

Construção de Um Teste de Visualização a Partir da Psicologia Cognitiva

Desenho Cognitivo de Um Teste de Visualização

Gerardo Prieto Adánez¹

Universidad de Salamanca – Salamanca - Espanha

Angela Dias Vêlasco

Universidade Estadual Paulista – Guaratinguetá – SP -Brasil.

Resumo

Neste estudo descreve-se os critérios teóricos que guiaram o processo de construção de um teste de Visualização, teste este com itens gerados a partir de uma tarefa de Desenvolvimento de Superfícies. Administrou-se a amostras de estudantes do primeiro ano de Engenharia da Escola Politécnica da Universidade de São Paulo e da Universidade Estadual Paulista (Campus de Guaratinguetá), estimando-se a dificuldade dos seus itens mediante o modelo de Rasch. Observou-se que a dificuldade dos itens está associada a uma condição da tarefa implicada na habilidade para transformar mentalmente as figuras e para manter ativa a informação espacial durante o processo de transformação. Também verificou-se a utilidade deste teste no diagnóstico do desempenho dos alunos examinados na disciplina Desenho Técnico.

Palavras Chave: Visualização espacial; Desenho cognitivo; Modelo de Rasch.

Construction of a Visualization Test Based on Cognitive Psychology

Abstract

In this work we describe the theoretical criteria that guided the process of the construction of a Visualization test, whose items were generated from a Surfaces Development task. This test was applied to samples of first-year Engineering students of the Polytechnic School of the University of Sao Paulo and the Paulista State University (Guaratinguetá Campus). The item difficulty was estimated using the Rasch model. We observed that the item difficulty is associated to a task condition involved in the ability of mentally transforming the objects and of maintaining active the spatial information during the transforming process. Also we verify the usefulness of this test in diagnosing the students performance in the Technical Drawing discipline.

KEY WORDS: Spatial visualization; Cognitive design; Rasch Model.

A aplicação da psicologia cognitiva à construção dos itens é uma das correntes atuais mais importantes na construção de testes de aptidões (Embretson, 1996; Irvine e Kyllonen, no prelo). A utilização de princípios cognitivos no projeto de um

teste requer partir de hipóteses nas quais se especifique a influência das características dos itens na complexidade cognitiva da tarefa (os processos cognitivos, as estratégias e as estruturas de conhecimento requeridas para a sua resolução). Conseqüentemente, a geração dos itens parte de um projeto experimental no qual, por um lado, selecionam-se as características dos estímulos que hipoteticamente induzem os processos relevantes para o construto e, por outro, controlam-se as características associadas aos irrelevantes.

As análises empíricas permitem constatar se os dados representam a peculiaridade cognitiva que se supõe, e se os itens são úteis para estimar e interpretar as diferenças individuais no construto. Um dos procedimentos heurísticos mais usados na atualidade

¹ Facultad de Psicología, Universidad de Salamanca.
Avda. de la Merced, 109-131.
37001
Salamanca
Espanha
Tel.: 923294500, ramal 3269
Fax: 923294608
E-mail: gprieto@usal.es

Agradecimentos:

À Faculdade de Engenharia da UNESP, Campus de Guaratinguetá e à Escola Politécnica da USP.
Aos alunos que realizaram os testes.
À CAPES

é o questionamento acerca das fontes de dificuldade da tarefa, no sentido de determinar a relevância ou não das mesmas para o propósito do teste (Lohman, 2000).

Este tipo de evidência aporta indubitáveis vantagens teóricas e tecnológicas. Constituem um meio de grande interesse para analisar a validade de construto das medidas, a partir da perspectiva teórica. Da perspectiva aplicada, é indubitável que conhecer as condições da tarefa vinculadas à dificuldade tem importantes conseqüências para a construção de testes, onde se destaca a possibilidade de: (i) equiparar o conteúdo dos testes com base na sua complexidade cognitiva, (ii) construir testes adaptados informatizados a partir de critérios de conteúdo e (iii) estabelecer regras e procedimentos para a geração automática de itens (Bejar, 1993; Irvine e Kyllonen, no prelo).

Do ponto de vista metodológico, está claro que a análise não deve realizar-se sobre os resultados globais (pontuações no teste), mas sobre a execução de cada item. Portanto, no âmbito da Teoria da Resposta ao Item - TRI, e em especial na família dos modelos de Rasch, estão os procedimentos privilegiados neste enfoque (van der Linden e Hambleton, 1997).

