
VANTAGENS DO USO DE TESTES ADAPTATIVOS COMPUTADORIZADOS PARA AVALIAÇÃO DA MATURIDADE DO SISTEMA DE GESTÃO AMBIENTAL DE INDÚSTRIAS

DÉBORA SPENASSATO, BLÊNIO C. SEVERO PEIXE, ANDRÉA C. TRIERWEILLER, ANTONIO C. BORNIA e RAFAEL TEZZA

RESUMO

O Teste Adaptativo Computadorizado (TAC) apresenta vantagens em relação aos métodos não-adaptativos aplicados via computador ou papel e lápis, por selecionar itens mais informativos para os respondentes. Este trabalho tem por objetivo apresentar as vantagens do uso de testes adaptativos computadorizados para avaliação da maturidade do Sistema de Gestão Ambiental (SGA) de indústrias. O teste aborda questões relativas à política ambiental, planejamento, implementação e operação, verificação e ação corretiva e, por fim, a análise

crítica pela administração para a melhoria contínua. Para tanto, um estudo de simulação foi realizado para avaliar o efeito de diferentes regras de finalização do TAC e os resultados foram comparados com a sua versão original. Os resultados demonstraram que, implantando-se um TAC, poderia haver uma redução de 71% no comprimento do teste sem comprometer a validade e precisão da medida, beneficiando tanto os respondentes quanto os responsáveis pelo desenvolvimento e aplicação dos testes.

Testes Adaptativos Computadorizados (TAC), em inglês *Computerized Adaptive Testing* (CAT), têm sido amplamente utilizados como uma alternativa aos testes de papel e lápis (*paper and pencil*; P&P), comumente utilizados em avaliações (Lee e Dodd, 2012). A precisão e eficiência de um teste são preocupações primordiais e, muitas vezes, concorrentes na avaliação

aplicada, incluindo contextos organizacionais, havendo uma tendência para usar testes (questionários) curtos como alternativa para redução do tempo do teste e dos custos (Makransky e Glas, 2013).

Normalmente, os pesquisadores selecionam um subconjunto de itens (perguntas ou questões) disponíveis, de uma escala já desenvolvida e validada, ao invés de apresentar o teste completo

(Montgomery e Cutler, 2013). Esta abordagem pode reduzir drasticamente a precisão da medida e introduzir viés, especialmente para os indivíduos localizados nos extremos de uma escala (Makransky e Glas, 2013; Montgomery e Cutler, 2013). Neste sentido, é importante considerar um método que seja capaz de diminuir o número de itens necessário e, consequentemente, o tempo para a realização

PALAVRAS CHAVE / ISO 14001 / Sistema de Gestão Ambiental / Teoria da Resposta ao Item / Teste Adaptativo Computadorizado /

Recebido: 06/07/2015. Aceito: 29/07/2015.

Débora Spenassato. Graduada em Matemática, Universidade de Passo Fundo, Brasil. Mestre em Modelagem Computacional, Universidade Federal do Rio Grande, Brasil. Doutoranda em Engenharia de Produção, Universidade Federal de Santa Catarina (UFSC), Brasil. Endereço: Programa de Pós-Graduação em Engenharia de Produção, Centro Tecnológico, UFSC, Brasil. e-mail: debospenassato@gmail.com

Blênio C. Severo Peixe. Graduado em Ciências Contábeis, Universidade Federal do Paraná (UFPR), Brasil. Mestre em Ciências Contábeis, Universidade do Estado do Rio de Janeiro, Brasil. Doutor em Engenharia de Produção, UFSC, Brasil. Doutor em Ciências Empresariais, Universidad del Museo Social Argentino. Professor, UFPR, Brasil. e-mail: blenio@ufpr.br

Andréa C. Trierweiller. Bacharel em Administração, Especialista, Mestre e Doutora em Engenharia de Produção, UFSC, Brasil. Pesquisadora pós-doutoranda em Engenharia de Produção, UFSC, Brasil. e-mail: andreatri@gmail.com

Antonio C. Bornia. Graduado em Engenharia Mecânica, UFPR, Brasil. Mestre e Doutor em Engenharia de Produção, UFSC, Brasil. Professor, UFSC, Brasil. e-mail: cezar.bornia@ufsc.br

Rafael Tezza. Graduado em Engenharia de Produção Elétrica, Mestre e Doutor em Engenharia de Produção, UFSC, Brasil. Professor, Universidade do Estado de Santa Catarina, Brasil. e-mail: rafael.tezza@udesc.br

do teste, porém, sem comprometer a precisão da medida e a validade dos resultados, aplicando adequadamente itens de todos os subdomínios (áreas ou conteúdos) do traço latente investigado.

Os TAC surgem como uma solução para este problema, uma vez que são capazes de minimizar o número de itens que cada indivíduo deve responder, preservando a acurácia e precisão da medida (Montgomery e Cutler, 2013). Neste método de aplicação de testes, os itens são selecionados de um conjunto de itens após cada resposta dada, assim, aqueles mais informativos sobre o nível de proficiência do respondente são aplicados e o respondente não precisa responder a itens muito difíceis ou muito fáceis, os quais não contribuem para melhorar a precisão da estimativa da proficiência (Economides e Roupas, 2007; Hol *et al.*, 2008; Lee e Dodd, 2012; Makransky e Glas, 2013).

O TAC é utilizado para avaliar o nível de proficiência, habilidade, traço latente, conhecimento ou desempenho, representado pelo escore do teste ou pela letra grega theta (θ). Aplicações do TAC são desenvolvidas nas mais variadas áreas, por exemplo, na educação (Huang *et al.*, 2009), psicologia (Hol *et al.*, 2008), na saúde para tomada de decisão sobre tratamento, reabilitação e diagnóstico (Bent *et al.*, 2013), no contexto organizacional para pesquisa de opinião pública (Montgomery e Cutler, 2013) e marketing (Singh *et al.*, 1990). O presente estudo aborda o uso do TAC para o contexto de avaliação do Sistema de Gestão Ambiental (SGA) de empresas.

A fim de apresentar as vantagens do uso de TAC para avaliação da maturidade do SGA de indústrias, dados do trabalho de Severo Peixe (2014) foram utilizados. O teste para mensurar o nível de maturidade do SGA é composto por 55 itens, tendo sido considerado um teste relativamente longo, havendo baixa adesão das empresas em responder ao questionário, uma vez que, do total de 2994 empresas selecionadas para participarem da pesquisa, 354 responderam o teste. O teste foi originalmente aplicado pelo método não-adaptativo, isto é, todos os respondentes receberam os mesmos itens e, também, a mesma quantidade de itens.

