, 51, 58, 59
COMUNICAÇÃO (COM)	6, 13, 14, 20, 27, 32, 35, 39, 43, 55, 56, 60, 61
ORGANIZAÇÃO E MÉTODO (ORGM)	2, 3, 7, 8, 11, 12, 15, 18, 23, 24, 28, 29, 30, 33, 40, 41, 42, 45, 46, 47, 50, 52, 54, 57, 63, 64, 65
CARACTERÍSTICAS MATERIAIS	4, 5, 19, 34, 53, 62

2.2.3. A análise dos resultados obtidos com a aplicação da versão final do pré-questionário

A aplicação da versão final do pré-questionário foi feita a uma amostra de 10 formadores, seleccionados aleatoriamente. Seguimos, na construção da escala de avaliação - conforme Raposo (1981), Toucinho da Silva (1994), Trindade (1991), Carioca (1998) - a técnica de Likert nos seus pontos essenciais:

- ao atribuir os valores 5-1 ou 1-5 consoante o carácter favorável ou desfavorável dos itens formulados (simpler method of scoring). Estes pesos de 1 a 5 são atribuídos em função das necessidades de medição e de se ter verificado que métodos mais complexos de medição não possuíam quaisquer vantagens;
- ao adicionar os pontos obtidos por todos os sujeitos em todos os itens da escala.

O quadro n° 2 seguinte representa o somatório de pontos obtidos pelos 10 formadores que participaram no pré-questionário.

Quadro n° 2

Formadores	Suporte Scripto
1	253
2	223
3	220
4	217
5	185
6	190
7	194
8	200
9	194
10	240

A aplicação do coeficiente de correlação r de Pearson permitiu agrupar os itens estruturantes do questionário final. Para a construção do questionário final foram utilizados os itens do pré-questionário, excluindo os que não obtiveram um coeficiente de correlação⁸ superior em valor absoluto a 0.33 (significativo para $p = 0.01$ de acordo com Ferguson, 1988)

Desta forma, a escala do questionário final ficou distribuída da seguinte forma:

Quadro n° 3
Distribuição dos itens do Questionário por Categorias
(Discurso scripto)

Categorias	Itens
CONTEÚDOS ©	2,3,7,8,13,15,16,22,24,27,28,29,34,39,41
COMUNICAÇÃO (COM)	4,9, 11, 12, 18, 30, 31, 32, 37, 38
ORGANIZAÇÃO E MÉTODO (ORGM)	5,6,10, 14,19,20,23,25,26,33,35,36,40
CARACTERÍSTICAS MATERIAIS	1,17,21

A análise do quadro mostra que houve uma distribuição relativamente equitativa em termos de rejeição de itens, embora a maior percentagem se tenha

situado nas categorias Organização e Método (ORGM) e Características Materiais (CM), principalmente na primeira, vertente em que possivelmente houve mais dúvidas de interpretação do conteúdo, o que, em nossa opinião, se poderá explicar pela sua natureza conceptual, eventualmente mais subjectiva (e.g. formas de avaliação; sentido pedagógico).

A escala final ficou estruturada da seguinte forma:

Itens positivos: 4, 6, 8, 10, 12, 13, 15, 16, 17, 20, 24, 25, 28, 30, 33, 34, 35, 37, 41;

Itens negativos: 1, 2, 3, 5, 7, 9, 11, 14, 18, 19, 21, 22, 23, 26, 27, 29, 31, 32, 36, 38, 39, 40.

Os itens n.º8, 16, 22, 27, 39 e 41 são os considerados para efectuar a indexação aos níveis de formação profissional. Os restantes estruturam as “características intrínsecas dos produtos”.