O modelo de Rasch (1960) é conhecido como modelo logístico de um parâmetro porque a probabilidade de que um sujeito j responda corretamente a um item i ($P(X_{ij} = 1)$) depende da diferença entre seu nível de habilidade (q_j) e o parâmetro de dificuldade do item (b_i): quanto maior for a diferença $q_j - b_i$ maior será a probabilidade de que o sujeito resolva corretamente o item. A equação (1) descreve a relação entre ambos valores.

$$P(X_{ij} = 1) = \exp(q_j - b_i) / 1 + \exp(q_j - b_i) \quad (1)$$

A simplicidade do modelo de Rasch é muito atrativa para a construção de itens baseada na Psicologia Cognitiva por várias razões:

1. Os parâmetros das pessoas e dos itens expressam-se nas mesmas unidades e se localizam no mesmo contínuo (q_j , b_i)

2. Basta conhecer a diferença entre o nível do sujeito e a dificuldade do item ($q_j - b_i$) para explicar a resposta. Portanto, inferir os componentes cognitivos, as estratégias e as estruturas de conhecimento associadas à dificuldade do item equivale a explicar teoricamente a variável medida.

3. Trata-se do único modelo TRI com a propriedade de objetividade específica (Rasch, 1977;

Embretson e Reise, 2000, p.143). Esta propriedade implica que a diferença de dificuldade entre itens não depende das pessoas específicas que se utilizem para quantificá-la. Na Teoria Clássica dos Testes - TCT, pelo contrário, a dificuldade depende da amostra de sujeitos. Em conseqüência, a obtenção de conclusões generalizáveis acerca das fontes da dificuldade dos itens é conseguida mais facilmente quando se usa um modelo que permite obter estimações livres da amostra.

O propósito geral deste trabalho foi construir um teste de Visualização a partir de princípios cognitivos que fosse útil para diagnosticar o nível neste construto dos estudantes de Engenharia. Este propósito prático é de grande interesse para otimizar o processo de ensino-aprendizagem de matérias como o Desenho Técnico, nas quais se requer dos alunos um adequado desenvolvimento da Visualização.

A aptidão espacial é uma das habilidades mais estudadas no âmbito da cognição humana. Não é de estranhar o interesse que suscita já que, junto com a aptidão verbal e o raciocínio, explica a maior parte da variância das pontuações obtidas mediante baterias multifatoriais da inteligência (Lohman, 1988; Carroll, 1993). Também não se pode esquecer sua eficiência para prever o rendimento em atividades específicas de tipo técnico (engenharia, projeto mecânico, pilotagem de aviões, etc.) e artístico (arquitetura, desenho gráfico e visual, etc.).

A Visualização, a subaptidão espacial mais importante, foi definida como a habilidade para gerar uma imagem mental, levar a cabo diversas transformações sobre a mesma e reter as alterações produzidas na representação (Lohman, 1979). Para avaliar a Visualização empregaram-se diversas definições operacionais (Prieto e colaboradores, 1996). Uma das tarefas mais utilizadas denomina-se *desenvolvimento de superfícies*. Em distintas variantes desta tarefa se fundamentam testes como *DAT-SR* (Bennet, Seashore e Wesman, 1947), *SLAT* (Embretson, 1992) e *Surface Development* (Thurstone e Thurstone, 1949).

A partir de uma perspectiva cognitiva, Lohman (1988) propôs que as transformações mentais da figura são os processos mais característicos da Visualização. A complexidade da tarefa viria determinada pelo número de transformações mentais requerido para solucionar um problema. Nos últimos anos deu-se muita ênfase à importância da memória de trabalho (Baddeley, 1986), considerando que o número de transformações se traduz na complexidade das demandas a este sistema. Portanto, a elevada capacida-

de de Visualização requer tanto a habilidade para transformar mentalmente a figura, como a capacidade para manter ativa a informação espacial durante o processo de transformação (Lohman, 2000).

A partir de diversos enfoques da psicologia cognitiva, realizaram-se estudos experimentais sobre o processamento das tarefas de desenvolvimento de superfícies. Estes estudos são muito úteis para inferir hipóteses acerca das características dos itens associadas ao construto. Por exemplo, Shepard e Feng (1972) estudaram a execução em tarefas similares às do teste DAT-SR. Apreciaram uma relação linear do número de dobras da figura que haviam de ser processadas com os tempos de reação e a precisão. Prieto e colaboradores (1993 e 1996) obtiveram resultados semelhantes, ainda que matizaram que a dificuldade da tarefa não está determinada necessariamente pelo número total de faces do estímulo posto que, no caso de figuras muito complexas, os sujeitos só efetuam transformações espaciais de partes da configuração. Após analisar, desde a perspectiva cognitiva, o teste SLAT, Embretson (1993) formulou conclusões similares: o número de dobraduras mentais suscitado pelo estímulo é o maior determinante da dificuldade.

Portanto, seguimos os estudos anteriormente citados para definir os princípios cognitivos que guiaram a construção do teste de Visualização utilizado nesta pesquisa.

