UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE CONTABILIDADE E ATUÁRIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CONTROLADORIA E CONTABILIDADE

RACIOCÍNIO CRÍTICO DE ALUNOS DE GRADUAÇÃO EM CIÊNCIAS
CONTÁBEIS: APLICAÇÃO DO MODELO INSTRUCIONAL DE RICHARD PAUL

Doutorando: Ivan Carlin Passos
Orientador: Prof. Dr. Edgard Bruno Cornachione Júnior

SÃO PAULO
2011

Prof. Dr. João Grandino Rodas
Reitor da Universidade de São Paulo

Prof. Dr. Reinaldo Guerreiro
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Edgard Bruno Cornachione Júnior
Chefe do Departamento de Contabilidade e Atuária

Prof. Dr. Luis Eduardo Afonso
Coordenador do Programa de Pós-Graduação em Controladoria e Contabilidade

IVAN CARLIN PASSOS

**RACIOCÍNIO CRÍTICO DE ALUNOS DE GRADUAÇÃO EM CIÊNCIAS
CONTÁBEIS: APLICAÇÃO DO MODELO INSTRUCIONAL DE RICHARD PAUL**

Tese apresentada ao Departamento de Contabilidade e Atuária da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito para a obtenção do título de Doutor em Ciências Contábeis.

Orientador: Prof. Dr. Edgard Bruno Cornachione Júnior

Versão Corrigida

(versão original disponível na Unidade que aloja o Programa)

SÃO PAULO

2011

FICHA CATALOGRÁFICA

Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Passos, Ivan Carlin

Raciocínio crítico de alunos de graduação em Ciências Contábeis:
aplicação do modelo instrucional de Richard Paul / Ivan Carlin Passos.
– São Paulo, 2011.
250 p.

Tese (Doutorado) – Universidade de São Paulo, 2011.
Orientador: Edgard Bruno Cornachione Júnior.

1. Contabilidade – Estudo e ensino 2. Raciocínio 3. Métodos de ensino 4. Estudantes universitários I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

CDD – 657.07

Versão Corrigida

**À minha esposa Neuria, à minha filha Julia
e a toda minha família e amigos.**

Agradeço primeiramente a Deus por ter me dado saúde, força, perseverança e fé durante toda a minha vida.

Agradeço imensamente a minha esposa, filha, família e amigos, por terem me apoiado durante todo o meu Doutorado.

Agradeço aos meus pais João Passos e Cristina Aparecida Carlin Passos que me ensinaram a dar valor ao que somos e não ao que temos.

Agradeço especialmente ao professor Edgard Bruno Cornachione Júnior por ter se mostrado sempre presente em todos os momentos desta pesquisa, como orientador e amigo.

Por suas valiosas críticas, contribuições e sugestões, sou muito grato aos professores Gilberto de Andrade Martins, Alan John Anthony Sangster, Carlos Renato Theophilo e Mauro Maia Laruccia.

Agradeço especialmente a professora Tania Regina Sordi Relvas que consentiu com a pesquisa em suas turmas e muito colaborou com as intervenções e aplicações dos instrumentos de avaliação.

Agradeço os coordenadores de cursos, Gerlando Augusto Sampaio Franco de Lima e Eduardo Gimenes Romero que autorizaram a realização desta pesquisa em seus cursos.

Agradeço imensamente minha irmã Cristiane Carlin Passos por ter me auxiliado nas correções dos instrumentos de avaliação de raciocínio crítico e ao meu irmão Juan Carlin Passos pelo auxílio na construção do abstract.

Agradeço com carinho minha filha Julia e o seu namorado Felipe que me ajudaram na tabulação dos dados da pesquisa demográfica.

Agradeço ao meu tio “Paulinho”, Paulo César Passos, que revisou as traduções das citações originais em inglês.

Agradeço aos amigos Luiz Eduardo Gaio, Nálbia de Araújo Santos, Manoel Raimundo Santana Farias e Kelly Teixeira Rodrigues Farias pelas valiosas críticas e sugestões desde o início desta pesquisa.

Agradeço à professora Gláucia Lucas Ramiros pelo trabalho de revisão textual.

Enfim, agradeço todos os alunos que consentiram em participar desta pesquisa.

**“O único lugar aonde o sucesso vem antes
do trabalho é no dicionário.”**

Albert Einstein

RESUMO

Devido à demanda do mercado e da nova Lei das S.A's que exigem um perfil de profissional da área de negócios com maior nível de raciocínio crítico, o objetivo geral desta tese foi conhecer, analisar e avaliar empiricamente a efetividade do modelo de raciocínio crítico de Richard Paul em alunos de disciplinas da área contábil no desenvolvimento de habilidades para pensar criticamente sobre contabilidade e atitudes de raciocínio crítico em geral. O quase-experimento piloto foi realizado em uma instituição pública durante o segundo semestre de 2010, teve uma abordagem geral (quatro intervenções de 30 min.) de aplicação do modelo e o autor da pesquisa foi monitor das classes. O segundo quase-experimento foi realizado em uma instituição particular durante o primeiro semestre de 2011, teve uma abordagem mista (aulas e atividades da turma experimento) de aplicação e o autor da pesquisa foi o professor das turmas. Em ambos os quase-experimentos os participantes preencheram três instrumentos de avaliação: dois deles aplicados no início e no final do semestre e avaliaram habilidades (*Ennis Weir Critical Thinking Essay Test* – EWCTET) e atitudes (*California Critical Thinking Dispositions Inventory* – CCTDI) de raciocínio crítico; e um deles foi uma pesquisa demográfica (aplicado no final do semestre). Além de análises descritivas das amostras, para testar as hipóteses foi aplicada a análise de covariância (ANCOVA) utilizando-se os pré-testes como covariáveis. Na instituição pública não foi constatado efeito significativo do modelo instrucional no desenvolvimento de habilidades ou atitudes de raciocínio crítico e nenhuma diferença considerando o efeito isolado das variáveis idade ou gênero. Foram confirmadas duas hipóteses na instituição particular: a turma experimento ($p = 0,040$ e $d = 0,56$) e os alunos mais velhos ($p = 0,047$ e $d = 0,58$) apresentaram maior desenvolvimento de habilidades de raciocínio crítico. Apesar de as médias das atitudes não terem sido significativamente diferentes, encontrou-se também efeito significativo do modelo instrucional no desenvolvimento da atitude curiosidade ($p = 0,047$ e $d = 0,65$). Também foram constatadas diferenças significativas nas habilidades para os efeitos isolados de idade ($p = 0,002$ e $d = 0,77$) e gênero ($p = 0,016$ e $d = 0,60$); e para as atitudes no efeito isolado do gênero ($p = 0,033$ e $d = 0,46$). Para futuras pesquisas, recomenda-se a aplicação do modelo em cursos de graduação ou pós-graduação da área de negócios e investigações mais profundas sobre formas de raciocínio, comportamento e aprendizagem, considerando idade ou gênero.

Palavras-chave: Contabilidade – Estudo e ensino; Raciocínio; Métodos de ensino.

ABSTRACT

Due to market demands and the new Corporate Law's that require a professional profile of the business area with higher level of critical thinking, the general objective of this thesis was to recognize, analyze and evaluate empirically the effectiveness of teaching Richard Paul's model of critical thinking in students of accounting disciplines in the development of abilities to think critically about accounting and of critical thinking disposition's in general. The pilot quasi-experiment was conducted in a public institution during the second half of 2010 and had a general approach (four interventions of 30 min.) of the model's use and the author of the study was the monitor's classes. The second quasi-experiment was conducted in a private institution during the first half of 2011 and had a mixed application approach (used in most classes and activities of the experimental class) and the author of the study was the teacher's classes. In both quasi-experiments the applicants filled three evaluation instruments. Two of them were applied at the beginning and end of the semester and evaluated skills (Ennis Weir Critical Thinking Essay Test - EWCTET) and attitudes (California Critical Thinking Dispositions Inventory - CCTDI) of critical thinking and a demographic survey was applied at the end of the semester. In addition to descriptive analysis of the samples, to test the hypotheses set was applied the analysis of covariance (ANCOVA) using pretest as covariates. In the public institution was not found significant effect of instructional model in developing skills and attitudes of critical thinking and no difference for the effect of isolated variables age or gender. In the particular institution two hypotheses were confirmed, the experimental class ($p = .040$ and $d = .56$) and the older students ($p = .047$ and $d = .58$) have presented higher development of critical thinking skills. Although the mean of the attitudes were not significantly different, also met significant effect of instructional model in the development of the curiosity attitude ($p = .047$ and $d = .65$). There were also significant differences in the abilities to the isolated effects of age ($p = .002$ and $d = .77$) and gender ($p = .016$ and $d = .60$) and in the attitudes to the isolated effect of gender ($p = .033$ and $d = .46$). For future researches, it is recommended to apply the same model in undergraduate or graduate business area and further investigation on ways of thinking, behavior and learning, taking into account age or gender.

Keywords: Accounting – Study and Teaching; Reasoning; Teaching Methods.

SUMÁRIO

LISTA DE QUADROS	3
LISTA DE TABELAS	4
LISTA DE GRÁFICOS	6
1 INTRODUÇÃO	7
1.1 Contextualização	7
1.2 Problema da Pesquisa	13
1.3 Justificativa	15
1.4 Objetivos do Estudo	16
1.5 Hipóteses	16
1.6 Estrutura do Trabalho	18
2 REVISÃO DA LITERATURA	19
2.1 Raciocínio Crítico – Introdução e História	19
2.2 Definições de Raciocínio Crítico	23
2.2.1 Robert Ennis	23
2.2.2 John McPeck	24
2.2.3 Matthew Lipman	26
2.2.4 Richard Paul	27
2.2.5 Harvey Siegel	29
2.2.6 “O Relatório <i>Delphi</i> ” – Peter A. Facione	30
2.2.7 Definição Proposta	42
2.3 Avaliando Raciocínio Crítico	42
2.4 Modelo Instrucional de Richard Paul	46
2.5 Estudos Anteriores	51
2.6 Estudos na Área Contábil	60
3 PROCEDIMENTOS METODOLÓGICOS	67
3.1 Instituições Participantes	67
3.2 Alunos Participantes	68
3.3 Instrumentos de Avaliação	81
3.3.1 <i>Ennis-Weir Critical Thinking Essay Test - EWCTET</i>	82
3.3.2 <i>California Critical Thinking Dispositions Inventory – CCTDI</i>	84
3.3.3 Pesquisa Demográfica	85
3.4 Aplicações do Modelo Instrucional e Atividades	86
3.5 Método do Trabalho	102
3.5.1 Modelo de Pesquisa	103
3.5.2 Procedimentos para Análise dos Dados	112
4 ANÁLISE DOS DADOS E RESULTADOS	117
4.1 Análises Descritivas das Amostras	117
4.1.1 FEA-USP	117
4.1.2 UNIFIG	122
4.2 Hipótese 1 – Habilidades de Raciocínio Crítico	125
4.2.1 FEA - USP	126
4.2.2 UNIFIG	131
4.3 Hipótese 2 – Atitudes de Raciocínio Crítico	136
4.3.1 FEA - USP	137
4.3.2 UNIFIG	141
4.4 Hipótese 3 - Idade	150
4.4.1 FEA - USP	151

4.4.2 UNIFIG	158
4.5 Hipótese 4 - Gênero.....	165
4.5.1 FEA - USP.....	166
4.5.2 UNIFIG	172
5 CONCLUSÕES.....	183
5.1 Principais Conclusões	183
5.2 Limitações	190
5.3 Sugestões para Futuras Pesquisas	193
REFERÊNCIAS.....	195
BIBLIOGRAFIA.....	203
APÊNDICES	205
ANEXOS.....	231

LISTA DE QUADROS

Quadro 1 - Lista consensual de habilidades e sub-habilidades de raciocínio crítico	33
Quadro 2 - Descrições consensuais das principais habilidades e sub-habilidades de raciocínio crítico	34
Quadro 3 - Atitudes de raciocínio crítico	41
Quadro 4 - Códigos das variáveis da pesquisa	82
Quadro 5 - Representação do banco de dados da pesquisa	82
Quadro 6 - Exemplo de aplicação do Guia para identificar os elementos de raciocínio crítico de um artigo	91
Quadro 7 - Exemplo de aplicação do Guia para identificar os elementos de raciocínio crítico de um exercício	92
Quadro 8 - Fatores que afetam a validade interna e externa de experimentos	105
Quadro 9 - Resumo dos resultados das IES por hipótese	184

LISTA DE TABELAS

Tabela 1 - Distribuição dos pré-testes na FEA-USP	118
Tabela 2 - Distribuição dos pré-testes e pós-testes na FEA-USP	119
Tabela 3 - Distribuição dos erros padrão dos pré-testes e pós-testes na FEA-USP	120
Tabela 4 - Distribuição dos pré-testes e pós-testes por idade na FEA-USP	121
Tabela 5 - Distribuição dos pré-testes e pós-testes por gênero na FEA-USP	121
Tabela 6 - Distribuição dos pré-testes na UNIFIG	122
Tabela 7 - Distribuição dos pré-testes e pós-testes na UNIFIG	123
Tabela 8 - Distribuição dos pré-testes e pós-testes na UNIFIG sem valores atípicos	123
Tabela 9 - Distribuição dos erros padrão dos pré-testes e pós-testes na UNIFIG	124
Tabela 10 - Distribuição dos pré-testes e pós-testes por idade na UNIFIG	125
Tabela 11 - Distribuição dos pré-testes e pós-testes por gênero na UNIFIG	125
Tabela 12 - Distribuição do EWCTET por método de instrução na FEA-USP	126
Tabela 13 - Interação entre a covariável e as turmas	127
Tabela 14 - ANOVA para habilidades de raciocínio crítico – FEA-USP	128
Tabela 15 - ANCOVA para habilidades de raciocínio crítico – FEA-USP	128
Tabela 16 - Distribuição do EWCTET por método de instrução na UNIFIG	132
Tabela 17 - Distribuição do EWCTET por método de instrução na UNIFIG sem valores atípicos	132
Tabela 18 - ANCOVA para habilidades de raciocínio crítico – UNIFIG	133
Tabela 19 - Distribuição do CCTDI por método de instrução na FEA-USP	137
Tabela 20 - ANCOVA para atitudes de raciocínio crítico – FEA-USP	138
Tabela 21 - Distribuição de cada atitude do CCTDI por método de instrução na FEA-USP	139
Tabela 22 - ANCOVA para a atitude maturidade – FEA-USP	140
Tabela 23 - Distribuição do CCTDI por método de instrução na UNIFIG	141
Tabela 24 - Distribuição do CCTDI por método de instrução na UNIFIG sem valores atípicos	142
Tabela 25 - ANCOVA para atitudes de raciocínio crítico – UNIFIG	143
Tabela 26 - Distribuição de cada atitude do CCTDI por método de instrução na UNIFIG	144
Tabela 27 - ANCOVA para a atitude mente aberta – UNIFIG	145
Tabela 28 - ANCOVA para a atitude curiosidade – UNIFIG	146
Tabela 29 - ANCOVA para a atitude maturidade – UNIFIG	147
Tabela 30 - Distribuição do EWCTET por método de instrução e idade na FEA-USP	151
Tabela 31 - Distribuição do EWCTET por método de instrução e idade na FEA-USP sem valores atípicos	151
Tabela 32 - ANCOVA para habilidades de raciocínio crítico considerando idade – FEA-USP	153
Tabela 33 - Distribuição do CCTDI por método de instrução e idade na FEA-USP	155
Tabela 34 - Distribuição do CCTDI por método de instrução e idade na FEA-USP sem valores atípicos	155
Tabela 35 - ANCOVA para atitudes de raciocínio crítico considerando idade – FEA-USP	157
Tabela 36 - Distribuição do EWCTET por método de instrução e idade na UNIFIG	158
Tabela 37 - ANCOVA para habilidades de raciocínio crítico considerando idade – UNIFIG	160
Tabela 38 - Distribuição do CCTDI por método de instrução e idade na UNIFIG	162
Tabela 39 - ANCOVA para atitudes de raciocínio crítico considerando idade – UNIFIG	163

Tabela 40 - Distribuição do EWCTET por método de instrução e gênero na FEA-USP	166
Tabela 41 - ANCOVA para habilidade de raciocínio crítico considerando gênero – FEA-USP	167
Tabela 42 - Distribuição do CCTDI por método de instrução e gênero na FEA-USP	169
Tabela 43 - Distribuição do CCTDI por método de instrução e gênero na FEA-USP sem valores atípicos	169
Tabela 44 - ANCOVA para atitude de raciocínio crítico considerando gênero – FEA-USP	171
Tabela 45 - Distribuição do EWCTET por método de instrução e gênero na UNIFIG	172
Tabela 46 - Distribuição do EWCTET por método de instrução e gênero na UNIFIG sem valores atípicos	173
Tabela 47 - ANCOVA para habilidade de raciocínio crítico considerando gênero – UNIFIG	174
Tabela 48 - Distribuição do CCTDI por método de instrução e gênero na UNIFIG	176
Tabela 49 - Distribuição do CCTDI por método de instrução e gênero na UNIFIG sem valores atípicos	177
Tabela 50 - ANCOVA para atitude de raciocínio crítico considerando gênero – UNIFIG	178

LISTA DE GRÁFICOS

Gráfico 1 - Gênero dos alunos da FEA-USP por turma	69
Gráfico 2 - Idade dos alunos da FEA-USP por turma	70
Gráfico 3 - Cor/Raça dos alunos da FEA-USP por turma	70
Gráfico 4 - Estado civil dos alunos da FEA-USP por turma	70
Gráfico 5 - Quantidade de filhos dos alunos da FEA-USP por turma	71
Gráfico 6 - Emprego dos alunos da FEA-USP por turma	71
Gráfico 7 - Tipo de escola no ensino fundamental e médio dos alunos da FEA-USP por turma	71
Gráfico 8 - Renda familiar dos alunos da FEA-USP por turma	72
Gráfico 9 - Frequência de leitura dos alunos da FEA-USP por turma	72
Gráfico 10 - Frequência de escrita dos alunos da FEA-USP por turma	72
Gráfico 11 - Frequência de cálculo dos alunos da FEA-USP por turma	73
Gráfico 12 - Fluência em língua estrangeira dos alunos da FEA-USP por turma	73
Gráfico 13 - Participação em curso de leitura dos alunos da FEA-USP por turma	73
Gráfico 14 - Participação em curso de escrita dos alunos da FEA-USP por turma	74
Gráfico 15 - Participação em curso de cálculo dos alunos da FEA-USP por turma	74
Gráfico 16 - Participação em curso de raciocínio crítico dos alunos da FEA-USP por turma	74
Gráfico 17 - Gênero dos alunos da UNIFIG por turma	76
Gráfico 18 - Idade dos alunos da UNIFIG por turma	76
Gráfico 19 - Cor/Raça dos alunos da UNIFIG por turma	76
Gráfico 20 - Estado civil dos alunos da UNIFIG por turma	77
Gráfico 21 - Quantidade de filhos dos alunos da UNIFIG por turma	77
Gráfico 22 - Emprego dos alunos da UNIFIG por turma	77
Gráfico 23 - Tipo de escola no ensino fundamental e médio dos alunos da UNIFIG por turma	78
Gráfico 24 - Renda familiar dos alunos da UNIFIG por turma	78
Gráfico 25 - Frequência de leitura dos alunos da UNIFIG por turma	78
Gráfico 26 - Frequência de escrita dos alunos da UNIFIG por turma	79
Gráfico 27 - Frequência de cálculo dos alunos da UNIFIG por turma	79
Gráfico 28 - Fluência em língua estrangeira dos alunos da UNIFIG por turma	79
Gráfico 29 - Participação em curso de leitura dos alunos da UNIFIG por turma	80
Gráfico 30 - Participação em curso de escrita dos alunos da UNIFIG por turma	80
Gráfico 31 - Participação em curso de cálculo dos alunos da UNIFIG por turma	80
Gráfico 32 - Participação em curso de raciocínio crítico dos alunos da UNIFIG por turma	81
Gráfico 33 - Pré-testes e pós-testes do EWCTET da UNIFIG	136
Gráfico 34 - Pré-testes e pós-testes da atitude curiosidade da UNIFIG	150
Gráfico 35 - Pré-testes e pós-testes do EWCTET considerando idade da UNIFIG	165
Gráfico 36 - Pré-testes e pós-testes do EWCTET considerando gênero da UNIFIG	180
Gráfico 37 - Pré-testes e pós-testes do CCTDI considerando gênero da UNIFIG	181

1 INTRODUÇÃO

Este capítulo é dividido em seis partes. Na primeira, o assunto é contextualizado para a área da educação contábil, em seguida é apresentado o problema da pesquisa. A terceira parte apresenta as justificativas para a realização deste trabalho e a quarta, seus objetivos gerais e específicos. A quinta parte apresenta as hipóteses fundamentadas na literatura revisada e em pesquisas anteriores, e a última resume a estrutura da tese.

1.1 Contextualização

Em 2008, entrou em vigor no Brasil a Lei 11.638/07 (Lei das Sociedades Anônimas) que harmoniza a contabilidade brasileira às internacionais (IFRS – *International Financial Reporting Standards*). Foi criado também o CPC (Comitê de Pronunciamentos Contábeis), que analisa as normas internacionais e as adapta aos pronunciamentos contábeis aqui no Brasil.

Diversos países, inclusive os Estados Unidos da América (EUA), iniciaram um processo de convergência ou planejam adotar os IFRS. Conforme um recente artigo do *Journal of Instructional Pedagogies*, na mudança dos Princípios Contábeis Geralmente Aceitos nos EUA (USGAAP – *United States General Accounting Accepted Principles*) para o IFRS existe a necessidade de que o profissional contábil atual desenvolva novas habilidades. De acordo com Kermis e Kermis (2010, p. 2):

Com a mudança dos Princípios Geralmente Aceitos (USGAAP) para as Normas Internacionais (IFRS), há um aumento na necessidade de contadores com forte “*soft skills*” por causa da necessidade do IFRS em aplicar princípios em vez de seguir regras.

As normas internacionais são baseadas em princípios norteadores ao invés de regras para cada caso específico; portanto, é necessário que o profissional contábil se adapte a este novo contexto. Nesta nova realidade são exigidos os “*soft skills*”, ou seja, as habilidades que não são exatamente relacionadas à atividade técnica contábil, as quais são denominadas “*hard*

skills”. Enquanto estas são as habilidades relacionadas às disciplinas específicas do curso, como: Auditoria, Perícia, Contabilidade Tributária etc., as *soft skills* estão relacionadas às habilidades interpessoais, ao raciocínio crítico, à abertura de novas ideias etc. (KERMIS; KERMIS, 2010).

Nos EUA, desde o século passado, muito tempo antes de saberem da adoção das normas internacionais, alguns dos principais organismos da Contabilidade, como *American Accounting Association* (AAA), *American Institute of Certified Public Accountants* (AICPA), *Institute of Management Accountants* (IMA) e as BIG5 (*Arthur Andersen, Deloitte & Touche, Ernst & Young, KPMG e Pricewaterhouse & Coopers*), juntamente com W. Steve Albrecht e Robert J. Sack realizaram um estudo sobre o modelo de ensino superior em Contabilidade nos EUA incentivados pelos seguintes problemas: diminuição da procura pelo curso; críticas dos profissionais da área dizendo que o modelo estava quebrado e obsoleto; e muitos egressos dizendo que, se pudessem, teriam escolhido outro curso (ALBRECHT; SACK, 2000).

Durante a leitura da pesquisa de Albrecht e Sack (2000), torna-se clara a importância de se buscar conhecimentos, competências e habilidades que o atual mercado almeja nos profissionais. Dentre muitas, a habilidade citada com maior frequência pelos entrevistados como essencial para o sucesso profissional é a do Raciocínio Crítico (*Critical Thinking*). O estudo encerra dizendo que por 11 anos tem se enfatizado a mudança na profissão contábil de menos operacional para mais analítica. A mensagem é clara: a habilidade de raciocínio crítico deve ser enfatizada, já que a profissão precisa subir na cadeia de valor do processo contábil (operações, dados, relatórios, informações e tomada de decisões).

Em 1990, a *Accounting Education Change Commission* (AECC), que é ligada ao AAA, publicou um documento com diretrizes para que os cursos de graduação em Ciências Contábeis se adequassem às novas demandas do mercado. Este documento enfatizou a utilização de métodos de instrução que ensinam os alunos a aprenderem por eles mesmos e definiu esta abordagem como Aprendendo a Aprender (*Learning to Learn*), a qual também defende a necessidade do desenvolvimento de habilidades e atitudes que podem ser consideradas de raciocínio crítico. (FRANCIS; MULDER; STARK, 1995).

No *CPA Vision 2011 and Beyond*, o qual possui como principal objetivo “*Making sense of a changing and complex world*”¹, o Raciocínio Crítico continua sendo uma das habilidades mais importantes para o sucesso profissional contábil (AICPA, 2010).

