



CORRELAÇÃO ENTRE AVALIAÇÃO DE REAÇÃO E DE APRENDIZAGEM, EM TREINAMENTO OPERACIONAL

CORRELATION BETWEEN REACTION AND LEARNING ASSESSMENT IN OPERATIONAL TRAINING

Recebido em: 30 abr. 2020

Aprovado em: 25 out. 2020

Versão do autor aceita publicada online: 25 out. 2020

Publicado online: 04 jun. 2021

Como citar esse artigo - American Psychological Association (APA):

Pereira, G. B., Correa, J. M. J. S., Heizer, I. H., Melo, A. B., & Silva, C. M. (2022, abr./jun.).
Correlação entre avaliação de reação e de aprendizagem, em treinamento operacional. *Exacta*,
20(2), 444-469. <https://doi.org/10.5585/exactaep.2021.17091>.

Submeta seu artigo para este periódico 



Dados Crossmark



CORRELAÇÃO ENTRE AVALIAÇÃO DE REAÇÃO E DE APRENDIZAGEM, EM TREINAMENTO OPERACIONAL

Correlation between reaction and learning assessment in operational training

 Gilberto Braga Pereira¹
 Joicy Mary Josep Silva Correa²
 Ionara Houry Heizer³
 Ana Beatriz Melo⁴
 Carlucia Maria Silva⁵

¹ Doutorado
Pontifícia Universidade Católica de Minas Gerais – PUC Minas.
Belo Horizonte, Minas Gerais – Brasil.
gilbertobp2018@gmail.com

² Graduação
Fundação Comunitária de Ensino Superior de Itabira – FUNCESI.
Itabira, Minas Gerais – Brasil.
joicy-mary@hotmail.com

³ Mestrado
Fundação Comunitária de Ensino Superior de Itabira – FUNCESI.
Belo Horizonte, Minas Gerais – Brasil.
ionarahoury@gmail.com

⁴ Doutorado
Pontifícia Universidade Católica de Minas Gerais – PUC Minas.
Cidade, Estado – País.
melo.ana.bh@gmail.com

⁵ Pós-doutorado
Universidade do Estado de Minas Gerais – UEMG.
Belo Horizonte, Minas Gerais – Brasil.
carlucia.maria@gmail.com

Recebido em: 30 abr. 2020
Aprovado em: 25 out. 2020

Resumo: Objetivou-se apurar o tipo de correlação existente entre os níveis reação e aprendizagem (Donald Kirkpatrick & Kirkpatrick, J., 2006), em treinamento operacional. Metodologia quantitativa e descritiva, a partir do método documental. Universo de 600 empregados do setor operacional. Critério de amostragem probabilístico aleatório simples, selecionando-se resultados de 228 operadores de equipamentos de grande porte. Tratamento de dados a partir da análise estatística descritiva e inferencial (software SPSS versão 21.0 e Microsoft Excel versão 2013). Nos resultados comprovou-se a validade do instrumento de satisfação utilizado pela empresa pesquisada; o nível de reação demonstrou aceitação positiva das ações de treinamento, com altos níveis de favorabilidade. Quanto ao nível de aprendizagem (pré e pós-teste), obtiveram-se ganhos médios de 72%. A correlação de Spearman não encontrou índice significativo entre os níveis. Sugere-se, portanto, a necessidade de pesquisas aprofundadas que atestem ao relação de interdependência direta entre os níveis averiguados, a fim de que o modelo proposto confirme-se como válido para apurar a efetividade de sua aplicação.

Palavras-chave: Treinamento operacional. Avaliação de reação. Avaliação de Aprendizagem.

Abstract: Evaluating the effectiveness of training programs has been a challenge for practitioners and organizations. This research aimed to identify the type of correlation existent between staff and company in operational training in a national steel company. The adopted methodology was configured as quantitative and descriptive, from the documentary method. The delimited scenario was 600 employees from the operating sector. The method for sampling was determined as simple random probability selecting results from 228 who operate large equipment. The descriptive and inferential statistical analysis was used to data analysis through the software SPSS version 21.0 and Microsoft Excel version 2013. Data proved the validity of the satisfaction instrument used in the company researched, in addition the verification of the level of response showed positive acceptance of training actions with high levels of favorability. Employees obtained on average 72% gains in regard to the level of learning once the score of pre and post tests were compared. On the other hand, no significant index between levels has been found by Spearman's rank-order correlation.

Keywords: Operational training. Reaction evaluation. Learning Assessment.



1 Introdução

O desenvolvimento de estratégias voltadas à potencialização das competências organizacionais e pessoais impõem a escolha de métodos capazes de aferir a eficácia dos resultados e elevam o Treinamento, Desenvolvimento e Educação (TD&E) à condição de diferencial competitivo. As inovações tecnológicas deliberadamente promovem modificações significativas nos processos produtivos, com consequentes exigências de capacitação de pessoas.

Assim, a avaliação da eficácia organizacional a partir da adoção de políticas e práticas, dentre as quais recorta-se, neste estudo, as relacionadas a programas de qualificação de pessoas, é condição precípua.

Em tal âmbito, o da eficácia de TD&E, a indagação de pesquisa norteadora se justifica e foi configurada como: Qual o tipo de correlação existente entre o nível de reação e o nível de aprendizagem em empregados submetidos a treinamento operacional, em uma empresa do ramo siderúrgico de atuação nacional?

O dinamismo do ambiente organizacional exige o desenvolvimento de ações voltadas à qualificação das pessoas, para torná-las mais eficazes em suas atribuições (Gil, 2009). Desseler (2014) acresce que é surpreendente a influência que o treinamento tem sobre o desempenho das pessoas, apresentando maior pontuação do que feedback e avaliação, localizando-se logo abaixo da fixação de metas, sem obviamente destacar, como o faz Bastos (1994), que o mesmo é meio que possibilita aquisição de experiências capazes de modificar comportamentos.

O treinamento cumpre várias funções, dentre as quais, destaca Desseler (2014), está a de evitar o imprevisto por parte dos funcionários que não sabem o que ou como realizar uma atividade. Outra função ressaltada pelo autor está a redução de riscos.

Milkovich e Boudreau (2009) atestam que os treinamentos são divididos segundo o critério de clientela em gerenciais e operacionais. O treinamento operacional é essencial em qualquer organização e é frequentemente o mais realizado. Quando os objetivos pretendidos pelo treinamento são o aperfeiçoamento ou aquisição de habilidades voltadas para a execução de tarefas, estes são classificados como técnico-operacionais ou apenas operacionais.

No Brasil, 77% das empresas utilizam o critério de reação, 30% aprendizado, 9,8% mudança de comportamento e 4,9% resultados como práticas de aferição de eficácia. Quanto ao tipo de treinamento, 42% das empresas investem no público gerencial e 58% das ações são relacionados ao setor produtivo, voltados para operação, formação técnica, comportamental e capacitação (Associação Brasileira de Treinamento e Desenvolvimento, 2015), dados que atestam a relevância do tema e delimitam um desafio ao engenheiro de produção, visto que a Associação Brasileira de Engenharia de

Produção (ABEPRO) especifica como função do engenheiro, promover a melhoria contínua do processo produtivo, considerando as pessoas, os materiais e os equipamentos implicados.

Considerando as assertivas de autores, há pelo menos duas razões básicas para a realização do treinamento operacional, foco desta pesquisa. A primeira está vinculada a sua própria função, que é a de produzir bens e serviços com qualidade. A segunda está relacionada à evolução tecnológica; vários processos e equipamentos são constantemente substituídos por outros mais modernos e eficientes, requerendo assim a adaptação dos trabalhadores aos novos modelos de produção.

Dentre os modelos de avaliação da eficácia de TD&E, o proposto por Kirkpatrick, D. e Kirkpatrick, J. (2006) é o mais difundido e presente nas políticas e práticas organizacionais contemporâneas, muito embora pouca pesquisa acadêmica seja desenvolvida para testar as afirmações e postulados preconizados por seus autores. Ocorre, pois, uma aceitação tácita dos postulados, sem que se verifique no todo ou em parte sua validade heurística no contexto organizacional atual. O modelo de quatro níveis preconiza não somente a interdependência intrínseca, como uma alta correlação entre eles. São apresentados com uma complexidade crescente na avaliação do impacto de TD&E sobre a eficiência e eficácia organizacionais. O recorte do presente trabalho foi o de testar a natureza da correlação existente entre os dois primeiros níveis, quais sejam: reação e aprendizagem. Parte-se, pois, da afirmação de que ao se obter um alto índice de reação positiva ao treinamento operacional ministrado, há uma relação direta de crescimento do nível de aprendizagem. A partir de uma experiência no uso do modelo em questão, buscou-se identificar evidências empíricas capazes de sustentar a natureza correlacional preconizada, antecipando-se que neste relato apresentam-se argumentos que a contestam. O uso da abordagem quantitativa de pesquisa, a partir da aplicação de técnicas próprias à estatística inferencial e descritiva o estudo contribui com dados representativos que apontam a necessidades de novas pesquisas.