No que respeita à organização da informação, necessária para a verificação das características intrínsecas do produto de forma a uma melhor optimização do mesmo em contextos de Formação Profissional (Secção III da Base de Dados Informática), considerou-se que a referida análise qualitativa deveria expressar o somatório das categorias de cada escala. Desta forma, e relativamente a este suporte, a apreciação resultaria dos seguintes elementos, obtidos nas respostas à base de dados:

i. categoria **Conteúdos** (15 itens): variações de pontuação entre 15 e 75 pontos, com amplitude de 60. Os níveis de classificação serão:

Nível 1 (amplitude 15-34): O roteiro de conteúdos é reduzido em termos de informação disponível, revelando necessidade de actualização científica;

Nível 2 (amplitude 35-54): O roteiro de conteúdos apresenta uma quantidade razoável de informação, bem como alguma actualização científica;

Nível 3 (amplitude 55-75): O roteiro de conteúdos é pormenorizado e abrangente, apresentando rigor e actualização científica.

ii. Categoria **Comunicação** (10 itens): variações de pontuação entre 10 e 50, com amplitude de 40. Os níveis de classificação serão:

Nível 1 (amplitude 10-23): O produto revela inadequações no que respeita ao tipo de linguagem utilizada e aos elementos comunicacionais;

Nível 2 (amplitude 24 - 37): O produto revela alguma clareza em termos da estruturação da mensagem e dos elementos comunicacionais;

Nível 3 (amplitude 38-50): O produto apresenta-se bem estruturado em termos de linguagem e elementos comunicacionais.

- iii. Categoria **Organização e Método** (13 itens): variações de pontuação entre 13 e 65, com amplitude de 52. Os níveis de classificação serão:

Nível 1(amplitude 13-30): A organização da informação, nas suas diferentes vertentes (estrutura do manual, o modo de exposição, a sequência dos assuntos, o sentido pedagógico e a forma de avaliação), revela-se pouco adequada a uma correcta utilização do manual;

Nível 2 (amplitude 31- 48): A informação apresentada no manual revela alguma lógica organizativa, nas suas diferentes vertentes (estrutura do manual, o modo de exposição, a sequência dos assuntos, o sentido pedagógico e a forma de avaliação);

Nível 3 (amplitude 49- 65): O método de organização da informação, nas suas diferentes vertentes (estrutura do manual, o modo de exposição, a sequência dos assuntos, o sentido pedagógico e a forma de avaliação), revela-se adequado a uma correcta utilização do manual.

- iv. Categoria **Características Materiais** (3 itens): variações de pontuação entre 3 e 15, com amplitude de 12. Os níveis de classificação serão:

Nível 1 (amplitude 3-7): O produto denota fragilidade e baixa ergonomia;

Nível 2 (amplitude 8-11): O produto apresenta alguma resistência e ergonomia;

Nível 3 (amplitude 12-15): O produto revela resistência e ergonomia.

2.2.4. A Técnica de Likert - razões da opção assumida

As escalas tiveram como base a técnica de Likert para a valoração dos itens e apuramento dos resultados. A nossa opção por esta técnica resultou por um lado "da observação da sua relativa simplicidade" (Raposo, 1972; 1981) e, por outro - para além dos estudos de Raposo (1972; 1981) sobre avaliação de atitudes em relação ao emprego do computador - da opção feita em estudos recentes relativos a atitudes de docentes face à introdução da informática nas escolas, e entre eles os de Woodrow (1990), Kay (1993), e Carioca (1991;1998).

Por outro lado, esta técnica tem vantagens relativamente a outras técnicas de construção de escalas de atitudes, que se prendem com um mais elevado grau de fidelidade (por ex., a escala de Thurstone)⁹.

Considerámos igualmente, como metodologia específica para a construção da escala, os estudos de Likert (1971), Raposo (1981), Trindade (1991), Castaño (1992), Toucinho da Silva (1994) e Carioca (1998), que nos permitiram circunscrever concretamente o âmbito e todo o processo de elaboração da mesma.

2.2.5. A forma e a estrutura da escala

A escala foi construída, conforme referimos, com base no modelo formal descrito por Likert (*apud* Raposo, 1981; Trindade, 1991; Toucinho da Silva, 1994). Optou-se, desta forma, pela utilização da escala ordenada e valorada com os valores 1, 2, 3, 4 e 5. A estes 5 pesos fez-se corresponder um conjunto de afirmações em função do carácter positivo ou negativo do item a ser ponderado.