Outro recente artigo da *Accounting Education News* defende o desenvolvimento do raciocínio crítico para os profissionais contábeis melhorarem os julgamentos sobre trabalhos ligados à área e assim ajudarem a sociedade de uma maneira mais explícita. O texto coloca que os recentes escândalos, a crise financeira e as controvérsias sobre aspectos contábeis e financeiros do mercado fazem com que os conceitos contábeis sejam discutidos abertamente na mídia. O problema é que os contadores normalmente não são convidados para estes debates, já que não são vistos como profissionais que raciocinam criticamente (PETER; WILSON, 2010).

No Brasil, as Diretrizes Curriculares Nacionais para o Curso de Graduação em Ciências Contábeis coloca em seu artigo 3º, parágrafo III, que um dos objetivos do curso é: “[...] revelar a capacidade crítico-analítica de avaliação, quanto às implicações organizacionais com o advento da tecnologia da informação” (RESOLUÇÃO CNE/CES 10, 2004). Observa-se que o atual cenário mundial é marcado por um volume extraordinário de informações, principalmente devido ao surgimento e expansão da *internet*. Neste sentido e a exemplo de outras profissões, Teixeira e Gomes (2004, p. 3) comentam sobre as novas exigências do atual mercado profissional:

Há uma nítida redução no número de empregos oferecidos e inovações tecnológicas transformaram profundamente o campo das ocupações profissionais. Por conseguinte, espera-se que o trabalhador seja mais flexível, apresentando maior repertório de habilidades e competências.

Aliada a esta nova realidade tecnológica, a Lei das S.A.s (Lei 11.638/07) traz ao profissional contábil brasileiro a necessidade de desenvolver seu raciocínio crítico. Como a contabilidade nacional diminuiu a distância com a internacional e flexibilizou muitas práticas contábeis que eram extremamente rígidas, o contador precisará desenvolver novas habilidades para conseguir analisar, interpretar, julgar e decidir sobre a aplicação de certos princípios a novas situações em que nem sempre existirá uma norma ou um regulamento detalhado sobre como proceder contabilmente.

¹ “Fazer sentido de um mundo mutável e complexo.”

Iudicibus (2007, p. 11) discute em um artigo as raízes profundas dos princípios contábeis e faz o seguinte comentário sobre a prevalência da essência sobre a forma:

A prevalência da essência sobre a forma, todavia, provoca sérios problemas para os contadores acomodados, acostumados às regrinhas de custo, da objetividade extremada, treinados mais para obedecer às normas do que para serem criativos.

A “forma”, além de ser de aplicação mais fácil nas culturas de direito romano, ainda leva a vantagem da objetividade, pois o fato por detrás dela, ou ato, está consubstanciado por farta documentação. Afinal, a forma está na própria documentação jurídica comprobatória de uma operação.

A essência econômica é uma das principais raízes que nutrem e sustentam toda a árvore contábil. Quando se souber entendê-la profundamente e aplicá-la com propriedade, se chegará, talvez, ao ponto mais alto da evolução contábil, no qual se saberá dosar relevância com objetividade ou subjetivismo responsável.

Logo, com a internacionalização da contabilidade, o profissional precisará cada vez mais desenvolver novas habilidades que já eram demandadas por áreas contábeis, como a contabilidade gerencial, e que agora passam a ser importantes também nas demais áreas contábeis. Os alunos dos atuais cursos de graduação em Ciências Contábeis precisam aprender por meio de metodologias que os ajudem a exercer estas novas atribuições do profissional contábil.

A literatura norte-americana sobre *Critical Thinking* (CT), que neste trabalho será chamado de Raciocínio Crítico (RC), é vasta e o termo é estudado há muitos anos. No Brasil não existe um consenso sobre a tradução do termo e muitas são as palavras utilizadas, como, por exemplo: Pensamento Crítico, Pensando Criticamente, Argumentação Lógica, Raciocínio Lógico etc.

Analisando-se a história é possível perceber que desde antes de Cristo a capacidade de raciocinar é enfatizada como habilidade fundamental para que as pessoas entendam melhor a realidade que as cerca. Lefrançois (2009, p. 3), ao discutir o processo de aprendizagem e de como o conhecimento é formado, comenta:

Alguns dos filósofos gregos, como Aristóteles (384-322 a.C.), responderam a essas perguntas com uma teoria que propunha que tudo que existe fora de nós, no mundo, tem uma cópia na mente. O que acontece, explicou ele, é que o ato de perceber alguma coisa resulta numa cópia de algo que, de algum modo, passamos a conhecer. Assim, não conhecemos a realidade diretamente; tudo o que conhecemos é de forma indireta, resultado da percepção das cópias da realidade. Nós conhecemos não tanto como resultado dos nossos sentidos, argumentava ele, porém mais como resultado de nossa razão. Dessa forma, as pessoas instruídas, cujas mentes presumivelmente são mais capazes de raciocinar, conhecem a realidade de forma mais precisa do que as pessoas menos cultas. “As raízes da educação são amargas”, disse Aristóteles, “mas o fruto é doce”.

No atual cenário mundial, os países que detêm maior conhecimento são aqueles que tendem a deter maior poder. As grandes potências têm conhecimento disso, uma prova é que importantes associações dos EUA e Canadá incentivam o desenvolvimento do raciocínio crítico nas escolas (ABRAMI *et al.*, 2008).

Uma recente meta-análise resume as evidências empíricas disponíveis em 117 estudos abrangendo 20.698 participantes sobre o impacto de instruções no desenvolvimento ou aprimoramento das habilidades ou atitudes de raciocínio crítico dos alunos. Essa pesquisa concluiu que o tipo de intervenção e metodologia de ensino está substancialmente relacionado com os níveis de desenvolvimento de raciocínio crítico. Analisando-se os resultados, fica claro que o aprimoramento ou desenvolvimento das habilidades (capacidade de: identificar, analisar, avaliar, julgar, interpretar, decidir etc.) ou atitudes (comportamento: curiosidade, pensar independente, integridade intelectual etc.) de raciocínio crítico dos alunos não podem ser uma mera expectativa implícita. Educadores devem começar a explicitar os objetivos de desenvolvimento de raciocínio crítico nos seus cursos, bem como treinar outros professores a fazerem o mesmo para que consigam aplicar as metodologias e desenvolver estas habilidades e atitudes (ABRAMI *et al.*, 2008).

No Brasil, um artigo recentemente publicado na área contábil explorou, em uma amostra de 65 estudantes de graduação e pós-graduação em Ciências Contábeis da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA-USP), as habilidades de Raciocínio Crítico, leitura e escrita. Por meio de análises estatísticas, como correlações e diferenças de médias entre grupos (gênero, paternidade, nível do programa e estágio do programa), o estudo analisou os dados e refletiu sobre os elementos para o aprimoramento do desempenho dos estudantes em relação as suas habilidades de Raciocínio Crítico e o papel dos educadores como agentes de mudanças no cenário da educação superior. Os resultados indicaram boa qualidade de escrita, evidência de níveis elevados de raciocínio crítico e não apontaram diferenças significativas entre os grupos citados acima. (CORNACHIONE; DUNCAN; JOHNSON, 2007).

Revisando a literatura internacional, verifica-se que, embora haja considerável existência de teorias e instrumentos relacionados ao desenvolvimento e avaliação do raciocínio crítico, faltam pesquisas empíricas sobre este tópico na área da educação de negócios. Braun (2004,

p. 235) comenta: “[...] dados empíricos sobre o sucesso da utilização de técnicas de desenvolvimento de habilidades de raciocínio crítico são mínimos”.

Em 1998, Jennifer H. Reed defendeu uma tese que investigou o efeito de integrar o modelo de Richard Paul (1995) para desenvolvimento de raciocínio crítico em um curso de História Norte-Americana para desenvolver: habilidades para pensar criticamente sobre História Norte-Americana e sobre assuntos diários; atitudes para pensar criticamente; e conhecimento sobre o conteúdo de História Norte-Americana. Sua tese examinou também se a idade ou o gênero influenciam a eficiência do método de instrução. Duas turmas foram aleatoriamente selecionadas para servir como grupo experimento e as outras duas turmas serviram como grupo controle. O grupo experimento ($n = 29$) recebeu aproximadamente 90 minutos de instrução explícita distribuída durante o semestre sobre como utilizar o modelo de raciocínio crítico de Richard Paul para analisar e interpretar documentos. Adicionalmente, foi integrada em uma série de atividades em sala de aula. Já o grupo controle ($n = 23$) foi ensinado de uma maneira mais tradicional. Os alunos fizeram pré-testes e pós-testes para avaliar as atitudes e habilidades de raciocinar criticamente. Os resultados apontaram diferenças significativas no desenvolvimento de habilidades, indicando eficiência do modelo instrucional; no entanto, não apontaram diferenças significativas nas atitudes de raciocínio crítico nem influência significativa da idade ou do gênero (REED, 1998).

O modelo de Richard Paul (1995) inclui elementos de raciocínio, padrões intelectuais universais usados para avaliar o raciocínio, as atitudes e habilidades de uma mente que raciocina. De acordo com seu modelo, os elementos de raciocínio são os aspectos que devem ser identificados em um pensamento, artigo ou livro, para que seja possível uma análise crítica: Ponto de Vista, Propósito do Pensamento, Questão/Problema, Informações, Interpretações e Inferências, Conceitos, Suposições e Implicações e Consequências. As atitudes são os aspectos comportamentais que um “pensador crítico” deve apresentar: Pensar Independente, Empatia Intelectual, Humildade Intelectual, Coragem Intelectual, Integridade Intelectual, Fé na Razão etc. Os padrões intelectuais universais são os balizadores da profundidade, abrangência e clareza da discussão de determinada linha de raciocínio: Clareza, Retidão, Precisão, Relevância, Profundidade e Alcance. Já as habilidades unem os elementos e os padrões à determinada ação, como: Identificar o Propósito Claramente, Analisar o Problema Corretamente, Avaliar Interpretações Precisamente etc.

Paul (1995) defende que seu modelo pode ser aplicado a qualquer problema ou assunto que requeira raciocínio. Sua teoria é a de que para desenvolver raciocínio crítico nos alunos é preciso ensiná-los a pensar criticamente (REED, 1998). O modelo de raciocínio crítico foi testado pela primeira vez no estudo de Reed, que o escolheu para testar se ele era apropriado para analisar documentos e desenvolver as habilidades de argumentação e de resolução de problemas. Considerando-se que o profissional da área contábil necessita desenvolver habilidades de raciocínio crítico e que Reed teve sucesso com seus alunos, o modelo de Richard Paul foi escolhido pelo autor desta pesquisa para ser aplicado a alunos de disciplinas da área contábil nos mesmos moldes da pesquisa de Reed, possibilitando analisar a eficácia do modelo para a área contábil e realizar comparações com os resultados de pesquisas anteriores, principalmente a de Reed.

Portanto, dentre os modelos de raciocínio crítico encontrados na revisão da literatura, escolheu-se o de Richard Paul principalmente pelo fato de ele ser generalista, ou seja, o modelo de desenvolvimento de raciocínio crítico de Richard Paul pode ser aplicado isoladamente ou como material de suporte em um curso, que foi o caso da pesquisa de Reed e também desta tese.

Espera-se que os resultados desta pesquisa tragam significantes contribuições às pesquisas sobre educação contábil e incentive a discussão sobre o assunto com os professores e pesquisadores da área contábil brasileira e internacional.

1.2 Problema da Pesquisa

Considerando-se o que já foi exposto na introdução deste trabalho, é possível dizer que a contabilidade no Brasil vive um momento de transição parecido com o que viveu em 1976, com a antiga Lei das S.A.s (Lei 4.604/76). Com a introdução desta lei, viu-se naquela época a mudança de uma contabilidade com maiores traços da escola italiana ser modificada para uma contabilidade mais parecida com a norte-americana. É claro que a mudança não foi instantânea. Foi um processo de adaptação que durou alguns anos, mas que foi impulsionada academicamente com o lançamento do livro *Contabilidade Introdutória dos Professores da FEA-USP*, que, tendo o método de ensino fundamentado na escola americana, fez com que a

maioria das Faculdades de Contabilidade do Brasil passasse a utilizá-lo, concluindo, assim, em alguns anos, um ciclo de instrução da classe contábil brasileira.

Analogicamente ao exposto acima, vive-se um momento de transição na contabilidade brasileira com a aprovação da nova Lei das S.A.s (Lei 11.638/07), a qual inicia um processo de convergência das normas brasileiras às internacionais. O problema é que desta vez a mudança não é apenas de procedimentos, normas e formatos demonstrativos, mas da forma de atuação do profissional contábil, visto que ele passa a ser responsável pela análise da prática contábil mais adequada, que melhor mensure e evidencie determinada operação. Aliado a isto, vive-se um novo contexto tecnológico e uma internacionalização dos mercados que ampliam as funções do contador em demais áreas e também exigem o desenvolvimento de novas habilidades e atitudes.

Para tanto, a academia precisa, como o fez, em 1976, pesquisar metodologias de ensino e formas de adequar nosso egresso a este novo perfil de profissional exigido pelo mercado. Uma das exigências é o aumento do nível de raciocínio crítico, possibilitando ao profissional analisar situações, argumentos e problemas para que consiga decidir com base em seu próprio julgamento. Nesse sentido, esta pesquisa visa desenvolver habilidades e atitudes de raciocínio crítico em alunos de disciplinas contábeis por meio de intervenções em sala de aula, as quais explicam o modelo de raciocínio crítico de Richard Paul e o aplicam em análises de artigos, livros e resolução de problemas. Logo, a questão de pesquisa a ser respondida é:

A aplicação do modelo instrucional² de raciocínio crítico de Richard Paul melhora o nível de habilidades³ e atitudes⁴ de raciocínio crítico dos alunos de disciplinas da área contábil?

² O termo instrucional nesta pesquisa é utilizado como: “de instrução”, “para instruir” ou “para ensinar” e não dispensa a participação de um professor ou tutor para ensinar o modelo aos alunos.

³ A habilidade está diretamente relacionada com a prática do trabalho, mas de modo especialmente transcendente, se comparada à mera ação motora, englobando a capacidade de saber como trabalhar o conhecimento que se possui, compreendendo a técnica, a capacidade e o saber como (DURAND, 1998).

⁴ A atitude possui três componentes: o componente cognitivo, que se refere a uma opinião ou crença, o componente afetivo, relacionado ao sentimento e às emoções, e o componente comportamental, que se refere à intenção de se comportar de uma determinada maneira com relação a alguma coisa ou alguém (ROBBINS, 2002).

1.3 Justificativa

Uma primeira justificativa para o presente trabalho deve-se ao fato de que não existe atualmente no Brasil uma tese na área contábil que trate do assunto. Internacionalmente foram encontradas teses em outras áreas, porém não especificamente na área contábil.

Uma segunda justificativa à realização desta pesquisa deve-se à nova Lei das S.A.s (Lei 11.638/07), pois ela harmoniza a contabilidade brasileira à internacional, exigindo um subjetivismo responsável⁵ por parte dos contadores, no qual a habilidade e as atitudes de raciocínio crítico são essenciais.

Uma terceira justificativa deve-se ao fato de outras áreas contábeis, como a gerencial, por exemplo, demandarem profissionais com estas habilidades, uma vez que são extremamente necessárias na solução de problemas, análise de alternativas e tomadas de decisões.

Uma quarta deve-se ao fato de não se ter encontrado uma pesquisa posterior à de Reed que tenha aplicado o modelo de Richard Paul para desenvolvimento de raciocínio crítico em outras áreas de ensino; logo, esta pesquisa preencherá uma lacuna importante na literatura internacional sobre o assunto.

E, por fim, em um momento em que os órgãos governamentais informatizam seus procedimentos de cobranças tributárias (SPED – Sistema Público de Escrituração Digital⁶), as grandes corporações possuem subsidiárias em diversos países e concorrem com as pequenas e médias empresas locais, o acesso à *internet* cresce em ritmo galopante, os clientes estão cada vez mais exigentes, esta pesquisa se justifica para compreender o nível de raciocínio crítico dos atuais alunos de contabilidade e como um modelo instrucional pode interferir no seu desenvolvimento para ajudar na formação de um profissional capaz de obter sucesso neste mundo de constantes e rápidas mudanças.

⁵ A partir da edição da nova Lei das S.A.s, a contabilidade nacional diminuiu a distância com a internacional e flexibilizou muitas práticas contábeis que eram extremamente rígidas e detalhadas em normas.

⁶ Sistema Público de Escrituração Digital. Disponível em: <<http://www1.receita.fazenda.gov.br/>>. Acesso em: 10/12/2008.

1.4 Objetivos do Estudo

O objetivo geral deste trabalho é conhecer, analisar e avaliar empiricamente a efetividade em ensinar o modelo de Richard Paul (1995) para raciocínio crítico em alunos de disciplinas da área contábil no desenvolvimento de habilidades para pensar criticamente sobre contabilidade e atitudes de raciocínio crítico em geral. Os objetivos específicos são:

- 1) avaliar o desenvolvimento de habilidades e atitudes de raciocínio crítico;
- 2) analisar possíveis diferenças no desenvolvimento de habilidades e atitudes de raciocínio crítico em relação ao gênero e à idade;
- 3) comparar os resultados obtidos da amostra brasileira com os do estudo de Reed (1998).

Uma conclusão de Edgar Morin será norteadora para o desenvolvimento desta pesquisa. O autor comenta que as certezas teóricas não podem ser consideradas certezas absolutas, elas devem ser abandonadas pela ideia de falibilismo de Popper. Morin (2002, p. 23) cita a seguinte frase de Popper:

A história das ciências, como a de todas as ideias humanas, é uma história de sonhos irresponsáveis, de teimosias e de erros. Porém, a ciência é uma das raras atividades humanas, talvez a única, na qual os erros são sistematicamente assinalados e, com o tempo, constantemente corrigidos.

1.5 Hipóteses

Com base no problema já definido e, principalmente, no trabalho de Reed (1998), esta pesquisa possui as seguintes hipóteses:

- 1) Uma classe que receber explicitamente treinamento sobre o modelo de desenvolvimento de raciocínio crítico de Richard Paul obterá melhor resultado em um teste (EWCTET⁷) que avalia habilidade em análise e resposta de argumentos sobre um assunto contemporâneo do que uma classe que não receber instruções explícitas sobre o modelo.

Reed constatou em sua tese que os alunos de graduação do curso de História que foram expostos ao modelo de raciocínio crítico de Richard Paul obtiveram melhor resultado no EWCTET do que os alunos de uma turma controle (REED, 1998).

- 2) Uma classe que receber explicitamente treinamento sobre o modelo de desenvolvimento de raciocínio crítico de Richard Paul obterá melhor resultado em um teste (CCTDI⁸) que avalia atitudes de raciocínio crítico do que uma classe que não receber instruções explícitas sobre o modelo.

Apesar de Reed não ter constatado em sua tese diferenças significativas entre os desempenhos da turma experimento e controle em relação às atitudes de raciocínio crítico, a meta-análise de Abrami apresentou evidências de desenvolvimento de habilidades e atitudes de raciocínio crítico de acordo com o modelo instrucional (REED, 1998; ABRAMI *et al.*, 2008). Além disso, a teoria de Richard Paul é a de que estas atitudes devem ser explicitadas para conscientizar os alunos de que não basta possuir as habilidades de raciocínio crítico, precisa-se também da intenção ou atitude para utilizá-las (PAUL, 1995).

- 3) Alunos mais velhos (22 anos e acima⁹) obterão melhor desempenho nos testes (EWCTET e CCTDI) de habilidades e atitudes de raciocínio crítico do que alunos mais novos (abaixo de 22 anos) de acordo com o método instrucional (com ou sem o modelo de Richard Paul).

Pesquisadores têm demonstrado que estudantes mais velhos diferem de estudantes mais novos ou de idade tradicional para determinado curso em uma variedade de formas, incluindo abordagens de estudo, atitudes em relação à escola e desempenho (KING; KITCHENER,

⁷ ENNIS, Robert H; WEIR, E. *The Ennis-Weir Critical Thinking Essay Test (EWCTET)*., CA, USA: Pacific Grove, 1985.

⁸ FACIONE, N.; FACIONE, Peter A. *Manual of the California Critical Thinking Disposition Inventory (CCTDI)*. CA, USA: California Academic Press, 2010.

⁹ Adotou-se o mesmo ponto de ruptura na idade do estudo de Reed (1998) para fins de comparabilidade dos resultados.

1994). O modelo de desenvolvimento intelectual e moral de Perry (1970 *apud* KING; KITCHENER, 1994), estabelece que as concepções das pessoas sobre a natureza do conhecimento e seus entendimentos sobre elas mesmas como conhecedoras, pensadoras e racionais geralmente se desenvolvem com o tempo.

- 4) Há diferenças significativas nos testes (EWCTET e CCTDI) de habilidades e atitudes de raciocínio crítico dos alunos de acordo com o método instrucional (com ou sem o modelo de Richard Paul) e gênero (feminino/masculino).

Estudos têm demonstrado que as “formas de aprendizagem” das mulheres diferem-se das dos homens (MILLER; FINLEY; MCKINLEY, 1990). Reed (1998), por sua vez, não constatou diferenças significativas em relação ao gênero e ao método instrucional.

1.6 Estrutura do Trabalho

Este trabalho teve início com a introdução do assunto e a apresentação do problema, dos objetivos, das justificativas e das hipóteses da pesquisa desenvolvida. O Capítulo 2 versa sobre a revisão da literatura, abordando aspectos históricos sobre raciocínio crítico, as linhas de pesquisas existentes, as definições, os estudos anteriores em áreas diversas e na contábil.

O Capítulo 3 aborda os procedimentos metodológicos, a amostra estudada, a descrição das atividades desenvolvidas em sala de aula, a aplicação do método instrucional, suas definições operacionais, os instrumentos de avaliação, suas variáveis e os aspectos de validade e confiabilidade.

O Capítulo 4 apresenta a análise dos dados, os instrumentos estatísticos utilizados, as primeiras interpretações de cada resultado quantitativo e qualitativo. No Capítulo 5 é discutido o problema de pesquisa com base nas hipóteses, nos resultados, nos dados e nas informações coletadas de acordo com a plataforma teórica existente e revisada durante o desenvolvimento desta pesquisa.

2 REVISÃO DA LITERATURA

Este capítulo apresenta uma introdução histórica sobre o tema, os principais filósofos, suas teorias, além de uma definição de raciocínio crítico proposta pelo autor deste trabalho, um painel geral sobre avaliação de habilidades e atitudes de raciocínio crítico, o modelo instrucional de Richard Paul, estudos anteriores na área contábil e outras áreas do conhecimento.

2.1 Raciocínio Crítico – Introdução e História

Na História o raciocínio crítico é o processo que todos os filósofos utilizam para criar suas teorias e julgar se um argumento tem sentido ou não. Voltando para a Grécia antiga, encontra-se Platão introduzindo a lógica como um processo de pensamento dialético, como foi modelado por seu professor, Sócrates. Aristóteles, como aluno de Platão, refinou a análise lógica de tal maneira que se encontrou inalterada até o século vinte. No século XX, Robert Ennis escreveu um artigo intitulado “Um conceito de Raciocínio Crítico”, no qual ele iniciou um debate para a definição do termo e com certeza se baseou em muitos filósofos que o precederam no século XIX como: John Dewey, William James e Charles Sanders Peirce (THAYER-BACON, 2000).

Na época de Sócrates, Platão e Aristóteles certamente a discussão não se aprofundava ao campo epistemológico das diferenças entre os termos: lógica, raciocínio crítico, raciocínio reflexivo, criativo e etc. como atualmente ocorrem, porém é interessante compreender como a habilidade de raciocinar criticamente já era valorizada e como estes filósofos contribuíram construindo os pilares das atuais teorias (THAYER-BACON, 2000).

Quando se pensa em uma pessoa que raciocina criticamente ou simplesmente no raciocínio crítico, qual a primeira imagem que se vem à mente? Para cada indivíduo existe um pré-conceito de pensador crítico, porém uma imagem conhecida mundialmente é a obra “O

Pensador” de Rodin. Ela é usada como um modelo de muitos materiais que tratam de raciocínio crítico, como, por exemplo: livros, conferências, congressos e seminários. Para explicar a concepção da imagem solitária e reflexiva da obra de Rodin é necessário entender um pouco da História (THAYER-BACON, 2000).