A fim de trazer contribuições concernentes a esta discussão, o presente relato de pesquisa quantitativa e descritiva está estruturado, inclusa esta introdução, em cinco seções. A segunda e terceira, trazem respectivamente os fundamentos essenciais referentes ao modelo em questão e a correlação afirmada por Kirkpatrick, D. e Kirkpatrick, J. (2006) acerca da avaliação de reação e aprendizagem. Em sequência, nas seções quatro e cinco, respectivamente, discorre-se sobre o percurso metodológico e a apresentação/discussão dos resultados, a fim de delimitar respostas ao problema de pesquisa.

2 Conceituação, processo e avaliação de treinamento

Considerando as assertivas de autores, há pelo menos duas razões básicas para a realização do treinamento operacional. A primeira está vinculada a própria função organizacional, que é a de produzir



bens e serviços com qualidade. A segunda está relacionada à evolução tecnológica; vários processos e equipamentos são constantemente substituídos por outros mais modernos e eficientes, requerendo assim a adaptação dos trabalhadores aos novos modelos de produção.

Milkovich e Boudreau (2009) acrescentam que o achatamento das estruturas hierárquicas permitiu a formação de lideranças entre os trabalhadores de níveis mais baixos na pirâmide organizacional, novas responsabilidades que exigem mais do que especialidade técnica, mas a aquisição de habilidades típicas de gerenciamento (Milkovich & Boudreau, 2009).

Visando auxílio às organizações, a International Organization for Standardization (ISO), desenvolveu a norma ISO 10015: Gestão da Qualidade - Diretrizes para Treinamento. A versão equivalente brasileira da norma é a NBR ISO 10015, publicada pela ABNT (2001). O objetivo desta norma é dispor de orientações para justificar a eficácia dos programas de TD&E, inclusas as etapas de elaboração de programas e os investimentos realizados, de toda ordem, no tocante à capacitação de trabalhadores. (ABNT, 2001).

Cumpridas as etapas precedentes, como sugere a NBR ISO 10015 (ABNT, 2011) de delimitação de necessidades, planejamento e execução das ações de capacitação, interessa particularmente a este relato de pesquisa, detalhar a quarta etapa: avaliação da eficácia ou apuração dos resultados auferidos, a qual constitui etapa crítica e, devido sua complexidade, na maioria das vezes é negligenciada. Entretanto, é a única capaz de confirmar se os objetivos foram atendidos. (Gil, 2009; Palmeira, 2009). Tanto tais procedimentos normativos quanto os livros didáticos nos quais se discutem tal temática, o modelo de quatro níveis de Kirkpatrick e Kirkpatrick (2006) figuram como um dos mais usuais e representativos.

O processo de avaliação, segundo Borges-Andrade (2002), pode ser definido como um meio de coleta de dados para inferir juízo de valor sobre um programa de treinamento. É a etapa principal, capaz de fornecer informações para o aperfeiçoamento do sistema, enquanto Gil (2009) afirma que compreende qualquer tentativa de reunir informações sobre os impactos de um programa.

Dentre as razões básicas para se avaliar, a melhoria de programas futuros é a mais comum. A segunda é auxiliar na tomada de decisão para continuar ou não o programa. A terceira é justificar o esforço e o investimento realizados. Além disso, proporcionará mais segurança aos profissionais responsáveis por demonstrar à gerência superior a importância estratégica de tal processo (Kirkpatrick, D. & Kirkpatrick, J., 2006). No entanto, Milkovich e Boudreau (2009) advertem que as avaliações são evitadas, por vezes, em face ao receio de que os resultados contrários ao esperado prejudiquem os procedimentos adotados ou acarretem em corte de investimento.

Métodos e modelos específicos são sugeridos por autores, o mais amplamente utilizado é o desenvolvido por Donald Kirkpatrick na década de cinquenta e posteriormente em continuidade com seu filho James Kirkpatrick, cuja primeira publicação se deu no *Jornal da Associação Americana de*

Treinamento e Desenvolvimento (ASTD), em 1959 (Kirkpatrick & Kirkpatrick, 2006). Desde então, tem sido referência internacional para avaliar a eficácia do treinamento.

Independentemente do tipo de treinamento ou o meio onde são realizados, o modelo de quatro níveis adota como premissa uma espécie de correlação entre os fatores. Segundo Kirkpatrick e Kirkpatrick (2006) a avaliação deve seguir a ordem, do nível 1 até o nível 4 sucessivamente. Cada nível tem seu impacto no nível posterior, por isso, nenhum nível deve ser ignorado. Os autores preconizam que a avaliação de reação mensura a satisfação dos treinandos. A aceitação do treinamento é o primeiro indicador de que o programa vai ser efetivo, pois revela a motivação para o aprendizado. Até porque, está reação será compartilhada repercutindo dentro e fora da organização. Atesta a aceitação do, considerando aspectos como conteúdo, recursos didáticos, qualidade da instrução e da infraestrutura. Entretanto, a mesma não é suficiente para identificar se os objetivos foram alcançados, vez que não comprova a aquisição de novos conteúdos ou conhecimentos, habilidades e atitudes transmitidos pelo instrutor (Gil, 2009; Palmeira, 2013).

Apesar dessa limitação implícitas, Kirkpatrick e Kirkpatrick (2006) afirmam que a reação positiva pode não garantir o aprendizado, mas a reação negativa certamente reduz a sua possibilidade de ocorrer, por esta razão assume relevância.

No segundo nível sugerido no modelo em discussão, os participantes serão avaliados quanto ao aprendizado. Mede o que o treinando aprendeu (Gil, 2009; Palmeira, 2013).

Sua relevância é justificada por Kirkpatrick e Kirkpatrick (2006) asseverando que é uma pré-condição para que se efetive a mudança de comportamento (terceiro nível), mas o contrário não é verdadeiro. Nessa perspectiva, avaliar a mudança de comportamento sem considerar a avaliação de aprendizado, pode-se concluir erroneamente que o aprendizado não ocorreu quando a mudança de comportamento não for constatada. (Kirkpatrick & Kirkpatrick, 2006).

No nível três (mudança de comportamento) pós ação de capacitação, deseja-se aferir a aplicação do aprendizado refletido nas atitudes e no modo de realizar as atividades (Gil, 2009; Palmeira, 2013). Na linguagem de Kirkpatrick e Kirkpatrick (2006) se houve transferência de aprendizado e o quanto, na dependência, porém, de que sejam oferecidas oportunidades para tanto. Bastos (1994), Gil (2009) e Palmeira (2013), corroboram que outros fatores influenciam a transferência, variando da falta de ferramentas, passando pela motivação do indivíduo e envolvendo as condições do ambiente. Como meio de minimização, Kirkpatrick e Kirkpatrick (2006) recomendam que a organização cuide das variáveis extra-organizacionais, organizacionais e individuais que podem ser impeditivas, gerando ambiente encorajador.

Por fim, no quarto nível, serão apurados os impactos gerados nos resultados organizacionais. Esses resultados podem ser identificados a partir de indicadores como: aumento de vendas, redução de custos, índice de satisfação dos clientes, aumento da qualidade, redução de retrabalho ou de produtos



fora de especificação e lucro, entre outros (Kirkpatrick, 2006; Palmeira, 2013). Como etapa mais complexa, é difícil distinguir quais efeitos globais realmente se podem ser atribuídos ao treinamento, porém, qualquer que seja o crescimento da organização em função dos objetivos do treinamento pode ser considerado como um resultado eficiente e eficaz desta ação. (Bastos, 1994; Gil, 2009; Kirkpatrick & Kirkpatrick, 2006).

Das várias formas de se medir a reação, nem todas são efetivas, alertam Kirkpatrick e Kirkpatrick (2006). Os autores disponibilizam oito orientações: determinar o que se deseja encontrar; elaborar instrumentos para quantificar as reações; encorajar os treinandos a escrever comentários e sugestões; obter cem por cento de respostas imediatas; adquirir respostas honestas; desenvolver padrões aceitáveis; comparar os resultados das avaliações com os padrões desenvolvidos e tomar decisões apropriadas; comunicar os resultados.