As cinco possibilidades de resposta relativamente aos itens da escala surgem então da seguinte forma:

<u>Itens positivos</u>		<u>Itens negativos</u>
5	Acordo total (AT)	1
4	Acordo parcial (AP)	2
3	Indecisão (I)	3
2	Desacordo parcial (DP)	4
1	Desacordo total (DT)	5

Considerando a escala, a variação absoluta situa-se entre 41 e 205 (41 itens). Por outro lado, a escala foi construída com base num conjunto de regras, que foram quer adaptadas em função das recomendações das leituras sobre esta temática. Fundamentalmente, na elaboração dos itens teve-se em atenção o seguinte:

- redacção dos itens de forma clara, breve e simples;
- a disseminação dos itens numa mesma categoria por cada metade da escala;
- a tentativa de equilíbrio entre itens positivos e itens negativos não só no conjunto geral mas em relação a cada categoria, aspecto já considerado para a elaboração da escala do pré-questionário;
- a preocupação em incluir itens simétricos, i.e., com o mesmo conteúdo mas com formulação oposta, de forma a poder elaborar testes de coerência nas respostas;
- a eliminação dos itens neutros e/ou provocadores de respostas extremas, de forma a evitar desvios estatísticos.

Refiramos, por último, a preocupação de manter um equilíbrio na distribuição pelas duas metades da escala de igual número de itens por categoria, aspecto que foi unicamente possível com algumas categorias, tendo em conta que o número de itens com um sentido positivo era diferente do número de itens com um sentido negativo. No entanto, não nos parece ser muito relevante, para o modelo genérico de construção da escala, algum desvio existente neste princípio.

3. Validação da Escala de Atitudes

3.1. Verificação da fidelidade da escala

Richards *et al.* (1986, referidos por Helena Pereira, 1990) afirmam que se torna imprescindível a existência de instrumentos fiéis e válidos quando se pretende realizar qualquer investigação científica. A determinação da fidelidade e da validade do instrumento utilizado numa investigação são critérios a ter prioritariamente em conta para garantir o rigor científico da mesma.

Várias técnicas existem para avaliar a fidelidade de um instrumento. Referindo-se a uma delas (coeficiente de correlação inter-classe), Lindzey e Aronson (1968) referem que, em virtude de se basear numa técnica de *teste-reteste*, este método é pouco utilizado, tendo em conta que a memória que os sujeitos têm em relação às respostas dadas no teste pode influenciar a correlação entre os *scores* do teste e do reteste. Os mesmos autores referem que a fidelidade pode também ser determinada a partir de um único teste. Este teste, conhecido como a fórmula 20 de Kuder-Richardson ou o coeficiente α de Cronbach, baseia-se na subdivisão do conjunto de itens em dois subconjuntos da mesma dimensão aleatoriamente escolhidos e na determinação de um índice de correlação entre esses dois subconjuntos.

No nosso caso, e tendo em conta que pretendíamos construir e validar um instrumento de medição, assumimos a lógica de Bisquerra (1987) e calculámos o coeficiente α de Cronbach¹⁰, que confere a este tipo de procedimentos um elevado rigor, tendo sido obtidos os valores seguintes, nas situações de pré-questionário e questionário definitivo:

- a) Pré-questionário ($\alpha = 0,9124$);
- b) Questionário definitivo ($\alpha = 0,9152$);

valores que se podem considerar bastante elevados, confirmando a fidelidade dos três instrumentos.

3.2. Verificação do processo de validação da escala

3.2.1. Introdução

Um questionário é válido “quando mede exactamente o que pretende medir” (Anastasi, 1992). Segundo este autor a validade depende necessariamente da fidelidade e a sua determinação só pode ser feita de forma indirecta. No nosso estudo, no que respeita à verificação da fidelidade da escala, constatámos que a mesma é bastante elevada, confirmando a consistência interna da mesma.

Oppenheim (1979) refere que a dificuldade principal em verificar a validade de um questionário de atitudes é a falta de critérios. Segundo este autor, o que necessitamos são grupos de pessoas com características atitudinais conhecidas (*criterion groups*) de forma a que possamos saber se as nossas questões os conseguem diferenciar.