Pesquisas mostram que testes longos, além de elevar custos financeiros, podem ocasionar o aumento de taxas de não-resposta, maior probabilidade de abandono (desistência) do teste, aumento da seleção de alternativas de respostas intermediárias, dentre outras implicações (Yammarino *et al.*, 1991; Burchell e Marsh, 1992; Crawford *et al.*, 2001; Krosnick *et al.*, 2002; Bowling, 2005;

Galesic e Bosnjak, 2009). No caso deste artigo, em que os respondentes foram gestores ou responsáveis pela área ambiental das indústrias, um dos motivos daqueles que deixaram de responder o teste, pode estar relacionado ao longo tempo a ser dispendido para realizar esta tarefa. Assim, a implantação do TAC seria uma alternativa para facilitar esse processo, trazendo vantagens, as quais serão aqui discutidas.

A seguir, apresenta-se o referencial teórico que aborda as principais características do SGA e do TAC, o método utilizado neste estudo, que inclui o desenvolvimento do banco de itens para o TAC, os procedimentos para a simulação com dados reais e métodos para avaliação dos resultados. Posteriormente, os resultados e discussão e, finalmente, as conclusões.

Referencial Teórico

Maturidade do SGA

O tema ‘maturidade’ está originalmente relacionado à gestão de projetos. Porém, em seu cerne, os modelos de maturidade preveem uma estrutura para a avaliação sistemática, que permita à empresa comparar seus projetos com as melhores práticas ou contra as práticas de seus concorrentes. Desta forma, pode ser contextualizado no âmbito da gestão ambiental.

Os modelos de maturidade definem uma rota estruturada para o melhoramento contínuo das ações, demonstrando o quanto uma empresa progrediu, sua eficiência e eficácia para concluir projetos (Andersen e Jessen, 2003; Cooke-Davies e Arzymanow, 2003; Grant e Pennypacker, 2006). O nível de maturidade não é proporcional ao tempo de existência da empresa (Dinsmore, 1999). Porém, não se encontra um processo totalmente maduro na organização, pois, para isso, é necessário atingir o desenvolvimento total de sua missão, em todas as suas operações (Webster, 1988).

Neste artigo, a avaliação da maturidade do SGA foi embasada em aspectos e princípios da norma ISO 14001 (ISO, 2004) e ajustada ao ciclo PDCA, que define a estrutura da norma (Halila e Tell, 2013): 1) Política Ambiental (PA), 2) Planejamento (*Plan*), 3) Implementação e Operação (*Do*), 4) Verificação e Ação Corretiva (*Chek*), e 5) Análise Crítica pela Administração (*Act*). Assim, o ciclo PDCA define a espiral da melhoria contínua do SGA e é utilizada, principalmente, por profissionais da área de qualidade para entendimento do padrão ISO 9001 e ISO 14001 (Balzarova e Castka, 2008). A seguir,

apresenta-se resumidamente, cada um desses componentes, segundo a norma ISO 14001 (ISO, 2004).

A política ambiental é definida pela ISO 14001:2004 como a declaração da organização, expondo suas intenções e princípios em relação ao seu desempenho ambiental global, que provê uma estrutura para a ação e definição de seus objetivos e metas ambientais. Além disso, também estão dispostos os fatores que devem ser considerados no momento do Planejamento do SGA, que envolvem aspectos ambientais, requisitos legais, objetivos, metas e programas. Esses fatores devem ser considerados para que o SGA realmente forneça benefícios para o meio ambiente. Destacam To e Lee (2014), a ISO 14001 fornece a estrutura de uma abordagem holística e estratégica para uma organização definir sua PA, planos e ações.

A implementação e operação se relaciona à disponibilidade de recursos, definição de funções, responsabilidades. É preciso desenvolver competência, treinamento e conscientização do pessoal, implementar, manter e melhorar o SGA. Para Zorpas (2010) a avaliação dos impactos ambientais das atividades, processos e serviços é um dos passos mais importantes da implementação do SGA.

A verificação e ação corretiva trata do monitoramento e medição das características principais das operações que possam ter um impacto ambiental significativo e compreende a avaliação periódica do atendimento a requisitos legais e outros requisitos; implementar e manter procedimentos para tratar as não conformidades para executar ações corretivas e preventivas e ainda, controlar, estabelecer e manter registros. Deve-se observar a condução de auditoria interna, em intervalos planejados para determinar se o SGA está em conformidade com os arranjos planejados para a gestão ambiental e fornecer informações sobre os resultados da auditoria.

A análise crítica do SGA pela administração deve ser realizada em intervalos planejados para garantir sua continuidade, pertinência e eficácia. As análises devem incluir a avaliação de oportunidades de melhoria e a necessidade de alterações no SGA, da PA e dos objetivos e metas ambientais. De acordo com Oliveira e Serra (2010: 436), “...a norma NBR ISO 14001 tem sido o instrumento mais utilizado para desenvolver a gestão ambiental em empresas industriais. No Brasil, sua adoção vem aumentando continuamente nos últimos anos, indicando amadurecimento das questões ambientais empresariais na direção de uma gestão sustentável.”

Tendo sido apresentada a estrutura da norma ISO 14001, considerada para a formulação dos itens, será abordado o TAC.

O Teste Adaptativo Computadorizado

O TAC assume que existe um traço latente a ser medido e que os itens que compõem o banco de itens são indicadores apropriados dessa característica e estão em uma escala validada psicometricamente (Montgomery e Cutler, 2013). Comumente, utiliza-se a escala normal com média igual a 0 e desvio padrão igual a 1.

O algoritmo TAC seleciona os itens do banco de itens (BI) que fornecem mais informações sobre a proficiência atual do respondente, de modo a maximizar a precisão do teste com base nos itens respondidos anteriormente (Hol *et al.*, 2008; Makransky e Glas, 2013). Desta forma, se um indivíduo responde a um item corretamente, um item mais difícil será administrado como um próximo item, caso a resposta seja incorreta, um item mais fácil será administrado. A aplicação de itens continua até que um critério de parada predefinido seja satisfeito (Hol *et al.*, 2008; Luecht e Sireci, 2011; Lee e Dodd, 2012), conforme Figura 1.

A linha tracejada na Figura 1 representa a verdadeira proficiência do respondente ($\theta = 0$) e os círculos correspondem às estimativas da proficiência a cada resposta dada. Estabelecendo uma regra de parada do teste de 55 itens, a proficiência estimada ($\hat{\theta}$) foi 0 e erro padrão (EP) foi 0,015. No entanto, pode-se observar que, com oito itens a estimativa é muito próxima da verdadeira ($\hat{\theta} = -0,001$ e EP = 0,133). Desta forma, o teste já poderia ser finalizado e o respondente teria uma redução de 85% no comprimento do teste, se comparado ao método não-adaptativo.