Sob outro ponto de vista, Palmeira (2013) orienta que os questionários de satisfação devem ser aplicados preferencialmente uma semana após o treinamento, para evitar que a euforia dos participantes ao final do programa interfira nos resultados. No entanto, quando há dificuldades, seja por falta de estrutura ou de interesse dos participantes, é melhor que seja feito ao final do treinamento do que deixar de fazê-lo. Quanto ao formato do questionário, é comum encontrá-los também na forma de ícones tradutores de graus de satisfação, para cada aspecto o treinando deve escolher a expressão que melhor traduz o seu sentimento em relação ao treinamento (Kirkpatrick & Kirkpatrick, 2006).

Todavia, é recomendável que as reações medidas através dos questionários sejam quantificáveis, podendo assim definir um padrão de classificação. O método apresentado por Kirkpatrick e Kirkpatrick (2006) é simples, constitui-se de três passos conforme as orientações abaixo: 1) Atribuir um peso para cada grau de satisfação: excelente (5); bom (4); médio (3); ruim (2); péssimo (1); 2) Multiplicar o número de respostas dadas em cada grau pelo seu peso. Somar todos os produtos; 3) Dividir pelo total de respostas obtidas.

Para as avaliações de aprendizagem podem ser aplicados testes ou questionários, abordando questões sobre o assunto, sendo elas discursivas ou de múltipla escolha. Além disso, é possível que sejam feitas antes e após o evento, evidenciando assim a o que e o quanto foi retido (Kirkpatrick & Kirkpatrick, 2006; Gil, 2009; Palmeira, 2013; GIL,2009). Borges-Andrade (2002) se opõe ao método, justificando que iniciar um treinamento com avaliações pode causar constrangimentos. Igualmente, a avaliação final é, também, considerada imprópria no contexto das organizações de trabalho, devendo-se restringir seu uso ao ensino formal.

A parte as controvérsias, cuidados são demandados, tais como: clareza na elaboração das questões e, preferencialmente, que primem por um caráter quantitativo, o que permitirá tratamento estatístico dos dados a fim de se estabelecer medidas de correlação e níveis de confiança (Scorsolini-Comin, Inocente, Miura, 2011).

Outra maneira de se avaliar o aprendizado é comparar o desempenho entre grupos que participaram do programa e grupos que não participaram do programa (Bastos, 1994; Kirkpatrick & Kirkpatrick, 2006).

Extrapolando o escopo desta pesquisa aprofundar nos dois níveis subsequentes, mudança de comportamento e impactos nos resultados, motivo pelo qual especificou-se fundamentos necessários apenas à compreensão da avaliação de reação e de aprendizagem.

3 Correlação entre as variáveis de avaliação reação e aprendizagem

À medida que a demanda de treinamentos nas empresas foi aumentando, diversas pesquisas foram desenvolvidas a fim de testar os modelos de avaliação existentes, quanto à eficácia e confiabilidade dos dados obtidos. Há uma necessidade legítima de compreender as variáveis que explicam a aprendizagem no contexto do trabalho (Mourão & Marins, 2009).

O relacionamento das variáveis-critério, conforme determina o modelo de Kirkpatrick e Kirkpatrick (2006), vem apresentando inconsistências, sem que seu uso seja comprometido. Não obstante os autores salientam que todos os níveis não somente são relevantes, se relacionam diretamente, uma vez que os níveis de reação contribuem para o aprendizado, que por sua vez tem impactos na mudança de comportamento e assim gera resultados para organização. (Kirkpatrick & Kirkpatrick, 2006).

Algumas premissas são contestadas no modelo de quatro níveis. Borges-Andrade (2002), Abbad, Gama e Borges-Andrade (2000), afirmam a inexistência de relação de causalidade direta sugerida. Outras variáveis são apontadas como interventoras no processo, as quais, se atribuídas ao treinamento, se tornariam dependentes de um conjunto cada vez mais amplo de fatores, dificultando sua identificação.

No entanto, Alliger e Janak (1989 *apud* Abbad; Gama; Borges-Andrade, 2000) relatam que poucos estudos sobre o relacionamento entre os quatro níveis têm sido realizados e concluem que os pesquisadores e profissionais apenas aceitam as premissas de correlação positiva, abstendo-se de investigar. Para os autores, a relação entre aprendizagem e reação nem sempre são significativas. Conforme pesquisa de Abbad, Gama, Borges-Andrade (2000) e Santos (2012), a relação significativa e positiva encadeada entre os níveis não se evidenciam suficientemente. Por sua vez, Alliger e Janak (1989 *apud* Abbad; Gama; Borges-Andrade, 2000), sugerem que as medidas de reação e aprendizagem, de maneira geral, são tomadas ao mesmo tempo, por isso, afirmar a causalidade das variáveis seria um equívoco.

Todavia, Kirkpatrick e Kirkpatrick (2006) explicam que a reação positiva em relação ao treinamento incorrerá em melhor aproveitamento de aprendizagem. Afirmam, também, que a



continuidade dos programas de treinamento é baseada, em sua maioria, na percepção de qualidade observada por grupos antecessores.

Apesar da controvérsia dos resultados de pesquisas em relação às ideias de Kirkpatrick e Kirkpatrick (2006), merece aprofundamento, haja a vista que o seu modelo continua sendo referência acadêmica nos livros didáticos e para as organizações, antes que se conclua pela sua inoperância.

4 Metodologia

O estudo configurou-se como de natureza quantitativa, descritiva e de campo (Barros & Leheld, 2007; Kauark, Manhães e Medeiros, 2010; Mascarenhas, 2012) e considerou como universo, 600 empregados de uma siderurgia de atuação nacional, submetidos a treinamento operacional, entre os anos de 2012 a 2014. A amostra probabilística aleatória simples ficou constituída por 228 operadores de equipamentos de grande porte (Barros & Samata, 2007; Malhotra, 2011), mediante fórmula proposta por Mattar (2011).

Os dados documentais (Gil, 2010; Mascarenhas, 2012) permitiu contabilizar resultados da avaliação de reação e de aprendizagem (pré-teste e pós-teste) de 278 empregados. A análise estatística descritiva e inferencial foi processada a partir do Microsoft Excel, versão 2013 e o software SPSS versão 2010.

A apuração e análise dos resultado seguiu três passos distintos, consoantes com os três objetivos específicos do estudo, a saber: Passo 1, referente ao primeiro objetivo específico (validação do instrumento de avaliação de reação): Aqui fez-se uso da técnica de validação de constructo e a análise de confiabilidade alfa de Cronbach (Valores desejáveis: $>0,7$), foram utilizados para responder ao primeiro objetivo específico, qual seja: Validação do Instrumento de Reação em uso na empresa. Analisou-se, também, a paridade da construção do instrumento com modelos e parâmetros preconizados pela literatura (MacDaniel & Gates, 2003; Malhotra, Rocha, Laudisio, Altherman & Borges, 2009; Mattar, 1995). No coeficiente alfa Cronbach, a apuração considerou a correlação inter-itens, a correlação total de item corrigida e o Alfa de Cronbach, se o item foi excluído. Os parâmetros de correlação inter-item considerado foi: moderada entre 0,15 e 0,50 (Clark & Watson, 1995 apud Welter, 2007). A correlação total de item corrigida representa o grau em que cada item do questionário está relacionado à soma total dos pontos, onde itens com valor inferior a 0,3 não são recomendados, pois não se relacionam com o total. Para o Alfa de Cronbach se o item for excluído que apresentar valor consideravelmente superior ao alfa global deve ser eliminado para aumentar a confiabilidade do questionário (Field, 2005).

Passo 2: Em resposta ao segundo objetivo específico, apuração dos resultados de avaliação do treinamento, a fim de determinar o nível de reação e o nível de aprendizagem, foram utilizados

parâmetros de estatística descritiva como: média, desvio padrão, máximo, mínimo, amplitude e coeficiente de variação de Pearson. A média (μ) representa o valor central para um conjunto de dados, onde há maior concentração, sendo influenciada por todos os valores. Das medidas de dispersão: o desvio padrão (σ) apresenta a variabilidade média dos dados em torno da média calculada, definindo um intervalo ($\mu \pm \sigma$) em que se concentram cerca de 60% a 80% dos dados; máximo e mínimo representam os dados extremos superiores e inferiores e a amplitude é dada pela diferença entre eles, identificando o intervalo total dos dados coletados; o coeficiente de variação de Pearson define a representatividade da média com base na dispersão relativa dos dados em que valores menores que 15% indicam baixa dispersão, entre 15% e 30% indicam média dispersão e iguais ou acima de 30% indicam elevada dispersão (Farber & Larson, 2012; Martins & Domingos, 2011). Quanto às análises inferências, foram utilizados os testes de hipótese para normalidade de Kolmogorov-Smirnov e Shapiro-Wilke, teste de amostras pareadas de Wilcoxon (Farber & Larson, 2012; Martins & Domingos, 2011; Mattar, 1995).