Outros autores (e.g., Henerson *et al.*, 1987), ao analisarem a questão da validade dum instrumento destinado a medir atitudes, distinguem 4 tipos ou espécies de validade de uma escala: validade preditiva, validade concorrente, validade de conteúdo e validade de construção.

Em relação ao estudo não considerámos os dois primeiros tipos de validade de Henerson *et al.* (1987), e situamos as nossas razões na mesma linha de pensamento de Raposo (1981). Efectivamente - e partindo da definição conceptual de atitude preditiva de Shaw e Wright (*apud* Raposo, 1981) segundo a qual "o processo habitual para o cálculo deste tipo de validade consiste em medir a atitude e em predizer o comportamento futuro com base no somatório de pontos" – não era matéria em análise quaisquer tipo de previsões acerca de comportamentos futuros, o que anulava esta possível opção. Refiramos também - conforme Raposo (1981) - que a definição de validade preditiva pressupõe a existência de um critério exterior para a sua avaliação.

A condicionante da existência de um critério exterior para a sua avaliação, levou-nos igualmente a não considerar a validade concorrente, em virtude da não existência desse factor.

3.2.2. Validade de conteúdo

Para validar a escala, utilizámos os outros dois tipos referidos por Henerson *et al.* (1987): a validade de conteúdo e a validade de construção. Quanto à primeira, estes autores consideram-na como um indicador que se refere "à representatividade da amostra de questões incluídas no instrumento". Outros, como Shaw e Wright (*apud* Raposo, 1981), referem que a validade de conteúdo é "avaliada, determinando

o grau de adequação entre os itens da escala e o conteúdo do domínio da atitude, isto é, o grau em que o conteúdo da escala de atitudes corresponde ao sistema de atitudes" (p.203).

Anastasi (1992), na mesma linha de pensamento, afirma que tal tipo de validade envolve fundamentalmente o exame sistemático do conteúdo do teste, a fim de verificar se abrange uma amostra representativa do campo do comportamento a ser medido. Os segundos consideram ainda que a mesma deve ser medida tendo em conta um conjunto de passos e, entre eles, a definição do objecto da atitude.

Estas considerações pressupõem que o instrumento deve:

- conter itens representativos do objecto em estudo;
- conter itens que representem cada uma das categorias apresentadas.

Relativamente ao primeiro aspecto, a metodologia que utilizámos no nosso estudo garante este critério de exigência. Efectivamente, na fase de elaboração do anteprojecto da escala foi feita uma profunda revisão da literatura e uma reflexão exaustiva acerca da problemática, tendo em conta todo o conhecimento empírico-especulativo que possuíamos. Com base nestes pressupostos, foi construído o conjunto total de itens. Para além disso, com a aplicação do coeficiente de correlação r de Pearson aos itens do anteprojecto, mantivemos apenas na versão definitiva da escala os itens cujas correlações existiam ao nível de significância .01. Os itens que se aproximavam do limiar de significância, e que foram igualmente incluídos nas escalas, foram por nós considerados relevantes para a caracterização do sistema de atitudes, verificando-se uma distribuição quase equitativa relativamente à rejeição de itens nas diferentes categorias.

Quanto ao segundo aspecto e dado que se mantiveram todas as categorias, do universo inicial de itens para o universo final da versão definitiva é lícito considerar que fica assim garantida esta exigência.

Considerando as técnicas utilizadas e demonstradas, é possível afirmar a validade de conteúdo.

3.2.3. Validade de construção

Relativamente à validade de construção,¹¹ Lindzey e Aronson (1968) referem que para a avaliação deste tipo uma das técnicas geralmente utilizada é a dos grupos conhecidos (*Known groups*) na qual o instrumento é administrado a dois grupos de sujeitos, um dos quais se assume com confiança possuir o atributo num grau mais elevado do que o outro. Foi esta a técnica utilizada por Raposo (1981). Shaw e Wright (*apud* Raposo 1981) referem que "se a definição de atitude subjacente leva a

esperar que dois grupos ou mais apresentarão atitudes diferentes em relação a um determinado objecto, deduz-se que uma escala válida para medir a atitude em questão deverá mostrar somatórios de pontos diferentes para esses grupos” (p. 198).