Uma das principais vantagens do TAC sobre os testes não-adaptativos é a eficiência, a qual está associada à precisão e redução no tempo de teste (Economides e Roupas, 2007; Luecht e Sireci, 2011). Esse método permite a monitoração longitudinal dos resultados, dado que itens diferentes podem ser administrados em momentos

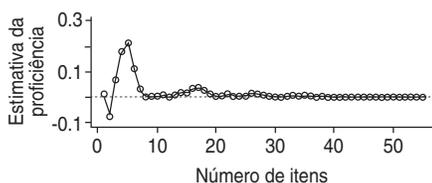


Figura 1. Estimativa da proficiência de um respondente submetido ao TAC com 55 itens.

diferentes, com base no nível de proficiência do respondente. E mesmo administrando itens diferentes, pela Teoria da Resposta ao Item (TRI), é possível a comparação dos escores passados com os escores atuais (Bent *et al.*, 2013). Além disso, o resultado do teste é divulgado após o término do teste, junto com um *feedback* individual detalhando os conteúdos que o respondente domina, o que é mais informativo do que simplesmente oferecer um escore final, pois permite conhecer melhor as características de cada respondente (Wainer, 2000).

Os testes adaptativos computadorizados baseados na TRI para seleção de itens e estimação da proficiência dos respondentes, apresentam algumas vantagens em relação aos outros métodos existentes na literatura. A TRI possui vários modelos matemáticos que possibilitam colocar itens e respondentes em uma escala comum (Luecht e Sireci, 2011; Thompson e Weiss, 2011). Desta forma, permite que diferentes conjuntos de itens sejam aplicados aos respondentes, mas que essas estimativas sejam comparáveis entre si (Hol *et al.*, 2008; Luecht e Sireci, 2011).

O sistema TAC seleciona os itens de um conjunto de itens calibrados pela TRI, ou seja, com os parâmetros dos itens já estimados (Economides e Roupas, 2007; Montgomery e Cutler, 2013), possibilitando determinar qual item dentre os disponíveis é mais informativo para um nível particular de proficiência (Hol *et al.*, 2008). Por ser um teste adaptativo, o TAC impõe algumas restrições para o respondente, não sendo possível: i) visualizar todos os itens e, em seguida, respondê-los; ii) ignorar um item sem respondê-lo; e iii) voltar atrás e rever ou alterar a sua resposta a um item anterior (Economides e Roupas, 2007). Essas ações alterariam toda a cadeia de itens administrados durante o teste (Wainer, 2000; Wang e Kolen, 2001).

O algoritmo básico de um TAC é composto por: i) método para iniciar o teste; ii) método de seleção dos itens; iii) método para estimar o escore; e iv) método para finalizar o teste (critério de parada). Várias restrições adicionais podem ser especificadas no algoritmo, sendo as mais comuns o balanceamento de conteúdo e o controle da taxa de exposição dos itens. Para cada um desses componentes, há uma infinidade de métodos disponíveis na literatura, além de vários artigos que buscam compará-los.

O controle da taxa de exposição dos itens geralmente é implantado em testes de alto impacto, como de acesso ao ensino superior ou pós-graduação, para aquisição de bolsas de estudo,

contratação, licenciamento, certificação profissional, dentre outros (Economides e Roupas, 2007). Esses testes exigem rigoroso controle da segurança dos itens para não comprometer a validade dos resultados. O balanceamento de conteúdo é utilizado quando itens de vários conteúdos devem ser administrados de forma equilibrada para avaliar o traço latente (Luecht *et al.*, 1998; Luecht e Sireci, 2011). No entanto, quando o balanceamento de conteúdo e controle da exposição dos itens são incorporados no algoritmo, os ganhos de eficiência associados a um TAC podem ser reduzidos (Luecht e Sireci, 2011).

Duas regras de parada comumente utilizadas são: 1) comprimento de teste fixo ou 2) comprimento variável, ou seja, com base em um nível mínimo de precisão. Quando é estabelecido um comprimento fixo, todos os respondentes recebem o mesmo número de itens, independentemente do erro de medida associado com o seu escore (Luecht e Sireci, 2011). Quando é baseado na precisão, o número de itens por respondentes pode variar, pois depende do nível da escala em que ele se encontra.

Para decidir sobre quais métodos são mais adequados em cada situação particular de testes, simulações *post-hoc* (com dados reais) são utilizadas. Essas simulações servem como um teste piloto e são necessárias para auxiliar os profissionais a avaliar a viabilidade de implantação de um TAC, o *design* do teste, o desempenho do TAC, entre outros fatores que podem influenciar no resultado (Wang e Kolen, 2001; Bjorner *et al.*, 2007; Thompson e Weiss, 2011).

Método

Os dados utilizados neste estudo foram obtidos do trabalho de Severo Peixe (2014). Nesta seção, é descrito como foi desenvolvido o banco de itens utilizado para o TAC, os procedimentos de simulação e as análises que serão utilizadas para comparação do teste original com a aplicação do TAC.

Desenvolvimento do banco de itens para TAC

Composição do teste: O teste é composto por 55 itens, assim distribuídos: política ambiental (10 itens); planejamento (9); implementação e operação (11); verificação e ação corretiva (11) e análise crítica pela administração para melhoria contínua (14 itens).

Os itens foram respondidos com base nas afirmações indicadas na escala ordinal, atribuindo valores de 1 a

5, conforme as seguintes categorias de resposta. 1: Não, a empresa não realiza prática ou ação, neste sentido. 2: Não, a empresa pretende realizar prática ou ação, neste sentido. 3: Não, a empresa está iniciando a implantação desta prática ou ação, e pretende formalizar. 4: Sim, a empresa realiza esta prática ou ação, e está sendo formalizada. 5: Sim, a empresa realiza esta prática ou ação, e está formalizada.

Aplicação do teste original. A aplicação do instrumento para a coleta de dados foi realizada por e-mail enviado a 2994 empresas, no período de outubro/2012 a maio/2013. Os respondentes poderiam acessar o link da pesquisa para responder os itens ou enviar suas respostas no documento em word. O foco era avaliar o SGA de empresas industriais de médio e grande porte (Severo Peixe, 2014).

Perfil dos respondentes e empresas. Os responsáveis por responder o teste tinham o seguinte perfil: técnico em gestão ambiental (38%), engenheiro (21%), administrativo (14%), qualidade (9%), recursos humanos (5%), outros (13%). Os segmentos das empresas respondentes foram: complexo agroindustrial (13%), complexo eletrometal mecânico (22%), complexo têxtil (19%), complexo florestal (16%), complexo tecnológico (7%), complexo químico (11%), outros (12%). No total, 354 empresas industriais da Região Sul do Brasil, responderam ao teste (Severo Peixe, 2014).

Parâmetros dos itens e nível de maturidade do SGA das empresas: o Modelo de Resposta Gradual (GRM; Samejima, 1969) da TRI foi utilizado para estimar os parâmetros dos itens e os escores das empresas. Este modelo é utilizado para dados politômicos com categorias de respostas ordenadas. O modelo pressupõe que um respondente com nível baixo terá uma probabilidade muito elevada de responder a categoria 1, ou seja, a categoria mais baixa. Conforme o nível de maturidade do SGA da empresa aumenta, a probabilidade de responder uma categoria mais elevada também aumenta, conforme mostra a Figura 2. Quanto maior a discriminação do item, mais inclinadas serão as curvas das categorias.