Passo 3: Finalmente, em resposta ao terceiro objetivo específico, cálculo da correlação entre as variáveis estudadas, nível de reação e nível de aprendizagem, foi utilizada a correlação de Spearman, cuja aplicação não exige que os dados sigam a distribuição normal. Segundo Farber e Larson (2012) o coeficiente de correlação pode variar de -1 a 1, sendo que quando mais próximo de 1 maior a correlação positiva e quanto mais próximo de -1, maior a correlação negativa entre as variáveis. Se não há correlação ou há uma correlação muito fraca, o coeficiente está próximo de 0. Complementando, Martins e Domingos (2011) definem os parâmetros do coeficiente de correlação como: 0,9 e 1,0 = ótima; 0,8 e 0,9 = boa; 0,6 e 0,8 = média; 0,4 e 0,6 = baixa e 0,0 a 0,4 péssima; 0,0 = não há. A mesma interpretação deve ser dada aos coeficientes negativos.

5 Discussão de resultados

A apresentação dos resultados segue os Passos 1, 2 e 3, especificado na metodologia e detalhados aqui, respondendo respectivamente ao primeiro objetivo específico (validação do instrumento de avaliação de reação), o segundo (apuração dos níveis de reação e aprendizagem) e o terceiro (cálculo da correlação entre os níveis reação e aprendizagem).

Para compor a base de dados deste estudo foram analisados três documentos: Avaliação de Reação (DOC 1), Avaliação de Aprendizagem Inicial (DOC 2) e Final, (DOC 3).

Quanto ao perfil da amostra, averiguou-se que 45% dos sujeitos estão na faixa de 41 e 60 anos, 41% entre 26 e 40 anos, 4% acima de 60 anos, 1% entre 18 e 25 anos e 9% não responderam. Faixa etária, pois, ampla. Intervalos entre 26 e 40 anos e 41 e 60 anos são as que concentram 86% dos treinandos, equitativamente. Idade mínima: 19 anos e idade máxima: 67 anos. Todos são do sexo



masculino e exercem atividade de operação de pá carregadeira. Escolaridade mínima: ensino fundamental incompleto e máxima de ensino técnico.

Iniciando (Passo 1/Primeiro Objetivo Específico) pela validação do instrumento de avaliação de reação em uso na empresa pesquisada, a partir da aplicação da técnica de validação de constructo e da análise de confiabilidade. (MacDaniel & Gates, 2003; Malhotra, Rocha, Laudisio, Altherman & Borges, 2009; Mattar, 1995), buscou-se averiguar o quanto o mesmo mostra paridade com o modelo teórico de referência, ou seja, os itens abordados e a escala utilizada devem ser justificados pelos autores que defendem a teoria relacionada ao constructo que está se medindo.

A Figura 1 apresenta o modelo de avaliação de reação utilizado pela empresa estudada. Sua estrutura contempla as variáveis: infraestrutura, instrução e recursos didáticos, com escala de quatro graus de reação cada um. O último aspecto, satisfação geral, além de possuir escala com 3 graus de reação, tem a opção para que o treinando deixe seus comentários e sugestões.

Figura 1

Avaliação de reação da empresa estudada

[opcional] Nome:				
Avaliação – Reação				
Infraestrutura				
As instalações, sala, ambiente, cortesia, atendimento.	OTIMO	BOM	ATENDE	RUIM
	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Instrutor:				
Domínio do conteúdo do treinamento:	OTIMO	BOM	ATENDE	RUIM
	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Facilidade em repassar o conhecimento:	OTIMO	BOM	ATENDE	RUIM
	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Recursos Didáticos:				
Material eletrônico:	OTIMO	BOM	ATENDE	RUIM
	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Apostila, pasta:	OTIMO	BOM	ATENDE	RUIM
	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Satisfação geral:				
O curso atendeu sua necessidade:	Superou	Atendeu	Não atendeu	
	<input type="text"/>	<input type="text"/>	<input type="text"/>	
Deixe sua sugestão:				

Fonte: Dados de pesquisa, 2016.

Através da Figura 1 é possível identificar que foram utilizados sorrisos, representando cada grau de satisfação. O questionário contém uma estrutura simples, compatível com a escolaridade dos respondentes e agiliza o preenchimento, condições que aumentam a confiabilidade dos resultados. Quanto à escala, apresentam números de opções equilibradas, evitando tendência central, assegurando objetividade ao instrumento. Donde se conclui que o instrumento de avaliação de reação utilizado na empresa estudada corresponde aos aspectos mencionados por Palmeira (2013), Gil (2009) e Kirkpatrick (2006), adquirindo assim a validade de constructo em uma das variáveis consideradas.

Contudo, dando mais consistência à validade, o instrumento submetido à análise de confiabilidade, a partir do método de consistência interna, representada pelo Alfa Cronbach. Valores



desejados estão entre 0,7 e 0,9 (MacDaniel & Gates, 2003; Malhotra, Rocha, Laudisio, Altherman & Borges, 2009; Mattar, 1995). Índice obtido: 0,73. Apurou-se, ainda, a correlação média inter-item (parâmetro: 0,15 e 0,50) (Clark & Warson, 1995 *apud* Welter, 2007), expressa na Tabela 1 e na Tabela 2 o resultado de item-total.

A partir da matriz de correlação (Tabela 1 adiante) identifica-se o grau de associação entre os itens do questionário. Todos os coeficientes são maiores do que zero, constatando a existência de correlação positiva, mas não uma correlação significativa. A maior correlação encontrada foi de 0,5 e a menor correlação foi de 0,2. A infraestrutura foi o item que mais se diferenciou de todos os demais.

Assim, os índices de correlação entre os itens revelam que os itens estão dentro do intervalo desejado e não comprometem a validade do instrumento. Portanto, pode-se dizer que cada item está mensurando uma variável distinta da reação e não há, pois, redundância. Mesmo os itens mais correlacionados mostram uma identidade de avaliação distinta.

A correlação de item total corrigida (Tabela 2 adiante) indica o grau em que cada item está relacionado com o alfa global, não sendo aceitáveis valores inferiores a 0,3. Adicionalmente, se a exclusão de um item resultar em um valor de alfa maior que o valor global, este deve ser eliminado para aumentar a confiabilidade (Field, 2005).

Tabela 1

Matriz de correlação entre os itens

	<i>Infraestrutura - As instalações, sala, ambiente, cortesia, atendimento</i>	<i>Instrutor- Domínio do conteúdo do treinamento</i>	<i>Instrutor - Facilidade em repassar o conhecimento</i>	<i>Recursos Didáticos - Material eletrônico</i>	<i>Recursos Didáticos - Apostila, pasta</i>	<i>Satisfação Geral - O curso atendeu sua necessidade</i>
<i>Infraestrutura - As instalações, sala, ambiente, cortesia, atendimento</i>	1,0	0,2	0,2	0,4	0,2	0,2
<i>Instrutor- Domínio do conteúdo do treinamento</i>	0,2	1,0	0,5	0,5	0,4	0,3
<i>Instrutor - Facilidade em repassar o conhecimento</i>	0,2	0,5	1,0	0,4	0,4	0,3
<i>Recursos Didáticos - Material eletrônico</i>	0,4	0,5	0,4	1,0	0,5	0,3
<i>Recursos Didáticos - Apostila, pasta</i>	0,2	0,4	0,4	0,5	1,0	0,3
<i>Satisfação Geral - O curso atendeu sua necessidade</i>	0,2	0,3	0,3	0,3	0,3	1,0

Fonte: Dados de pesquisa, 2016.

Observando a Tabela 2 (adiante), conclui-se que todos os itens se correlacionam positivamente com o alfa global e apresentam valores iguais ou superiores a 0,3, com atenção especial aos itens de infraestrutura e satisfação geral que exibiram menor correlação com a escala total, sendo eles de 0,35 e 0,34 respectivamente. O único item que apresentou valor de alfa se eliminado acima do alfa global foi infraestrutura, com 0,74. Entretanto, o nível de alteração do coeficiente alfa é consideravelmente pequeno, não acarretando em uma alteração significativa na confiabilidade da escala, assim, não se faz necessária a retirada do item do questionário. Pode-se afirmar também que a presença de correlação dos itens com a escala de alfa global indica que eles estão medindo o mesmo constructo.