Sócrates (470 – 399 a.C.) é um marco na história da filosofia, a ponto de os que vieram antes dele receberem todos, apesar de suas diferenças, a denominação de “pré-socráticos” (ABRÃO, 1999). Aluno de Sócrates, Platão (427 – 347 a.C.) é considerado o primeiro filósofo da antiga Europa Ocidental e a primeira pessoa a estabelecer uma Academia no Mundo Antigo. No diálogo de Platão, o qual ele escreve para os cidadãos da Grécia, seu principal debatedor é Sócrates, o famoso filósofo Grego e seu professor. Com os diálogos de Platão sobre o que é virtude ou o que é uma boa forma de governo, há uma teoria epistemológica que molda como ele aborda seu problema, o que considera como evidência e como determina se chegou à uma solução ou não (THAYER-BACON, 2000; ABRÃO, 1999). Logo, nota-se que Platão já utilizava uma metodologia bem definida na identificação de um problema, identificação de evidências e maneira ou caminho pelo qual ele chegava a conclusões.

Quando Platão descreve como é que nós sabemos a resposta para algo, ele fala sobre “a alma” de uma maneira singular. Isto é porque ele assume que cada um de nós possui uma alma individual que é imortal e sabe tudo. Esse pressuposto é a raiz da ideia da Antiga Europa Ocidental, logo os sábios eram vistos como pessoas solitárias, em vez de pessoas sociáveis. Então, se chegou à imagem do Pensador de Rodin, e assim a explicação do porque que quando se pensa na imagem de um pensador, vem-se a mente uma pessoa solitária em vez de uma comunidade de investigadores (THAYER-BACON, 2000; ABRÃO, 1999).

Atualmente, o pensador crítico não pode ser visto como alguém solitário e isolado devido às rápidas e constantes mudanças do mundo que exigem que o mesmo esteja atento, trabalhe em grupo e tenha um olhar crítico do mundo que o cerca (PAUL, 1995).

Pensando na evolução do raciocínio crítico, se Platão é considerado o primeiro filósofo da antiga Europa Ocidental, seu aluno Aristóteles (384 – 324 a.C.) é considerado o segundo. Aristóteles é frequentemente chamado de “O pai da Ciência” por ele ter desenvolvido um

sistema de lógica no qual ele trabalhou classificando o mundo ao seu redor (THAYER-BACON, 2000, ABRÃO, 1999).

Aristóteles inventou um método lógico chamado silogismo para ajudar os sábios a testarem as suas próprias ideias e até mesmo criarem um novo conhecimento. Logo, com seu método se pode dizer que há uma grande evolução na história do raciocínio crítico. Um argumento pode ser válido, de acordo com a lógica de Aristóteles, e mesmo assim sentenças com ele podem não ser verdadeiras. Um exemplo: Se há nuvens no céu e nenhum sol é visível, então está chovendo. O argumento é válido, mas não é necessariamente verdadeiro, já que muitas áreas do mundo possuem dias que estão nublados, o sol não aparece e não há chuva (THAYER-BACON, 2000; ABRÃO, 1999).

De acordo com a lógica de Aristóteles, para estabelecer uma verdadeira conclusão, as premissas de um argumento devem guiar a uma conclusão necessária, o argumento deve ser válido e cada uma das premissas do argumento deve ser verdadeira. A grande contribuição que Aristóteles fez a filosofia foi oferecer uma maneira de descobrir o que é verdade, baseado na verdade de premissas existentes (THAYER-BACON, 2000; ABRÃO, 1999).

Outro elemento importante no desenvolvimento do raciocínio crítico é o debate ou diálogo. O debate já existia na sociedade Grega. Nos diálogos de Platão, Sócrates frequentemente se refere aos Sofistas como uma ameaça à Filosofia, porque os Sofistas ensinavam os filhos de ricos Gregos como serem bons debatedores, baseados na esperteza e na efetividade de seus argumentos e não no correto e na verdade. Aristóteles compartilhava com seu professor a preocupação sobre os Sofistas e por isso criou uma forma de lógica que guiaria para a Verdade. Ela é conhecida como a primeira lógica “verdadeira”, pois ele ofereceu um caminho sistemático para detectar qualquer falácia (THAYER-BACON, 2000; ABRÃO, 1999).

Richard Paul também utiliza o termo Sofista de forma pejorativa fazendo uma distinção entre “fraco senso” e “forte senso” de raciocínio crítico:

- Fraco senso / Sofistas: aqueles que usam suas habilidades de raciocínio para defender seus próprios argumentos e não usam os mesmos critérios para atacar os argumentos dos outros, pensando sempre em se sobressair em um debate, independentemente da Verdade (PAUL, 1995);

- Forte senso / Verdadeiros Pensadores: aqueles que procuram a Verdade aplicando em seus próprios argumentos os mesmos critérios de raciocínio crítico utilizados na análise dos argumentos dos outros (PAUL, 1995).

O debate dos filósofos do século XX, que são analisados em tópicos deste trabalho, foi bastante influenciado por três filósofos da escola pragmática americana do século XIX, são eles: Charles Sanders Peirce, William James e John Dewey.

Charles Sanders Peirce (1834-1914) tentou escrever uma filosofia sistemática como a de Platão e Aristóteles, a qual cobrisse todas as áreas de conhecimento, uma tarefa que muitos filósofos da atualidade acreditam ser impossível de se fazer. Peirce foi primeiro um cientista, depois passou a ser um lógico e ultimamente passou a ser conhecido como um dos maiores filósofos da História dos Estados Unidos da América. Ele originou o conceito de pragmatismo, em 1878, e muitas pessoas nunca tinham ouvido falar no termo, até que seu amigo, William James (1842-1950) publicou sobre o conceito e creditou Peirce. De acordo com James, Peirce formulou o conceito de pragmatismo na época em que os dois eram alunos de graduação e participavam de um Clube de Metafísica que eles fundaram na Universidade de Harvard. Embora James tenha ficado conhecido como professor de Filosofia e Psicologia na Universidade de Harvard, Peirce não foi capaz de permanecer fixo em nenhuma universidade e se aposentou, retirando-se à uma fazenda em Mildford, Pennsylvania, quando ainda estava por volta dos seus 40 anos de idade (THAYER-BACON, 2000).

Outro filósofo clássico do pragmatismo americano foi John Dewey (1859-1952), ele não se referia a si mesmo como um pragmático, pois preferia os termos instrumentalista ou experimentalista, embora não tenha sido muito feliz na utilização destes termos. Dewey teve uma longa e distinta carreira, seus primeiros anos foram passados na Universidade de Michigan, depois foi para a Universidade de Chicago e seus últimos anos foram na Universidade da Columbia (THAYER-BACON, 2000).

Um marco na História do desenvolvimento do raciocínio crítico, agora já no século XX, foi o artigo de Robert Ennis intitulado “Um conceito de Raciocínio Crítico”, o qual teve como objetivo definir e considerar alguns aspectos e dimensões do raciocínio crítico. Ennis começou a escrever sobre raciocínio crítico durante um tempo na História Americana em que

a maioria dos americanos estava preocupada com o comunismo e com a necessidade de proteger a democracia. O raciocínio crítico foi visto como uma ferramenta valiosa para proteger os conceitos de verdade e justiça. Os debates sobre a teoria do raciocínio crítico foram evoluindo e abordando assuntos referentes à procura de uma definição, como ensinar e encorajar os outros a usarem, como testar e assegurar que eles aprenderam (ENNIS, 1993).

De Sócrates até a atualidade muitas foram as contribuições ao desenvolvimento do raciocínio crítico. Porém, não se trata do objetivo principal desta revisão de literatura percorrer todos os filósofos, pensadores e suas teorias, mas sim os principais que possibilitam compreender o modelo instrucional que é utilizado na pesquisa empírica.

2.2 Definições de Raciocínio Crítico

Neste tópico são abordadas definições ou teorias de raciocínio crítico tendo como ponto inicial a obra do Prof. Robert Ennis *Um conceito de Raciocínio Crítico*, publicada em 1962. Ele iniciou um debate entre diversos estudiosos que durou muitos anos, culminando em uma pesquisa organizada pela Associação Americana de Filosofia e liderada pelo Prof. Peter A. Facione para se chegar a uma definição consensual de raciocínio crítico. Em 1990, o produto final deste trabalho ficou conhecido como *O Relatório Delphi* (por utilizar o método Delphi), que contou com a participação de 46 especialistas de diferentes áreas.

2.2.1 Robert Ennis

Ennis é o coautor do *Teste de Raciocínio Crítico de Cornell* (*Cornell Critical Thinking Test*) (ENNIS; MILLMAN; TOMKO, 2005) e do *Teste Redacional de Raciocínio Crítico de Ennis-Weir* (EWCTET – *Ennis-Weir Critical Thinking Essay Test*) (ENNIS; WEIR, 1985). Ennis liderou o Projeto de Raciocínio Crítico da Universidade de *Illinois* até sua recente aposentadoria. Sua teoria é a de que raciocínio crítico é ensinado aos alunos ensinando lógica.

A primeira definição de Ennis sobre raciocínio crítico, em 1962, foi: “[...] a correta avaliação de afirmações” (1993, p. 179). A definição original de Ennis não toma conhecimento da importância de pensadores críticos possuírem a tendência de usar habilidades de raciocínio

crítico em suas vidas (atitudes de raciocínio crítico). Após participar e se beneficiar de muitos debates sobre raciocínio crítico, Ennis revisou sua definição original e dissertou que raciocínio crítico é (1993, p. 180): “[...] pensamento reflexivo racional que é focado em decidir em que acreditar ou fazer”. Para Ennis, o raciocínio crítico é uma atividade prática que possui crenças racionais ou ações como objetivo. Esta definição buscou incluir habilidades e atitudes que são necessárias a um pensador crítico, afirmando que não é suficiente ter as habilidades, é preciso possuir tendência ou atitude para utilizá-las.

De acordo com Ennis (1993, p. 180), para decidir racionalmente no que acreditar e o que fazer durante um debate, uma análise de um artigo, de um livro ou da resolução de um problema é necessário:

1. Julgar a credibilidade dos recursos.
2. Identificar conclusões, razões e suposições.
3. Julgar a qualidade de um argumento, incluindo a aceitabilidade de suas razões, suposições e evidências.
4. Desenvolver e defender a posição de um assunto.
5. Perguntar questões clarificadoras e apropriadas.
6. Planejar experimentos e julgar modelos de experimentos.
7. Definir termos de maneira apropriada ao contexto.
8. Ser mente-aberta.
9. Tentar ser bem informado.
10. Chegar a conclusões quando seguro, mas com cuidado.

Esta lista de habilidades e atitudes interdependentes pode ser utilizada para avaliar raciocínio crítico, servir como uma série de objetivos de um currículo inteiro ou de apenas uma atividade. É possível utilizá-la também como uma tabela de especificações para a construção de um teste de raciocínio crítico (ENNIS, 1993). Um dos testes (EWCTET) elaborado pelo Prof. Ennis é aplicado na parte empírica desta pesquisa.

2.2.2 John McPeck

Em 1981, John McPeck, filósofo canadense, publicou a obra *Raciocínio Crítico e Educação* (*Critical Thinking and Education*) e iniciou um caloroso debate entre os teóricos de raciocínio crítico sobre a natureza do pensamento. De acordo com McPeck (1981, p. 12) o raciocínio crítico é um componente do pensamento racional e este é: “[...] o uso da inteligência sobre todas as evidências disponíveis para a solução de algum problema”. Logo,

o autor considera o pensamento racional como algo bem abrangente, sem entrar na discussão do que é inteligência, pois, se assim o fizesse, entraria no campo da psicologia cognitiva.

O filósofo (1981) aponta que raciocínio crítico é somente utilizável em um assunto específico, não é universalmente generalizável; portanto, ele não pode ser ensinado como um assunto geral, mas somente dentro de uma área específica do conhecimento. Para o autor, pensar é sempre pensar sobre alguma coisa. Ele afirma que alguém não pode ensinar raciocínio crítico sem possuir um objeto de estudo. O autor entende que para ser crítico em relação a determinado assunto, primeiro se deve ter profundo conhecimento sobre ele. Afirma ainda que “[...] o principal significado do raciocínio crítico é a propensão e habilidade de se engajar em uma atividade com ceticismo reflexivo” (p. 8).

McPeck (1981, p. 15) critica os que veem a lógica como um componente fundamental do raciocínio crítico, pois, para ele, “[...] Lógica pode ajudar a eliminar hipóteses, conjecturas e soluções plausíveis, mas não pode provê-las”. O autor distingue lógica de informação e argumenta que a avaliação de boas razões para a aceitabilidade de determinada crença é largamente determinada por informações. Assim, ele conclui que com o apropriado ceticismo reflexivo, o raciocínio crítico é necessariamente ligado a áreas específicas de conhecimento.

McPeck iniciou um debate entre teóricos que argumentam que raciocínio crítico é uma habilidade geral, não necessariamente ligada a um assunto específico. Existem também os que argumentam que raciocínio crítico poderia ser ambos, geral ou relacionado a determinado assunto (SIEGEL, 1988).

Na análise das abordagens para ensinar raciocínio crítico, Ennis (1993) classifica quatro tipologias:

- Geral: independentemente do conteúdo, as habilidades e atitudes de raciocínio crítico são ensinadas aos alunos.
- Infusão: requer profunda reflexão sobre um assunto e, adicionalmente, princípios gerais de raciocínio crítico são ensinados e aplicados nas atividades desenvolvidas.

- Imersão: os alunos são incentivados a pensar criticamente sobre determinado assunto; porém, os princípios gerais de raciocínio crítico não são apresentados.
- Misto: uma mistura do Geral com Infusão ou Imersão, os alunos estão envolvidos em um curso em que analisam criticamente um assunto com ou sem os conceitos gerais de raciocínio crítico explícitos; porém, há uma parte separada do curso que tem como objetivo ensinar princípios gerais de raciocínio crítico.

Em 1990, McPeck em conjunto com Stephen Norris, Richard Paul e Harvey Siegel publicaram *Ensinando Raciocínio Crítico: Diálogo e Dialética* (*Teaching Critical Thinking: Dialogue and Dialectic*). Nesta obra McPeck defende sua posição e encoraja futuros diálogos.

2.2.3 Matthew Lipman

Matthew Lipman fundou o Instituto para o Avanço da Filosofia para as Crianças na Universidade Estadual de *Montclair* em Nova Jersey (*Institute for The Advancement of Philosophy for Children*) e escreveu muitos livros infantis para serem usados no seu programa.

De acordo com Lipman (2003, p. 43),

Algumas versões de raciocínio crítico parecem estar melhor colocadas sobre a rubrica de filosofia aplicada. Este poderia ser o caso da abordagem educacional conhecida como “Filosofia para Crianças”. É um exemplo claro de filosofia aplicada à educação com o propósito de produzir alunos com maior proficiência em raciocínio e julgamento.

Para o autor, ensinar filosofia para as crianças fundamenta-se na teoria de que esta disciplina tem sido a única a prover critério (princípios lógicos) que torna possível distinguir um bom raciocínio de um mau raciocínio (LIPMAN, 2003).

Analisando as correntes definições de raciocínio crítico, Lipman (2003) as critica como sendo mais focadas nos resultados do que em suas características essenciais. Ele enfatiza que se devem ampliar os resultados, identificando as principais características que definem raciocínio crítico e mostrando as conexões entre elas. Lipman (1988, p. 211-212) define

raciocínio crítico como “[...] pensamento que facilita julgamento porque ele se baseia em critérios, autocorreção e é sensível ao contexto”.

Para Lipman (2003), pensar significa a posse de um conceito, uma opinião ou uma crença, e isso não implica o uso de bom julgamento. Qualquer pensamento que envolve critério, comparação e se baseie em padrões é raciocinar. Em outras palavras, pensamento hábil é raciocinar, e, desde que raciocínio crítico é pensamento hábil, então raciocínio crítico é raciocinar. As opiniões são suportadas com razões e os julgamentos são feitos com base em critérios.

Pode-se dizer que por meio do raciocínio crítico se espera chegar a um bom raciocínio. Racionalidade é uma categoria lógica mais abrangente do que raciocínio crítico. Ennis aponta para os limites da teoria do raciocínio crítico quando analisa a maioria das definições de outros autores que se limitam a atos de julgamento. Não são incluídas outras atividades, as quais Ennis rotula como “atividades de pensamento contribuintes”, ou seja, observar, inferir, conceber alternativas e oferecer uma bem organizada linha de raciocínio (ENNIS, 1990).

2.2.4 Richard Paul

Para Paul (1995, p. 2), existem dois tipos de raciocínio crítico:

Se o pensamento é disciplinado para servir os interesses de um indivíduo ou grupo para a exclusão de outras pessoas ou grupos, nós chamamos de raciocínio crítico sofisticado ou fraco senso. Se o pensamento é disciplinado para levar em conta os interesses de diversas pessoas ou grupos, nós podemos chamar de raciocínio crítico justo ou forte senso.

Richard Paul (1995) trouxe à discussão o assunto sobre as atitudes de pensadores críticos com o conceito de “fraco senso” e “forte senso” de raciocínio crítico. Para o autor, ensinar os alunos a melhor suportar seus argumentos é o mesmo que apenas ajudá-los a se tornarem mais entrosados com seus próprios sistemas de crenças e suas próprias referências. Ao contrário, deve-se ensinar os alunos a questionar seus próprios pontos de vista, encorajando-os a examinar mais proximamente os de outros.

Paul (1995, p. 127) lista dezessete habilidades de raciocínio crítico:

1. Refinar generalizações e evitar simplificações.
2. Comparar situações análogas, transferindo conhecimentos para novos contextos.
3. Desenvolver a perspectiva de alguém: criando ou explorando as implicações das crenças, argumentos ou teorias.
4. Clarificar assuntos, conclusões ou crenças.
5. Clarificar e analisar os significados das palavras e frases.
6. Desenvolver critério para avaliação: clarificando valores e padrões.
7. Avaliar a credibilidade das fontes de informações.
8. Questionar profundamente: perseguindo e aprofundando-se na raiz de questões significantes.
9. Analisar ou avaliar argumentos, interpretações, crenças ou teorias.
10. Gerar ou avaliar soluções.
11. Analisar ou avaliar ações ou políticas.
12. Raciocinar com diálogo: comparando perspectivas, interpretações ou teorias.
13. Racionar dialeticamente: avaliando perspectivas, interpretações ou teorias.
14. Ler criticamente: construindo uma interpretação correta de um texto, entendendo os elementos de pensamento dentro dele e seu raciocínio.
15. Ouvir criticamente: construindo uma interpretação correta de uma comunicação oral, entendendo os elementos de pensamento dentro dela e seu raciocínio.
16. Escrever criticamente: criando, desenvolvendo, clarificando e transmitindo de forma escrita a lógica do pensamento de alguém.
17. Falar criticamente: criando, desenvolvendo, clarificando e transmitindo de forma falada a lógica do pensamento de alguém.

Para o autor, habilidades como estas possuem uma função principal no conceito de raciocínio crítico. Elas são essenciais na abordagem racional de assuntos atuais, problemas e situações, por exemplo, para entender os direitos e deveres de um cidadão. Isto requer que alguém tenha pelo menos a habilidade de comparar perspectivas e interpretações, ler e ouvir criticamente, analisar e avaliar políticas. Na verdade, não há uma macro habilidade na lista que não seria relevante ou até mesmo crucial para se pensar profundamente sobre os direitos e deveres de um cidadão. Da mesma forma, a capacidade de tomar decisões, de participar reconhecidamente no local de trabalho, de funcionar como parte da economia global, de dominar o conteúdo de qualquer coisa tão complexa como disciplinas acadêmicas, de aplicar os conhecimentos destes assuntos em situações reais da vida, de fazer conexões entre disciplinas, de comunicar-se efetivamente, cada uma delas depende de uma maneira fundamental em possuir um número significativo das habilidades listadas (PAUL, 1995).

Para Paul (1995, p. 129), um pensador crítico deve possuir também certas atitudes características, como:

1. Pensar independente.
2. Exercitar pensamento justo.
3. Desenvolver discernimento sobre egocentricidade e sociocentricidade.
4. Desenvolver humildade intelectual e suspensão de julgamento.
5. Desenvolver coragem intelectual.
6. Desenvolver boa-fé e integridade intelectual.
7. Desenvolver perseverança intelectual.

8. Desenvolver confiança na razão.
9. Explorar pensamentos sobre sentimentos e sentimentos sobre pensamentos.
10. Desenvolver curiosidade intelectual.

Para o autor, não é suficiente possuir habilidades de raciocínio crítico; é necessário possuir as atitudes para utilizá-las. O centro da teoria de Paul é a noção de que as pessoas por natureza são egocêntricas e etnocêntricas e que a tendência delas é serem irracionais. O autor afirma ainda que se deve trabalhar muito duro para ser tornar um pensador crítico, pois isto não vem fácil. Outro foco da teoria de Paul é a autocrítica.

Paul insiste que raciocínio crítico pode ser definido de várias maneiras, as quais não devem ser vistas como mutuamente excludentes. Dessa forma, ele se recusa a limitar o conceito de raciocínio crítico a uma única definição, já que seu interesse é desenvolver um conceito inclusivo de raciocínio crítico que permita desenhos e ideias de uma variedade de campos e perspectivas (PAUL, 1995).

Paul disserta que raciocínio crítico requer uma integração de campos cognitivos e afetivos. Para ele, o conteúdo em qualquer disciplina deveria ser visto e ensinado como um modelo de pensamento (por exemplo, história como pensamento histórico, biologia como pensamento biológico) e seu modelo instrucional defende que se deve ensinar os alunos a avaliar seus próprios pensamentos, seja por meio da leitura, escrita, audição ou fala, já que alguém incapaz de avaliar seu próprio pensamento não pode ser considerado um pensador crítico (PAUL, 1995). O modelo de Paul será mais detalhado e analisado no item 2.4 (Modelo Instrucional) desta pesquisa.

2.2.5 Harvey Siegel

Na obra de Harvey Siegel, *Educando Razão: Racionalidade, Raciocínio Crítico e Educação* (*Educating Reason: Rationality, Critical Thinking and Education*), seu argumento é o de que, para procurar razões, alguém tem que reconhecer e se comprometer com os princípios, pois estes são necessários para determinar a relevância e a força das razões (SIEGEL, 1988).

Siegel (1988, p. 23) define as características de raciocínio crítico como:

[...] o foco nas razões e no poder das razões para assegurar ou justificar crenças, alegações e ações. Um pensador crítico, então, é alguém que é apropriadamente movido pelas razões: ele tem a propensão ou atitude de acreditar e agir de acordo com razões; e ele tem a habilidade apropriada para avaliar a força das razões em muitos contextos nos quais elas desempenham uma função.

Siegel (1988) afirma que um pensador crítico deve ter bom entendimento de princípios de assuntos específicos e de assuntos neutros (lógicos) que governam a avaliação das razões e que o deixam apto a utilizá-los. Logo, para Siegel, o debate sobre a generalização ou não de habilidades de raciocínio crítico não é necessário.

O autor disserta que para ser um pensador crítico é necessário possuir (1988, p. 39): “[...] certas atitudes, temperamentos, hábitos mentais e traços característicos [...]”. O autor continua sua obra enfatizando que não é o bastante saber avaliar razões, deve-se também estar disposto a isto. No mesmo sentido, também não é o bastante possuir habilidades de raciocínio crítico, deve-se estar disposto a usá-las (SIEGEL, 1988).

2.2.6 “O Relatório Delphi” – Peter A. Facione

Na década de 1980, a discussão sobre o papel do raciocínio crítico na área educacional dos Estados Unidos da América aumentou. Em dezembro de 1987, a Associação Americana de Filosofia (*American Philosophical Association*) pediu para que o Prof. Peter A. Facione investigasse sistematicamente os conceitos e as avaliações de raciocínio crítico na opinião de diversos especialistas sobre o assunto.

A pesquisa utilizou uma metodologia conhecida como método Delphi, o qual requer a formação de um painel interativo de especialistas. Enfatiza-se que os participantes devem estar dispostos a compartilhar seu conhecimento e a trabalhar em busca de uma resolução consensual sobre assuntos e opiniões. Para tanto, foram convidados 46 especialistas sobre raciocínio crítico, dos quais 52% eram da área da Filosofia, 22% da área da Educação, 20% das Ciências Sociais e 6% das Ciências Físicas (FACIONE, 1990).

Dos autores abordados nos tópicos anteriores da revisão bibliográfica desta pesquisa, participaram do denominado *O Relatório Delphi* (*The Delphi Report*): Robert Ennis, Matthew Lipman e Richard Paul. Como na pesquisa empírica desta tese o modelo instrucional e os testes de avaliação de habilidades e atitudes de raciocínio crítico utilizados estão em harmonia

com os resultados deste relatório, suas principais definições e recomendações serão expostas neste tópico.