Na Figura 2 é possível observar que uma empresa com proficiência média (igual a 0), tem maior probabilidade de responder a categoria 4 (~0,48). Por sua vez, uma empresa com proficiência igual a 1, tem maior probabilidade de responder a categoria 5 (~0,79).

O software Multilog (Thissen *et al.*, 2003) foi utilizado para estimação dos parâmetros dos itens e do

nível de maturidade do SGA das empresas. A Tabela I apresenta os valores estimados dos parâmetros dos itens. Esses itens apresentam bons parâmetros de discriminação, variando entre 0,86 a 3,98. Itens com um bom poder de discriminação possuem valores >0,7.

Função de informação do teste (FIT): A FIT é dada pela soma das funções de informação dos itens que compõem o teste. Assim, é possível determinar quanto cada item contribui para a precisão do escore e qual item é mais informativo para determinado nível de proficiência (Luecht e Sireci, 2011). A análise da informação do teste demonstra a região da escala, em que o instrumento é mais confiável para medir a 'maturidade do SGA das indústrias'.

Na Figura 3, a linha contínua representa a curva de informação do teste e a linha pontilhada representa o

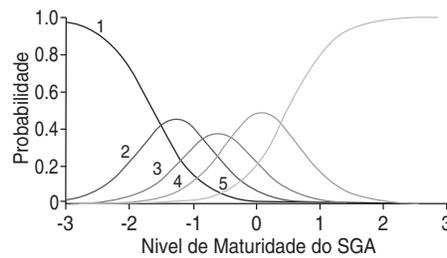


Figura 2. Função de resposta ao item correspondente ao item 13 do teste.

EP da medida; quanto maior a informação, menor é o erro. Observa-se que a escala desenvolvida demonstra ser mais precisa para o intervalo de -2,3 a 1,7 do nível de maturidade do SGA. Assim, nota-se que, conforme o nível de maturidade da empresa se aproxima dos extremos da escala, a precisão do escore tende a diminuir. O nível 0 significa a média, e valores positivos e negativos significam desvios acima ou abaixo da média para a escala (0,1).

A FIT é muito importante para um teste adaptativo, pois a partir dela é possível verificar onde há mais informação (mais itens disponíveis no BI para determinado nível) e onde é preciso desenvolver novos itens afim de melhorar a informação naquele ponto (geralmente nos extremos da escala). Quanto maior a gama de cobertura da escala, mais precisa será a estimação da proficiência nos diversos níveis.

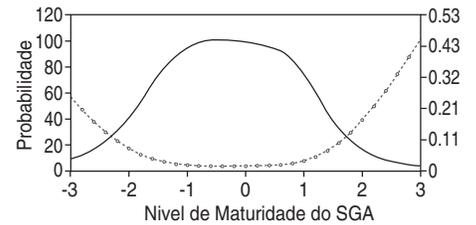


Figura 3. Função de informação do teste e erro de medida para mensurar a maturidade do SGA das empresas.

TABELA I
PARÂMETROS DOS ITENS QUE COMPÕEM O BANCO DE ITENS PARA O TAC

Item	a	b1	b2	b3	b4	Grupo	Item	a	b1	b2	b3	b4	Grupo
1	2,64	-1,48	-0,77	-0,17	0,63	PA	29	0,86	-2,64	-1,03	0,14	2,02	Do
2	2,54	-1,74	-0,90	-0,42	0,57	PA	30	1,33	-2,91	-2,05	-1,44	-0,19	Do
3	1,27	-0,97	-0,30	0,38	1,68	PA	31	2,58	-1,57	-0,72	-0,14	0,60	Do
4	3,04	-1,29	-0,71	-0,19	0,33	PA	32	2,37	-1,57	-0,79	-0,26	0,42	Chek
5	2,84	-1,46	-0,79	-0,23	0,48	PA	33	1,61	-1,25	-0,57	0,09	0,94	Chek
6	3,39	-1,19	-0,56	-0,06	0,64	PA	34	1,12	-0,98	-0,20	0,42	1,16	Chek
7	1,90	-2,00	-1,34	-0,82	-0,08	PA	35	2,26	-1,34	-0,58	0,01	0,56	Chek
8	1,11	-1,01	-0,24	0,57	1,40	PA	36	2,60	-1,36	-0,66	-0,05	0,50	Chek
9	1,23	-0,97	-0,29	0,28	1,53	PA	37	2,80	-1,16	-0,44	0,27	0,94	Chek
10	1,84	-1,29	-0,50	0,04	0,89	PA	38	2,48	-0,76	-0,04	0,58	0,92	Chek
11	2,33	-1,40	-0,78	-0,38	0,44	Plan	39	1,55	-1,11	-0,12	0,50	1,45	Chek
12	1,66	-2,29	-1,31	-0,51	0,63	Plan	40	2,26	-1,04	-0,25	0,38	0,97	Chek
13	2,68	-1,63	-0,90	-0,31	0,49	Plan	41	3,09	-0,88	-0,15	0,39	0,88	Chek
14	1,95	-1,66	-0,97	-0,26	0,63	Plan	42	3,39	-0,86	-0,14	0,44	1,07	Act
15	1,88	-2,23	-1,41	-0,89	0,29	Plan	43	2,74	-1,06	-0,26	0,28	1,13	Act
16	2,94	-1,48	-0,69	-0,18	0,81	Plan	44	2,47	-1,05	-0,34	0,25	1,22	Act
17	2,75	-1,75	-0,93	-0,39	0,42	Plan	45	1,51	-1,83	-0,82	-0,01	1,24	Act
18	2,48	-1,34	-0,47	0,14	1,05	Plan	46	1,35	-2,16	-1,20	-0,52	1,08	Act
19	3,55	-1,23	-0,59	0,07	0,58	Plan	47	2,31	-1,20	-0,57	0,10	1,08	Act
20	3,53	-1,34	-0,57	0,05	0,81	Do	48	2,98	-1,00	-0,53	0,00	0,64	Act
21	2,60	-1,68	-0,83	-0,28	0,70	Do	49	3,98	-0,97	-0,41	0,13	0,73	Act
22	3,14	-1,43	-0,62	-0,18	0,64	Do	50	3,90	-0,95	-0,43	0,08	0,81	Act
23	2,34	-1,32	-0,36	0,35	1,19	Do	51	2,08	-1,68	-0,66	-0,01	0,98	Act
24	2,39	-1,38	-0,53	0,13	0,97	Do	52	2,94	-0,96	-0,35	0,41	1,03	Act
25	2,09	-2,19	-1,04	-0,54	0,43	Do	53	2,97	-1,34	-0,45	0,08	0,89	Act
26	2,15	-1,72	-0,81	-0,19	0,60	Do	54	3,01	-0,88	-0,24	0,51	1,09	Act
27	1,22	-2,41	-1,38	-0,35	1,33	Do	55	1,77	-1,30	-0,57	0,05	0,98	Act
28	1,02	-2,53	-1,42	-0,45	1,65	Do							

Amostra utilizada no estudo de simulação: Dentre as 354 empresas que responderam o teste original, somente as 330 que responderam a todos os 55 itens no teste original foram utilizadas para a simulação do TAC. A média do EP das estimativas do nível de maturidade do SGA obtidas com o teste original, composto por 55 itens, foi de 0,135, variando entre 0,087 e 0,458.