Tabela 2

Estatísticas de item-total

	Correlação de item total corrigida	Alfa de Cronbach se o item for excluído
Infraestrutura - As instalações, sala, ambiente, cortesia, atendimento	0,35	0,74
Instrutor- Domínio do conteúdo do treinamento	0,55	0,69
Instrutor - Facilidade em repassar o conhecimento	0,51	0,68
Recursos Didáticos - Material eletrônico	0,64	0,63
Recursos Didáticos - Apostila, pasta	0,53	0,67
Satisfação Geral - O curso atendeu sua necessidade	0,34	0,72

Fonte: Dados de pesquisa, 2016.

Finalmente, a Tabela 3, apresenta o resultado das estatísticas básicas do item.

Tabela 3

Estatísticas básicas do item

	Média	Mínimo	Máximo	Amplitude	Variância	N de itens
Médias de item	3,6	2,8	3,9	1,2	0,19	6

Fonte: Dados de pesquisa, 2016.

Na Tabela 3 é possível identificar as médias estatísticas básicas dos itens. A pontuação média foi de 3,6 e a variância de 0,19. O mínimo médio encontrado foi de 2,8 e máximo de 3,9, o que caracterizou amplitude de 1,2. Os dados estão uniformemente distribuídos, com baixa dispersão relação à média.

Pode-se abstrair, em síntese, que o questionário de reação revela validade de constructo, corroborando a teórica. O instrumento correspondeu positivamente ao teste de confiabilidade, embasando os resultados obtidos. Entretanto, é possível fazer melhorias no instrumento para alcançar maiores níveis de confiabilidade, como por exemplo, aumentar o número de questões e utilizar uma escala de níveis mais ampla.

No Passo 2/Segundo Objetivo Específico (apuração do índice de reação e de aprendizagem), foi realizada a apuração dos resultados de avaliação do treinamento, objetivando determinar o nível de

reação e o nível de aprendizagem, a partir de parâmetros descritivos e inferenciais. Aplicou-se o teste de normalidade *Kolmogorov-Smirnov* e *Shapiro-Wilk*.

Considerou-se de início a normalidade de distribuição dos dados. Os testes de aderência foram aplicados com o auxílio do *software* SPSS. Os resultados do teste de aderência são apresentados na Tabela 4.

Tabela 4

Resultado do teste de aderência à normalidade

Teste de Normalidade	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estatística	Df	Sig.	Estatística	Df	Sig.
Infraestrutura - As instalações, sala, ambiente, cortesia, atendimento	0,42	260	0,00	0,63	260	0,00
Instrutor- Domínio do conteúdo do treinamento	0,54	260	0,00	0,28	260	0,00
Instrutor - Facilidade em repassar o conhecimento	0,50	260	0,00	0,46	260	0,00
Recursos Didáticos - Material eletrônico	0,47	260	0,00	0,53	260	0,00
Recursos Didáticos - Apostila, pasta	0,49	260	0,00	0,48	260	0,00
Satisfação Geral - O curso atendeu sua necessidade	0,47	260	0,00	0,54	260	0,00
Nota inicial	0,05	260	0,06	0,99	260	0,49
Nota final	0,12	260	0,00	0,94	260	0,00

Fonte: Dados de pesquisa, 2016.

Segundo Torman, Birck e Riboldi (2012), se o valor de significância do teste for menor que o nível de significância escolhido, os dados não seguem uma distribuição normal. Considerando um nível de significância de 5% (0,05), através da Tabela 4 percebe-se que apenas a amostra de Nota Inicial não rejeita a hipótese de normalidade para a variável com distribuição normal. Para todos os demais itens, ambos os resultados de *Kolmogorov-Smirnov* e *Shapiro-Wilk* rejeitam esta hipótese nula, indicando que os dados não seguem uma distribuição normal. Sendo assim, neste estudo, só deverão ser aplicados testes inferenciais cujos requisitos não exijam que os dados estejam distribuídos normalmente, os testes não paramétricos.

A reação ao treinamento foi mensurada através de cinco itens por uma escala tipo Likert de quatro pontos equilibrados, sendo que quanto mais próximo de quatro, mais favorável foi a reação do treinando, e quanto mais próximo de um, menos favorável foi a reação. O questionário conta ainda com um sexto item, mensurado por uma escala tipo Likert de três pontos, sendo que quanto mais próximo



de três, mais favorável foi a reação do treinando, e quanto mais próximo de um, menos favorável foi a reação

Os resultados obtidos através do DOC 1, sobre a reação dos empregados ao treinamento em termos de valores médios, coeficiente de variação, desvio padrão, mínimo, máximo e amplitude são apresentados na Tabela 5.

Tabela 5

Estatística descritiva da avaliação de reação

Estatísticas descritivas	Média	Coeficiente de variação (%)	Desvio padrão	Min	Max	Amplitude
Infraestrutura - as instalações, sala, ambiente, cortesia, atendimento	3,6	17	0,6	1	4	3
Instrutor- domínio do conteúdo do treinamento	3,9	8	0,3	2	4	2
Instrutor - facilidade em repassar o conhecimento	3,8	10	0,4	2	4	2
Recursos didáticos - material eletrônico	3,8	13	0,5	1	4	3
Recursos didáticos - apostila, pasta	3,8	10	0,4	1	4	3
Satisfação geral - o curso atendeu sua necessidade	2,8	14	0,4	2	3	1
Média geral	3,6	11	0,4	1,5	3,8	2,3

Fonte: Dados de pesquisa, 2016.

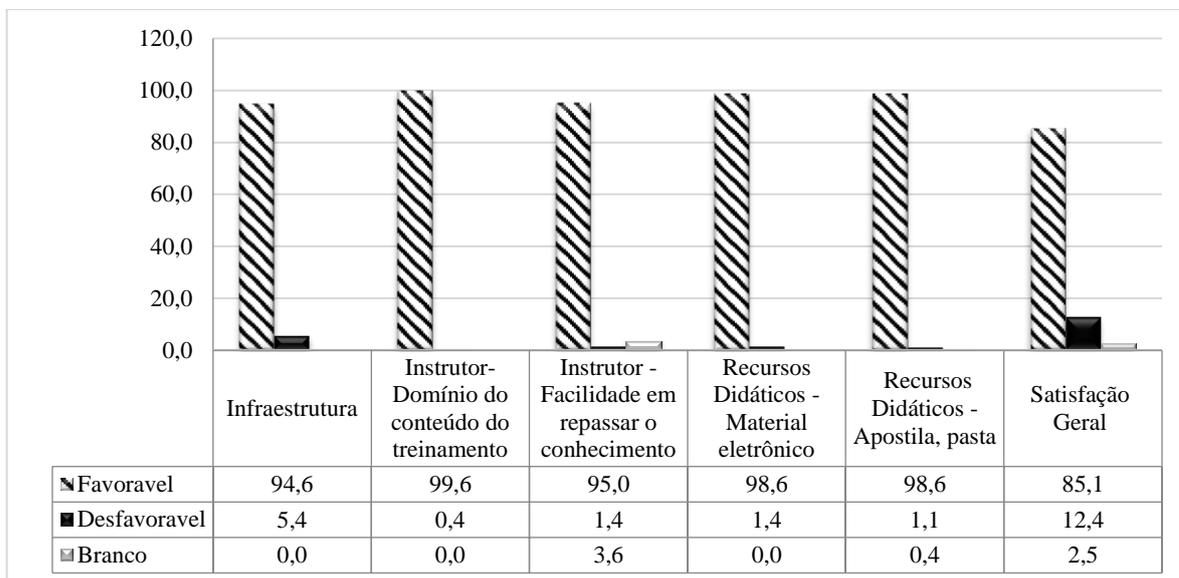
A reação ao treinamento (Tabela 5) mostrou-se bastante positiva, apresentando nota média geral de 3,6. O domínio do instrutor sobre o conteúdo do treinamento apresentou a melhor avaliação, com a média de 3,9, aproximando-se mais da nota ótima. Já a infraestrutura foi o item que mais sofreu restrições, obtendo a média que mais se afastou da escala ótima, com nota de 3,6. Conclui-se que, em geral, os treinandos apresentaram satisfação alta para os itens avaliados (valores de 3,6 a 3,9 para os itens com 4 graus e de 2,8 para o item com 3 graus). O Gráfico 1 apresenta as médias da avaliação de reação de cada dimensão.

A satisfação geral, especificamente, o indicador de favorabilidade compreende o número de respostas positivas (3) somadas à metade da resposta neutra (2) sobre o número total de respostas, enquanto que o indicador desfavorável, compreende o número de respostas negativas (1) somadas a metade da resposta neutra (2), sobre o número total de respostas.