O relatório está organizado em quatro seções, sendo que na primeira ele aborda a necessidade do estudo e conclui com uma definição consensual sobre raciocínio crítico e sobre “o Pensador Crítico Ideal”. Na segunda, ele descreve a metodologia Delphi de pesquisa. Na terceira, aborda as habilidades de raciocínio crítico e, na última, as atitudes. Além disso, o autor tece quinze recomendações relacionadas ao ensino e à avaliação do raciocínio crítico. Na seção I do relatório, Facione (1990, p. 2) faz a seguinte definição consensual sobre o raciocínio crítico e o Pensador Crítico Ideal:

Nós entendemos raciocínio crítico como julgamento objetivo e auto regulatório que resulta em interpretação, análise, avaliação e inferência, e também como uma explicação das considerações evidenciais, conceituais, metodológicas, criteriológicas ou contextuais sobre o que o julgamento está baseado. Raciocínio crítico é essencial como uma ferramenta de investigação. Tal como, raciocínio crítico é uma força liberal na educação e um recurso poderoso na vida pessoal e cívica de alguém. Enquanto não sinônimo de bom pensamento, raciocínio crítico é um fenômeno humano presente e de autocorreção. O pensador crítico ideal é habitualmente investigador, bem informado, confiante na razão, mente aberta, flexível, justo na avaliação, honesto em enfrentar vieses pessoais, prudente em fazer julgamentos, disposto a reconsiderar, claro sobre assuntos, organizado em problemas complexos, especialista em procurar informação relevante, racional na seleção de critérios, focado na investigação e persistente em procurar resultados que são tão precisos quanto os assuntos e as circunstâncias da investigação permitam. Então, educar bons pensadores críticos significa trabalhar em busca deste ideal. Isto combina o desenvolvimento de habilidades de raciocínio crítico com a criação destas atitudes, as quais produzem constantemente ideias úteis que são à base da sociedade racional e democrática.

Na seção II do relatório é exposta a metodologia Delphi de pesquisa, na qual relata que os trabalhos tiveram início em fevereiro de 1988 e terminaram em novembro de 1989. Os especialistas, que participaram de seis rodadas de questões, trabalharam em busca de um consenso compartilhando suas opiniões racionais, dispostos a reconsiderá-las de acordo com comentários, objeções e argumentos oferecidos pelos outros especialistas. Para não gerar tipo algum de problema relacionado ao status ou à área dos especialistas, cada rodada de questões foi iniciada pelo diretor do projeto e todas as respostas foram coordenadas por ele (FACIONE, 1990).

Facione (1990, p. 3) relatou a história do projeto na seção II:

Rodada 1 (11 de fevereiro de 1988) e Rodada 2 (14 de março de 1988) iniciado o processo Delphi. Em ambas as rodadas os participantes foram convidados a nominar outros especialistas em raciocínio crítico para se unirem a este projeto de pesquisa. Os especialistas chegaram a um consenso sobre a suposição do trabalho que “o conceito de raciocínio crítico deveria ser operacional para tornar possível que partes importantes do raciocínio crítico pudessem ser

mensuradas válida e confiavelmente”. Os especialistas concordaram em começar suas análises de raciocínio crítico “identificando os principais elementos de raciocínio crítico que devem racionalmente ser esperados nos calouros e estudantes de faculdades em geral”.

O raciocínio para esta decisão foi que o nível de construção teórica da faculdade para raciocínio crítico poderia ser racionalmente utilizado para guiar o que poderia ser dito sobre raciocínio crítico no ensino fundamental. Os especialistas também notaram que a maioria dos participantes tinha mais experiência na educação em nível superior do que fundamental.

Rodada 3 (4 de maio de 1988) foi um convite aberto-fechado para especialistas escreverem sua própria lista de operações as quais eles entendem como centrais para raciocínio crítico. A primeira síntese deste input foi apresentada para revisão dos especialistas na Rodada 4 (23 de setembro de 1988). Esta síntese focou na dimensão de habilidades de raciocínio crítico. A Rodada 4 convidou respostas considerando cada habilidade e sub-habilidade identificada, um proposto (e ultimamente rejeitado) modelo de input/output de operações de raciocínio crítico, uma lista de operações cognitivas, as quais podem ou não serem distintas de raciocínio crítico, uma afirmação geral considerando o que uma habilidade é e como ela pode ser ensinada e uma lista de precauções relacionadas à instrução e avaliação de raciocínio crítico.

Rodada 5A (28 de fevereiro de 1989) revisão das definições e classificações das habilidades de raciocínio crítico à luz das respostas dos especialistas da Rodada 4. Rodada 5B (também em 28 de fevereiro de 1989) afirmações propostas de acordo com a dimensão de atitude de raciocínio crítico e sobre suas possíveis conotações normativas. Rodada 5C (10 de março de 1989) requisitada para específicas recomendações sobre instrução e avaliação de raciocínio crítico e oferecida uma revisão das afirmações gerais sobre ensinar e avaliar uma habilidade cognitiva. Rodada 5 incluiu muitas citações, comentários e reações sobre as últimas respostas reunidas pelos especialistas.

Os comentários dos especialistas sobre as várias citações incluídas em cada rodada ajudaram muito o diretor do projeto a entender as visões dos especialistas sobre o assunto. Com estas e com as respostas específicas das Rodadas 5A, 5B e 5C o diretor do projeto elaborou uma primeira versão do relatório com todos os achados, inclusive com recomendações. Rodada 6 (25 de setembro de 1989) circulou o relatório e foi dada aos especialistas de raciocínio crítico a oportunidade de expressar suas visões ou tecer comentários para inclusão no relatório final, o qual passou por sua última revisão em novembro de 1989.

Como comentado na definição consensual sobre raciocínio crítico, os especialistas definem as habilidades e atitudes de um pensador crítico. Na seção III do relatório é exposta uma tabela com seis habilidades de raciocínio crítico e suas sub-habilidades. Cada uma das seis habilidades é considerada central para o desenvolvimento de raciocínio crítico. Uma explicação detalhada de cada uma delas é também realizada com vistas a facilitar a elaboração de modelos de instrução e avaliação. Com base em Facione (1990), segue Quadro 1 com a lista consensual das habilidades e sub-habilidades de raciocínio crítico:

Quadro 1 - Lista consensual de habilidades e sub-habilidades de raciocínio crítico.

Habilidades	Sub-habilidades
Interpretação	Categorização
	Decodificando Significância
	Elucidando Significado
Análise	Examinando Ideias
	Detectando Argumentos
	Analizando Argumentos
Avaliação	Avaliando Alegações
	Avaliando Argumentos
Inferência	Questionando Evidência
	Conjecturando Alternativas
	Esboçando Conclusões
Explicação	Afirmando Resultados
	Justificando Procedimentos
	Apresentando Argumentos
Autorregulação	Autoexame
	Autocorreção

Fonte: FACIONE, 1990, p. 6.

Os especialistas entraram em grande consenso na descrição de cada uma dessas habilidades e sub-habilidades de raciocínio crítico. Algumas descrições são acompanhadas por exemplos, porém, o autor ressalta que eles servem para clarificar os itens e que não foram objeto de análise dos especialistas. Facione (1990) ressalta que analisando cada uma das descrições pode-se chegar a estratégias de instrução e avaliação de raciocínio crítico.

Como o modelo de desenvolvimento de raciocínio crítico de Richard Paul utilizado na parte empírica desta pesquisa fundamenta-se no *Relatório Delphi* (FACIONE, 1990, p. 6-11), será disponibilizada no Quadro 2 a seguir a lista de descrições e exemplos de cada uma das habilidades e sub-habilidades de raciocínio crítico. Sua leitura para compreensão das habilidades e sub-habilidades e elaboração de atividades em sala de aula tem se mostrado de extrema importância. Como a maioria dos itens possui exemplos, torna-se mais simples a criação de atividades relacionadas à área contábil que contemplem estas habilidades e sub-habilidades.

Quadro 2 - Descrições consensuais das principais habilidades e sub-habilidades de raciocínio crítico.

1. INTERPRETAÇÃO: Compreender e expressar o significado ou significância de uma larga variedade de experiências, situações, dados, eventos, julgamentos, convenções, crenças, regras, procedimentos ou critérios.

1.1. CATEGORIZAÇÃO:

- Entender ou apropriadamente formular categorias, distinções ou estruturas para entendimento, descrição ou caracterização de informações.
- Descrever experiências, situações, crenças, eventos etc., os quais remetam a compreensíveis significados em termos de categorizações, distinções ou estruturas apropriadas.

Por exemplo: reconhecer um problema e definir suas características sem prejudicar a investigação; determinar uma maneira útil de classificar e subclassificar informações; elaborar um relatório compreensível sobre a experiência de alguém em determinada situação; classificar dados, achados ou opiniões utilizando um determinado esquema de classificação.

1.2. DECODIFICANDO SIGNIFICÂNCIA:

- Detectar, assistir e descrever o conteúdo informativo, as pretensões, as funções diretivas, as intenções, os motivos, os propósitos, a significância social, os valores, as visões, as regras, os procedimentos, os critérios ou as relações inferenciais expressas em sistemas de comunicação baseados em convenções, tais como na linguagem, comportamentos sociais, desenhos, números, gráficos, tabelas, mapas, sinais e símbolos.

Por exemplo: detectar e descrever os propósitos de uma pessoa em perguntar uma determinada questão; apreciar a significância de uma particular expressão facial ou gesto utilizado em determinada situação social; discernir o uso de ironia ou questões retóricas em um debate; interpretar os dados apresentados utilizando uma forma particular de instrumentação.

1.3. ELUCIDANDO SIGNIFICADO:

- Parafrasear ou tornar explícito por meio de descrições, analogias ou expressões figuradas os significados contextuais, convencionais ou pretendidos de palavras, ideias, conceitos, afirmações, comportamentos, desenhos, números, sinais, mapas, gráficos, símbolos, regras, eventos ou cerimônias.
- Usar descrições, analogias ou expressões figuradas para remover incertezas ou ambiguidades confusas ou para esboçar um procedimento racional que faça isto.

Por exemplo: reafirmar o que uma pessoa disse utilizando palavras ou expressões diferentes preservando os significados que a pessoa atribuiu; encontrar um exemplo que ajude explicar alguma coisa a alguém; desenvolver uma distinção que deixe claras as diferenças conceituais ou remova os problemas de ambiguidade.

- 2. ANÁLISE:** Identificar os relacionamentos inferenciais pretendidos e realizados dentre afirmações, questões, conceitos, descrições ou outras formas de representações que pretendem expressar crenças, julgamentos, experiências, razões, informações ou opiniões.

2.1. EXAMINANDO IDEIAS:

- Determinar a função que várias expressões exercem ou é esperado que exerçam no contexto de um argumento, raciocínio ou persuasão.
- Definir termos.
- Comparar ou contrastar ideias, conceitos ou afirmações.
- Identificar assuntos ou problemas e determinar suas partes componentes e também identificar os relacionamentos conceituais entre elas e com o todo.

Por exemplo: identificar uma frase que pretende causar uma resposta emocional simpática, a qual deve induzir uma plateia a concordar com a opinião; examinar bem de perto propostas relacionadas a determinado problema e determinar seus pontos de similaridade e divergência; dada uma complicada tarefa, determinar como ela pode ser quebrada em partes menores, mais gerenciáveis; definir um conceito abstrato.

2.2. DETECTANDO ARGUMENTOS:

- Dado um conjunto de afirmações, descrições, questões ou representações gráficas, determinar se o conjunto expressa ou pretende expressar uma razão ou razões que suportam ou contestam alguma alegação, opinião ou ponto de vista.

Por exemplo: dado um parágrafo, determinar se uma leitura padrão dele no contexto de como e onde ele está publicado, sugeriria que ele apresenta uma alegação ou uma razão ou razões que suportam aquela alegação; dada uma passagem de um editorial de um jornal, determinar se o autor pretendeu que a mesma fosse uma expressão das razões ou contra uma determinada alegação ou opinião; dado um anúncio comercial, identificar qualquer alegação que está junta com razões apresentadas em seu suporte.

2.3. ANALISANDO ARGUMENTOS:

- Dada a expressão de uma razão ou razões que pretendem suportar ou contestar alguma alegação, opinião ou ponto de vista, identificar e diferenciar: (a) a pretendida conclusão principal, (b) as premissas e razões colocadas como suporte da conclusão principal, (c) outras premissas e razões colocadas como suporte das premissas e razões que suportam a conclusão principal, (d) elementos adicionais não expressados sobre aquele raciocínio, tais como conclusões intermediárias, suposições não afirmadas ou pressupostas, (e) a estrutura geral do argumento ou pretendida corrente de raciocínio, e (f) qualquer item contido no corpo das expressões sendo examinadas, o qual não se pretendeu como parte do raciocínio expressado ou do seu suporte.

Por exemplo: dado um breve argumento, um argumento do tamanho de um parágrafo ou a posição de um artigo sobre um assunto social controverso, identificar a alegação principal do autor, as razões e premissas que o autor utiliza em defesa da alegação, as bases das informações utilizadas para suportar aquelas razões e premissas, suposições cruciais implícitas no raciocínio do autor; dadas diversas razões ou correntes de razões em suporte a uma particular alegação, desenvolver uma representação gráfica que caracterize plenamente o fluxo inferencial daquele raciocínio.

3. AVALIAÇÃO: Avaliar a credibilidade de afirmações ou outras representações que são contas ou descrições da percepção, experiência, situação, julgamento, crença ou opinião de uma pessoa; e avaliar a força lógica dos relacionamentos inferenciais realizados ou esperados entre afirmações, descrições, questões ou outras formas de representação.

3.1. AVALIANDO ALEGAÇÕES:

- Reconhecer os fatores relevantes para avaliar o grau de credibilidade para atribuir a uma fonte de informação ou opinião.
- Avaliar a relevância contextual de questões, informações, princípios, regras ou procedimentos.
- Avaliar a aceitabilidade, o nível de confiança da probabilidade ou verdade de qualquer representação de uma experiência, situação, julgamento, crenças ou opinião dada.

Por exemplo: reconhecer os fatores que tornam uma pessoa testemunha confiável em relação a um determinado evento ou autoridade confiável em relação a um determinado tópico; determinar se um dado princípio de conduta é aplicável para decidir o que fazer em determinada situação; determinar se uma dada alegação pode ser verdadeira ou falsa baseada no que alguém sabe ou pode racionalmente descobrir.

3.2. AVALIANDO ARGUMENTOS:

- Julgar se a aceitabilidade assumida de premissas de um determinado argumento de alguém justifica ou não como verdade (dedutivamente certo) ou muito provavelmente verdade (indutivamente justificado), a conclusão expressa daquele argumento.
- Antecipar ou criar questões ou objeções e avaliar se elas apontam para significantes fraquezas dos argumentos que estão sendo avaliados.
- Determinar se um argumento se baseia ou não em suposições ou pressupostos falsos ou duvidosos e, então, determinar o quão crucialmente eles afetam sua força.
- Julgar entre racionais e falsas inferências.
- Julgar a força probatória das premissas ou suposições de um argumento com uma visão de determinar a aceitabilidade do argumento.
- Determinar e julgar a força probatória esperada ou não de um pretendido argumento, com uma visão de julgar a aceitabilidade do argumento.

- Determinar a extensão que cada possível informação adicional pode reforçar ou enfraquecer um argumento.

Por exemplo: dado um argumento, julgar se as suas conclusões seguem com certo grau de certeza e confiança as suas premissas; checar por identificáveis falácias formais e informais; dada uma objeção a um argumento, avaliar a força lógica daquela objeção; avaliar a qualidade e aplicabilidade de argumentos analógicos; julgar a força lógica de argumentos fundamentados em situações hipotéticas ou raciocínios causais; julgar se um determinado argumento é relevante, aplicável ou tem implicações para a situação em questão; determinar como possíveis novos dados poderiam guiar logicamente para futuras confirmações ou não confirmações de uma dada opinião.

- 4. INFERÊNCIA:** Identificar e assegurar elementos necessários para elaborar conclusões racionais; elaborar conjecturas e hipóteses; considerar informações relevantes e extrair os resultados advindos dos dados, afirmações, princípios, evidências, julgamentos, crenças, opiniões, conceitos, descrições, questões ou outras formas de representações.

4.1. QUESTIONANDO EVIDÊNCIA:

- Em particular, reconhecer premissas que requerem suporte e formular uma estratégia para procurar e levantar informação que pode subsidiar aquele suporte.
- Em geral, julgar a informação relevante para decidir a aceitabilidade, plausibilidade ou méritos relativos de uma determinada alternativa, questão, assunto, teoria, hipótese ou afirmação requerida e determinar estratégias de investigação plausíveis para adquirir aquela informação.

Por exemplo: ao se tentar desenvolver um argumento persuasivo em suporte à opinião de alguém, julgar qual base de informações seria útil ter e desenvolver um plano que produziria resposta clara tendo ou não determinadas informações disponíveis; após julgar que a falta de certas informações seria crucial em determinar se dada opinião é mais ou menos racional do que uma opinião competidora, planejar uma procura que revelará se aquela informação está disponível.

4.2. CONJECTURANDO ALTERNATIVAS:

- Formular múltiplas alternativas para resolver um problema, postular uma série de suposições de acordo com uma questão, projetar hipóteses alternativas de acordo com um evento, desenvolver uma variedade de diferentes planos para atingir determinado objetivo.
- Elaborar pressupostos e projetar a área de possíveis consequências de decisões, posições, políticas, teorias ou crenças.

Por exemplo: dado um problema com ramificações técnicas, éticas ou orçamentárias, desenvolver uma série de opções para direcionar e resolver aquele problema; dada uma série de prioridades com as quais alguém pode ou não concordar, projetar dificuldades e benefícios que são normalmente esperados se aquelas prioridades forem adotadas em um processo decisório.

4.3. ESBOÇANDO CONCLUSÕES:

- Aplicar modos de inferência apropriados para determinar qual posição, opinião ou ponto de vista deveria ser escolhido em um determinado problema ou assunto.
- Dada uma série de afirmações, descrições, questões ou outras formas de representações, deduzir, com alto grau de força lógica, seus relacionamentos inferenciais e as consequências ou os pressupostos aos quais elas suportam, garantem ou implicam.
- Empregar com sucesso várias subespécies de raciocínio, como, por exemplo, raciocinar analogicamente, aritmeticamente, dialeticamente, cientificamente etc.
- Determinar qual de diversas possíveis conclusões é mais fortemente garantida ou suportada pelas evidências em mãos, ou qual deveria ser rejeitada ou considerada como menos plausível de acordo com determinada informação.

Por exemplo: desenvolver experimentos e aplicar técnicas apropriadas de inferências estatísticas em ordem de confirmar ou não uma hipótese empírica; dado um assunto controverso examinar diferentes opiniões, considerar várias visões opostas e razões decorrentes das mesmas, levantar informação relevante e formular uma opinião própria sobre o assunto em questão; deduzir um teorema advindo de axiomas utilizando regras prescritas de inferência.

- 5. EXPLANAÇÃO:** Afirmar os resultados do raciocínio de alguém; justificar aquele raciocínio em termos de considerações evidenciais, conceituais, metodológicas, criteriológicas e contextuais sobre as quais aquele resultado foi baseado; apresentar o raciocínio de alguém de forma que fique claramente exposto.

5.1. AFIRMANDO RESULTADOS:

- Produzir afirmações, descrições, ou representações com acurácia dos resultados das atividades de raciocínio de alguém para que seja possível analisar, avaliar, inferir ou monitorar estes resultados.

Por exemplo: afirmar as razões de alguém para manter uma determinada visão; escrever sobre seu próprio futuro utilizando um pensamento atual sobre um importante ou complexo problema; afirmar os achados da pesquisa de alguém; buscar a análise e julgamento de alguém sobre um trabalho de arte; afirmar a opinião de alguém sobre um problema prático de emergência.

5.2. JUSTIFICANDO PROCEDIMENTOS:

- Apresentar as considerações evidenciais, conceituais, metodológicas, criteriológicas e contextuais, as quais alguém utiliza para formar suas interpretações, análises, avaliações ou inferências para que alguém possa com acurácia gravar, avaliar, descrever ou justificar aqueles processos para ele mesmo ou para outras pessoas, como também para remediar deficiências de uma maneira geral.

Por exemplo: manter um diário dos passos seguidos em um trabalho durante um longo e difícil problema ou procedimento científico; explicar a escolha de alguém por um particular teste estatístico para propósitos de análise de dados; afirmar os padrões que alguém utilizou para avaliar um fragmento da literatura; explicar como alguém compreende um conceito-chave quando clareza

conceitual é crucial para futuros progressos de um determinado problema; mostrar que os pré-requisitos para o uso de uma determinada técnica metodológica foram satisfeitos; informar a estratégia utilizada na tentativa de tomar uma decisão de forma racional; desenhar um gráfico que represente a informação quantitativa ou espacial utilizada como evidência.

5.3. APRESENTANDO ARGUMENTOS:

- Dar razões para aceitar uma alegação.
- Encontrar objeções para o método, conceitualizações, evidência, critério ou contextualizações utilizadas para os julgamentos inferenciais, analíticos e avaliadores.

Por exemplo: escrever um artigo no qual alguém argumenta sobre uma dada posição ou política; antecipar e responder a críticas racionais que alguém poderia esperar que fossem geradas contra suas visões políticas; identificar e expressar evidência e contra evidência que pretende contribuir dialeticamente para seu próprio pensamento ou de outra pessoa sobre um assunto de profunda preocupação pessoal.

6. AUTORREGULAÇÃO: Monitorar conscientemente suas próprias atividades cognitivas, os elementos utilizados nestas atividades e resultados elucidados, particularmente por meio de habilidades de análise e avaliação de julgamentos inferenciais com uma visão de questionar, confirmar, validar ou corrigir seu raciocínio ou resultados.

6.1. AUTOEXAME:

- Refletir sobre seu próprio raciocínio e verificar ambos, o resultado produzido e a correta aplicação e execução das habilidades cognitivas envolvidas.
- Fazer uma objetiva e refletida auto avaliação meta-cognitiva de suas opiniões e razões para mantê-las.
- Julgar a extensão que o pensamento de alguém é influenciado pelas deficiências de seu conhecimento, ou por estereótipos, preconceitos, emoções ou qualquer outro fator, o qual coage a objetividade ou racionalidade de alguém.
- Refletir sobre as motivações, valores, atitudes e interesses de alguém com a visão de que este alguém tem sido tentado a ser livre de vieses, justo no pensar, minucioso, objetivo, respeitador da verdade e racional em suas análises, interpretações, avaliações, inferências ou expressões.

Por exemplo: examinar a visão de alguém sobre um assunto controverso com sensibilidade para possíveis influências de vieses pessoais ou interesses próprios; revisar a metodologia de alguém ou calcular com uma visão de detectar aplicações erradas ou inadvertidos erros; ler fontes para assegurar que alguém não deixou para trás informação importante; identificar e revisar a aceitabilidade dos fatos, opiniões ou suposições que caem sobre determinado ponto de vista; identificar e revisar as razões de alguém e seu processo de raciocínio em chegar à determinada conclusão.

6.2. AUTOCORREÇÃO:

- Onde autocorreção revela erros e deficiências, desenhar procedimentos racionais para remediar ou corrigir, caso possível, estes enganos e suas causas.

Por exemplo: dado um erro metodológico ou deficiência fatural no trabalho de alguém, revisá-lo com o objetivo de corrigir o problema e determinar se as revisões garantem mudanças em qualquer posição, achado ou opinião baseada no problema em questão.

Fonte: FACIONE, 1990, p. 6-11.

Analisando-se o Quadro 2 é possível verificar que o instrumento utilizado para avaliar as habilidades (EWCTET) está em sintonia com as definições e descrições das habilidades e sub-habilidades de raciocínio crítico. O Capítulo 3 desta pesquisa detalha os instrumentos utilizados para avaliar as habilidades e atitudes de raciocínio crítico dos alunos.

Na seção IV do relatório é abordada a dimensão relativa às atitudes de raciocínio crítico. A ideia central é a de que não adianta determinada pessoa possuir as habilidades de raciocínio crítico se não possuir as atitudes certas para utilizá-las. Portanto, Facione (1990, p. 12-13) comenta que os especialistas definiram as atitudes de raciocínio crítico como as atitudes de um bom pensador crítico:

Para os especialistas, um bom pensador crítico, o caso paradigma, é habitualmente disposto a se engajar em julgamento crítico e encorajar os outros a fazerem o mesmo. Está apta a fazer tais julgamentos em uma larga área de contextos e para uma larga variedade de propósitos. Embora talvez nem sempre seja o mais importante na cabeça de alguém, as justificativas racionais para cultivar estas atitudes afetivas as quais caracterizam o paradigma do pensador crítico são altamente baseadas em valores pessoais e cívicos de raciocínio crítico. Raciocínio crítico é conhecido por contribuir para as análises de pensamento justo e resoluções de questões. Raciocínio crítico é uma poderosa ferramenta na procura por conhecimento. Raciocínio crítico pode ajudar as pessoas a superar a cega, sofista ou irracional defesa da intelectualidade defeituosa ou de opiniões enviesadas. Raciocínio crítico promove autonomia racional, liberdade intelectual e a investigação objetiva, racional e baseada em evidências de uma área muito ampla de assuntos e preocupações pessoais e sociais.