Os objetivos concretos deste trabalho foram:

1. Construir um teste de Visualização, denominado TVZ-2001, baseado em uma tarefa clássica de Visualização: o desenvolvimento de superfícies.
2. Desenhar os itens desde princípios cognitivos manipulando uma condição da tarefa associada a transformações mentais espaciais: a distância do objetivo.
3. Estimar a dificuldade dos itens mediante o modelo de Rasch.
4. Analisar empiricamente a relação entre a distância do objetivo e a dificuldade dos itens.
5. Determinar a utilidade do teste para o diagnóstico precoce do rendimento dos alunos de Engenharia em matérias que requerem um desenvolvimento apropriado da Visualização.

Método

Instrumentos

Teste TVZ-2001:

Com o objetivo de minimizar as estratégias de resolução não-espaciais (analítico-verbais), optou-

se por usar em todos os itens uma figura regular. Quando se empregam figuras irregulares, facilita-se o etiquetado verbal de características distintivas da figura (ângulos, tamanho, etc) que permite o emprego de processos que não são de caráter espacial. Assim, a tarefa consiste em um cubo no qual todas suas faces estão identificadas com letras. À direita do cubo, aparece seu desdobramento no plano, com uma de suas faces identificada e outra marcada com uma interrogação (?). Pede-se ao sujeito que identifique a letra da face marcada com a interrogação e sua aparência. O observador deve escolher a resposta correta entre as 10 opções fornecidas (ver a Figura 1).

Condições do domínio:

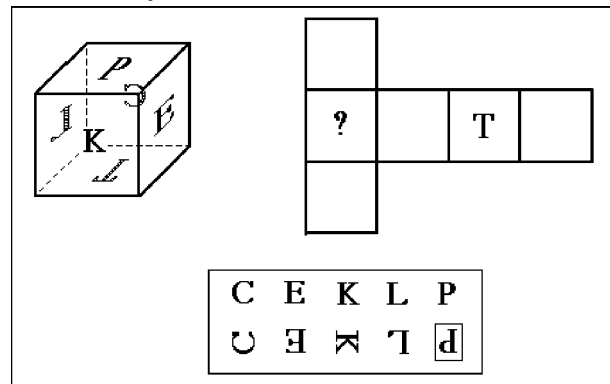


Figura 1. Características da tarefa do teste TVZ. A opção correta é a marcada pelo retângulo.

Distância do objetivo (face marcada com a interrogação) em relação à face identificada.

Para operacionalizar esta condição, designam-se as faces do cubo da seguinte forma:

C0: Face identificada com uma letra no cubo desdobrado.

C1: Face adjacente a C0.

C2: Face adjacente a C1.

C3: Face adjacente a C2.

C4: Face adjacente a C3.

C5: Face adjacente a C4.

Incluem-se 32 itens com a interrogação nas faces C2, C3, C4 e C5 (8 itens em cada condição).

Finalidade: induzir o número de transformações mentais da figura. Assume-se que o número de transformações mentais se incrementará de acordo com a seguinte ordem das condições: C5 > C4 > C3 > C2. Observe na Figura 2 um item com distância do objetivo C4.

Crítérios:

1. *Mental Cutting Test* (MCT), que é uma adaptação de um subteste do CEEB Special

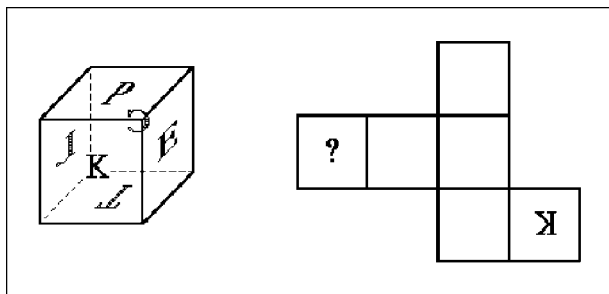


Figura 2. Exemplo de um item do teste TVZ. Condição: Distância do objetivo C4.

Aptitude Test in Spatial Relations efetuada por K. Suzuki e que tem sido empregada em vários países para avaliar a Aptidão Espacial dos estudantes de Engenharia (Suzuki e colaboradores, 1992).

Os estudos de validação do MCT revelam que é um excelente indicador do construto de Visualização (Saito e colaboradores, 1998).

2. Qualificações no primeiro exame acadêmico da disciplina de Desenho Técnico. As qualificações dos alunos em Desenho Técnico foram obtidas a partir das notas das primeiras provas. Estas provas foram individuais, com similares procedimentos de aplicação, tendo como conteúdo, na FEG-UNESP, a execução das vistas ortográficas de objetos, tanto a partir de suas perspectivas como de outras vistas e na EPUSP, Geometria Cotada, com problemas práticos sobre Superfícies Topográficas. Essa diferença deve-se a desigualdade dos programas de ensino das duas Universidades. De qualquer maneira, os dois tipos de provas aplicados versam sobre temas básicos da área gráfica para a Engenharia.