No estudo não seria possível esta via de validação, pois não foram consideradas quaisquer hipóteses e não era objectivo comparar grupos de sujeitos.

Outros autores (e.g., Abdel-Gaid *et al.*, 1986) referem que é possível validar o construto de uma escala de forma convergente. Igualmente, e no que se refere a este tipo de validação, não era objectivo do estudo construir escalas semelhantes, de modo a comparar se os resultados médios das mesmas eram convergentes. Não nos pareceu de importância crucial considerar esta vertente, porque as variantes de validação do construto da escala previstas nos parecerem suficientes para assegurar esse critério.

No entanto, Henerson *et al.* (1987) indicam uma metodologia a utilizar que nos permite em muitos casos demonstrar a validade de construção de um instrumento e que designam por *apelo à lógica*. A demonstração, segundo este método, passa pela verificação de dois factores:

- dar tempo suficiente para a administração do instrumento de forma a que os sujeitos não sejam pressionados no sentido de responder rapidamente;
- ausência de pressões que levem a fazer surgir determinado tipo de respostas. Estas pressões podem surgir em função do formato e/ou instruções do instrumento, do ambiente em que ele é administrado e da personalidade de quem o aplica.

Efectivamente, no nosso caso, consideramos ter sido suficiente o tempo dado para o preenchimento do pré-questionário. As instruções de qualquer um dos instrumentos não condicionam de forma alguma as respostas, pois são claras e objectivas, e o seu conteúdo revela total sentido de imparcialidade relativamente ao conteúdo manifesto do instrumento. O formato parece-nos estar perfeitamente enquadrado nos parâmetros normais. O ambiente em que foram aplicados não era por nós estabelecido, o que inviabilizava qualquer tipo de pressão. E, finalmente, não foram preenchidos na presença de quem o aplicou, de forma que quaisquer características de personalidade deste não influíram nas respostas.

Outros autores (e.g. Abdel-Gaid *et al.*, 1986) referem que é possível validar o construto de uma escala através do poder discriminatório da mesma.

Neste sentido, utilizou-se o Teste Z para proceder à comparação de médias entre a metade dos indivíduos com atitudes mais favoráveis e a outra metade,

tomando como referência os totais obtidos. Os valores referentes aos cálculos efectuados apresentam-se nos quadros seguintes:

Quadro nº4
Suporte scripto-poder discriminatório

Atitude	N	Mean	Std Deviation	Std Error Mean
Menos Favorável	5	118,8000	5,4955	2,4576
Mais favorável	5	151,2000	12,9112	5,7741

O Teste Z aplicado permitiu obter os seguintes valores:

- escala em suporte scripto: obtido o valor de $Z = 2,974$, significativo com 95% de confiança, Sig (2-tailed) = 0.001, comprovando elevado valor discriminatório da escala;

Em função do conjunto de pressupostos considerados, parece-nos lícito afirmar a validade de construção do instrumento.

Notas

1. O presente artigo insere-se no quadro do projecto TransEuropean Environmental Vocational Training, financiado pelo Programa Leonardo da Vinci, com referência P/00/B/F/PP-125047.
2. A linha de pensamento segundo a qual “a atitude é uma predisposição apreendida para responder de forma consistente, favorável ou desfavoravelmente, a um objecto” (Shrigley, Koballa e Simpson, 1988, citados por Trindade, 1993), constituiu-se como vector de orientação do estudo que desenvolvemos, no que respeita ao significado do conceito que assumimos, o que se traduziu, em termos práticos, na validação de uma escala para efectuar a indexação de materiais aos níveis e contextos de formação profissional, de forma a otimizar a sua utilização.
3. A definição conceptual das categorias apresenta-se no ANEXO I.
4. Conforme Raposo (1981), Trindade (1991) e Carioca (1998) a definição de categorias pode ser feita a *anteriori* ou a *posteriori*. No nosso caso, em função das características da investigação, pareceu-nos ser conveniente agrupar inicialmente os itens dentro da sua constelação conceptual, i.e., a categoria.
5. Este pressuposto, da homogeneidade numérica entre as diferentes categorias da escala, implicou logicamente novas leituras acerca do assunto, de forma a

elaborar outros itens para as categorias em que se verificava um maior desequilíbrio.