Na seção IV, Facione (1990, p. 13) apresenta um quadro (assim como o Quadro 3, a seguir) das atitudes de raciocínio crítico divididas em abordagens para a vida em geral e abordagens para assuntos, questões ou problemas específicos:

Quadro 3 - Atitudes de raciocínio crítico.

ABORDAGENS PARA A VIDA EM GERAL

- Curiosidade em relação a uma vasta área de assuntos.
- Preocupação em ficar e permanecer geralmente bem informado.
- Alerta para oportunidades de utilizar raciocínio crítico.
- Confiança no processo de investigação racional.
- Autoconfiança na própria habilidade de raciocinar.
- Mente aberta de acordo com divergentes visões do mundo.
- Flexibilidade em considerar alternativas e opiniões.
- Entendimento das opiniões de outras pessoas.
- Pensamento justo em apreciar raciocínio.
- Honestidade em encarar suas próprias tendências de vieses, preconceitos, estereótipos, egocentrismo e sociocentrismo.
- Prudente em suspender, fazer ou alterar julgamentos.
- Disposto a reconsiderar e a revisar visões onde uma reflexão honesta sugere que a mudança é garantida.

ABORDAGENS PARA ASSUNTOS, QUESTÕES OU PROBLEMAS ESPECÍFICOS

- Clareza em afirmar a questão ou preocupação.
- Organização em trabalhar com complexidade.
- Diligente em procurar por informação relevante.
- Racionalidade em selecionar e aplicar critério.
- Cuidado em focar atenção na preocupação em mãos.
- Persistência quando as dificuldades são encontradas.
- Precisão em relação ao grau permitido pelo assunto e à circunstância.

Fonte: FACIONE, 1990, p. 13.

Em relação ao Quadro 3, a maioria dos especialistas (61%) concordou com as atitudes como parte da contextualização de raciocínio crítico. O consenso (83%) é que bons pensadores críticos podem ser caracterizados por exibir estas atitudes (FACIONE, 1990).

Pode-se dizer que desde a.C. o raciocínio crítico é enfatizado como uma habilidade importante. Analisou-se a história da evolução do debate sobre o assunto até os dias atuais, visitando os principais teóricos e especialistas da área. Analisou-se e reportou-se à lista consensual sobre a definição de raciocínio crítico e de suas habilidades e atitudes, resultado de anos de debate e de uma metodologia que envolveu quarenta e seis estudiosos do assunto, dentre eles alguns citados e estudados na presente obra. O modelo instrucional utilizado na pesquisa está em harmonia com a maior parte do que foi estabelecido como consensual no *Relatório Delphi*, o que ajudará na sua aplicação e interpretação. As teorias descritas e abordadas, além de importantes para contextualizar e melhor compreender o assunto, servirão como base teórica para análise dos dados coletados e elaboração das principais conclusões.

2.2.7 Definição Proposta

Considerando-se a definição consensual do *Relatório Delphi*, as definições dos autores citados, uma forte influência do modelo de Richard Paul e da definição de Matthew Lipman, a seguinte definição foi proposta aos participantes desta pesquisa:

Raciocínio crítico é pensamento que facilita julgamento, pois identifica, analisa e julga os elementos de um raciocínio. Os elementos de um raciocínio são: objetivo, problema, informações, interpretações, conceitos, suposições, implicações e ponto de vista.

Os elementos de raciocínio crítico são detalhados e explicados no item 2.4 desta pesquisa.

2.3 Avaliando Raciocínio Crítico

A necessidade de se avaliar o raciocínio crítico nasceu da questão de como saber se uma intervenção é mais benéfica do que outra no desenvolvimento de habilidades e atitudes. Existem testes padrões e muitas pesquisas sobre a validade e confiabilidade destes testes (ABRAMI *et al.*, 2008).

Ennis (1993), em um de seus artigos, descreve justificativas e definições de raciocínio crítico e lista uma série de testes de avaliação publicados e comercializados com diversos propósitos, dentre eles o da pesquisa com alunos. O autor disserta sobre as instruções para desenvolver raciocínio crítico e as ferramentas para avaliá-lo. Também faz uma crítica dizendo que tanto as instruções quanto as ferramentas foram negligenciadas por um bom período, embora a importância de desenvolver habilidades venha sendo discutida desde o início do século XX.

De acordo com Ennis (1993, p. 180-181), os principais propósitos dos testes de raciocínio crítico são:

1. Avaliar o nível de raciocínio crítico dos alunos. Se nós conhecemos onde focar nossa instrução, nós devemos “começar onde eles estão” em específicos aspectos de raciocínio crítico. Testes podem ajudar bastante mostrando específicas áreas de força e fraqueza (por exemplo, habilidade de identificar suposições).
2. Dando aos alunos resposta sobre seu progresso relacionado ao raciocínio crítico. Se os alunos conhecem suas específicas fraquezas e forças, suas tentativas de melhorar podem ser focadas.
3. Motivando os alunos a serem melhores em relação ao raciocínio crítico. Embora frequentemente não utilizados como um mecanismo de motivação, testes podem e motivam os alunos a aprenderem sobre a matéria que vai ser coberta pelo teste. Se raciocínio crítico é omitido dos testes, baterias de testes, ou outros procedimentos de avaliação, os alunos tenderão a negligenciar os mesmos (Smith, 1991; Shepard, 1991).
4. Informar os professores sobre o sucesso de seus esforços para ensinar os alunos a pensar criticamente. Professores podem utilizar testes para obter respostas sobre suas instruções sobre raciocínio crítico.
5. Fazer pesquisa sobre assuntos e questões de instrução de raciocínio crítico. Sem cuidadosa comparação de uma variedade de abordagens, a dificuldade relacionada a assuntos de instrução e organização curricular referentes a raciocínio crítico não podem ser respondidas. Mas se a pesquisa requer avaliação, então as comparações podem ser feitas.
6. Prover ajuda em decidir se um aluno deveria entrar em programa educacional. Pessoas em diversos campos já utilizaram a avaliação de raciocínio crítico para ajudar nas decisões de admissões. Exemplos são: medicina, enfermagem, direito e faculdades em geral. A ideia parece boa, mas a eficácia dos esforços existentes em selecionar melhores pensadores críticos não tem sido estabelecida. Pesquisas precisam ser realizadas nesta área.
7. Provendo informação para as escolas proprietárias responsáveis sobre o progresso de raciocínio crítico de seus alunos. Um atual e popular propósito dos testes inclui a pressão que escolas e professores recebem de avaliar os resultados de seus alunos e informar a parte proprietária e responsável pela instituição como um todo.

Pensando no propósito desta pesquisa, o item 5 se encaixa perfeitamente; porém, os anteriores também fazem parte da aplicação do modelo instrucional de desenvolvimento de raciocínio crítico. Já os itens 6 e 7 referem-se a uma realidade de escolas, faculdades ou universidades dos EUA.

Ligadas aos propósitos existem algumas armadilhas na utilização dos testes para avaliação de raciocínio crítico, as quais estão ligadas principalmente às questões de validade e confiabilidade dos testes e da forma que são aplicados. Sobre este assunto, Ennis (1993, p. 181) lista e explica as principais armadilhas que os educadores ou pesquisadores precisam estar atentos:

1. Resultados de testes podem ser comparados com normas e a alegação de diferença ou similaridade é o resultado da instrução. Há normalmente outras possíveis explicações do resultado, tais como influências dos vizinhos. A atual responsabilidade e cobrança de prestar responsabilidade sobre os testes nos atraem para esta armadilha.
2. Um pré-teste e um pós-teste podem ser dados sem comparar a classe com um grupo controle. A falta de um grupo controle torna os resultados da pesquisa dúbios, já que muitas outras coisas além da instrução aconteceram aos alunos e podem contar para os resultados.
3. O uso do mesmo teste para o pré-teste e pós-teste tem o problema de alertar os alunos para as questões do teste. Por outro lado, o uso de diferentes formas do mesmo teste para comparações do pré-teste e pós-teste, dado que o teste é para raciocínio crítico, é provavelmente pior, já que diferentes formas são na verdade diferentes testes. Comparabilidade é sempre suspeita, já que muito depende do conteúdo específico do teste.
4. A maioria dos testes de raciocínio crítico não é abrangente, especialmente aqueles que são mais fáceis de utilizar, os testes de múltipla escolha. Estes testes tipicamente perdem muito do que é importante para avaliar raciocínio crítico.
5. Outro problema no uso de testes de múltipla escolha está na diferença de crenças e suposições entre o elaborador do teste e o participante do teste. Já que um pensador crítico emprega um bom conhecimento da situação, crenças diferentes sobre esta podem algumas vezes resultar em justificáveis respostas diferentes para as questões do teste (Norris & Ennis, 1989).
6. Resultados significativos podem ser esperados em um curto período de tempo. Aprender a pensar criticamente requer um longo tempo. É requerida muita prática reflexiva com muitos exemplos em uma variedade de situações.
7. Propósitos de admissão ou avaliação oficial geralmente interferem na validade do teste. Isto acontece parcialmente por causa de escolas que ensinam seus alunos como se sair bem sem a preocupação de realmente medir o progresso do raciocínio crítico dos mesmos. Os alunos geralmente aprendem truques para fazer os testes.
8. Escassez de recursos (indicado pelo baixo orçamento para os testes e por professores com sobrecarga de trabalho) geralmente leva a concessões que afetam a validade do teste. Por causa das despesas do teste ou das despesas relacionadas ao tempo necessário dos professores para avaliar os testes, muitos testes em última instância tem aplicado questões de múltipla escolha, as quais são menos válidas do que pequenas respostas, redações e testes de desempenho de raciocínio crítico.

No Capítulo 3 da presente pesquisa são detalhados os procedimentos metodológicos da parte empírica, nos quais se torna evidente a preocupação com as questões referentes à validade e confiabilidade dos resultados esperados. As armadilhas relacionadas anteriormente serviram de guia para o desenho da pesquisa de campo e escolha dos instrumentos de avaliação de habilidades e atitudes de raciocínio crítico. Ainda em se tratando da confiabilidade, validade e imparcialidade dos instrumentos de avaliação de raciocínio crítico, o *Relatório Delphi* (1990, p. 16) diz: “Avaliando a aceitabilidade de uma estratégia ou instrumento de raciocínio crítico

se deve considerar a validade do conteúdo, a validade da construção, a confiabilidade e a imparcialidade.”

Facione (1990, p. 16-17) detalha todos os itens que devem ser considerados na aceitabilidade de uma estratégia ou instrumento de raciocínio crítico:

1. Validade do Conteúdo: A estratégia ou instrumento deveria ser baseado em um apropriado conceito de raciocínio crítico e um claro entendimento de quais aspectos de raciocínio crítico a avaliação objetiva. Cada tarefa ou questão deveria ser avaliada para assegurar que respondendo corretamente aquele item não é uma questão de rota de aprendizado ou repetição de uma informação tanto para a sala de aula ou para propósitos educacionais mais abrangentes, a avaliação de raciocínio crítico deveria incluir estratégias que objetivassem a dimensão de suas atitudes e habilidades.
2. Validade da Construção: Em uma avaliação de raciocínio crítico aceitável cada tarefa ou questão deveria ter sido avaliada para assegurar que alunos que respondessem corretamente fizessem isso nas bases de um bom raciocínio crítico e que respostas inadequadas ou incorretas são o resultado de um fraco ou inadequado raciocínio crítico. Estratégias inteiras ou específicos itens nos quais bom raciocínio crítico leva a respostas incorretas ou pobre raciocínio crítico leva a respostas corretas, não deveriam ser utilizados.
3. Confiabilidade: Em mensurações aceitáveis de raciocínio crítico cada tarefa ou questão deveria ter sido avaliada para assegurar que bons pensadores críticos geralmente se sairão melhor naquele item do que fracos pensadores críticos. Se diferentes pessoas estão envolvidas na avaliação dos resultados, por exemplo, atribuindo notas às redações ou julgando apresentações, as avaliações de diferentes juizes deveriam ser cruzadas e checadas para assegurar que seus achados fossem confiáveis, ou seja, no geral fossem consistentes um com o outro. Entretanto, trata-se de uma questão em aberto se níveis de desempenho, atitudes e sub-habilidades de raciocínio crítico são positivamente correlacionados. Pesquisas empíricas sobre como sub-habilidades de raciocínio crítico se correlacionam entre si e com várias atitudes têm ainda sido prometidas. Portanto, até este momento, devida precaução deveria ser exercida na interpretação de medidas técnicas da confiabilidade do formato do teste no caso de instrumentos de avaliação de raciocínio crítico em papel e lápis.
4. Imparcialidade: mensurações de raciocínio crítico não deveriam parcialmente trazer vantagens ou desvantagens a grupos de alunos em relação às suas bases de habilidades de leitura, conhecimento de uma área específica (amplamente entendido com incluindo as considerações evidenciais, conceituais, metodológicas, criteriologicas ou familiaridade como vocabulário técnico), gênero ou idade relacionada à experiência de vida, etnicidade ou status socioeconômico, diferenças em normas sociais ou em suposições culturais. Avaliação de raciocínio crítico localiza tarefas e questões em um contexto assumido, como também em um assunto específico, do cotidiano ou fictício. Entretanto, garantir que todos os alunos, em se tratando de suas bases individuais, virão para a avaliação de raciocínio crítico em uma perfeita equidade em termos de bases de conhecimento, habilidade de leitura, experiência de vida, etc. é impossível. Logo, examinar a estratégia ou instrumento de avaliação para ter certeza que estes fatores não parcialmente influenciem os resultados é prudente e racional. Embora alguém não possa eliminar a influência destas variáveis, alguém pode estar apto a neutralizar ou controlar seus efeitos.

Estas considerações, mesmo que desatualizadas em alguns casos (como no item 3, no momento em que se diz que as pesquisas precisam ser realizadas para testar e verificar possíveis correlações entre desempenho, atitudes e sub-habilidades de raciocínio crítico), são

cuidadosamente consideradas nas escolhas dos instrumentos de avaliação de raciocínio crítico e nos procedimentos metodológicos desta pesquisa (Capítulo 3).

2.4 Modelo Instrucional de Richard Paul

O modelo de Richard Paul (1999) para desenvolvimento de raciocínio crítico é utilizado como base para o tratamento experimental desta pesquisa. No Apêndice A encontra-se a apostila de raciocínio crítico disponibilizada e ensinada aos alunos do grupo experimento em intervenções e atividades desenvolvidas em sala de aula.

O modelo de Paul inclui elementos de raciocínio, padrões intelectuais universais usados para avaliar raciocínio, atitudes e habilidades dos alunos. Paul apresenta sua abordagem para raciocínio crítico como um modelo geral de raciocínio que pode ser aplicado a qualquer problema ou assunto que requer raciocínio. Sua teoria é a de que para desenvolver raciocínio crítico nos alunos é preciso ensiná-los a pensar criticamente (REED, 1998). Ele foi escolhido para esta pesquisa principalmente por ser generalista, fato que possibilita ser aplicado a qualquer contexto. Outra justificativa à escolha do modelo é que Reed (1998) constatou resultados positivos no desenvolvimento de habilidades de raciocínio crítico em alunos de História, logo este estudo analisa se o mesmo pode ocorrer com alunos de Contabilidade.

A Figura 1 a seguir demonstra graficamente os elementos de raciocínio:

UM PENSADOR CRÍTICO

Considera os Elementos de Raciocínio

Figura 1 - Elementos de raciocínio crítico.

Fonte: Adaptada de Paul e *Foundation for Critical Thinking*, 1999.

Em cada atividade, resolução de um problema, debate, discussão, análise de um artigo, de um livro, os oito elementos devem ser considerados no processo de raciocinar. Paul (1999, p. 3-9) define e comenta cada um dos elementos:

1. **PONTOS DE VISTA:** Pensamento humano é relacional e seletivo. É impossível entender qualquer pessoa, evento ou fenômeno de todos os pontos de vista simultaneamente. Raciocínio crítico requer que este fato seja levado em consideração quando se analisa ou mensura pensamento. Isto não é dizer que o pensamento humano é incapaz de verdade e objetividade, mas apenas que a verdade humana, objetividade e discernimento são sempre virtualmente limitados e parciais, nunca virtualmente totais e absolutos. As ciências exatas são bons exemplos disto, visto que realidades qualitativas são sistematicamente ignoradas em favor de realidades quantificáveis.
2. **PROPÓSITO DO PENSAMENTO:** A intenção, objetivo ou fim de um documento, discussão, atividade, relacionamento, etc.
3. **QUESTÃO SOBRE O ASSUNTO:** Um problema, situação ou pessoa que é difícil de compreender, lidar ou resolver. Problemas e questões podem ser divididos em muitos tipos, incluindo monológicos (problemas que podem ser resolvidos por meio do raciocínio exclusivamente dentro de uma disciplina, ponto de vista ou estrutura de referência) e multi-lógicos (problemas que podem ser analisados e abordados por mais de uma disciplina e geralmente conflitando pontos de vista ou estruturas de referência).

4. **INFORMAÇÃO:** Os dados (fatos, figuras ou informações) nos quais um julgamento ou conclusão pode estar baseado ou por qual uma prova ou probabilidade pode ser estabelecida. Pensadores críticos distinguem a evidência ou os dados primários sobre os quais eles baseiam suas interpretações ou conclusões das inferências ou suposições que conectaram os dados às conclusões. Pensadores não críticos tratam suas conclusões como alguma coisa gerada por suas experiências, como alguma coisa que eles diretamente observaram no mundo. Como resultado, eles encontram dificuldade em enxergar a razão pela qual qualquer pessoa poderá discordar de suas conclusões.
5. **INTERPRETAÇÃO & INFERÊNCIA:** Uma inferência é um passo da mente, um ato intelectual pelo qual alguém conclui que alguma coisa é ou parece ser claramente algo. Se você vem em minha direção com uma faca na mão, eu provavelmente irei inferir que você pretende me ferir. Inferências podem ser fortes ou fracas, justificadas ou injustificadas. Inferências são baseadas em suposições.
6. **CONCEITOS:** Ideias ou pensamentos, especialmente uma ideia generalizada de uma coisa ou de uma classe de coisas. Humanos pensam dentro de conceitos ou ideias. Nós nunca teremos controle sobre nossos pensamentos a não ser que nós aprendamos como comandar nossos conceitos ou ideias. Então nós devemos aprender como identificar os conceitos ou ideias que nós estamos usando, contrastar eles com conceitos ou ideias alternativas e clarificar o que nós incluímos ou excluímos por meio deles. Por exemplo, a maioria das pessoas diz que acreditam fortemente na democracia, mas poucas podem clarificar com exemplos o que esta palavra implica ou não.
7. **SUPOSIÇÕES:** Afirmações aceitas ou supostas como verdade sem prova ou demonstração; uma não confirmada premissa ou crença. Todo pensamento e experiência humana são baseados em suposições. Nosso pensamento deve iniciar com alguma coisa que tomamos como verdade em um particular contexto. Nós tipicamente não temos conhecimento do que nós assumimos e assim raramente questionamos nossas suposições. Muito do que está errado no pensamento humano pode ser encontrado nas suposições não críticas ou não examinadas que o baseiam. Por exemplo, nós geralmente experimentamos o mundo de tal maneira que assumimos que o que nós estamos observando é realmente daquele jeito, como pensamento nós estamos vendo o mundo sem o filtro de um ponto de vista. As pessoas com as quais nós discordamos, nós reconhecemos como tendo um ponto de vista. Uma das atitudes chave de raciocínio crítico é o senso que nós como humanos sempre pensamos dentro de uma perspectiva, que nós virtualmente nunca experimentamos coisas absolutamente ou totalmente. Há uma conexão, portanto, entre pensamento estando consciente das nossas suposições e estando intelectualmente limitado.
8. **IMPLICAÇÕES & CONSEQUÊNCIAS:** Uma alegação ou verdade que é seguida por outras alegações ou verdades. Uma das mais importantes habilidades de raciocínio crítico é a habilidade de distinguir entre o que é na verdade implicado por uma afirmação ou situação do que pode ser inferido descuidadamente pelas pessoas. Pensadores críticos tentam monitorar suas inferências mantendo elas em linha com que está sendo na verdade implicado pelo o que eles sabem. Quando falando, pensadores críticos tentam utilizar palavras que implicam apenas o que eles podem legitimamente justificar. Eles reconhecem que há usos estabelecidos das palavras que geram estabelecidas implicações. Dizer que um ato é assassinato, por exemplo, implica em dizer que é intencional e não justificado.

Outro item que deve ser considerado no modelo abrange os padrões intelectuais, os quais devem ser aplicados nos momentos em que se queira checar a qualidade do raciocínio de alguém sobre um problema, assunto ou situação. Pensar criticamente envolve ter comando destes padrões. Para ajudar os alunos, os professores deveriam expor questões que os instigassem a investigar, questões que os fizessem pensar por um determinado tempo. Estas questões, consistentemente utilizadas pelos professores na sala de aula, são internalizadas pelos alunos, os quais sentem necessidade de reproduzi-las a eles próprios (PAUL, 1999).

Paul (1999, p. 10-11) disserta e analisa as questões referentes aos padrões intelectuais:

1. CLAREZA: Você poderia elaborar mais sobre determinado ponto de vista? Você poderia expressar aquele ponto de outra maneira? Você poderia me dar uma ilustração? Você poderia me dar um exemplo? Clareza é um padrão. Se uma afirmação não é clara, nós não podemos determinar se ela é precisa ou relevante. De fato, nós não podemos dizer nada sobre isto porque nós não sabemos ainda o que isto está dizendo. Por exemplo, a questão “O que pode ser feito sobre o sistema educacional no Brasil? não é clara. Para endereçar adequadamente a questão, nós necessitaríamos compreender mais claramente qual o problema considerado pela pessoa que formulou a questão. Uma questão mais clara seria “O que os educadores podem fazer para assegurar que seus alunos aprenderão habilidades que os ajudarão a atuar com sucesso em suas carreiras profissionais e nas decisões da vida cotidiana?”.
2. RETIDÃO: É aquilo realmente verdade? Como poderíamos checar aquilo? Como poderíamos descobrir se aquilo é verdade? Uma afirmação pode ser clara, mas não correta, como “a maioria dos cachorros pesa mais do que 135 kg”.
3. PRECISÃO: Você poderia me dar mais detalhes? Você poderia ser mais específico? Uma afirmação pode ser ambas, clara e correta, mas não precisa, como em “Pedro está acima do peso”. (Nós não sabemos o quanto Pedro está acima do peso, um quilo ou 100 quilos).
4. RELEVÂNCIA: Como aquilo está conectado com a questão? Como aquilo confirma o assunto? Uma afirmação pode ser clara, correta e precisa, mas não relevante ao assunto em questão. Por exemplo, alunos geralmente pensam que a quantidade de esforço que eles colocaram em um curso deveria ser usada na nota do mesmo. Geralmente, entretanto, “esforço” não mede a qualidade do aprendizado do aluno, e quando isto acontece, esforço é irrelevante para a nota apropriada.
5. PROFUNDIDADE: Como suas respostas tratam as complexidades inerentes à questão? Como você está levando em conta os problemas da questão? Está lidando com os fatores mais significativos? Uma afirmação pode ser clara, correta, precisa e relevante, mas superficial. Por exemplo, a frase “Apenas diga Não” que é geralmente utilizada para desencorajar crianças e adolescentes a usarem drogas, é clara, correta, precisa e relevante. Porém, ela não é profunda, porque trata de uma forma superficial um assunto extremamente complexo.
6. ALCANCE: Nós precisamos considerar outro ponto de vista? Há outra maneira de olhar para esta questão? Como isto pareceria de um ponto de vista conservador? Como isto pareceria do ponto de vista de ...? Uma linha de raciocínio pode ser clara, correta, precisa, relevante e profunda, mas com falta de alcance (como em um argumento tanto do ponto de vista conservador ou liberal que se aprofunda no assunto, mas reconhece apenas os discernimentos de um lado da questão).
7. LÓGICA: Isto realmente faz sentido? Aquilo segue o que você disse? Como é que aquilo segue? Mas antes você implicou aquilo e agora você está dizendo isto, eu não vejo como ambos podem ser verdadeiros. Quando nós pensamos, nós trazemos uma variedade de pensamentos juntos em alguma ordem. Quando as combinações não se suportam mutuamente, são contraditórias em certo sentido ou “não fazem sentido”, a combinação “não é lógica”.

Paul (1995) afirma que um pensador crítico deve ser guiado pelos padrões intelectuais universais independentemente da área ou dos assuntos em discussão. Para ele, habilidades e atitudes intelectuais são adicionadas para acabar com os vieses e infundadas suposições que as pessoas trazem para um problema. Portanto, o modelo relaciona os oito elementos de raciocínio crítico, os padrões intelectuais universais, as habilidades e atitudes de raciocínio crítico.

Figura 2 - Relacionamento entre os elementos, habilidades, atitudes e padrões.
 Fonte: Adaptada de Paul e *Foundation for Critical Thinking*, 1999.