Participantes

Participaram no estudo 163 alunos do primeiro ano de Engenharia. 92 procediam da Universidade Estadual Paulista, Campus de Guaratinguetá (UNESP), e 71 da Escola Politécnica da Universidade de São Paulo (EPUSP). A idade média da amostra era de 19 anos e 6 meses e o desvio padrão de 2 anos e 2 meses, com 19% dos participantes do sexo feminino.

Procedimento

Os testes foram administrados a cada participante em uma única sessão, no mês de novembro de 2000, mediante uma aplicação informatizada construída com o programa MetaCard (Metacard, 1997).

O primeiro exame da disciplina de Desenho Técnico ocorreu no mês de abril de 2000.

Análises

A análise do ajuste dos dados ao modelo de Rasch e a estimação conjunta dos parâmetros dos itens e dos sujeitos foi feita pelo programa Quest (Adams e Khoo, 1996).

Resultados e Discussão

Medição conjunta

A maior vantagem do modelo de Rasch em relação à TCT é a propriedade de medição conjunta: os parâmetros dos itens e das pessoas estão na mesma escala. Na Figura 3 aparecem uma representação gráfica e dados descritivos do escalonamento conjunto.

Pode-se observar uma representação conjunta dos alunos e itens em uma escala de intervalo que oscila entre -4 e 4 logit. O ponto 0 da escala se situa na dificuldade média dos itens. Portanto, a média de dificuldade dos itens é 0 e seu desvio padrão, 0,95. A maior parte dos itens se agrupam em uma faixa média de dificuldade (entre $-1,51$ e $2,35$), com uma distribuição assimétrica positiva, indicativa de que alguns itens são relativamente difíceis para a amostra de sujeitos. Os alunos se distribuem em uma ampla faixa de aptidão (entre $-3,77$ e $3,12$). A média e desvio padrão são respectivamente $-0,69$ e $1,32$. Novamente estes dados indicam que o teste é relativamente difícil para os alunos avaliados. Não obstante, a Figura 3 mostra uma sobreposição adequada entre os níveis de aptidão e de dificuldade, podendo-se considerar o teste apropriado para avaliar com suficiente precisão o nível de Visualização da maioria dos alunos. Esta conclusão da análise gráfica é confirmada pelas estatísticas que se usam no modelo de Rasch para avaliar a fidedignidade das estimações dos itens e das pessoas, denominados fidedignidade da separação das pessoas e fidedignidade da separação dos itens (Wright e Masters, 1982). Ambas estatísticas são estimações da qualidade da medida para diferenciar entre pessoas e entre extratos de dificuldade. Sua interpretação apresenta certas semelhanças com o coeficiente Alpha de Cronbach (Bond e Fox, 2001): a proporção da variância observada, que é reproduzível a partir do modelo. Nestes dados, a fidedignidade da separação das pessoas e dos itens foi muito elevada: de 0,87 e 0,95, respectivamente.

Ajuste dos dados ao modelo de Rasch

As vantagens do modelo de Rasch (invariância, unidimensionalidade e aditividade)