O Gráfico 1 apresenta o grau de favorabilidade para cada item de satisfação.

Gráfico 1

Indicadores de favorabilidade



Fonte: Dados de pesquisa, 2016.

Conclui-se a partir do Gráfico 1 que todos os itens apresentaram altos índices de favorabilidade, entre 85,1% para satisfação geral e 99,6% para o domínio do conteúdo do instrutor.

Os resultados concordam com as premissas de Kirkpatrick e Kirkpatrick (2006), afirmando que a aceitação positiva sobre o treinamento é o primeiro indício de sua efetividade. Entretanto, existem alguns aspectos a serem melhorados, conforme parâmetros de Palmeira (2013) e Gil (2009), qual seja: infraestrutura. Observando os resultados de satisfação geral, há um potencial de melhoria de 12,4%.

Seguindo o método de avaliação de Kirkpatrick e Kirkpatrick (2006), após realizar a avaliação de reação, deve ser avaliada a aprendizagem, considerando ainda o Passo 2/Segundo Objetivo Específico. Os resultados obtidos através dos documentos DOC 2 (pré, nota inicial) e DOC 3 (pós-teste, nota final), sobre a aprendizagem, estão expostos na Tabela 6.

Considerou-se para avaliar a aprendizagem dos participantes do programa de treinamento o que o aluno assimilou, isto é, a diferença entre nota final e nota inicial (Gil, 2009; Kirkpatrick & Kirkpatrick, 2006; Palmeira, 2013).



Tabela 6

Estatística descritiva das avaliações de aprendizagem

Estatísticas descritivas	Média	Coefficiente de Variação	Desvio padrão	Min	Max	Amplitude
Nota inicial	53,8	28,6	15,4	6	90	84
Nota final	92,6	6	5,6	75	100	25
Diferença de teste inicial e teste final	38,8	39,4	15,3	-2	82	84

Fonte: Dados de pesquisa, 2016.

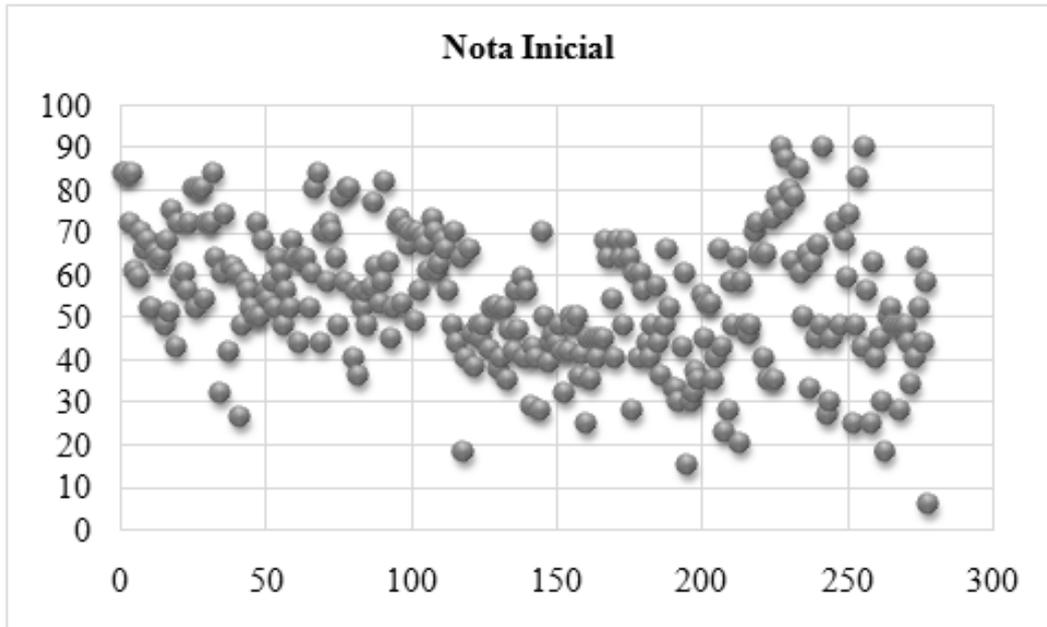
O valor médio obtido por meio da diferença das notas final e inicial dos participantes foi de 38,8, o que significa que os treinandos responderam mais eficazmente às provas depois que participaram dos treinamentos. A nota média antes do curso foi 53,8, enquanto a nota média após o curso foi 92,6. Salto equivalente a 72%, aproximadamente, em aprendizagem. O resultado vai ao encontro das assertivas de Kirkpatrick e Kirkpatrick (2006).

A representatividade da média para a população pode ser observada através das medidas de dispersão: máximo, mínimo, amplitude, desvio padrão e coeficiente de variação. Apesar da média das diferenças entre DOC 3 e DOC 2 ter sido avaliada em 38,8, o valor máximo atingiu a pontuação de 82 em ganho de aprendizagem, bastante significativo. A amplitude de 84 pontos permaneceu inalterada em relação ao teste inicial, o que pode explicar a tendência de nivelamento de conhecimento dos empregados. Os indicadores seguintes explicam esta asserção.

O desvio padrão foi de 15,4 para nota inicial; 5,6 para nota final e 15,3 para diferença entre as duas notas. As notas do pré-teste se encontram mais dispersas do que as notas do teste final, quase três vezes menor. Considerando que as notas do pós-teste estavam praticamente niveladas, isso explica o desvio padrão da diferença entre as notas estar muito semelhante ao encontrado na nota inicial. Ainda sobre a variabilidade, o coeficiente de variação para a nota inicial foi de 6, indicando maior concentração dos dados em torno da média. Em contrapartida a nota final foi 28,6 e a diferença entre as notas foi 39,4, confirmando alta dispersão dos dados e pouca representatividade da média.

A seguir, os gráficos de dispersão das notas obtidas no pré-teste e no pós-teste consolidam os resultados obtidos.

Gráfico 2

Dispersão das notas de pré-teste

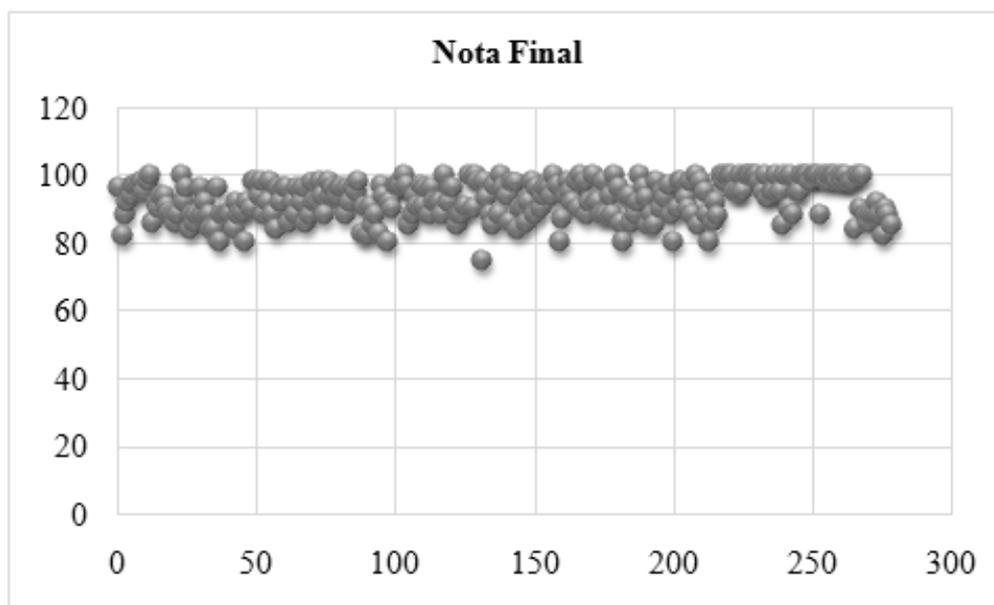
Fonte: Dados de pesquisa, 2016.

No primeiro momento (Graf. 2), é possível perceber que há uma grande dispersão das notas, ou seja, mínimo de 6 e máxima de 90 pontos, confirmando uma amplitude acentuada de 84 pontos, conforme apresentado na Tabela 6.



Gráfico 3

Dispersão de notas de pós-teste



Fonte: Dados de pesquisa, 2016.

Após a realização do treinamento (Graf. 3), observou-se claramente que houve uma tendência de nivelamento de conhecimento dentre os empregados. A dispersão é notoriamente menor, concentrada na região: nota mínima = a 75 e máxima = 100. A nova amplitude alcançada foi de 25 pontos (30% menor).