O modelo de Paul consiste em raciocinar sobre um campo de estudo, assunto, documento, problema etc., de acordo com os oito elementos. Cada elemento de raciocínio e cada um dos padrões intelectuais, juntamente com as atitudes, ajudarão no desenvolvimento das habilidades e sub-habilidades de raciocínio crítico por meio do cumprimento de três fases: processo, objeto e padrão, conforme detalhadas na Figura 2. A aplicação destes conceitos em sala de aula pode se dar por meio de atividades como: resoluções de exercícios, leituras, seminários, debates etc., as quais devem ser previamente elaboradas. Por exemplo: para a atividade de análise de um artigo ou livro identificando os elementos de raciocínio, é exposto um guia na apostila fornecida aos alunos na parte empírica desta pesquisa.

2.5 Estudos Anteriores

Abrami, Bernard, Borokhovski, Wade, Surkes, Tamim e Zhang descrevem uma meta-análise que resume as evidências empíricas disponíveis sobre o impacto de instruções no desenvolvimento ou aprimoramento das habilidades ou atitudes de raciocínio crítico dos alunos. A meta-análise encontrou 117 estudos fundamentados em 20.698 participantes. A principal conclusão da pesquisa foi que o tipo de intervenção e metodologia de ensino está substancialmente relacionado com os níveis de desenvolvimento de raciocínio crítico. Os resultados desta pesquisa deixam claro que o aprimoramento ou desenvolvimento das habilidades ou atitudes de raciocínio crítico dos alunos não podem ser uma mera expectativa implícita (ABRAMI *et al.*, 2008).

A visão do artigo é a de que os educadores devem começar a explicitar os objetivos de desenvolvimento de raciocínio crítico nos seus cursos, bem como treinar outros professores para que façam o mesmo e consigam aplicar as metodologias para desenvolver estas habilidades e atitudes. O texto justifica que desenvolver raciocínio crítico nos alunos ajuda a perpetuar a democracia, já que eles serão os futuros cidadãos e tenderão a se basear em evidências, na lógica e na análise das suas ações. Abrami *et al.* (2008, p. 1103) comentam:

Em um nível social abrangente, uma democracia composta de cidadãos que podem pensar por eles mesmos considerando como base a evidência e análise concomitante, em vez de emoção, preconceito ou dogma é algo a mais – de fato, isto sustenta, constrói e perpetua a democracia.

Serão apresentados a seguir outros estudos que abordam o tema raciocínio crítico ou que investigaram algum tipo de relacionamento entre os modelos instrucionais ou metodologias de ensino e o desenvolvimento ou aprimoramento de habilidades ou atitudes de raciocínio crítico.

Bledsoe (1955) desenvolveu um estudo utilizando 46 alunos de graduação, foram pesquisadas as habilidades de raciocínio crítico antes e depois de um curso intensivo de metodologia de pesquisa. Nenhuma diferença significativa foi encontrada (habilidades avaliadas pela escala *Allport-Vernon-Lindzey*). Uma diferença significativa ocorreu na média do *Critical Thinking Appraisal*.

Annis e Annis (1979) realizaram um estudo para analisar o impacto da Filosofia na habilidade de raciocínio crítico dos alunos. Alunos de universidade matriculados em diversos tipos de cursos de Filosofia e um grupo controle de alunos cursando cursos que não eram da área de Filosofia foram testados no início e no final de um semestre por meio do *Watson-Glaser Critical Thinking Appraisal* para determinar se o efeito do curso de Filosofia aumentava o raciocínio crítico. Foi encontrado que o curso de Lógica obteve um significativo impacto comparado com o grupo controle enquanto os outros cursos não apresentaram diferenças significativas. Implicações dos achados para aplicações em outras disciplinas e estudos foram discutidas.

Vockell e Rivers (1984) realizaram um estudo com simulações em computadores na área de ciências para estimular a habilidade de resolver problemas em alunos do ensino médio na disciplina de biologia. Resultados indicaram que os alunos que utilizaram as simulações atingiram os objetivos das atividades da mesma forma que os alunos do grupo controle.

McMillan (1987) revisou 27 estudos que investigaram o efeito sobre o raciocínio crítico de alunos de graduação com a utilização de métodos de instrução, cursos, programas e experiências em faculdades em geral. Da sua amostra, apenas dois estudos utilizaram como modelo de pesquisa o experimento. Os resultados falharam em suportar que o uso de métodos de instrução ou cursos aumenta o raciocínio crítico; no entanto, suportaram que a presença escolar melhora o raciocínio crítico. O que faltou na pesquisa foi uma definição consensual de raciocínio crítico, boa instrumentação de avaliação e uma clara descrição teórica da natureza de uma experiência que deveria aumentar o raciocínio crítico.

Bangert-Drowns e Bankert (1990) fizeram uma meta-análise para sintetizar os resultados do impacto de instruções explícitas sobre o desenvolvimento de raciocínio crítico. A abordagem coletou 250 estudos (livros, artigos, dissertações e resumos) do banco de dados ERIC e do *Resumos de Dissertações Internacional (Dissertation Abstracts International)*. Para serem incluídos na meta-análise, os estudos tinham de ter comparado o desempenho de dois grupos de alunos em alguma medida de raciocínio crítico. Um total de 20 estudos, dos quais 19 eram teses de doutorado, foi considerado adequado para a meta-análise. Resultados consistentemente favoreceram programas que utilizaram métodos de instrução explícita. Programas intensivos provaram maior efetividade do que programas periódicos de treinamento em raciocínio crítico.

Unrau (1991) realizou um estudo para verificar o efeito de instrução explícita em ler criticamente e escrever com argumentação utilizando um procedimento denominado TASK (*Thesis Analysis and Synthesis Key*). Os resultados mostraram que dentre os alunos que tiraram as menores notas em suas avaliações, os do grupo experimento obtiveram maiores ganhos nas habilidades de ler e argumentar criticamente (de acordo com o *Ennis Weir Critical Thinking Essay Test* – EWCTET) do que os alunos do grupo controle.

Riesenmy, Mitchell, Hudgins e Ebel (1991) ensinaram algumas habilidades de raciocínio crítico diretamente para 38 alunos de 4ª e 5ª séries do ensino fundamental de escolas públicas de *Saint Louis* nos Estados Unidos. Os autores esperavam que os alunos aos quais foram ensinadas as regras de quatro modos de pensar (definidor de tarefas, estrategista, monitor e desafiador) se saíam melhor em um pós-teste de resolução de um problema que exigiria transferência de habilidades de raciocínio crítico. Os resultados foram de acordo com as expectativas dos autores e os grupos experimento superaram os grupos controle.

Colbert (1993) analisou os efeitos da participação em debates sem argumentação ou com agressão verbal. Os resultados indicaram que alunos com experiência em debates competitivos diferem dos que não tiveram experiência de acordo com dois testes que avaliam raciocínio crítico.

Terenzini, Springer, Pascarella e Nora (1995) estimaram os efeitos relativos sobre as mudanças no raciocínio crítico de acordo com três dimensões das experiências de alunos de graduação: a exposição curricular, experiências na sala de aula e atividades acadêmicas e experiências fora da sala de aula. Em todos os casos os ganhos de raciocínio crítico foram significativos e estes resultados apoiam as especulações de diversos teóricos sobre as influências acadêmicas e não acadêmicas no desenvolvimento do raciocínio crítico dos alunos.

Allegretti e Frederick (1995) realizaram um estudo com estudantes do ensino superior na utilização de um modelo de raciocínio crítico baseado na teoria de Toulmin et al. (1984) aplicando-o à assuntos éticos relacionados à psicologia, filosofia, enfermagem e negócios. Durante este curso interdisciplinar, 24 estudantes identificaram uma alegação, suas evidências e seus contra-argumentos para ganhar habilidade na utilização do modelo. Scores no *Cornell*

Critical Thinking Tests foram significativamente altos e a utilização do modelo em outras pesquisas é incentivado.

Halpern (1998) disserta sobre um modelo empírico para guiar o ensino e aprendizagem em raciocínio crítico. A autora descreve que o modelo é dividido em quatro partes: um componente de treinamento para tornar os alunos mais dispostos ao trabalho cognitivo, instruções sobre habilidades de raciocínio crítico, treinamento em aspectos estruturais de problemas e argumentos para transferir habilidades de raciocínio crítico entre contextos e um componente metacognitivo que permite checar a acurácia e o progresso para o atingimento do objetivo.

Reed (1998) foi pioneira em investigar o efeito de integrar o modelo de Richard Paul para desenvolvimento de raciocínio crítico em um curso. Ela o fez em uma disciplina de História Norte-Americana em uma universidade para desenvolver: a) habilidades para pensar criticamente sobre História Norte-Americana e sobre assuntos diários; b) atitudes para pensar criticamente; e c) conhecimento sobre o conteúdo de História Norte-Americana. O estudo também examinou se idade ou gênero influenciam a eficiência do método de instrução. Quatro turmas da disciplina “História Americana de 1877 até o Presente” participaram do estudo de um semestre. Duas turmas foram aleatoriamente selecionadas para servir como grupo experimento e as outras duas turmas serviram como grupo controle. O grupo experimento (n = 29) recebeu aproximadamente 90 minutos de instrução explícita distribuída durante o semestre sobre como utilizar o modelo de raciocínio crítico de Paul para analisar e interpretar documentos primários. Adicionalmente, o modelo foi integrado em uma série de atividades em sala de aula. O grupo controle (n = 23) foi ensinado de uma maneira mais tradicional. Estudantes fizeram três pré-testes e quatro pós-testes para avaliar a efetividade do modelo de instrução: o “*Documents Based Question*” (DBQ) do “*Advanced Placement Examination*”, o “*Ennis-Weir Critical Thinking Essay*” (EWCTET), o “*California Critical Thinking Dispositions Inventory*” (CCTDI) e um Exame de conteúdo de História. O grupo experimental obteve resultado significativamente maior no DBQ e EWCTET e não houve diferenças significativas no CCTDI e no Exame de conteúdo de História. Nenhuma diferença significativa foi encontrada entre idade e gênero.

Allen, Berkowitz, Hunt e Loudon (1999) realizaram uma meta-análise considerando o impacto sobre raciocínio crítico de vários métodos para melhorar as habilidades de

comunicação em público. Os resultados mostraram que as instruções sobre comunicação melhoraram as habilidades de raciocínio crítico dos participantes.

Walsh e Hardy (1999) analisaram as diferenças de atitudes de raciocínio crítico relacionadas à gênero e curso de graduação. O *California Critical Thinking Disposition Inventory* (CCTDI) foi aplicado à 334 estudantes de graduação. Os cursos de Inglês, Psicologia e Enfermagem obtiveram os maiores escores. As disciplinas não práticas como: Inglês, História e Psicologia obtiveram maior escore do que disciplinas práticas como: Enfermagem, Educação e Negócios. As mulheres obtiveram maiores escores nas atitudes: mente aberta e maturidade.

Kuhn (1999) aborda em seu estudo que a maioria dos programas para ensinar raciocínio crítico não possui modelos desenvolvidos por meio de pesquisa empírica. A autora propõe um modelo de raciocínio crítico que deriva de pesquisa empírica sobre o desenvolvimento intelectual de crianças e adolescentes.

Tsui (1999) investigou como diferentes tipos de cursos e técnicas de instrução afetaram uma auto avaliação de estudantes sobre crescimento em raciocínio crítico. Os dados foram coletados do Programa de Pesquisa de Cooperação Institucional de 1989 de 24.837 alunos que entraram na faculdade, em 1985. Os resultados revelaram que cursos de escrita, interdisciplinares, de história, de ciências, de estudos da mulher, de matemática, de línguas estrangeiras e de estudos étnicos foram positivamente associados com a auto avaliação de crescimento em raciocínio crítico. O resultados sobre as técnicas revelaram que possuir um artigo criticado por um instrutor, conduzir uma pesquisa, trabalhar em grupo, apresentar trabalhos tiveram também associação positiva com a auto avaliação de crescimento em raciocínio crítico e que realizar testes de múltipla escolha foram negativamente associados.

Adams (1999) analisou 20 pesquisas sobre ensino de Enfermagem entre 1977 e 1995 e encontrou que o teste Watson Glaser de Apreciação de Raciocínio Crítico (*Watson Glaser Critical Thinking Appraisal*) foi utilizado em 18 deles. A principal conclusão de sua revisão foi a de que não existiu evidência consistente de que a educação em Enfermagem aumenta raciocínio crítico. Imperfeições nos modelos de pesquisa, definições inconsistentes de raciocínio crítico e falta de ferramentas apropriadas de avaliação são as possíveis causas desse resultado.

Rudd, Baker e Hoover (2000) pesquisaram a possível existência de um relacionamento entre os estilos de aprendizagem de alunos de graduação em Agricultura e suas atitudes de raciocínio crítico. Embora os autores tenham identificado diferenças significativas entre as atitudes de raciocínio crítico de homens e mulheres, não houve diferença significativa entre alunos que tiveram a estratégia de ensino de campo independente e os que tiveram a de campo dependente. Muitos dos alunos da amostra estudada apresentaram baixo nível de atitude de raciocínio crítico (30,5%), enquanto 1,7% dos estudantes apresentaram um alto nível de atitudes de raciocínio crítico.

Scheffer e Rubenfeld (2000) realizaram um estudo para encontrar uma definição consensual sobre raciocínio crítico na enfermagem. Utilizaram a técnica *Delphi* com 5 rodadas de reuniões com enfermeiros de nove países: Brasil, Canadá, Inglaterra, Islândia, Japão, Coréia, Holanda, Tailândia e de 23 estados dos Estados Unidos da América, entre 1995 e 1998. Os resultados foram: uma definição consensual de raciocínio crítico na enfermagem; 10 hábitos da mente (componentes afetivos) e 7 competências (componentes cognitivos) do raciocínio crítico em enfermagem. Os hábitos da mente do raciocínio crítico em enfermagem são: confiança, perspectiva contextual, criatividade, flexibilidade, curiosidade, integridade intelectual, intuição, mente aberta, perseverança e reflexão. As habilidades de raciocínio crítico em enfermagem são: análise, aplicação de normas, discriminar, buscar informação, raciocínio lógico, prever e transformar conhecimento.

Maudsley e Strivens (2000) dissertaram sobre a importância em relacionar a teoria educacional contemporânea para alunos de medicina com situações práticas da profissão. Os autores concluem que a aprendizagem experimental é importante para a aquisição de conhecimento profissional, raciocínio crítico e capacidade de resolução de problemas clínicos.

Pithers e Soden (2000) realizaram uma revisão da literatura sobre os métodos e as concepções de ensino que inibem ou aumentam as habilidades de raciocínio crítico. Um guia com maneiras de ensinar raciocínio crítico e ajudar os alunos a pensar bem e por eles mesmos é descrito e discutido.

Solon (2001) dissertou sobre sua pesquisa que teve como objetivo constatar possíveis desenvolvimentos de raciocínio crítico em alunos de um curso introdutório em Psicologia. Os resultados demonstraram que os alunos do grupo experimento (que tiveram sessões ensinando

as habilidades de raciocínio crítico) obtiveram melhor desempenho no curso do que os alunos do grupo controle (aula tradicional), porém não foram realizadas análises estatísticas para verificar a significância dos achados.

Chau, Chang, Lee, Lee e Wootton (2001) descreveram um estudo que teve como objetivo determinar se houve efeito na utilização de vinhetas gravadas em vídeo tapes para promover raciocínio crítico de alunos de enfermagem e habilidades para gerenciar diferentes situações clínicas. Os resultados mostraram que os escores nos pós-testes dos alunos foram mais altos do que nos pré-testes, porém não foram estatisticamente significativos. Uma pesquisa de satisfação mostrou que os alunos ficaram satisfeitos com a nova abordagem de ensino.

Anderson, Howe, Soden, Halliday e Low (2001) realizaram um estudo reportando um programa educacional no qual habilidades de raciocínio crítico são ensinadas. O principal objetivo do estudo foi desenvolver, implementar e avaliar um programa para ensinar os alunos a justificarem baseando-se em evidências. Uma amostra de 48 estudantes participaram de dez sessões de intervenções. Os resultados das atividades mostraram na sua maioria fracas justificativas (fraco senso), utilizando analogias ou experiências gerais e poucas fortes justificativas (baseada em evidências) apareceram nos relatórios dos alunos.

Lederer (2001) disserta sobre um estudo de caso na *Colorado State University*. A autora descreve sobre estratégias de pesquisa e assistências da bibliotecária que podem ensinar aos alunos habilidades como avaliar e analisar livros, artigos e páginas da *internet* verificando se são úteis às suas pesquisas.

Rauen (2001) realizou um estudo sobre a utilização de simulações na área de enfermagem para ensinar raciocínio crítico. Seu trabalho descreve as simulações e defende a utilização do método para o desenvolvimento das habilidades de raciocínio crítico nos alunos da área.

Parkinson e Ekachai (2002) realizaram um estudo sobre a utilização do método Socrático em um curso introdutório de relações públicas. As variáveis independentes do estudo foram o método tradicional (*Lecture*) e o método Socrático. As hipóteses traçadas para o grupo experimento (método Socrático) em comparação ao grupo controle (*Lecture*) foram: (1) reteriam mais informação baseadas em evidências sobre relações públicas; (2) sentir-se-iam mais confiantes sobre seu conhecimento e suas habilidades necessárias para trabalhar com

relações públicas; (3) reportariam mais oportunidade de praticar raciocínio crítico; (4) reportariam mais oportunidades de praticar habilidades de resolução de problemas; (5) reportariam mais aspiração em trabalhar com relações públicas; e (6) reportariam maiores níveis de satisfação com o curso. Houve duas diferenças significativas a favor do grupo experimento, eles reportaram mais oportunidades em utilizar as habilidades de raciocínio crítico e resolver problemas.

Jeong (2003) em seu estudo objetivou examinar a interação de um grupo e seu raciocínio crítico em discussões *online* segmentadas. A Ferramenta de Análise de Discussões (DAT) foi usada para identificar padrões nas interações e determinar quais promoviam raciocínio crítico. Com o DAT transcrições das discussões foram codificadas em doze eventos e as probabilidades de transição entre os eventos foram computadas usando o método de análise sequencial (Bakeman e Quera, 1995). Os resultados mostraram que as interações envolvendo pontos de vista conflitantes promoveram mais discussão e raciocínio crítico e que a avaliação dos argumentos era mais provável de ocorrer quando as conclusões estavam sendo desenhadas do que quando os argumentos estavam sendo apresentados.

Gelder (2005) disserta sobre seis lições chave da ciência cognitiva para professores de raciocínio crítico. As lições são: adquirir experiência em raciocínio crítico é duro; praticar as habilidades de raciocínio crítico aumentam as mesmas; a transferência de habilidades deve ser praticada; algum conhecimento teórico é necessário; diagramas de argumentos (mapas) desenvolvem habilidades e alunos são propensos a acreditar em crenças. O estudo apresenta alguns guias para ensinar de acordo com estas lições.

Brooke (2006) descreve sobre a utilização de estudos de caso para ensinar alunos em cursos *online*. A autora disserta que os estudos de caso promovem habilidades de raciocínio crítico e diálogos Socráticos. As atividades desenvolvidas são detalhadas e a autora incentiva que sejam utilizadas em outros estudos.

Tiwari, Lai, So e Yuen (2006) realizaram um experimento com alunos de graduação em enfermagem para comparar os efeitos de dois diferentes métodos de ensino, o Aprendizado Baseado em Problemas (PBL) e o da Leitura (Lecture) no desenvolvimento de atitudes de raciocínio crítico avaliadas pelo *California Critical Thinking Dispositions Inventory* (CCTDI). Os resultados demonstraram que as médias não foram estatisticamente diferentes,

porém que a turma que utilizou o PBL apresentou maior desenvolvimento das atitudes: Buscando a verdade, Analiticidade e Autoconfiança.

Harrel (2007) comenta em seu estudo que há substancial evidência de muitas áreas que representações visuais geram formas de cognição. Seu estudo teve como objetivo determinar se representações visuais de estruturas de argumentação aumentavam o desenvolvimento de habilidades de raciocínio crítico no contexto de um curso de filosofia introdutória. A autora encontrou efeito significativo no uso dos diagramas. De acordo com a autora os resultados sugerem que pequenas e naturais modificações no padrão dos cursos de raciocínio crítico proveem substanciais aumentos no aprendizado e desempenho dos estudantes.

Quitadamo, Faiola, Johnson e Kurtz (2008) dissertaram sobre a utilização de métodos de ensino que incentivam o aluno do curso de graduação em Biologia a investigar, a ser mais ativo do que em métodos de ensino tradicionais e a desenvolver o raciocínio crítico. De acordo com os autores, algumas pesquisas mostram que enquanto professores e alunos julgam raciocínio crítico essencial para o sucesso acadêmico e profissional, apenas uma pequena fração demonstra possuir esta habilidade. Os autores comentam que muitos acadêmicos estão considerando métodos de ensino não tradicionais que incorporam pesquisa na graduação porque eles acreditam estarem mais alinhados com o processo da ciência investigativa. O estudo compara um método de ensino focado na pesquisa chamado “*Community-based inquiry*” (CBI) com o método tradicional de leitura e laboratório no ensino da Biologia. Os resultados mostram ganhos significantes de raciocínio crítico nos grupos que usaram o CBI em comparação aos tradicionais e a um grupo que usou parte tradicional e parte CBI.

Existem milhares de trabalhos sobre ensino de raciocínio crítico (ABRAMI *et al.*, 2008), e os citados aqui demonstram que o assunto é pesquisado há muitos anos no exterior. As pesquisas experimentais, quase-experimentais, de revisões e de meta-análise são abundantes e denotam a importância dada ao tema. Nota-se que os resultados são inconclusivos em algumas dimensões, o que sugere novas investigações. No Brasil, não se encontraram pesquisas com desenho quantitativo buscando evidências nas metodologias de ensino em relação ao desenvolvimento de habilidades de raciocínio crítico. Na verdade, a própria definição do termo raciocínio crítico não é amplamente debatida. Sendo assim, este trabalho também procura contribuir com a abertura de uma grande porta para futuras pesquisas no país.

2.6 Estudos na Área Contábil

Paisey e Paisey (2004) realizaram uma revisão da pesquisa em educação contábil e constataram a necessidade de pesquisas em áreas relacionadas ao desempenho e à avaliação do aluno. Para melhorar a educação na área de negócios, é preciso melhorar o entendimento do raciocínio crítico e as habilidades de comunicação dos alunos. Diversos trabalhos enfatizam a necessidade de observar mais atentamente o desenvolvimento de habilidades de raciocínio crítico, analítico e de comunicação nos alunos de contabilidade (CORNACHIONE, DUNCAN; JOHNSON, 2007).

Cornachione, Duncan e Johnson (2007) realizaram um estudo que explorou uma amostra de 65 estudantes de graduação e pós-graduação em Ciências Contábeis da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, Brasil (FEA-USP), avaliando as habilidades de raciocínio crítico, leitura e escrita por meio dos seguintes instrumentos: *Motivated Strategies for Learning Questionnaire* (MSLQ), *Ennis-Weir Critical Thinking Essay Test* (EWCTET), *Flesch Reading Ease* (FRE) e FRE Brasil (modelo adaptado para o português pela USP - São Carlos). Por meio de análises estatísticas de correlações e diferenças de médias entre grupos (gênero, paternidade, nível do programa e estágio do programa), o estudo analisou os dados e refletiu sobre elementos para o aprimoramento do desempenho dos estudantes em relação às suas habilidades de raciocínio crítico e o papel dos educadores como agentes de mudanças no cenário da educação superior. Os resultados demonstraram um alto nível de raciocínio crítico da amostra; porém, algumas ressalvas foram feitas em relação à amostra original que serve de comparação (por exemplo, a amostra era formada por apenas graduandos e sua realidade era de 20 anos atrás).

Voltando à década de 1990, Baldwin e Ingram (1991) realizaram uma pesquisa para repensar os objetivos e conteúdos de cursos introdutórios de contabilidade. Para os autores, qualquer proposta que pretenda significativamente melhorar a educação contábil nos EUA deve iniciar com uma análise dos cursos introdutórios de contabilidade. Eles descrevem que os alunos ideais de contabilidade são pessoas com habilidades criativas de resolução de problemas, que podem trabalhar com sucesso em situações adversas com problemas não familiares e desestruturados.

Stoiber (1991) estudou os efeitos de dois tipos de instrução em uma disciplina de negócios, uma que enfatizou princípios e habilidades técnicas e a outra que enfatizou uma abordagem reflexiva e construtivista. Comparada a turma controle (técnica), os participantes da turma experimento (reflexiva) utilizaram estratégias para a resolução de problemas com maior frequência. Diferenças do contexto instrucional e os resultados positivos da turma mais reflexiva são discutidos.