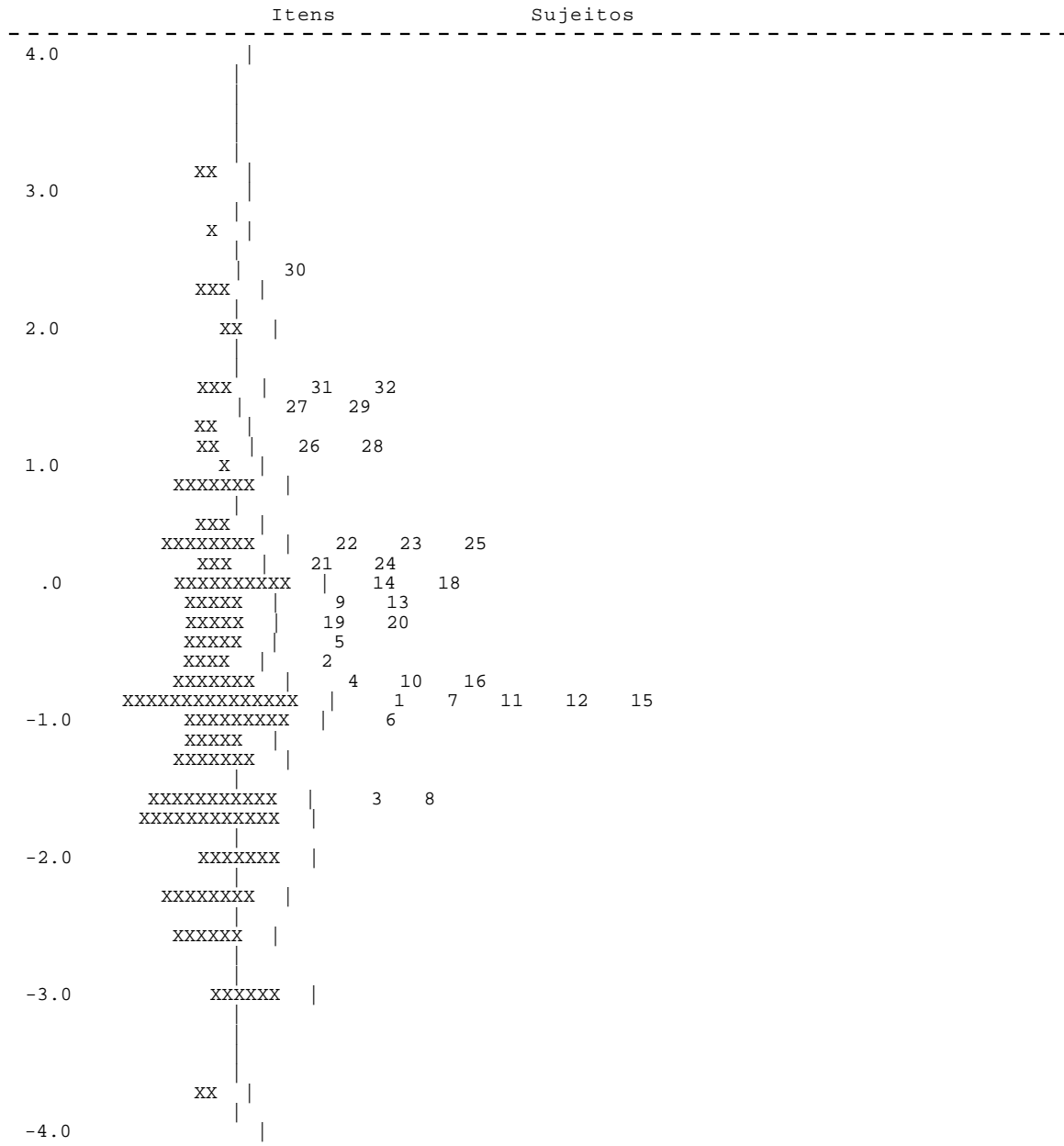


Figura 3. Medição conjunta de itens e alunos. Cada X representa 1 aluno.

só podem ser obtidas se os dados empíricos se ajustam ao modelo. De acordo com a equação (1), a probabilidade de resposta a um item depende somente dos níveis da pessoa e o item no atributo medido. A presença de respostas aberrantes tais como pessoas pouco competentes resolvendo corretamente itens difíceis, indicaria que os parâmetros de sujeitos e itens são meros numerais, carentes de significado teórico. A falta de ajuste poderia dever-se a diversos fatores: multidimensionalidade ou viés dos itens, falta de precisão no enunciado ou nas opções, respostas ao azar, falta de motivação ou cooperação, erros ao anotar a resposta, cópia da solução correta etc (Karabatsos, 2000).

Os procedimentos de análise permitem detectar os itens e as pessoas que não se ajustam ao modelo. Propuseram-se diversas estatísticas para avaliar o ajuste dos dados (Karabatsos, 2000).

A estatística de ajuste mais empregada se denomina *Infit*, que é a média dos resíduos quadráticos ponderados por sua variância. *Infit* tem uma distribuição entre 0 e infinito, com um valor esperado de 1. Por convenção, considera-se que os valores superiores a 1,3 indicam desajuste (Bond e Fox, 2001). Os programas computacionais fornecem representações gráficas que facilitam a interpretação. As respostas foram codificadas dicotomicamente: 1 = correto, 0 = incorreto ou não respondido.

Em primeiro lugar, apresentam-se os resultados da análise do ajuste ao modelo dos itens e dos participantes. As estatísticas descritivas dos valores de *Infit* aparecem na Tabela 1.

da dificuldade dos itens foi a variável dependente (b) e a variável independente, a distância do objetivo (C2= 2, C3= 3, C4= 4, C5= 5). A equação obtida foi: $Y = 0,72X - 2,52$. A

Tabela 1. Estatísticas de ajuste dos itens e dos alunos (*Infit*)

Estatística	Itens-TVZ	Alunos-TVZ	Itens-MCT	Alunos-MCT
Média	0,99	1,00	1,00	1,00
D.P.	0,10	0,21	0,90	0,19
Máximo	1,25	1,65	1,17	1,59
% > 1,3	0,00	4,97	0,00	6,13

Os dados manifestam um bom ajuste ao modelo em ambos os testes (TVZ-2001 e MCT). Por um lado, as médias e os desvios padrões dos valores são os esperáveis quando não há divergências substanciais entre as previsões do modelo e os dados empíricos. Por outro, nenhum dos itens de ambos os testes, e pequenas porcentagens dos alunos apresentam um valor superior a 1.3. Portanto, a variabilidade das respostas deve-se fundamentalmente à variação conjunta dos participantes e dos itens no atributo medido por ambas as provas.