6. No subconjunto de formadores considerado para a validação das escalas foi salvaguardada a diversidade de perfis, em função de indicadores como a experiência profissional na técnica de construção tipo Likert e as áreas de formação em questão ((Educação Ambiental e Tecnologias da Comunicação)
7. Importa igualmente considerar o facto da maior parte dos itens estarem já validados em estudos efectuados, o que confere maior consistência às escalas que utilizámos.
8. Para este efeito foi utilizado o cálculo do coeficiente de correlação de Pearson, que nos pareceu ser o mais ajustado aos objectivos que nos propúnhamos atingir, tendo em conta os estudos de Raposo (1981), Trindade (1991), Toucinho da Silva (1994) e Selwyn (1997).
9. De acordo com Seiler (L.H.) e Hough (R.), regra geral o método de medição de Likert apresenta um mais elevado índice de fidelidade que o de Thurstone. Segundo estes autores, em 27 estudos comparativos realizados, unicamente em uma situação se confirmou uma igualdade de valor e nunca um índice mais elevado do método de Thurstone (em *Empirical Comparisons of the Thurstone and Likert Techniques*, in Summers (Gene), Attitude Measurement, Mc Nally, Chicago, 1971, p. 163).
10. Critério igualmente adoptado por Trindade (1991) e Selwyn (1997), para verificar a consistência interna do pré-questionário e do questionário de avaliação de atitudes do seu estudo
11. Segundo Bohrnstedt (1970, referido por Abdel-Gaid *et al.*, 1986) o investigador valida a escala investigando se ela confirma ou nega as hipóteses propostas pela teoria que, por sua vez, se baseia num construto.

BIBLIOGRAFIA

ANASTASI, A. (1992): "What Counselors Should Know about the Use and Interpretation of Psychological Tests". *Journal of Counseling and development*. Vol. 70 (5); 610-615.

BISQUERRA, R. (1987): *Introducción a la Estadística Aplicada a la Investigación Educativa: un enfoque informático com los paquetes BMDP y SPSS*. 1ª ed. Barcelona: PPU.

CABERO, J. (1992): Los medios audiovisuales en España. IN ICE de la Universidad de Cantabria: *Las nuevas tecnologías en la educación*, Santander: ICE de Universidad Cantabria, pp.65-114.

CABERO, J. (1993): "Esfuerzo mental y percepciones sobre la televisión/vídeo y el libro. Replicando un estudio de Salomón". *Bórdon*, 45 (2), pp.143-153.

CARIOCA, V. (1991). *Avaliação de atitudes de docentes predispostos paraa utilização da informática em ambiente educativo*. Tese de Mestrado, apresentada à Faculdade de Psicologia e de Ciências da Educação da Universidade de Lisboa. Documento original.

CARIOCA, V. (1998). *Validação de uma escala de atitudes de docentes relativamente à utilização da informática educativa na sua formação contínua*. Tesis Doctoral apresentada à Universidade da Extremadura-Espanha. Documento inédito.

CASTAÑO, C. (1992): *Análisis y evaluación de las actitudes de los profesores hacia los medios de enseñanza*, Sevilla, Facultad de Filosofía y CC. de la Educación. Tesis Doctoral inédita.

ESTRELA, A. (1986): *Teoria e Prática de Observação de Classes: uma estratégia de formação de professores*. 2ª ed. Lisboa: INIC.

GÉRARD, F. e ROEGIERS, X. (1998): *Conceber e Avaliar Manuais Escolares*. Porto: Porto Editora, Lda.

HELENA PEREIRA, M. (1990). *Atitudes de alunos sobre Ensino Assistido por Computador*. Tese de Mestrado apresentada à faculdade de Psicologia e de Ciências da Educação de Lisboa. Inédita.

HENERSON, M.; MORRIS, L. e FITZ-GIBBON, C. (1987): *How to measure Attitudes*. Newbury Park: Sage Publications.

KAY, R. (1993). A critical evaluation of gender differences in computer-related behaviour. *Computers in schools*. Vol. 9 (4), pp.81-93.