A Figura 4 apresenta o nível de maturidade das 330 empresas utilizadas para simulação do TAC. Para que fosse apresentado somente valores de escores positivos, a escala (50,10) foi utilizada, ou seja, a média zero na escala (0,1) é equivalente a 50. Observa-se que 48,7% das empresas estão abaixo da média 50 da escala e a maioria se encontra no nível 40. Nível elevado de maturidade indica que a empresa pratica ações constantes em todos, ou na maioria dos itens do teste. A interpretação detalhada dos níveis pode ser encontrada em Severo Peixe (2014).

Características do algoritmo TAC: Para realizar a simulação do TAC, utilizou-se o pacote catR (Magis e Raiche, 2012) do software R (R Core Team, 2014). Foram utilizadas as respostas reais emitidas ao teste original não-adaptativo das 330 empresas. O TAC inicia com a aplicação de três itens com dificuldade mediana. Após, é efetuada a estimativa da proficiência provisória via método EAP (média da distribuição posterior). Depois de cada resposta dada, a estimativa é atualizada e o próximo item a ser aplicado é selecionado utilizando o método de máxima informação de Fisher (Magis e Raiche, 2012).

As seguintes regras de parada para o TAC foram utilizadas para comparação dos resultados: banco de itens completo (55 itens); TAC com comprimento variável: com base na definição de uma precisão mínima de 0,3 e 0,4 ($EP(\theta) < 0,3$ e $EP(\theta) < 0,4$); TAC com comprimento fixo: teste com 16 itens (para que fossem aplicados pelo menos três itens de cada conteúdo) e 27 itens

(aproximadamente metade do teste original completo).

As regras de comprimento fixo e variável foram simuladas em conjunto com a inserção ou não da estratégia de balanceamento de conteúdo (BC) para a seleção dos itens. Desta forma, um total de nove cenários foram testados.

Análise dos resultados

Os escores dos diferentes cenários de simulações do TAC foram comparados com os escores obtidos do teste original (não-adaptativo) com o banco de itens completo. Estatísticas como correlações de Pearson, raiz do erro quadrático médio (RSME), viés e média do EP das estimativas dos escores via TAC foram utilizadas para análise dos resultados. Além disso, resultados em relação a utilização dos itens do BI nos diferentes cenários são investigados.

Resultados e Discussão

Diferentes regras de finalização do teste foram utilizadas afim de verificar o impacto na estimativa da proficiência e compará-las com os resultados do teste original não-adaptativo. O BC corresponde a administrar itens referentes à política ambiental, ao planejamento, à implementação e operação, à verificação e ação corretiva e à análise crítica pela administração para melhoria contínua, de forma proporcional a quantidade de itens em cada área no BI, para mensurar a maturidade do SGA das empresas.

A restrição de BC em conjunto com uma regra de finalização do teste que estabelece um número fixo de itens a serem aplicados, foi utilizada para garantir que pelo menos três itens de cada área seriam administrados aos respondentes, pois conforme Luecht e Sireci (2011) e Makransky e Glas (2013), testes adaptativos muito curtos contendo, por exemplo, 10 itens, poderiam não cobrir adequadamente o conteúdo suficiente para tomar decisões ou utilizar escores válidos. Este

fato ocorreu quando foi utilizada a regra de finalização com base na precisão, sendo administrado aos respondentes, em média, 6 a 9 itens (Tabela II).

A regra de parada com base na precisão garante que todos os respondentes irão finalizar o teste após atingirem uma precisão mínima na estimativa do escore, contanto que o BI dê suporte para isso. A qualidade de um BI é essencial para mensurar com precisão o escore dos respondentes. Um BI pode ser considerado bom quando possui validade de conteúdo (cobertura de todos os aspectos relacionados ao traço latente a ser medido) e tem itens suficientes para obter alta precisão da medida em todos os níveis da escala (Bjorner *et al.*, 2007). Ou seja, itens com vários níveis de dificuldade, permitindo que o algoritmo selecione itens com nível de dificuldade próximo ao nível de proficiência estimada do respondente (Veldkamp e Matteucci, 2013). O nível de precisão aceitável para determinados níveis da escala pode variar de acordo com o objetivo do teste.

É preciso tomar cuidado ao estabelecer um critério de precisão muito baixo como regra de parada se o BI não é capaz de estimar com alta precisão todos os níveis da escala, uma vez que respondentes localizados nos extremos da escala poderão nunca atingir este critério. Por isso, esta regra é combinada com um número máximo de itens a serem aplicados, prevendo que esta situação possa ocorrer.

Precisão das estimativas do nível de maturidade do SGA das empresas

Neste estudo, utilizando a regra $EP(\theta) < 0,3$, com ou sem BC, a proporção de respondentes que atenderam ao critério de parada foi de 0,94. Somente 6% (19 respondentes) não atingiram a regra de precisão estabelecida, os quais estavam localizados nos extremos da escala. Os respondentes que satisfizeram a regra pertenciam ao intervalo de -2,93 a 2,29, a região da FIT (Figura 3) com mais informação. Utilizando a regra $EP(\theta) < 0,4$, com

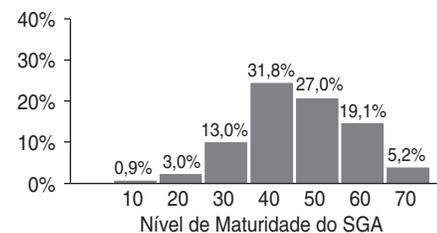


Figura 4. Frequência do nível de maturidade do SGA das 330 indústrias utilizadas para simulação do TAC.

TABELA II
EFEITOS DA MANIPULAÇÃO DA REGRA DE PARADA E BALANCEAMENTO DE CONTEÚDO

Estratégia de BC	Regra de parada	quantidade média de itens	Correlação (θ teste original \times θ TAC)	média do EP	RMSE	viés
sem BC	$EP(\theta) < 0,3$	9,1	0,975	0,278	0,266	0,008
	$EP(\theta) < 0,4$	5,7	0,966	0,345	0,312	-0,006
	16 itens	16	0,990	0,183	0,175	-0,006
	27 itens	27	0,995	0,145	0,127	-0,001
com BC	$EP(\theta) < 0,3$	9,2	0,980	0,285	0,238	-0,015
	$EP(\theta) < 0,4$	5,7	0,967	0,349	0,307	0,000
	16 itens	16	0,992	0,188	0,157	-0,007
	27 itens	27	0,994	0,146	0,129	-0,012

ou sem BC, a proporção de respondentes que atenderam ao critério de parada foi de 0,99. Apenas quatro respondentes não atingiram a precisão estabelecida, sendo assim responderam aos 55 itens (BI completo). Os respondentes que não atenderam esta regra, apresentaram proficiência >2,75, a região da FIT com menos informação.