O teste de hipótese Wilcoxon para amostras pareadas (amostras: DOC 2 e DOC 3), segundo Farber e Larson (2012), trata-se de teste não paramétrico usado para comparar as diferenças entre duas amostras dependentes.

Adaptando-se aos objetivos dessa pesquisa, as hipóteses podem ser traduzidas como: H_0 : as notas de aprendizagem não se alteram após o treinamento; H_1 : as notas de aprendizagem se alteram após o treinamento.

Inicialmente, o teste apresenta uma tabela de classificações (*ranks*) para a quantidade de dados que apresentou redução, elevação ou se mantiveram inalterados após a aplicação do pós-teste.

Tabela 7

Resultado da classificação do teste de Wilcoxon

Classificações		N	Mean Rank	Sum of Ranks
Nota final -	Negative Ranks	1 ^a	1,5	1,5
Nota inicial	Positive Ranks	275 ^b	139	38224,5
	Ties	0 ^c		
	Total	276		

a. Nota final < Nota inicial

b. Nota final > Nota inicial

c. Nota final = Nota inicial

Fonte: Dados de pesquisa, 2016.

Na Tabela 7 verifica-se que 275 sujeitos obtiveram ganho de aprendizagem (nota final > nota inicial). Em apenas um caso o resultado se inverteu. Assim, se sig. $\geq 0,05$ a hipótese nula não deve ser rejeitada, assumindo igualdade entre as amostras e se sig. $< 0,05$ rejeita-se a hipótese de igualdade entre as amostras (Farber & Larson, 2102; Martins & Domingues, 2011; Mattar, 1995).

Na Tabela 8 o resultado da estatística do teste e a probabilidade de significância.

Tabela 8

Estatístico do teste de Wilcoxon

Estatística do teste ^a	
	Nota final - Nota inicial
Z	-14,401 ^b
Sig. Assint. (2 caudas)	0

a. Wilcoxon Signed Ranks Test

b. Com base em postos negativos.

Fonte: Dados de pesquisa, 2016.

Os resultados do teste Wilcoxon para amostras pareadas, considerando o pré e o pós-teste, apontaram diferenças significativas entre as notas. Neste caso, como o valor da significância (sig.assint.) ficou abaixo de 5% (0,05), comparado ao nível de significância adotado para esta amostra, rejeita-se a hipótese nula, ou seja, o resultado dos dois testes não podem ser considerados equivalentes. Houve, pois, aprendizagem.

A fim de concluir, resta considerar o cálculo da correlação entre as variáveis estudadas (Passo 3/Terceiro Objetivo Específico. Relembrando, coeficiente de correlação varia de -1 a 1, sendo correlação

alta entre 0,9 e 1,0; boa entre 0,8 e 0,9; média entre 0,6 e 0,8; baixa entre 0,4 e 0,6 e ruim entre 0,0 e 0,4. Índices negativos recebem a mesma interpretação (Larson & Farber, 2012).

A Tabela 9 expressa os resultados da correlação de Spearman. Devido ao tamanho, a tabela foi reduzida, condensando as informações principais para análise deste objetivo específico.

Tabela 9

Matriz de correlação de Spearman

	Descritores	Nota final	Diferença de nota final e inicial
Infraestrutura - As instalações, sala, ambiente, cortesia, atendimento	Correlações de coeficiente	0,01	0,13*
	Sig. (2 extremidades)	0,82	0,02
	N	277	278
Instrutor- Domínio do conteúdo do treinamento	Correlações de coeficiente	-0,06	-0,04
	Sig. (2 extremidades)	0,33	0,53
	N	277	278
Instrutor - Facilidade em repassar o conhecimento	Correlações de coeficiente	0,03	0,08
	Sig. (2 extremidades)	0,59	0,18
	N	267	268
Recursos Didáticos - Material eletrônico	Correlações de coeficiente	0,08	-0,00
	Sig. (2 extremidades)	0,19	0,95
	N	277	278
Recursos Didáticos - Apostila, pasta	Correlações de coeficiente	0,13*	0,01
	Sig. (2 extremidades)	0,03	0,81
	N	276	277
Satisfação Geral - O curso atendeu sua necessidade	Correlações de coeficiente	0,11	0,01
	Sig. (2 extremidades)	0,07	0,80
	N	270	271
Média de satisfação	Correlações de coeficiente	0,08	0,03
	Sig. (2 extremidades)	0,16	0,62
	N	277	278

** . A correlação é significativa no nível 0,01 (2 extremidades)

* . A correlação é significativa no nível 0,05 (2 extremidades)

Fonte: Dados de pesquisa, 2016.

Na Tabela 10 encontram-se os coeficientes de correlação bivariada das variáveis de aprendizagem, definida como as notas obtidas por 278 participantes no pré-teste, pós-teste e a diferença entre as duas medidas. O nível de reação encontra-se representado pelos 6 aspectos avaliado pelos empregados no questionário de satisfação: Infraestrutura - as instalações, sala, ambiente, cortesia, atendimento; Instrutor - facilidade em repassar o conhecimento; Instrutor- domínio do

conteúdo do treinamento; Recursos Didáticos - material eletrônico; Recursos Didáticos - apostila, pasta; Satisfação Geral - o curso atendeu sua necessidade e

Ainda na Tabela 10 é possível identificar o coeficiente de correlação entre as variáveis, a significância obtida e o número de dados N que foram utilizados. Os resultados do coeficiente de correlação das variáveis de reação e nota inicial variaram entre -0,06 e 0,13; em relação à diferença de médias variaram de -0,04 e 0,13. Conclui-se que em sua maioria, os resultados estão mais próximos de zero, indicando ausência de correlação entre reação e aprendizagem ou péssima correlação.

Destaque pode ser dado aos resultados de “Recursos Didáticos - apostila, pasta” e “Nota final e “Infraestrutura” e “Diferença de médias”, pois foram os únicos itens a apresentarem correlação positiva e significativa de 0,13, entretanto, o índice não é satisfatório. Informações que corroboram Abbad, Gama e Borges-Andrade (2000) e Santos (2012), para os quais não é possível comprovar a causalidade da relação entre reação e aprendizagem, como ainda relativamente aos resultados apurados Bastos (2012), Alvarce (2014), Mourão e Marins (2009).

Não se confirmam, então, as premissas do modelo dos quatro níveis de Kirkpatrick (2006) na amostra em questão. Independente do grau de satisfação, os empregados alcançaram altos índices de aprendizagem, tanto na nota final, quanto na diferença entre pós-teste e pré-teste. Não há um padrão nas pesquisas que estudam o relacionamento entre as variáveis, tanto quanto ao público e que resultados são esperados em treinamento operacional, o que torna difícil e complexa a compreensão do fenômeno, requerendo, assim, que novas pesquisas, com esta proposição, avancem.

6 Considerações e conclusões

O modelo de quatro níveis interdependentes, de Kirkpatrick e Kirkpatrick (2006), inobstante controvérsias relativas a inconsistências internas, continua sendo o mais difundido e utilizado no Brasil, sem que se identifiquem pesquisas acadêmicas que confirmem seus postulados. A lacuna existente entre o uso corrente (aplicação prática) e pesquisas acadêmicas que confirmem ou refutam os postulados que dão fundamento ao modelo em questão é que serviu de mote para o presente estudo.

Buscar elucidar a natureza da correlação das variáveis de reação e de aprendizagem, pretendeu contribuir para que se critique o uso do modelo de avaliação de eficácia de treinamento. Atende, pois, ao pressuposto do interesse social desta questão, tanto quanto daqueles dos próprios trabalhadores em específico e, ainda relevantemente ao estudo acadêmico, no âmbito da gestão da capacitação de pessoas ou da avaliação organizacional.

Sinteticamente o estudo apresenta resultados concernentes a: 1) comprova a validade do instrumento de avaliação de reação parametrizado e preconizado na literatura especializada; 2) Os dados igualmente apontaram grau de favorabilidade de 94,6% para “infraestrutura”; 99,6% para



“Instrutor - domínio do conteúdo do treinamento”; 95% para “Instrutor – facilidade em repassar o conhecimento”; 98,6% para “Recursos didáticos – material eletrônico”; 98,6% para “Recursos didáticos – apostila, pasta” e 85,1% para satisfação geral, como indicadores válidos e confiáveis à avaliação do nível de reação; 3) As análises do pré-teste e pós-teste revelaram efetividade e ganho de aprendizagem correspondente a 72% em treinamento operacional, contribuindo para o alcance de nivelamento de conhecimento; 4) Contudo, os índices apurados no cálculo de correlação existente entre as variáveis de reação e aprendizagem indica ausência ou correlação muito fraca, contrariando a proposição de Kirkpatrick e Kirkpatrick (2006), de que os quatro níveis possuem entre si um forte relacionamento.