McEwen (1994) realizou uma pesquisa com 67 professores da área de negócios e identificou que os melhores métodos para ensinar raciocínio crítico são: estudos de caso; resolver problemas; simulações; projetos e discussões em grupo.

Sormunen e Chalupa (1994) revisaram métodos para avaliar raciocínio crítico e sugeriram maneiras para modificar testes comerciais. Para as autoras, é preocupante o fato de as organizações estarem demandando habilidades de raciocínio crítico de seus funcionários. De acordo com elas, existe uma crescente pressão para produzir empregados que possam pensar criticamente.

Kimmel (1995) desenvolveu uma estrutura para incorporar o raciocínio crítico na educação contábil. O autor faz sugestões para a implantação da estrutura com considerações sobre as implicações relativas aos estágios de desenvolvimento intelectual dos alunos.

Henninger e Hurlbert (1996) descrevem a inclusão de raciocínio crítico e informações pesquisadas em um programa de graduação em negócios utilizando técnicas de aprendizado ativo. Os autores apresentam um modelo de atividades, uma lista bibliográfica e oferecem uma avaliação do projeto, discutindo com os bibliotecários formas de colaboração e interação para criar uma atividade apropriada ao desenvolvimento de raciocínio crítico.

Muir (1996) disserta em seu trabalho que há mais em uma comunicação bem sucedida do que aprender estilos e procedimentos frequentemente encontrados no programa tradicional de comunicação empresarial. O autor discute os benefícios de utilizar uma abordagem de raciocínio crítico para ajudar os alunos a compreenderem o complexo ambiente social e político em que a comunicação empresarial é praticada.

Stone e Shelley (1997) compararam a efetividade de um método instrucional denominado “Projeto de Descoberta (PD)” (*Project Discovery*) com um programa tradicional de contabilidade para graduação na Universidade de *Illinois* (UIUC), USA. O PD enfatiza adquirir habilidades intelectuais e melhorar as atitudes sem perder o conhecimento tradicional de contabilidade. As conclusões apontaram resultados positivos na utilização do método.

Bonk e Smith (1998) dissertaram sobre estratégias alternativas de instrução para desenvolvimento de raciocínio criativo e crítico nos currículos dos cursos de graduação em Ciências Contábeis. De acordo com os autores, no meio de numerosas reformas na contabilidade declarando que a memorização de fatos contábeis não irá bastar, a economia global tem aumentado a pressão sobre as Universidades para o desenvolvimento de habilidades de pensamento de alto nível. O trabalho dos autores apresenta um modelo de ensino para atender a estes desafios.

Baril, Cunningham, Fordham, Gardner e Wolcott (1998) realizaram uma pesquisa sobre o raciocínio crítico na profissão do contador público, debatendo sobre as aptidões e atitudes necessárias ao profissional. Os achados revelam que a profissão utiliza o termo raciocínio crítico para se referir a uma larga série de competências, as quais incluem atributos, atitudes e comportamentos cognitivos e não cognitivos.

Roever (1998) descreve uma atividade em uma disciplina de comunicação empresarial, na qual os estudantes construíram *portfolios* com artigos do *The Wall Street Journal* explicitando e explicando como os artigos estão relacionados com os conceitos discutidos em sala. A autora argumenta que a atividade incentiva o desenvolvimento das habilidades de raciocínio crítico e de escrita e desenvolve uma compreensão dos conceitos da área de comunicação empresarial.

Macpherson (1999) relata a eficácia de um módulo utilizando avaliação aos pares no desenvolvimento de habilidades de raciocínio crítico de alunos do segundo ano da área de negócios. Os resultados indicaram que as estratégias para desenvolver habilidades de raciocínio crítico podem influenciar os pares ou o tutor na avaliação das atividades desenvolvidas em sala.

Crumbley e Smith (2000) realizaram um estudo utilizando a técnica de contar pequenas histórias para ensinar raciocínio crítico e habilidade de comunicação para alunos de contabilidade tributária. Os autores descrevem a prática em seu trabalho listando as vantagens na adoção do método. A técnica consiste em, resumidamente, os alunos desenvolverem pequenas histórias nas quais os contadores são os principais personagens ou heróis. A melhor história é colocada na *internet* para que os outros alunos possam revê-la.

Gold, Holman e Thorpe (2002) dissertaram que raciocínio crítico é um importante objetivo da educação da área de negócios. O trabalho dos autores relatou a avaliação de um módulo de desenvolvimento pessoal que combinou análises de argumentos, baseadas na teoria de argumentação de Toulmin, com contar histórias. Os resultados mostraram que a combinação das técnicas podem promover muitos aspectos de raciocínio crítico.

Wilson (2002) realizou um estudo sobre as pesquisas em educação contábil fazendo uma retrospectiva dos últimos dez anos e traçando algumas projeções futuras. O autor possui uma visão da educação contábil como um grupo multidisciplinar de competências, que inclui habilidades de pensamento técnicas, interpessoais e criativas.

Smith (2003) disserta em seu estudo que ensinar os alunos como pensar é um objetivo universal das instituições educacionais. De acordo com o autor as escolas de negócio tem dirigido para este objetivo injetando em seus programas atividades de raciocínio crítico e oferecendo cursos de gerenciamento para tomada de decisões. Seu estudo revê os esforços nesta área e conclui que eles não são adequados para o desafio de ensinar os alunos da área de negócio a pensar criticamente. O autor propõe um programa que inclui raciocínio crítico e tomada de decisões, discute os assuntos abordados no programa e sugere atividades para ensinar os alunos da área de negócios a como pensarem criticamente.

Athanassiou, McNett e Harvey (2003) discutem o uso da taxonomia de Bloom como um modelo metacognitivo para aulas de negócios centradas nos alunos. O estudo avalia empiricamente o uso do modelo, o desempenho dos alunos e sugere maneiras de aplicá-lo.

Braun (2004) disserta que educadores da área de negócios são atualmente cobrados pela tarefa de melhorar o raciocínio crítico nas faculdades. Em seu estudo a autora investigou os passos que os professores estão tomando para melhorar o raciocínio crítico de seus alunos e a

efetividade de tais métodos. Ela reporta as abordagens comuns para o desenvolvimento de raciocínio crítico nos currículos de negócios, os resultados publicados e propõe adicionais esforços necessários para alcançar esta meta. Para a autora, embora haja considerável existência de teorias e instrumentos relacionados ao desenvolvimento e à avaliação do raciocínio crítico, há falta de pesquisas sobre este tópico na área da educação de negócios.

Lima, Koehler e Spiro (2004) discutiram como conjuntamente o Método *Harvard* de estudos de casos, as Tecnologias Interativas de Comunicação e a Teoria da Flexibilidade Cognitiva podem contribuir para a aprendizagem baseada em casos para tomada de decisão na área de negócios. Em particular, abordaram a concepção de ambientes de aprendizagem que facilitem os raciocínios críticos e criativo com base em análises organizacionais em múltiplas perspectivas. Um exemplo desse tipo de ambiente é a interface Panteon, um sistema desenvolvido para a *Internet* que permite análises organizacionais colaborativas. Os autores propuseram um modelo de implementação e avaliação baseado nas categorias de Aprendizagem Significativa de David Jonassen, bem como no modelo de Raciocínio Integrado da Secretaria de Educação do estado de *Iowa*. Com base nesse referencial, descreveram a concepção do Panteon e os resultados obtidos com o uso de entrevistas de grupos focais, observação participativa e questionários aplicados para investigar a avaliação dos participantes sobre a metodologia hipertextual do Panteon em comparação com o Método *Harvard* de estudos de casos. Os resultados indicaram que tanto os estudantes quanto os *experts* que acompanharam os experimentos favorecem o Panteon devidos às suas características de estímulo à habilidade de raciocínio crítico.

Turner (2006) explora o discurso de raciocínio crítico dentro da prática do Ensino Superior e avalia as experiências e os desempenhos de estudantes chineses no ensino da área de negócios em um curso de pós-graduação na Inglaterra. As principais conclusões são que as definições de raciocínio crítico são muitas vezes pouco claras e as diferenças culturais podem afetar a definição do que é ser crítico em relação à determinado assunto.

Ladyshevsky (2006) utilizou um método de aprendizagem em seu experimento denominado *Peer Coaching* (PC). De acordo com o autor o PC pode ser usado para aumentar a profundidade da aprendizagem na educação gerencial e desenvolver raciocínio crítico. Seu relato descreve a participação de 43 alunos em um programa de PC como parte de sua

educação gerencial. Poderosos efeitos da aprendizagem são relatados pelos participantes e os resultados bem sucedidos do método são analisados.

Cornachione, Duncan e Johnson (2007), dissertando sobre a importância do desenvolvimento e da avaliação do raciocínio crítico dos alunos de contabilidade, alegam que isto ajudará os alunos a terem maiores chances de sucesso profissional. Também ajudará os professores na identificação de ações pedagógicas que, em vez de apenas enfatizar conhecimento e compreensão com ocasionais aplicações de exercícios, consigam ensinar e avaliar raciocínio crítico.

Springer e Borthick (2007) realizaram um quase-experimento sobre o aumento do desempenho nos cursos de graduação em Ciências Contábeis utilizando tarefas de conflito cognitivo. As evidências da pesquisa demonstram que se deve insistir em tais tarefas, já que, apesar de contínuas demandas por habilidades de pensamento de alto nível em cursos de graduação de contabilidade, educadores contábeis têm resistido em enfatizar estas habilidades baseados na suposição de que fazendo isto eles estariam arriscando o domínio dos estudantes sobre o conhecimento tradicional contábil. Porém, as evidências da pesquisa mostraram o contrário, ou seja, os estudantes que participaram do grupo experimento, no qual as habilidades de pensamento de alto nível foram enfatizadas, apresentaram melhores resultados do que o grupo controle.

Nentl e Zietlow (2008) dissertaram sobre a utilização da taxonomia de Bloom para ensinar raciocínio crítico à alunos da área de negócios. O estudo primeiramente descreve a taxonomia de Bloom e como o modelo pode ser utilizado para desenvolver o raciocínio crítico dos alunos. Posteriormente, o artigo explica como métodos secundários de pesquisa podem ser utilizados para levar os alunos à mais altos níveis de pensamento de acordo com a taxonomia de Bloom. Por último é utilizado o critério de Herrick para analisar e avaliar a qualidade dos dados secundários.

Carrithers e Bean (2008) descrevem um método para avaliar as habilidades de raciocínio crítico de alunos da área de finanças. De acordo com os autores os alunos foram convidados a recomendar e justificar um curso de ação para um cliente que estava enfrentando problemas financeiros. A análise das redações dos alunos revelou fraquezas de pensamento crítico que podem ser amenizadas por mudanças nas estratégias de ensino. As redações revelaram quatro

tipos de problema de raciocínio: incapacidade de resolver o problema do cliente, soluções que não aplicam ferramentas financeiras, incapacidade de compreensão de conceitos da área financeira e impossibilidade de construir gráficos. As implicações curriculares do estudo são discutidas.

Zabit (2010) realizou uma revisão dos estudos utilizando PBL (*Problem-Based Learning*) nos ensino da área de negócios na Malásia para o desenvolvimento de habilidades de raciocínio crítico nos alunos. O autor concluiu que há grande potencial de aplicação do método em novos programas com o intuito de desenvolver as habilidades de raciocínio crítico.

Garvey e Buckley (2011) realizaram um estudo no qual eles utilizam uma inovadora tecnologia de previsão de mercado no ensino de negócios para encorajar os alunos a pensarem sobre risco de uma maneira aberta e flexível. O artigo explica como o modelo pode ser aplicado aos alunos e como pode desenvolver as habilidades de raciocínio crítico necessárias para as análises de risco do mercado.

Detmering e Johnson (2011) são dois bibliotecários da *University of Louisville* que desenvolveram um módulo de informação literária para um curso de introdução à área de negócios que enfatiza raciocínio crítico sobre a informação em vez da simples aplicação de ferramentas de procura da biblioteca.

Os trabalhos citados demonstraram a evolução cronológica do debate e das pesquisas sobre raciocínio crítico, principalmente, no ensino da contabilidade nos EUA. No Brasil, o trabalho pioneiro a tratar especificamente do assunto foi o artigo de Cornachione, Duncan e Johnson (2007). Busca-se por meio deste trabalho incentivar o aumento de pesquisas e discussões sobre o tema.

3 PROCEDIMENTOS METODOLÓGICOS

Neste capítulo uma visão geral da pesquisa empírica será apresentada em cinco tópicos. O primeiro descreve as instituições participantes. O segundo tópico aborda os alunos participantes e quais os critérios utilizados para selecionar as amostras. O terceiro detalha os instrumentos de avaliação utilizados nesta pesquisa. O quarto aborda o modelo instrucional e as atividades desenvolvidas com os participantes e o último tópico disserta sobre o modelo da pesquisa e os procedimentos de análise dos dados.

Neste estudo, o modelo de Richard Paul foi utilizado em disciplinas da área contábil em duas instituições para o desenvolvimento de habilidades e atitudes de raciocínio crítico. A primeira pode ser considerada um estudo piloto, já que um dos objetivos era um maior conhecimento do modelo de desenvolvimento de raciocínio crítico de Richard Paul e dos instrumentos de avaliação que foram aplicados (EWCTET, CCTDI e Pesquisa Demográfica). Em ambas havia duas turmas, uma controle e uma experimento, com disciplinas de idêntico conteúdo e com o mesmo professor. Elas foram testadas no início e no final do curso em relação ao desenvolvimento de habilidades e atitudes de raciocínio crítico. Também foram analisadas as possíveis diferenças neste desenvolvimento relacionadas ao gênero e à idade.

3.1 Instituições Participantes

As duas Instituições de Ensino Superior que participaram desta pesquisa possuem o curso de graduação em Ciências Contábeis, sendo que uma delas é pública e a outra, particular. Ambas consentiram a realização da pesquisa (Carta de Autorização do Coordenador do Curso – Anexo A).

O estudo piloto foi realizado na instituição pública que foi a Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA-USP), localizada na

cidade de São Paulo – SP. O curso é o de graduação em Ciências Contábeis. O autor é atual aluno do curso de pós-graduação da instituição e atuou como monitor das classes participantes. Ambas as turmas tiveram a mesma professora, a qual ajudou na aplicação do modelo para a turma experimento no segundo semestre de 2010. A abordagem de aplicação do modelo de desenvolvimento de raciocínio crítico de Richard Paul foi a Geral, com quatro intervenções de 30 minutos explicando uma apostila sobre o modelo e como utilizá-lo.

A instituição particular é o Centro Universitário Metropolitano de São Paulo (UNIFIG), localizado no município de Guarulhos – SP. O curso de graduação em Ciências Contábeis possui 39 anos e seu coordenador é Doutor em Controladoria e Contabilidade pela FEA-USP. Como o curso de Ciências Contábeis não possui duas turmas no mesmo semestre, não foi possível aplicar a pesquisa para seus alunos; logo, a pesquisa empírica foi desenvolvida em uma disciplina contábil para os alunos do curso de graduação em Administração. O pesquisador ministra na instituição desde 2005 e foi o professor das classes participantes durante o primeiro semestre de 2011. A abordagem de aplicação do modelo de desenvolvimento de raciocínio crítico de Richard Paul foi a Mista, seguindo o mesmo protocolo de Reed (1998), ou seja, realizando quatro intervenções para explicar a apostila sobre o modelo e aplicando o mesmo em diversas atividades em sala de aula durante o semestre.

As instituições foram escolhidas principalmente pelo fato de estarem dispostas a colaborar com a pesquisa, de possuírem tradição no ensino da área de negócios e pela facilidade de atuação do pesquisador.

3.2 Alunos Participantes

No estudo piloto da FEA-USP os participantes pertenciam a duas classes da disciplina Contabilidade Intermediária I do 2º semestre do curso de graduação em Ciências Contábeis. Ambas as classes cursaram esta disciplina às segundas-feiras, com carga horária de 4 horas-aula semanais, sendo que uma delas é do período matutino e a outra, do noturno. Os possíveis vieses relacionados às diferenças de perfis dos alunos que atenderam o curso matutino e o

curso noturno foram analisados por meio dos dados de pesquisa demográfica e foram tratados cuidadosamente na aplicação e análise das ferramentas estatísticas.

A classe experimento (aleatoriamente escolhida) foi a do período matutino, com 52 alunos matriculados. A classe controle foi a do período noturno, composta por 56 alunos matriculados. A participação na pesquisa foi voluntária (Apêndice B – Consentimento Explícito) e o controle dos participantes da pesquisa foi realizado por meio de uma planilha de Excel com nome, número (Registro Acadêmico) e data do teste realizado¹⁰. Consentiu e iniciou a participação na pesquisa um total de 90 alunos; porém, ao final do semestre, 52 alunos completaram todos os aspectos da pesquisa e foram utilizados nas análises.

A seguir estão alguns gráficos com dados demográficos (Pesquisa Demográfica – Apêndice C) dos alunos que completaram a pesquisa.

Gráfico 1 - Gênero dos alunos da FEA-USP por turma.

¹⁰ Estas informações são para controle, e os nomes dos alunos não foram mencionados na pesquisa.

Gráfico 2 - Idade dos alunos da FEA-USP por turma.

Gráfico 3 - Cor/Raça dos alunos da FEA-USP por turma.

Gráfico 4 - Estado civil dos alunos da FEA-USP por turma.

Gráfico 5 - Quantidade de filhos dos alunos da FEA-USP por turma.

Gráfico 6 - Emprego dos alunos da FEA-USP por turma.

Gráfico 7 - Tipo de escola no ensino fundamental e médio dos alunos da FEA-USP por turma.

Gráfico 8 - Renda familiar dos alunos da FEA-USP por turma.

Gráfico 9 - Frequência de leitura dos alunos da FEA-USP por turma.

Gráfico 10 - Frequência de escrita dos alunos da FEA-USP por turma.

Gráfico 11 - Frequência de cálculo dos alunos da FEA-USP por turma.

Gráfico 12 - Fluência em língua estrangeira dos alunos da FEA-USP por turma.

Gráfico 13 - Participação em curso de leitura dos alunos da FEA-USP por turma.

Gráfico 14 - Participação em curso de escrita dos alunos da FEA-USP por turma.

Gráfico 15 - Participação em curso de cálculo dos alunos da FEA-USP por turma.

Gráfico 16 - Participação em curso de raciocínio crítico dos alunos da FEA-USP por turma.

O autor da pesquisa participou como monitor, e dentre as quatro abordagens (Geral, Infusão, Imersão e Mista) de instrução de raciocínio crítico definidas por Ennis (1993) e discutidas no Capítulo 2 desta pesquisa, utilizou-se a Geral, com quatro intervenções de 30 minutos durante o semestre para explicar os conceitos e a apostila (Apêndice A) disponibilizada com o modelo de Richard Paul.

Houve uma programação por aula, que foi fielmente seguida pela professora em ambas as turmas. A única diferença entre elas foi a intervenção na turma experimento. O tempo utilizado pelo monitor na turma experimento foi disponibilizado para a resolução de exercícios e atendimento às dúvidas da elaboração do artigo em grupo que foi entregue à professora na última aula. O autor da pesquisa acompanhou a turma experimento em todas as aulas e trabalhou como coorientador na elaboração dos artigos.

Na UNIFIG os participantes pertenciam à duas classes do 3º semestre do curso de graduação em Administração no primeiro semestre de 2011. Eles cursaram a disciplina Contabilidade e Análise Financeira I com o autor desta pesquisa como professor. A classe aleatoriamente escolhida como experimento contava com 42 alunos matriculados e a turma controle, 57. Ambas atenderam esta disciplina às sextas-feiras no período noturno, com carga horária de 2 horas-aula semanais. A participação na pesquisa foi voluntária (Apêndice B – Consentimento Explícito) com o incentivo de um ponto adicional na média de quem participasse de todos os pré-testes e pós-testes de avaliação do raciocínio crítico. O controle dos participantes da pesquisa foi realizado por meio de uma planilha de Excel com nome, número (Registro Acadêmico) e data do teste realizado¹¹. Consentiram e iniciaram a participação na pesquisa um total de 73 alunos; porém, ao final do semestre 40 alunos completaram todos os aspectos da pesquisa e foram utilizados nas análises.

A seguir estão alguns gráficos com dados demográficos (Pesquisa Demográfica – Apêndice C) dos alunos que completaram a pesquisa.

¹¹ Essas informações são para controle, e os nomes dos alunos não foram mencionados na pesquisa.

Gráfico 17 - Gênero dos alunos da UNIFIG por turma.

Gráfico 18 - Idade dos alunos da UNIFIG por turma.

Gráfico 19 - Cor/Raça dos alunos da UNIFIG por turma.

Gráfico 20 - Estado civil dos alunos da UNIFIG por turma.

Gráfico 21 - Quantidade de filhos dos alunos da UNIFIG por turma.

Gráfico 22 - Emprego dos alunos da UNIFIG por turma.

Gráfico 23 - Tipo de escola no ensino fundamental e médio dos alunos da UNIFIG por turma.

Gráfico 24 - Renda familiar dos alunos da UNIFIG por turma.

Gráfico 25 - Frequência de leitura dos alunos da UNIFIG por turma.

Gráfico 26 - Frequência de escrita dos alunos da UNIFIG por turma.

Gráfico 27 - Frequência de cálculo dos alunos da UNIFIG por turma.

Gráfico 28 - Fluência em língua estrangeira dos alunos da UNIFIG por turma.

Gráfico 29 - Participação em curso de leitura dos alunos da UNIFIG por turma.

Gráfico 30 - Participação em curso de escrita dos alunos da UNIFIG por turma.

Gráfico 31 - Participação em curso de cálculo dos alunos da UNIFIG por turma.

Gráfico 32 - Participação em curso de raciocínio crítico dos alunos da UNIFIG por turma.

Um cronograma das atividades foi desenvolvido e fielmente seguido para que os mesmos conteúdos fossem aplicados para ambas às classes. Porém, como neste caso o pesquisador foi o próprio professor das turmas, dentre as quatro categorias de modelo instrucional de raciocínio crítico definidas por Ennis (1993), a abordagem aplicada foi a Mista (segundo o mesmo protocolo de Reed (1998) para fins de comparabilidade dos resultados), com a incorporação do modelo de raciocínio crítico em diversas atividades da disciplina, as quais eram aplicadas na turma controle de uma forma mais tradicional. A apostila com a explicação dos conceitos e com o modelo de Richard Paul foi disponibilizada e apresentada para a turma experimento.

3.3 Instrumentos de Avaliação

Este estudo trabalhou com resultados obtidos a partir de dois instrumentos: um teste redacional para avaliar habilidades de raciocínio crítico denominado *Ennis-Weir Critical Thinking Essay Test* (EWCTET) e um teste comercial em escala *Likert* para avaliar as atitudes de raciocínio crítico denominado *California Critical Thinking Dispositions Inventory* (CCTDI) (para o EWCTET, o Prof. Ennis autorizou formalmente a utilização, e para o CCTDI as licenças foram adquiridas). Os dois instrumentos foram aplicados tanto no início (pré-testes) dos cursos quanto no final (pós-testes), sendo que o modelo instrucional só começou a ser ensinado para as turmas experimento após as aplicações dos pré-testes. Estes dois instrumentos de avaliação são padronizados e foram aplicados por Reed (1998) em sua tese de doutorado para medir habilidades e atitudes de raciocínio crítico em alunos de

História. Adicionalmente aos dois instrumentos citados, esta pesquisa ainda considerou os dados de uma pesquisa demográfica (Apêndice C).

As variáveis da pesquisa foram identificadas pelos seguintes códigos:

Quadro 4 - Códigos das variáveis da pesquisa.

Variável	Código
<i>Ennis-Weir Critical Thinking Essay Test</i>	EW
<i>California Critical Thinking Disposition Inventory</i>	CT
Gênero	GE
Idade	ID

Para as variáveis que possuem pré-teste e pós-teste foi adicionada a palavra “pré” ou “pós” subscrito ao código, por exemplo: EW_{pré}. Após a apuração dos dados obteve-se um banco de dados com a seguinte representação:

Quadro 5 - Representação do banco de dados da pesquisa.

Elementos	EW_{pré}	EW_{pós}	CT_{pré}	CT_{pós}	GE	ID
1	EW _{pré1}	EW _{pós1}	CT _{pré1}	CT _{pós1}	GE ₁	ID ₁
2	EW _{pré2}	EW _{pós2}	CT _{pré2}	CT _{pós2}	GE ₂	ID ₂
3	EW _{pré3}	EW _{pós3}	CT _{pré3}	CT _{pós3}	GE ₃	ID ₃
4	EW _{pré4}	EW _{pós4}	CT _{pré4}	CT _{pós4}	GE ₄	ID ₄
5	EW _{pré5}	EW _{pós5}	CT _{pré5}	CT _{pós5}	GE ₅	ID ₅

3.3.1 Ennis-Weir Critical Thinking Essay Test - EWCTET

O EWCTET foi utilizado para testar a habilidade dos estudantes em avaliar um argumento e gerar uma resposta crítica por escrito em relação a um assunto em questão. Este instrumento mensura a habilidade de responder a argumentos quando eles aparecem naturalmente em uma discussão, disputa ou debate do mundo real (ENNIS; WEIR, 1985; CORNACHIONE, DUNCAN; JOHNSON, 2007; REED, 1998).