LIKERT, R. (1971): "A Technique for the Measurement of Attitudes". In SUMMERS, F. (Ed.). *Attitude Measurement*. Chicago: Rand Mc.Nally Co; 149-157.

LINDZEY, G. e ARONSON, E. (Eds.) (1968):. *The Handbook of Social Psychology*. (Vol.II). 2ª ed. Massachussets: Addison-Wesley Publishing Company.

MOREIRA, D. (1994): *Planeamento e Estratégias de Investigação Social*. Lisboa: Instituto Superior de Ciências Sociais e Políticas.

OPPENHEIM, A. (1979): *Questionnaire Design and Attitude Measurement*. 7ª ed. London: Heinemann Educational Books Lda.

RAPOSO, N. (1981). *O computador e a avaliação da aprendizagem*. 1ª ed. Coimbra: Coimbra Editora Lda.

RAPOSO, N. (1992). *Essai d' évaluation de l'attitude des étudiants à l'égard de l'apprentissage audio-tutorial appliqué à l'étude de la Physique*. Dissertação de Licenciatura em Ciências de Educação (policopiada). Lovaina.

RICHARDS; P; DAVID, J. e JOHNSON, R. (1986). A scale for Assessing Students Attitudes towards Computers: Preliminary Findings. *Computers in the schools*. Vol.3 (2), pp.31-38.

SELWYN, N.(1997). Students attitudes Towards Computers: validation of a computer attitude scale. *Computers and education*. Vol.28 (1), pp.35-41.

TOUCINHO DA SILVA, A. (1994). *Atitudes de alunos finalistas dos cursos de professores da Escola Superior de Educação de Beja no ano lectivo 1993/94*. Provas Públicas inéditas para acesso à categoria de professor-adjunto.

TRINDADE, A. (1990): *Introdução à Comunicação Educacional*. Lisboa: Universidade Aberta.

TRINDADE, V. (1991): *Contributos para o Estudo da Atitude Científica dos professores de Ciências*. Tese de Doutoramento (policopiada). Évora: Universidade de Évora.

TRINDADE, V. (1993): As atitudes e a Educação Escolar. Em *Actas do 3º e 4º Seminários sobre a Componente de Psicologia na Formação de Professores e outros Agentes Educativos*. Évora: Universidade de Évora, pp.111-123.

TUCKMAN, B. (2000): *Manual de Investigação em Educação*. 1ª ed. Lisboa. Fundação Calouste Gulbenkian.

WOODROW. J. (1990). Locus of Control and Student Teacher Computer Attitudes. *Computer and Education*. Vol.14 (5),pp. 421-432.

<http://www.deb.min-edu.pt/1ciclo/manuaiscolares.htm>

ANEXO I
PRODUTOS EM SUPORTE SCRIPTO
(Categorias de análise)

- **CONTEÚDOS (C)**

Categoria que visa examinar o conjunto de mensagens, numa lógica de identificar a qualidade e quantidade de informações propostas, como a sua cientificidade, a actualidade da informação.

- **COMUNICAÇÃO (COM)**

Categoria que visa examinar não o significado da mensagem mas a sua forma (significante), analisando um conjunto de aspectos estruturantes do todo comunicacional (e.g. relação emissor-receptor; tipo de mensagens visuais; densidade da informação, etc).

- **ORGANIZAÇÃO E MÉTODO (ORGM)**

Categoria que visa analisar a metodologia utilizada, de forma explícita ou implícita, e a lógica organizativa da informação (e.g. estrutura do manual; o modo de exposição; sequência de assuntos, o sentido pedagógico e a forma de avaliação).

- **CARACTERÍSTICAS MATERIAIS (CM)**

Categoria que visa analisar os aspectos materiais e físicos do produto.

ANEXO II
QUESTIONÁRIO FINAL
PRODUTOS EM SUPORTE SCRIPTO

O presente instrumento é constituído por uma única secção, que inclui uma listagem de itens, que exprimem opiniões acerca da avaliação de produtos em suporte scripto.

Assegurando-se a confidencialidade das respostas obtidas, solicita-se a resposta a todos os itens. Obrigado pela colaboração.