A Tabela II apresenta os principais resultados das estatísticas geradas para comparação entre as diferentes regras para o TAC e o teste original. De uma forma geral, verifica-se que as correlações entre os escores do teste original e escores TAC de comprimento fixo e variável foram >0,96, sendo consideradas altas. As médias do EP das estimativas variaram de 0,145 a 0,349 e na maioria dos casos, as estimativas dos escores foram subestimadas (viés negativo) e muito próximas a 0.

Estabelecendo a regra de parada com base na precisão, a média dos EP foi maior do que TAC com comprimento fixo, o que é esperado, uma vez que TAC com base na precisão administrou um número menor de itens. O RMSE mostra resultados muito parecidos para a regra de parada com base na precisão, com e sem BC, sendo que TAC com BC apresentou resultados levemente melhores. Da mesma forma ocorre para TAC com regra de parada com comprimento fixo e BC.

A Figura 5 representa a dispersão dos escores obtidos no teste original e os escores obtidos com a aplicação do TAC considerando como regra de parada 55 itens, ou seja, o BI completo foi aplicado aos respondentes. Na Figura 5 pode-se observar que as estimativas de -1,5 a 1,5 tiveram uma pequena variação, mas nos extremos, a precisão das estimativas foi muito boa. Os resultados mostraram uma correlação de 0,998 entre os testes, RMSE de 0,078 e viés de 0,006. A média do EP foi de 0,116, inferior a 0,135, que foi obtida nas estimativas utilizando o teste original. Esses

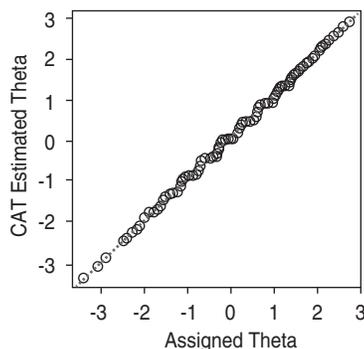


Figura 5. Gráfico de dispersão dos escores originais vs TAC com 55 itens.

resultados demonstram que o método permite obter resultados muito próximos dos obtidos via método tradicional (não-adaptativo) de aplicação de testes.

TAC com comprimento variável: O número médio de itens aplicado variou entre 6 ($EP(\theta) < 0,4$) e 9 ($EP(\theta) < 0,3$) na Tabela II, tanto com BC ou sem BC, com mediana de 3 e 5 itens, respectivamente. Isso implica que o teste é capaz de estimar o escore das empresas com boa precisão aplicando poucos itens, gerando uma redução de pelo menos 83% em relação ao teste original. Entretanto, quando tão poucos itens são aplicados, conforme mencionado, pode comprometer a validade dos resultados. Em relação ao RMSE, verifica-se que os valores mais baixos são obtidos quando a restrição de BC é estabelecida (0,238 e 0,307), obtendo-se assim resultados mais precisos, conforme mostra a Figura 6, assim como correlações mais altas.

TAC com comprimento fixo: De acordo com os resultados apresentados na Tabela II, verifica-se que, com a aplicação de 16 itens já é possível obter uma correlação alta com o teste original, de 0,99. Além disso, as médias do EP das estimativas do escore das empresas são baixas, <0,19 e com RMSE <0,17. Os resultados em relação a presença ou não da restrição de BC para esta regra de parada foram muito próximos. Por isso, sugere-se o uso da restrição de BC para obter uma melhor medida da proficiência investigada, pois quando 16 itens são aplicados, os respondentes precisarão responder a 3

itens de cada uma das áreas sobre política ambiental, planejamento, implementação e operação, verificação e ação corretiva, e 4 itens da área de melhoria contínua.

A estratégia de BC controla o número de itens selecionado de cada área, além de maximizar a informação dentro das áreas para selecionar o próximo item a ser aplicado (Luecht *et al.*, 1998). Com a aplicação de 16 itens, há uma redução de 71% em relação ao teste original, e o TAC é capaz de estimar o nível de maturidade do SGA das empresas com uma boa precisão. Aplicando 27 itens, o que corresponde a uma redução de ~50% em relação ao teste original, os resultados são muito próximos aos obtidos com 16 itens. A Figura 7 mostra que há uma alta precisão de medida, mesmo com uma grande redução de itens. Isso é possível porque o TAC aplica itens ideais para o nível dos respondentes.

Por fim, características relacionadas a regra de parada e restrição de BC não diferiram substancialmente em relação a precisão do escore em comparação ao teste original. Estes procedimentos apresentaram um bom desempenho, estando de acordo com o estudo de Luecht *et al.* (1998), que concluíram que os escores e precisão não foram afetados significativamente, mas a validade de conteúdo pode ser comprometida quando ignorado o BC.

Uso do banco de itens

Embora os itens para mensurar a maturidade do SGA das

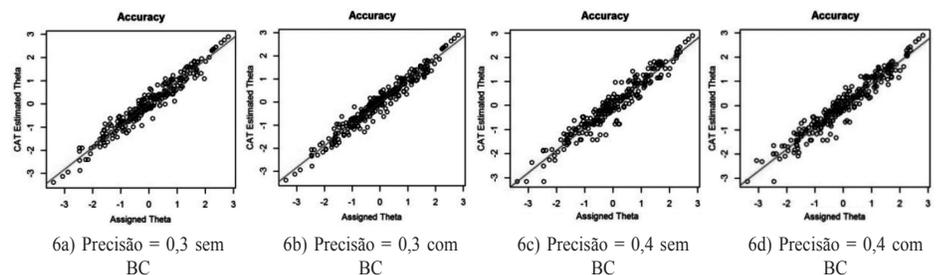


Figura 6. Gráfico de dispersão dos escores obtidos pelo teste original e pelo TAC com regras de parada baseadas na precisão das estimativas, com e sem a presença da restrição de balanceamento de conteúdo.