Efeitos na dificuldade das condições da tarefa

Como já se indicou, a seleção dos itens do TVZ-2001 esteve orientada a indagar os efeitos na dificuldade das condições dos itens.

Na Tabela 2 e na Figura 4, aparecem os resultados de uma análise de variância na qual as estimações da dificuldade dos itens foi a variável dependente (b) e o fator principal a distância do objetivo (C2, C3, C4 e C5).

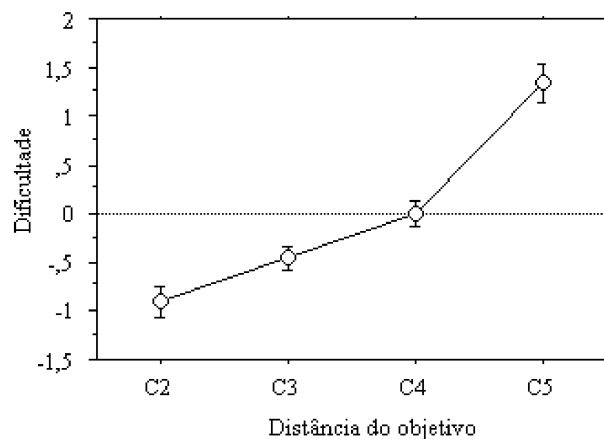


Figura 4. Efeitos da distância do objetivo na dificuldade dos itens. Barras de erro: ± 1 erro padrão

correlação de Pearson entre a distância do objetivo e a dificuldade, muito elevada (0,86) e estatisticamente significativa ($P < 0,0001$), revela que 74% da variância da dificuldade é explicada por esta condição da tarefa.

Tabela 2. Análise de variância: efeitos da distância do objetivo na dificuldade dos itens. As médias que não compartilham um mesmo subíndice diferem significativamente a 95% em um contraste de Fisher.

Distância do objetivo	C2	C3	C4	C5	F	GL	P
Média	- 0,90a	- 0,45b	0,00c	1,35d	39,61	3	0,0001
D.P.	0,43	0,35	0,38	0,55			
Nº itens	8	8	8	8			

Pode-se observar que são significativos os efeitos da distância da solução sobre a dificuldade. Os contrastes a posteriori mostraram que são significativas as diferenças de dificuldade entre todas as condições da variável independente.

Com o objetivo de que estes resultados possam ser comparados com outros estudos similares, também analisou-se os dados por meio de uma regressão simples, na qual as estimativas

Estes dados são convergentes com nossos estudos acerca de outros testes construídos com a tarefa de desenvolvimento de superfícies (Prieto e colaboradores 1993 e 1996). Portanto, pode-se concluir que é bastante robusto o efeito da característica que denominamos distância do objetivo sobre a dificuldade do item.

Estas descobertas têm repercussões na construção e na seleção dos itens para futuras

aplicações deste material de estímulo: a manipulação da característica distância do objetivo está associada a distintas faixas da aptidão medida.

Quais são as demandas cognitivas características de diferentes faixas de aptidão? Desde uma perspectiva clássica, propôs-se que a precisão em levar a cabo as transformações necessárias para reconstruir mentalmente a figura constituem o núcleo específico da Visualização (Lohman, 1985). Nos últimos anos, não obstante, pôs-se a ênfase não só na habilidade para manipular mentalmente a figura, mas na importância da memória de trabalho visual. Desde este ponto de vista, o armazenamento temporal, a manutenção ativa e o controle das transformações teriam um papel relevante nas tarefas de Visualização (Logie, 1995; Miyake e Shah, 1999). Parece razoável supor que a manipulação da distância do objetivo afetaria as demandas da memória de trabalho visual.

Em consequência, a análise e a interpretação das fontes de dificuldade dos itens destacam que a variável medida pelo teste TVZ-2001 converge com a definição teórica do construto de Visualização. Esta conclusão está balisada pelos dados correlacionais, dado que é elevado o coeficiente de validade das pontuações-Rasch do teste TVZ-2001 e do teste MCT: 0,55.

Utilidade do TVZ-2001 para prever o rendimento em Desenho Técnico

Um dos objetivos deste trabalho foi analisar a relação entre as pontuações no TVZ-2001 e o rendimento em Desenho Técnico dos alunos do primeiro ano de Engenharia.