Às organizações em geral, tais dados põem em questão o uso indiscriminado e acrítico do modelo de quatro níveis para apurar a eficácia de treinamento e a adoção do mesmo como efetivo na avaliação organizacional, muito embora considerou-se somente os dois primeiros níveis iniciais. Minimamente alerta que merece maior atenção e pesquisas que atestem sua consistência interna e seus postulados. Se se confirma por um lado a eficácia, em separado, de um e outro nível, põe em dúvida a interdependência dos mesmos.

Aos trabalhadores o estudo traz um contributo no sentido de que se repense as práticas e as políticas organizacionais atinentes ao desenvolvimento de programas de capacitação em nível operacional, para que aqueles não sejam submetidos a ações de formação e treinamento que não redundem em ganhos de aprendizagem e otimização de desempenho, minimizando desgastes.

Para a academia, o estudo aponta a lacuna existente entre a prática e a teorização acerca do modelo de quatro níveis de avaliação de eficácia organizacional. Modelo este proposto nos anos 1950 e amplamente aplicado nas organizações contemporâneas, sem que se tenha investido em pesquisas de atualização, ajuste contextual e circunstancial de seus pressupostos.

Em face disso, novas pesquisas que levem em conta os quatro níveis e outros postulados advindos das proposição de Kirkpatrick e Kirkpatrick (2006) são bem vindas. Igualmente outros recortes de pesquisas futuras parecem pertinentes, como por exemplo o comportamento das mesmas variáveis em distintos níveis de treinamento, para além do operacional, tanto quanto em outros setores empresariais, quer se considere o porte ou segmento de atuação.

Referências

- Abbad, A. L. G. (1999). *Um modelo integrado de avaliação do impacto do treinamento no trabalho-IMPACT*. Instituto de psicologia. Brasília: Universidade de Brasília, 1999.
- Abbad; A. L. G.; Borges-Andrade, J. E. (2000). *Treinamento: análise do relacionamento da avaliação nos níveis de reação, aprendizagem e impacto no trabalho*. Revista de Administração Contemporânea.

- Alvarce, D. C. (2014). *Desenvolvimento e avaliação da reação, aprendizagem e impacto de treinamento online para profissionais da saúde*. 263 f. Tese (Doutorado) - Curso de Enfermagem, Universidade de São Paulo, São Paulo.
- Associação Brasileira de Normas Técnicas – ABNT (Brasil) (2001). *NBR ISO 10015: gestão da qualidade - diretrizes para treinamento*. Rio de Janeiro: Copyright.
- Associação Brasileira de Engenharia de Produção - ABEPRO (Brasil). (n.d.). *Um panorama da engenharia de produção*. Disponível em: <http://www.abepro.org.br/interna.asp?ss=1&c=924> . Acesso em: 30 maio 2016.
- Associação Brasileira de Treinamento e Desenvolvimento – ABTD (Brasil). (2016). *Panorama do treinamento no Brasil: fatos, indicadores, tendências e análises*. Disponível em: <http://www.integracao.com.br/pesquisa-panorama-do-treinamento-no-brasil-2015.pdf> . Acesso em: 09 mai. 2016.
- Barros, A. J. S.; Lehfel, N. A. S. (2007). *Metodologia Científica*. 3. ed. São Paulo: Pearson Prentice Hall.
- Bastos, L. F. L. (2012). *Avaliação de reação, aprendizagem e impacto de treinamento em um hospital do município de São Paulo*. 2012. 142 f. Dissertação (Mestrado) - Curso de Enfermagem, Escola de Enfermagem da Universidade de São Paulo, São Paulo.
- Bastos, O. P. M. Diagnóstico e avaliação de T&D: processo de T&D. In: BOOG, Gustavo G. (Coord). (1984). *Manual de treinamento e desenvolvimento ABTD*. São Paulo: Makron Books.
- Bongani, P. S. Treinamento técnico-operacional. In: Boog, G. (Coord). (1995). *Manual de treinamento e desenvolvimento ABTD*. 2. ed. São Paulo: Makron Books.
- Boog, G. G.; BOOG, M. (Coord). (2013). *Manual de treinamento e desenvolvimento: processos e operações*. São Paulo: Pearson.
- Borges-Andrade, Jairo Eduardo. (1982). *Avaliação somativa de sistemas instrucionais: integração de três propostas*. Tecnologia Educacional.
- Borges-Andrade, Jairo Eduardo. (2002). *Desenvolvimento de medidas em avaliação de treinamento*. Estudos de Psicologia.
- Decenzo, D.; Robbins, S. (2001). *Administração de recursos humanos*. Rio de Janeiro: LTC.
- Decenzo, D.; Robbins, S. (2004). *Fundamento de administração: conceitos essenciais e aplicação*. São Paulo: Pearson Prentice Hall.
- Desseler, G. (2014). *Administração de recursos humanos*. 3. ed. São Paulo: Pearson Education do Brasil.
- Field, A. (2013). *Descobrimo a estatística usando o SPSS*. 2. ed. Porto Alegre: Artmed Editora.
- Gil, A. C. (2002). *Como elaborar projetos de pesquisa*. 4. ed. São Paulo: Atlas.
- Gil, A. C. (2008) *Métodos e Técnicas de Pesquisa Social*. 6. ed. São Paulo: Atlas.
- Gil, A. C. (2009) *Gestão de pessoas: enfoque nos papéis profissionais*. São Paulo: Atlas.



- Kauark, F. S.; Manhaes, F. C.; Medeiros, C. H. (2010). *Metodologia da Pesquisa: um guia prático*. Itabuna: Via Litterarum.
- Kirkpatrick, D. L.; Kirkpatrick, J. D. (2006). *Evaluating training programs: the four levels*. 3. ed. San Francisco: Berrett-koeherPublishers.
- Kirkpatrick Partners (2016). *The official site of Kirkpatrick model*. Disponível em: <http://www.kirkpatrickpartners.com/Home/tabid/38/Default.aspx>
- Larson, R.; Farber, B. (2012). *Estatística aplicada*. 4. ed. São Paulo: Pearson Prentice Hall.
- Malhotra, N. K. (2011). *Pesquisa de Marketing: foco na decisão*. 3. ed. São Paulo: Pearson Prentice Hall.
- Malhotra, N. K.; Rocha, I.; Laudisio, M. C.; Altherman, É.; Borges, F. M. (2009). *Introdução à pesquisa de marketing*. São Paulo: Pearson Prentice Hall.
- Martins, G. A.; Domingues, O. (2011). *Estatística geral e aplicada*. 4. ed. São Paulo: Editora Atlas.
- Mascarenhas, S. A. (2012). *Metodologia Científica*. São Paulo: Pearson.
- Mattar, F. N. (1995) *Pesquisa de marketing*. 2. ed. São Paulo: Editora Atlas.
- Mattar, F. N. (2011). *Pesquisa de marketing*. 4. ed. – 3. São Paulo: Atlas.
- Milkovich, G. T.; Boudreau, J. W. (2009). *Administração de recursos humanos*. São Paulo: Atlas.
- Mourão, L; Marins, J. (2009). *Avaliação de treinamento e desenvolvimento nas organizações: resultados relativos ao nível de aprendizagem*. Revista Psicologia: Organizações e Trabalho.
- Palmeira, C. (2013). Avaliação de resultados - retorno do investimento. In: BOOG, Gustavo G.; BOOG, Magdalena (Coord). (2013). *Manual de treinamento e desenvolvimento: Processos e operações*. São Paulo: Pearson Education do Brasil, Cap. 3.
- Scorsolini-Comin, F.; Inocente, D. F.; Miura, I. K. (2011). *Avaliação de programas de treinamento, desenvolvimento e educação no contexto organizacional: Modelos e perspectivas*. Revista Psicologia: organizações e trabalho.
- Torman, V. B. L.; Birck, A. R.; Riboldi, J. (2012). *Normalidade de variáveis: métodos de verificação e comparação de alguns testes não-paramétricos por simulação*. Revista HCPA, Porto Alegre, v. 32, n. 2, p.227-234, ago.
- Welter, G. M. R. (2007). *Human Guide: evidência de validade da versão brasileira*. 2007. 172 f. Dissertação (Mestrado) - Curso de Psicologia, Universidade São Francisco, Itatiba.