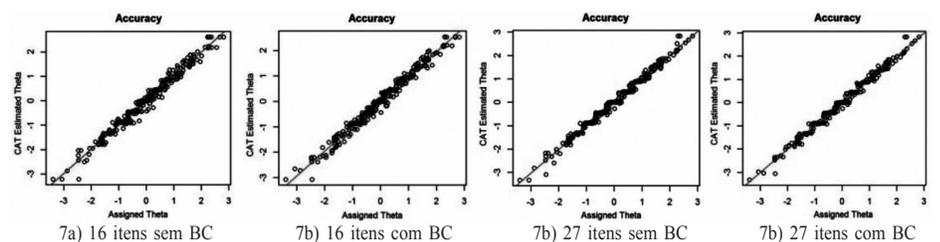


Figura 7. Gráfico de dispersão dos escores obtidos pelo teste original e pelo TAC de comprimento fixo.

empresas não sejam sigilosos, buscou-se verificar as taxas de exposição dos itens e a distribuição do uso dos itens do BI nos diferentes cenários avaliados neste estudo. Para avaliar o uso dos itens do BI, a taxa de exposição do item foi verificada, dada pelo número de respondentes que responderam o item dividido pelo número total de respondentes; já, a taxa de sobreposição de itens se refere ao número de itens comuns entre os respondentes.

Em todos os cenários, três itens apresentaram taxa máxima de exposição, ou seja, foram apresentados a todos os respondentes (itens 17, 44, 49, os quais possuem alto poder de discriminação). O item 17 aborda o planejamento (A empresa define os responsáveis pela implementação dos programas para atingir as metas e objetivos?) e os itens 44 e 49 abordam questões de melhoria contínua (44: A empresa analisa as ameaças e oportunidades em relação aos *stakeholders*? e 49: Periodicamente, é realizada análise para definir melhoria contínua na gestão ambiental?).

A Tabela III mostra as taxas mínimas de exposição obtidas para cada cenário e o número de itens com mínima exposição. Quando a regra de parada foi baseada na precisão, todos os itens do BI foram aplicados pelo menos uma vez. Para o TAC de comprimento de 16 itens, sem BC, 6 itens não foram apresentados a nenhum respondente; com BC, 8 itens não foram aplicados. Para a regra de 27 itens, sem BC, todos os itens foram aplicados; com BC, o item 11 nunca foi aplicado. O item 11, que não foi aplicado quando a regra de parada do TAC é comprimento fixo, relaciona-se ao planejamento: Na definição dos objetivos e metas, a empresa leva em consideração a opinião dos *stakeholders*?

Os resultados referentes à taxa de sobreposição de TAC com BC na Tabela III são muito próximos aos resultados sem esta restrição. Ao se analisar a regra de parada com comprimento variável ou fixo, percebe-se que o TAC de comprimento variável obteve uma taxa menor de sobreposição de itens do que o TAC de comprimento fixo. Isto é

esperado, já que o TAC de comprimento variável apresentou melhor uso do BI como um todo e um número menor de itens são aplicados.

Por fim, pode-se afirmar que, o TAC apresenta vantagens sobre o teste original, podendo ser utilizado com eficiência para mensurar o nível de maturidade do SGA das empresas. Este método é utilizado por organizações educacionais para reduzir o custo dos testes, carga de trabalho, erros humanos, assim como para aumentar a precisão e a eficiência do teste e atraso na divulgação dos resultados (Economides e Roupas, 2007). Em contextos organizacionais, TACs podem ser utilizados para gerar informações mais detalhadas de cada respondente, auxiliando na tomada de decisões de promoção e seleção de funcionários (Makransky e Glas, 2013). Em comparação aos testes de P&P, TAC proporciona, entre outras vantagens, a redução de custos com materiais para o desenvolvimento de testes, armazenamento e correção (Wainer, 2000).

Conclusões

Este artigo teve por objetivo apresentar as vantagens do uso de TACs para avaliação da maturidade do SGA de indústrias. A partir dos resultados, foi possível verificar os benefícios em termos de eficiência e precisão que podem ser obtidos com o uso de testes adaptativos computadorizados para estimar o escore das empresas. Todos os cenários avaliados em comparação aos resultados obtidos por meio da aplicação do teste original, composto por 55 itens, mostraram-se precisos, com pequeno viés e RMSE, além de altas correlações entre os escores nos diferentes métodos.

A implantação deste método de avaliação poderia trazer uma redução de ~71% em relação ao teste original, estabelecendo a aplicação de 16 itens com a restrição de balanceamento de conteúdo, cenário sugerido para utilização futura no desenvolvimento e aplicação efetiva do TAC às empresas, garantindo resultados precisos e mais adequados para o traço latente investigado. A restrição de

balanceamento de conteúdo mostrou não interferir na precisão das estimativas.

Esta redução no comprimento do teste pode incentivar os gestores a responderem ao teste e a monitorar seus resultados ao longo do tempo, demonstrando oportunidade para a melhoria contínua do SGA das indústrias, além de gerar um resultado detalhado sobre suas práticas. Ademais, novos trabalhos poderão ser implementados com o uso de TACs para garantir resultados com redução de tempo e custos na aplicação, tanto na área de gestão ambiental quanto em outras áreas.

AGRADECIMENTOS

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro.

REFERÊNCIAS

- Andersen ES, Jessen SA (2003) Project maturity in organizations. *Int. J. Proj. Manag.* 21: 457-461.
- Balzarova MA, Castka P (2008) Underlying mechanisms in the maintenance of ISO 14001 environmental management system. *J. Cleaner Product.* 16: 1949-1957.
- Bent LM, Mulcahey MJ, Kelly EH, Calhoun CL, Tian F, Ni P, Vogel LC, Haley SM (2013) Validity of computer adaptive tests of daily routines for youth with spinal cord injury. *Top Spinal Cord Inj. Rehabil.* 19: 104-113.
- Bjorner JB, Chang CH, Thissen D, Reeve BB (2007) Developing tailored instruments: item banking and computerized adaptive assessment. *Quality of Life Research.* 16: 95-108.
- Bowling A (2005) Mode of questionnaire administration can have serious effects on data quality. *J. Public Health* 27: 281-291.
- Burchell B, Marsh C (1992) The effect of questionnaire length on survey response. *Qual. Quant.* 26: 233-244.
- Cooke-Davies TJ, Arzymanow A (2003) The maturity of project management in different industries: an investigation into variations between project management model. *Int. J. Proj. Manag.* 21: 471-478.
- Crawford SD, Couper MP, Lamias MJ (2001) Web surveys: Perceptions of burden. *Soc. Sci. Comput. Rev.* 19: 146-162.
- Dinsmore PC (1999) *Transformando estratégias empresariais em resultados através da gerência de projetos*. Qualitymark. Rio de Janeiro, Brasil. 268 pp.
- Economides AA, Roupas C (2007) Evaluation of computer adaptive testing systems. *Int. J. Web-Based Learn. Teach. Technol.* 2: 70-87.
- Galesic M, Bosnjak M (2009) Effects of questionnaire length on participation and indicators of response quality in Web surveys. *Public Opin. Quart.* 73: 349-60.
- Grant KP, Pennypacker JS (2006) Project management maturity: an assessment of project management capabilities among and