O primeiro ano dos estudos de Engenharia apresenta grandes desafios e problemas de adaptação aos estudantes. Em alguns casos, estas dificuldades têm consequências negativas para o rendimento acadêmico dos alunos e podem ser a causa do abandono dos estudos. Os problemas de adaptação ao novo currículo são mais graves nas maté-

rias nas quais não existe uma preparação específica e extensa durante a educação secundária. Sem dúvida, este é o caso do Desenho Técnico. A nosso juízo, as atividades escolares realizadas durante a educação primária e secundária não promovem suficientemente o desenvolvimento das aptidões espaciais. Por isso, é necessário o atendimento do professor a alunos com limitações que poderiam ser sanáveis com um tratamento diferenciado. Do nosso ponto de vista, o processo de ensino-aprendizagem se veria muito favorecido se ao começo do curso se levasse a cabo um diagnóstico da aptidão espacial dos alunos, de maneira que se pudesse administrar-lhes um tratamento didático diferenciado ou um programa específico de treinamento das habilidades espaciais. A primeira proposta é uma aplicação dos clássicos desenhos ATI propostos por Cronbach e Snow (1981), segundo a qual deve-se adaptar a metodologia didática ao nível de aptidão do aluno. A segunda, refere-se ao uso de recursos para o treinamento e a melhora da aptidão espacial. Alguns destes recursos foram utilizados com êxito em amostras de estudantes de Engenharia (Hsi, Linn e Bell, 1997).

Com respeito a este objetivo, obteve-se uma correlação de Pearson entre as pontuações-Rasch do TVZ-2001 e as qualificações no primeiro exame de Desenho Técnico de 0.31 ($p < 0.0003$), o que revela uma associação moderada similar à obtida por Suzuki e colaboradores (1992) e Saito e colaboradores (1998) em amostras de estudantes japoneses de Engenharia.

Para esclarecer a utilidade de previsão do teste, classificou-se os alunos em quatro categorias em função das seguintes faixas de qualificação (\bar{X}) no primeiro exame de Desenho Técnico: insuficiente ($\bar{X} < 5$), aceitável ($5 = \bar{X} < 7$), bom ($7 = \bar{X} < 9$) e excelente ($\bar{X} = 9$). Posteriormente contrastou-se a significação das diferenças entre as médias dos quatro grupos nas pontuações-Rasch do TVZ-2001, mediante ANOVA e contrastes *a posteriori* de Fisher. Os resultados aparecem na Tabela 3 e na Figura 5.

Tabela 3. Análise de variância: diferenças em TVZ-2001 entre grupos de alunos com distinto rendimento em Desenho Técnico.

Rendimento	Insuficiente	Aceitável	Bom	Excelente	F	GL	P
Média	- 1,02a	- 0,76a	0,18b	0,18b	3,27	3	0,02
D.P.	1,48	1,14	1,22	1,43			
N	31	47	37	17			

As médias que não compartilham um mesmo subíndice diferem significativamente a 95% em um contraste de Fisher.

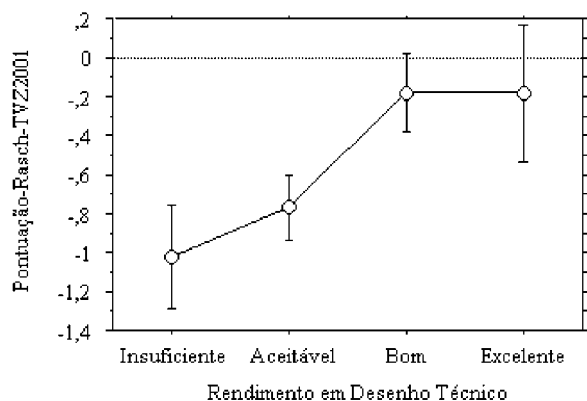


Figura 5. Diferenças no TVZ-2001 entre grupos de diferente rendimento.

Barras de erro: ± 1 erro padrão.

Pode-se observar que o valor F é significativo. Os contrastes *a posteriori* revelam que as médias no TVZ-2001 dos alunos dos grupos de menor rendimento em Desenho Técnico (insuficiente e aceitável) são significativamente menores que as médias dos grupos de melhor rendimento (bom e excelente).

Portanto, pode-se afirmar que o TVZ-2001 poderia ser um instrumento de diagnóstico útil para identificar precocemente os alunos com distinta capacidade para enfrentar o Desenho Técnico.

Em resumo, consideramos que a incorporação de testes psicométricos de Visualização à metodologia de ensino-aprendizagem certamente facilitaria a obtenção de dados importantes para o planejamento da mesma, otimizando seu desenvolvimento mediante novas e diversas estratégias didáticas que permitam superar os obstáculos que se apresentam atualmente.