Instruções:

- Leia, com atenção, todos os itens apresentados;
- Assinale a sua resposta, marcando uma única opção em cada item, de acordo com a escala que se apresenta nos mesmos.

Graus da escala tipo LIKERT

Atitude favorável (positiva)

5	4	3	2	1
-----	-----	-----	-----	-----
Estou completamente de acordo	Estou de acordo	Não sei	Discordo	Discordo bastante

Atitude desfavorável (negativa)

A pontuação da escala é inversa da apresentação

1	2	3	4	5

LISTAGEM DE ITENS

1. A capa não tem uma qualidade que permite resistir à manipulação dos utilizadores

1	2	3	4	5

Discordo bastante	Discordo	Não sei	Estou de acordo	Estou completamente de acordo

2. Os objectivos gerais não apontam para o desenvolvimento de atitudes importantes para a vida, no quadro da defesa do ambiente.
3. Os objectivos do manual não estão adequados ao público-alvo.
4. O manual apresenta situações ligadas à vida quotidiana.
5. O manual não apresenta espaços em branco para notas, questões, auxiliares de memória.
6. O manual ajuda o aluno a desenvolver bons hábitos na organização do seu trabalho.
7. Ao longo do manual as noções abstractas não estão ilustradas com exemplos
8. O produto contempla situações que permitem desenvolver competências de autonomia.
9. A linguagem utilizada não é adequada ao nível de compreensão da população a que se destina.
10. A apresentação dos conteúdos é feita de forma correcta.
11. O código linguístico utilizado na apresentação dos conteúdos não facilita a leitura e interpretação do manual.
12. Os conceitos científicos importantes estão bem definidos.
13. O manual apresenta propostas para trabalhos práticos, que não se reduzem à simples observação.
14. Não existem exercícios de aplicação suficientemente diversificados.
15. O manual propõe actividades exploratórias para a realização de aprendizagens.
16. Os conteúdos do produto explicitam, em rigor, os fundamentos científicos da área temática em questão.
17. As ilustrações, o colorido, a forma e o tamanho do manual são adequadas.
18. As ilustrações do manual são meramente decorativas e não ajudam a compreender o texto.
19. As actividades não cobrem os diferentes domínios taxonómicos (saber, saber-fazer cognitivo, saber-fazer gestual, saber-ser, saber - tornar-se).
20. A forma de apresentação dos conteúdos permite encontrar ideias para discutir questões controversas.
21. A ilustração da capa não faz sobressair a natureza do conteúdo.
22. O nível de profundidade do conteúdo é manifestamente muito limitado.
23. O manual não sugere formas de trabalho em grupo.
24. O manual apresenta informação em quantidade que permite a compreensão das partes e conceitos mais difíceis.
25. O manual tem uma organização que propicia para a preparação e concepção de projectos de intervenção em educação ambiental.
26. O corpo de caracteres não foi bem escolhido pois não coloca em evidência os títulos, subtítulos, parágrafos, etc.
27. O produto não contempla situações que permitem desenvolver a autonomia ao nível da profissão, mas unicamente ao nível técnico.
28. O conteúdo do manual faz referência a situações relacionadas com o ambiente do formando.

29. O manual não apresenta situações que propiciam para a auto-iniciativa e pesquisa individual.
30. O tamanho e o tipo de letra permitem uma leitura fácil do manual.
31. A ilustração utilizada não explicita a informação que se quer transmitir.
32. As abreviaturas utilizadas no texto não estão bem definidas e não clarificam a mensagem.
33. O índice do manual faz sobressair claramente a organização dos conteúdos.
34. A informação contida no manual é suficiente para a resolução das situações práticas propostas.
35. A capa e contracapa sintetizam a informação do manual.
36. As actividades são altamente estruturadas não permitindo a liberdade ao formando.
37. As ilustrações adicionam informações suplementares em relação ao conteúdo do manual.
38. As ilustrações do manual não estimulam o desejo de saber mais acerca do assunto.
39. O roteiro de conteúdos não integra actividades que preparem para o desempenho de funções de coordenação.
40. A forma como os conteúdos são apresentados não permite identificar facilmente o tema central.
41. O produto abrange uma grande diversidade de assuntos.