TABELA III
USO DO BANCO DE ITENS E EXPOSIÇÃO DE ITENS DO TESTE

	sem BC				com BC			
	EP(θ)<0,3	EP(θ)<0,4	16 itens	27 itens	EP(θ)<0,3	EP(θ)<0,4	16 itens	27 itens
Taxa mínima de exposição	0,0576	0,0121	0	0,0273	0,0576	0,0121	0	0
Número de itens com mínima exposição	3	3	6	1	4	1	8	1
Taxa de sobreposição de itens	0,466	0,5698	0,6847	0,757	0,4711	0,5636	0,6937	0,764

- between industries. *IEE Trans. Eng. Manag.* 53: 59-68.
- Halila F, Tell J (2013) Creating synergies between SMEs and universities for ISO 14001 certification. *J. Cleaner Product.* 48: 85-92.
- Hol AM, Vorst HCM, Mellenbergh GJ (2008) Computerized adaptive testing of personality traits. *J. Psychol.* 216: 12-21.
- Huang YM, Lin YT, Cheng SC (2009) An adaptive testing system for supporting versatile educational assessment. *Comput. Educ.* 52: 53-67.
- ISO (2004) *NBR ISO 14.001: Sistemas de Gestão Ambiental - Especificação e Diretrizes para Uso.* International Organization for Standardization. ABNT. Rio de Janeiro, Brasil.
- Krosnick JA, Holbrook AL, Berent MK, Conway M (2002) The impact of "no opinion" response options on data quality: Non-attitude reduction or an invitation to satisfice? *Public Opin. Quart.* 66: 371-403.
- Lee HY, Dodd BG (2012) Comparison of exposure controls, item pool characteristics, and population distributions for CAT using the partial credit model. *Educ. Psychol. Meas.* 72: 159-175.
- Luecht RM, Sireci SG (2011) *A Review of Models for Computer-Based Testing.* Research Report 2011-12. College Board. EEUU. 50 pp.
- Luecht RM, De Champlain A, Nungester RJ (1998) Maintaining content validity in computerized adaptive testing. *Adv. Health Sci. Educ.* 3: 29-41.
- Magis D, Raiche G (2012) Random generation of response patterns under computerized adaptive testing with the R package catR. *J. Stat. Softw.* 48: 1-31.
- Makransky G, Glas CAW (2013) The applicability of multidimensional computerized adaptive testing for cognitive ability measurement in organizational assessment. *Int. J. Testing.* 13: 123-139.
- Montgomery JM, Cutler J (2013) Computerized adaptive testing for public opinion surveys. *Polit. Anal.* 21: 172-192.
- Oliveira OJ, Serra JR (2010) Benefícios e dificuldades da gestão ambiental com base na ISO 14001 em empresas industriais de São Paulo. *Produção* 20: 429-438.
- R Core Team (2014) *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. Vienna, Austria. www.R-project.org/
- Samejima FA (1969) *Estimation of Latent Ability Using a Response Pattern of Graded Scores.* Psychometric Monograph N° 17. Psychometric Society. Richmon, VA, EEUU. 100 pp.
- Severo Peixe BC (2014) *Mensuração da Maturidade do Sistema de Gestão Ambiental de Empresas Industriais Utilizando a Teoria da Resposta ao Item.* Tese. Universidade Federal de Santa Catarina. Brasil. 381 pp.
- Singh J, Howell RD, Rhoads GK (1990) Adaptive designs for Likert-type data: An approach for implementing marketing surveys. *J. Market. Res.* 27: 304-321.
- Thissen D, Chen WH, Bock RD (2003) *Multilog* (ver. 7) Scientific Software International. Lincolnwood, IL, EEUU.
- Thompson NA, Weiss DJ (2011) A framework for the development of computerized adaptive tests. *Pract. Assess. Res. Eval.* 16: 1-9.
- To WM, Lee PKC (2014) Diffusion of ISO 14001 environmental management system: global, regional and country-level analyses. *J. Cleaner Product.* 66: 489-498.
- Veldkamp BP, Matteucci M (2013) Bayesian computerized adaptive testing. *Ensaio* 21: 57-82.
- Wainer H (2000) CATs: Whither and whence. *Psicológica* 21: 121-133.
- Wang T, Kolen MJ (2001). Evaluating comparability in computerized adaptive testing: Issues, criteria and an example. *J. Educat. Meas.* 38: 19-49.
- Webster (1988) *The New Lexicon Webster's Dictionary of the English Language.* Lexicon. New York, EEUU.
- Yammarino FJ, Skinner SJ, Childers TL (1991) Understanding mail survey response behavior: A metaanalysis. *Public Opin. Quart.* 55: 613-639.
- Zorpas A (2010) Environmental management systems as sustainable tools in the way of life for the SMEs and VSMEs. *Bioresource Technol.* 101: 1544-1557.

ADVANTAGES OF THE USE OF COMPUTERIZED ADAPTIVE TESTING FOR MATURITY EVALUATION OF THE ENVIRONMENTAL MANAGEMENT SYSTEM IN INDUSTRIES

Débora Spennassato, Blênio C. Severo Peixe, Andréa C. Trierweiler, Antonio C. Bornia and Rafael Tezza

SUMMARY

The Computer Adaptive Test (CAT) selects the most informative items for respondents, which is an advantage over non-adaptive methods that use computer or paper and pencil. This paper aims to present the advantages of using CAT for maturity evaluation of the Environmental Management System in industries. The test considers aspects relating to environmental policy, planning, implementation and operation, checking and corrective action, and review by management, focus-

ing to continuous improvement. In this sense, we conducted a simulation study to evaluate the effect of different CAT stopping rules and compared the results with the original version of the test. Results show that there could be a reduction of 71% in the test length in deploying a CAT without compromising the validity and accuracy of the measure, benefiting both respondents and those responsible by development and application of tests.

VENTAJAS DEL USO DE TESTS ADAPTATIVOS INFORMATIZADOS PARA EVALUACIÓN DE LA MADUREZ DEL SISTEMA DE GESTIÓN AMBIENTAL DE INDUSTRIAS

Débora Spennassato, Blênio C. Severo Peixe, Andréa C. Trierweiler, Antonio C. Bornia y Rafael Tezza

RESUMEN

El Test Adaptativo Informatizado (TAI) presenta ventajas en comparación con los métodos no adaptativos aplicados por computadora o papel y lápiz, por lo que selecciona ítems más informativos para los encuestados. Ese trabajo tiene por objetivo presentar las ventajas del uso de los tests adaptativos computadorizados para análisis de la madurez del Sistema de Gestión Ambiental (SGA) de industrias. El test trata cuestiones relativas a la política ambiental, planeamiento, implementación y

operación, verificación y acción continua. Para eso, realiza un estudio de la simulación, para evaluar el efecto de diferentes reglas de finalización del TAI y los resultados se comparan con su versión original. Estos resultados demuestran que implantándose un TAI se puede lograr una reducción del 71% en la extensión del test sin comprometer la validez y precisión de la medida, generando un beneficio tanto a los encuestados como a los responsables del desarrollo y aplicación de los tests.