Referências

- Adams, R.J. e Khoo, S. (1996). *Quest: The interactive Test Analysis System*. Victoria: ACER.
- Baddeley, A.D. (1986). *Working memory*. Oxford: Clarendon Press.
- Bejar, I.I. (1993). A Generative Approach to Psychological and Educational Measurement. Em N. Frederiksen, R.J. Mislevy e I.I. Bejar (Org.), *Test Theory for a new generation of tests*. (pp. 323-358). Hillsdale, N.J.: LEA.
- Bennet, G., Seashore, H., & Wesman, A. (1947) *Differential aptitude test*. New York: Psychological Corporation.
- Bond, T.G. e Fox, C.M. (2001). *Applying the Rasch Model*. Mahwah, NJ: LEA.
- Carroll, J.B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. Cambridge: Cambridge University Press.
- Cronbach, L.J. e Snow, R.E. (1981). *Aptitudes and instructional methods: A Handbook for research on interactions*. New York: Irvington P.
- Embretson, S.E. (1992). *Technical manual for The Spatial Learning Ability Test*. Technical Report 9201. Lawrence, KS: Department of Psychology, University of Kansas, Lawrence, Kansas.
- Embretson, S. E. (1993). Psychometric models for learning and cognitive processes. Em N. Frederiksen, R.J. Mislevy e I.I. Bejar (Org.), *Test Theory for a new generation of tests*. (pp. 125-150). Hillsdale, N.J.: LEA.
- Embretson, S.E. (1996). Cognitive Design Principles and the Successful Performer: A Study on Spatial Ability. *Journal of Educational Measurement*, 33, 29-39.
- Embretson, S.E. e Reise, S.P. (2000) *Item Response Theory for Psychologists*. Mahwah, NJ: LEA.
- Hsi, S., Linn, M.C. e Bell, J.E. (1997). The Role of Spatial Reasoning in Engineering and Design of Spatial Instruction, *Journal of Engineering Education*, April, 151-158.
- Irvine, S.H. e Kyllonen, P. (no prelo). *Item generation for test development*. Mahwah, NJ: LEA.
- Karabatsos, G. (2000). A Critique of Rasch Residual Fit Statistics. *Journal of Applied Measurement*, 1, 152-176.
- Logie, R.H. (1995). *Visuo-Spatial Working Memory*. Hove: LEA.
- Lohman, D.F. (1979). *Spatial ability: a review and reanalysis of the correlational literature* (Tech. Rep. No. 8). Stanford, CA: Stanford University Press.
- Lohman, D.F. (1988). Spatial abilities as traits, processes and knowledge. Em R.J. Sternberg (Org.), *Advances in the psychology of human intelligence*, Vol. 4. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lohman, D.F. (2000). Complex Information Processing and Intelligence. Em R.J. Sternberg (Org.) *Handbook of Intelligence*. (pp. 285-340). Cambridge, UK: Cambridge University Press.
- METACARD Corporation (1997). *MetaCard users Guide*, Boulder, CO, U.S.A.
- Miyake, A. & Shah, P. (1999). Toward unified theories of working memory. Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake and P. Shah (Org.), *Models of working memory. Mechanisms of active maintenance and executive control* (pp. 442-481). Cambridge, UK: Cambridge University Press.
- Prieto, G., Carro, J., Orgaz, B. e Pulido, R.F. (1993). Análisis cognitivo de un test informatizado de Visualización Espacial. *Psicothema*, 5, 2, 293-301.
- Prieto, G., Carro, J., Pulido, R.F., Orgaz, B., Delgado, A.R. e Loro, P. (1996). Medición de la Visualización Espacial mediante tests informatizados. *Estudios de Psicología*, 55, 41-59.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: The Danish Institute for Educational Research.
- Rasch, G. (1977). On specific objectivity: An attempt at formalizing the request for generality and validity of scientific statements. Em M. Glegvad (De.). *The Danish Yearbook of Philosophy* (pp. 59-94). Copenhagen: Munksgaard.
- Saito, T., Suzuki, K. e Jingu, T. (1998). Relations between Spatial Ability evaluated by a Mental Cutting Test and Engineering Graphics Education. *Proceedings Eighth International Conference on Engineering Computers Graphics and Descriptive Geometry*. Austin, Texas. 231-235.
- Shepard, R.N. e Feng, C. (1972). A chronometric study of mental paper folding. *Cognitive Psychology*, 3, 228-243.

- Suzuki, K., Shiina, K., Makino, K., Saito, T., Jingu, T., Tsutsumi, N., Kashima, S., Shibata, M., Maki, H., Tsutsumi, E. e Isoda, H. (1992). Evaluation os Student's Spatial Abilities by a Mental Cutting Test. *Proceedings 5th ASEE International Conference ECDGD*. Melbourne, Australia. 277-281.
- Thurstone, T.G. e Thurstone, L.L. (1949). *Mechanical Aptitude II; Description of Group Tests*. Chicago, Ill.: Univ.of Chicago, Psychometric Laboratory, Report 54.
- van der Linden, W.J. e Hambleton, R.K. (1997). *Handbook of Modern Item Response Theory*. New York: Springer.
- Wright, B.D. e Masters, G.N. (1982). *Rating scale analysis*. Chicago: MESA Press.

Recebido em 05/01/2002

Aceito em 08/04/2002