O teste é composto de uma página escrita pelo editor de um jornal que procura convencer o público sobre a adoção de uma lei que proíba estacionar em ruas públicas durante o período noturno. A carta possui oito parágrafos numerados mais um parágrafo sumário. Os participantes tinham de analisar cada parágrafo da carta teste com o objetivo de escrever uma

curta redação concordando ou discordando de cada argumento exposto nos parágrafos. Os responsáveis por desenvolver o teste proveram uma planilha com os critérios para avaliar cada um dos nove parágrafos escritos em resposta à carta. De acordo com a planilha, os alunos podem atingir notas de -9 a 29, já que para as análises dos primeiros oito parágrafos pode-se atingir de -1 (analisou incorretamente) até 3 (analisou e justificou corretamente) para cada um e no último parágrafo os critérios de correção permitem avaliá-lo de -1 até 5. O tempo máximo de duração recomendado é de 40 minutos (ENNIS; WEIR, 1985).

O Prof. Robert H. Ennis é um dos grandes especialistas no assunto, com diversos artigos publicados na área, além da participação no *Delphi Report*. De acordo com Reed (1998), o EWCTET tem sido utilizado com sucesso em uma variedade de situações e tem recebido forte apoio de especialistas. No artigo de Cornachione, Duncan e Johnson (2007), a utilização do instrumento teve a autorização dos Professores Ennis e Weir (o Prof. Ennis autorizou a utilização do instrumento nesta pesquisa).

Para analisar a confiabilidade do teste foram utilizados dois diferentes avaliadores para 27 redações escritas por estudantes de um curso de lógica informal e 28 redações de alunos do oitavo ano do curso de inglês. Foram obtidas correlações de 0,86 e 0,82, respectivamente, as quais são bastante altas para um teste de redação deste tipo. Na literatura sobre testes, normalmente discutem-se quatro tipos de validade: conteúdo, preditiva, concordância e construção. A validade alegada de maneira mais apropriada para o EWCTET é a de conteúdo (ENNIS; WEIR, 1985).

Na presente pesquisa, o próprio autor aplicou o teste para ambas as turmas da FEA-USP e da UNIFIG. As notas de cada teste saíram da média dos dois avaliadores. Antes de iniciar as correções os avaliadores se reuniram duas vezes; no primeiro encontro, para estudar e discutir o manual do teste, e, no segundo, para compreender, traduzir e discutir a ficha de avaliação dos testes. Após a avaliação das redações, qualquer diferença maior do que três pontos fez com que a redação fosse revisada e discutida pelos avaliadores, até chegar-se à diferença máxima de três pontos. As revisões das duas turmas (controle e experimento) foram realizadas em uma única sessão para cada quase-experimento (REED, 1998).

Após a primeira correção (antes das revisões) dos avaliadores foi analisada a confiabilidade do teste para os dois quase-experimentos. Na instituição pública o coeficiente de correlação (r de Pearson) encontrado foi de 0,76, e na instituição particular, 0,80.

3.3.2 California Critical Thinking Dispositions Inventory – CCTDI

O CCTDI foi desenvolvido por Noreen C. Facione e Peter A. Facione com o apoio de uma divisão da *California Academic Press* chamada *Insight Assessment*, que o comercializa (FACIONE; FACIONE, 2010).

O CCTDI, que utiliza a definição consensual de raciocínio crítico produzida pelo *Delphi Report*, foi criado para avaliar a atitude de raciocínio crítico dos alunos. O teste possui 75 afirmações a partir das quais os alunos devem indicar seu nível de concordância ou discordância em uma escala *Likert* de seis pontos (FACIONE; FACIONE, 2010; REED, 1998).

O teste, que leva de 15 a 20 minutos para ser aplicado, tem seus itens divididos em sete escalas que representam diferentes atitudes de um pensador crítico: buscando a verdade, mente aberta, analiticidade, sistematicidade, autoconfiança, curiosidade e maturidade cognitiva. Para cada atitude existem aproximadamente onze afirmações, e como se trata de um teste comercial, o gabarito não é disponibilizado. Sendo assim, após a aplicação, os formulários de respostas devem ser enviados para a *Insight Assessment*, que, após atribuir as pontuações, envia os dados tabulados em *Excel* via *e-mail* (FACIONE; FACIONE, 2010; REED, 1998).

A pontuação máxima de cada escala é 60, e do teste completo, 420. De acordo com os autores, uma pontuação menor do que 40 em uma escala indica que o indivíduo é fraco para aquela atitude; os que ficarem acima de 50 são fortes para aquela atitude. Dessa forma, uma pontuação acima de 350 indica significativa força, e menor que isso, uma fraqueza em relação à presença das atitudes de um pensador crítico (FACIONE; FACIONE, 2010; REED, 1998).

Em relação à confiabilidade do teste, o alfa de *Cronbach* entre alunos de ensino médio e superior foi de 0,90 e 0,91, respectivamente, e a confiabilidade das escalas ficou entre 0,72 e 0,80 (FACIONE; FACIONE, 2010; REED, 1998).

A alegação da validade do conteúdo do teste se fundamenta no argumento de os desenvolvedores terem participado do *Delphi Report*. O manual do teste e a revisão bibliográfica sobre o assunto demonstram que o teste já foi utilizado com sucesso por muitos pesquisadores (FACIONE; FACIONE, 2010; REED, 1998).

Dentre estes pesquisadores é possível citar: Albert, Albert e Radsma (2002), que analisaram a relação entre a pessoa ser bilíngue, suas habilidades e atitudes de raciocínio crítico; Bartlett e Cox (2002), que utilizaram o CCTDI para avaliar mudanças nas habilidades e atitudes de raciocínio crítico de alunos da área da saúde; e Lederer (2007), que analisou as atitudes de raciocínio crítico entre alunos de terapia ocupacional.

O teste, que foi importado pelo autor desta pesquisa, está disponível em dezessete línguas diferentes, inclusive na portuguesa. Foi aplicado para as turmas da FEA-USP e da UNIFIG na versão “papel e lápis”, e os formulários de respostas padronizados foram digitalizados (uma prevenção para o caso de extravio) e reenviados a *Insight Assessment*, que tabulou as respostas e enviou a planilha de *Excel* via *e-mail* para o autor deste trabalho. Embora exista a versão *online* do teste, optou-se pela versão “papel e lápis” em função do número de participantes (a suposta necessidade de deslocá-los até um laboratório de informática preocupou o autor da pesquisa).

3.3.3 Pesquisa Demográfica

Com base no trabalho de Reed (1998), o pesquisador elaborou uma pesquisa demográfica (Apêndice C). O objetivo foi utilizar dados primários para responder as questões relacionadas às influências de gênero e idade na aplicação do modelo instrucional. Também existem questões com o objetivo de levantar mais dados socioeconômicos (por exemplo: renda, etnia e ensino fundamental em escola particular ou pública) e sobre conhecimentos anteriores de cursos de raciocínio crítico (com diferentes nomenclaturas, como: lógica, raciocínio lógico etc.).

3.4 Aplicações do Modelo Instrucional e Atividades

O modelo de Richard Paul para desenvolvimento de raciocínio crítico, já apresentado no Capítulo 2, é utilizado como base para o tratamento experimental desta pesquisa. No Apêndice A encontra-se a apostila de raciocínio crítico disponibilizada, ensinada e aplicada aos alunos dos grupos experimentos em intervenções, orientações e atividades desenvolvidas em sala de aula.

No quase-experimento piloto da FEA-USP, considerando-se as classificações de Ennis (1993), a abordagem de instrução de raciocínio crítico foi a Geral; logo, o monitor aplicou o modelo para o grupo experimento em quatro intervenções de em média 30 minutos cada uma:

- Ensinando o modelo explicitamente: apresentação da parte conceitual do modelo de desenvolvimento de raciocínio crítico de Richard Paul em *Datashow*.
- Provendo uma apostila sobre o modelo: a estrutura da apostila, com ênfase nos guias de análise de artigos e livros, foi enviada por *e-mail* (lista de *e-mails* individuais obtida em sala) e apresentada em *Datashow*.
- Exemplificando a análise de um artigo utilizando o guia da apostila.
- Exemplificando a resolução e análise de um exercício utilizando o guia da apostila.

No início do segundo semestre de 2010, o cronograma de aulas para ambas as turmas (controle e experimento) foi definido e disponibilizado aos alunos. Os textos entre aspas refletem exatamente o que foi apresentado aos discentes, em seguida o autor explica as atividades desenvolvidas em cada data:

“Dia 02 de agosto: apresentação da disciplina, discussão do programa e Revisão das Demonstrações Contábeis”. A professora apresentou o plano de ensino com objetivos, conteúdo, estratégias de ensino, critérios de avaliação, bibliografia e programação por aula. Realizou-se a revisão das demonstrações, abordaram-se tópicos referentes a impostos (IPI e ICMS), desenvolveu-se um exercício em razonetes relembrando seus cálculos, pediu-se a

resolução de um exercício para o próximo encontro e a leitura do Pronunciamento 28 do CPC e do Capítulo 9 do Manual da Lei das S.A.s¹² (Propriedade para Investimento).

Ações referentes ao quase-experimento: A professora apresentou o monitor (autor desta pesquisa) à turma e explicou a importância do comparecimento nas datas de aplicação dos testes. Ao comentar sobre a necessidade de entrega de um artigo final, lembrou que o monitor ajudaria nas orientações dos artigos e também nas suas avaliações.

“Dia 09 de agosto: Investimentos e Propriedades para Investimento.” A professora apresentou uma aula expositiva (que tomou aproximadamente 90% do tempo) sobre os conceitos, aplicando-os em exercícios. Ela interrompeu perguntando se a turma estava gostando e ressaltou a necessidade de ler a matéria antes para conseguir debater os assuntos. Durante a sessão a professora provocou a classe com questões sobre classificação e contabilização, como por exemplo: a possibilidade de imobilização de um contrato de arrendamento operacional (*leasing*), pois, em sua essência, trata-se de um bem que está sendo adquirido.

“Dia 16 de agosto: Orientação para realização da pesquisa pelos alunos. Aplicação do EWCTET em ambas as turmas.”

Ações referentes ao quase-experimento: Neste dia, os pré-testes do EWCTET foram aplicados em ambas as turmas no início das aulas. Primeiramente, foram dadas aos alunos folhas almaço para as respostas e depois as folhas do teste e de consentimento explícito. As instruções do teste foram lidas para os alunos e dúvidas pertinentes foram solucionadas sem que houvesse interferência na interpretação da carta. Conforme orientação do manual, o teste durou certa de 40 minutos.

Após a aplicação do EWCTET, a professora assumiu a turma e explicou como seria o artigo que eles produziram em grupo (máximo quatro alunos) durante o semestre (entrega até a data da última prova). Todos os artigos deveriam ter o objetivo de contrastar as exigências de alguma norma do CPC contra publicações de empresas de determinado segmento (mínimo de

¹² FIPECAFI. **Manual de Contabilidade Societária:** Aplicável a todas as Sociedades de acordo com as Normas Internacionais e do CPC. 1. ed. São Paulo: Atlas, 2010.

oito empresas) listadas na Bovespa (foram apresentados os *links* do *site* para baixar os demonstrativos). A professora também explicou as partes de um artigo e a metodologia de construção. O autor desta pesquisa colocou-se à disposição da turma experimento para a coorientação e disponibilizou seu *e-mail* para contato.

“Dia 23 de agosto: Propriedades para Investimento. Continuação do conteúdo com correção de exercícios.” A professora iniciou a aula verificando quem tinha lido os materiais disponibilizados pela *internet*, checou se alguém, ao comparar sua resolução de exercícios com o gabarito, ficou com alguma dúvida e iniciou uma aula expositiva abordando os seguintes assuntos: Gastos com pesquisa e desenvolvimento de novos produtos, mensuração de ativos (Custo ou Valor Justo – *Fair Value*) e Redução ao valor recuperável do ativo (*Impairment*). Após o intervalo, corrigiu um exercício que os alunos estavam com dúvida e pediu para que realizassem uma pesquisa sobre EBITDA (*Earnings Before Interest Taxes Depreciation and Amortization*) para o próximo encontro. Na programação inicial a aplicação dos pré-testes do CCTDI em ambas as turmas e a primeira intervenção do modelo de Richard Paul na turma experimento estavam previstos para este dia. Porém, os testes levaram quase 30 dias para chegar dos Estados Unidos, pois ficaram mais de 15 dias na Alfândega Brasileira; logo, ambas foram realizadas no dia 13 de setembro de 2010.

“Dia 30 de agosto: Coligadas, *Joint Ventures* e Controladas.” A professora iniciou a aula corrigindo um exercício e depois desenvolveu os conteúdos relacionados na programação por meio de aula expositiva aplicando os conceitos em exercícios. Ficou como tarefa para a próxima aula ler o Pronunciamento 18 do CPC (Investimento em Coligada e em Controlada). A professora ressaltou que a turma da manhã (experimento) não estava dando tanta importância ao desenvolvimento do artigo quanto à turma da noite (controle).

“Dia 13 de setembro: Método de Equivalência Patrimonial. Aplicação do CCTDI em ambas as turmas.” A professora abriu a aula com uma sessão de dúvidas, depois apresentou os principais conceitos de acordo com a programação e passou a turma para o monitor aplicar os pré-testes do CCTDI.

Ações referentes ao quase-experimento: Primeiramente o monitor entregou aos alunos lápis nº 2, os quais são necessários para o preenchimento da folha de respostas (para a leitura do computador), depois entregou as folhas com as afirmações a serem analisadas e, por último,

as folhas de respostas. Ressaltou o cuidado com a folha de respostas e fez a leitura das instruções do teste, tirando dúvidas que não viessem a interferir no resultado. Após as devoluções do teste, o qual levou em média 20 minutos, o monitor e autor desta pesquisa apresentou o modelo do Richard Paul em *slides*. **A primeira intervenção durou aproximadamente 30 minutos.**

Primeira Intervenção – Apresentação do modelo de desenvolvimento de raciocínio crítico de Richard Paul: Primeiramente introduziu o assunto explicando a atual necessidade do desenvolvimento de habilidades e atitudes de raciocínio crítico nos profissionais na área contábil, em seguida apresentou o modelo de Richard Paul projetando a Figura 1 (Elementos de Raciocínio Crítico) e explicando cada um dos oito elementos necessários nas análises de problemas, discussões, artigos, livros etc. Depois apresentou a Figura 2 (Relacionamento entre os Elementos, Habilidades, Atitudes e Padrões) e explicou que para desenvolver as habilidades é necessário identificar os elementos, possuir as atitudes e aplicar os padrões intelectuais universais. Por fim, abriu espaço para as supostas dúvidas sobre o modelo e, após respondê-las, o monitor encerrou a intervenção pegando uma lista de *e-mails* dos participantes para enviar a apostila sobre raciocínio crítico (Apêndice A). Na turma controle, o tempo que se utilizou para a intervenção da turma experimento foi disponibilizado para dúvidas sobre exercícios e orientação na elaboração do artigo.

“Dia 20 de setembro: Método de Equivalência Patrimonial e Aspectos Complementares. Resolução de Exercícios.” Basicamente foi uma aula com resolução de exercícios e com atendimento às dúvidas. A professora ressaltou durante a aula que sempre antes de iniciar os exercícios os alunos deveriam pensar sobre a situação exposta para depois tomar a decisão de como proceder. Faltando 30 minutos para o encerramento da aula a professora passou a palavra para o monitor, que realizou a **segunda intervenção na turma experimento.**

Segunda Intervenção – Apresentação da apostila de desenvolvimento de raciocínio crítico (Apêndice A): O autor iniciou a apresentação da apostila enviada na semana anterior e ressaltou novamente a importância do desenvolvimento de habilidades e atitudes de raciocínio crítico expondo algumas citações de pesquisas sobre o tema e opiniões de importantes órgãos internacionais de Contabilidade. Em seguida, projetou a apostila em *Datashow* e percorreu rapidamente as páginas explicando a essência de cada tópico e a definição de raciocínio crítico defendida pelo autor desta pesquisa. Enfatizou os guias de análise de artigos ou livros

que identificam os elementos de raciocínio crítico e pediu para que os alunos os utilizassem nos fichamentos que realizariam para a elaboração do artigo exigido pela disciplina. Finalizou respondendo algumas questões, pegando *e-mails* de poucos alunos que haviam faltado no último encontro e disponibilizando novamente o *e-mail* pessoal para contato. Após esta intervenção, cinco alunos procuraram pessoalmente o monitor após a aula para tirar dúvidas de como utilizar o guia para auxiliar os fichamentos e dois grupos começaram a pedir orientações por *e-mail*. Novamente na turma controle, o tempo que se utilizou na turma experimento para a intervenção foi disponibilizado para dúvidas sobre exercícios e orientação na elaboração do artigo.

“Dia 27 de setembro: Avaliação 1.” Nesta data foi aplicada a primeira avaliação individual para ambas as turmas abordando os conceitos expostos até a data. De acordo com a professora, a média de ambas as turmas nesta primeira avaliação foi de 4,6 (abaixo da média 5,0 para aprovação). A professora chamou a atenção das duas turmas em relação à importância de lerem os materiais indicados antes das aulas e, para a turma experimento, enfatizou que o modelo para raciocínio crítico exposto pelo monitor poderia ajudar na confecção do artigo, o qual representa 40% da média total da disciplina.

“Dia 18 de outubro: Patrimônio Líquido, Reservas de capital, de lucros e ajustes de avaliação patrimonial. Conceitos gerais e exercícios.” A professora corrigiu a Avaliação 1 e abordou os conceitos por meio de aula expositiva e resolução de exercícios. Faltando 30 minutos para o encerramento da aula, passou a palavra ao monitor, que iniciou a **terceira intervenção**.

Terceira Intervenção – Apresentação do guia para análise de um artigo: Nesta intervenção o monitor apresentou em *slides* o exemplo de análise de um artigo¹³ (indicado pela professora como modelo para elaboração do artigo em grupo) utilizando o guia para identificar os principais elementos de raciocínio disponibilizado na apostila (Apêndice A) enviada por *e-mail*.

O Quadro 6 a seguir demonstra o exemplo da aplicação do guia:

¹³ PONTE, Vera Maria Rodrigues; LUCA, Márcia Martins Mendes de; OLIVEIRA, Marcelle Colares; AQUINO, Luiz Damásio Pereira de. Análise do grau de cumprimento das práticas de divulgação definidas pelo Pronunciamento Técnico CPC 13 no âmbito das companhias abertas. In: CONGRESSO USP DE CONTROLADORIA E CONTABILIDADE, 10., 2010, São Paulo. **Resumo...** São Paulo: USP, 2010.

Quadro 6 - Exemplo de aplicação do Guia para identificar os elementos de raciocínio crítico de um artigo.

A LÓGICA DE UM ARTIGO	
Análise do grau de cumprimento das práticas de divulgação definidas pelo Pronunciamento Técnico CPC 13 no âmbito das companhias abertas	
PONTE; LUCA, OLIVEIRA; AQUINO.	
<ul style="list-style-type: none"> • O principal propósito deste artigo é levantar o grau de cumprimento das orientações de divulgação definidas no CPC 13 pelas companhias listadas na Bovespa. • A questão-chave é: Existe uma maior observância pelas empresas dos níveis diferenciados de governança corporativa (NDGC)? • A informação mais importante no artigo é o conjunto de práticas de divulgação do CPC 13. • As principais conclusões do artigo são: baixo grau de cumprimento das orientações de divulgação do CPC 13 pelas companhias abertas e maior observância das empresas dos níveis diferenciados de governança corporativa (NDGC) do que as do segmento tradicional. • O(s) conceito(s)-chave que nós precisamos para entender o artigo é (são): Primeiras demonstrações contábeis de acordo com a 11.638/07, Disclosure e Práticas de Divulgação. • A(s) principal(is) suposição(ões) suportando o pensamento do autor é (são) que as empresas dos níveis diferenciados de governança corporativa (NDGC) observam mais as normas do que as empresas do segmento tradicional. • Se as pessoas adotarem esta linha de raciocínio seriamente, as implicações são: considerar que as empresas dos níveis diferenciados de governança corporativa (NDGC) são mais transparentes do que empresas do segmento tradicional e que empresas abertas possuem baixo grau de <i>disclosure</i> (evidenciação). • Se as pessoas falharem em adotar esta linha de raciocínio seriamente, as implicações são: acreditar que empresas abertas possuem alto grau de observância das normas ou que empresas dos NDGC cumprem todos os requisitos das normas de divulgação. • O(s) principal(is) ponto(s) de vista apresentado(s) no artigo é (são): Empresas dos NDGC tendem a observar mais as normas de divulgação. 	

Após a apresentação, foi aberta uma sessão para dúvidas ou comentários e um aluno comentou que a análise por meio do guia proposto torna o trabalho mais objetivo, pois acaba guiando para os principais elementos (objetivo, questão, informações, conclusões, conceitos, suposições, implicações e ponto de vista). O monitor procurou abrir um debate sobre o assunto, porém não houve interesse dos demais alunos em sala. O monitor finalizou a intervenção apresentando novamente o *e-mail* para dúvidas ou orientações. Os mesmos alunos que mostraram interesse na última participação procuraram mais informações ao final da aula e apenas dois *e-mails* sobre o artigo e três pedindo o envio da apostila foram respondidos até a data da intervenção seguinte (dia 8 de novembro).

“Dia 25 de outubro: Dividendos e Juros sobre o capital próprio, conceitos, critérios de cálculo e contabilização. Resolução de Exercícios.” A professora apresentou os conceitos por meio de aula expositiva e resolveu cinco exercícios de uma lista anteriormente disponibilizada pela *internet*.

“Dia 08 de novembro: Dividendos e Juros sobre o capital próprio, conceitos, critérios de cálculo e contabilização. Resolução de Exercícios.” A professora continuou a resolução de exercícios e as sessões para tirar dúvidas sobre a elaboração do artigo. Nos últimos 30 minutos do encontro, a professora passou a palavra ao monitor que iniciou a **quarta e última intervenção** antes da aplicação dos pós-testes.

Quarta Intervenção – Apresentação do guia para análise de um exercício: O monitor apresentou em *slides* a análise da solução de um exercício da lista disponibilizada pela professora. A intenção foi reforçar que os elementos de raciocínio crítico podem ser identificados em outros lugares e situações, como: na solução de problemas, em discussões, debates, na leitura de outros tipos de documentos etc.

O Quadro 7 a seguir demonstra um exemplo da aplicação do guia:

Quadro 7 - Exemplo de aplicação do Guia para identificar os elementos de raciocínio crítico de um exercício.

Exercício 21 - Cia “Duvidosa”	
Provisão ou Reserva para Contingência	
•	A questão-chave é: contabilizar como provisão ou reserva para contingência?
•	O(s) conceito(s)-chave que nós precisamos entender é (são): Provisão – ligado ao risco do negócio e à geração de uma obrigação; Reserva para Contingência – mais voltada a eventos que não são inerentes ao negócio, por exemplo, eventos naturais.
•	As informações mais importantes no problema são: as que permitem identificar se determinado fato deve ser provisionado ou gerará uma reserva.
•	As principais conclusões dos problemas são: as situações a) e c) referem-se a riscos inerentes ao negócio; logo, foram provisionadas; e as situações b) e d) referem-se a eventos futuros e problemas de gestão; logo, nem a reserva nem a provisão foram realizadas.
•	Se as pessoas adotarem esta linha de raciocínio seriamente, a implicação é: uma adequada classificação, contabilização e evidenciação do fato.
•	Se as pessoas falharem em adotar esta linha de raciocínio seriamente, a implicação é: uma inadequada contabilização, classificação e evidenciação do fato.

- O(s) principal(is) ponto(s) de vista apresentado(s) da resolução do problema é (são): **a correta evidencição da situação financeira, econômica e patrimonial da entidade em questão.**

“**Dia 22 de novembro:** Dividendos e Juros sobre o capital próprio, conceitos, critérios de cálculo e contabilização. Resolução de Exercícios. Aplicação do **EWCTET** em ambas as turmas.” A professora continuou desenvolvendo os conceitos e os aplicando na resolução de exercícios.

Ações referentes ao quase-experimento: Nos últimos 40 minutos da aula, em ambas as turmas, houve a aplicação dos pós-testes do EWCTET. O monitor explicou para ambos os grupos a importância de se aplicar a mesma seriedade nas respostas dos testes, visto que o objetivo é avaliar a evolução dos participantes durante o curso. Os procedimentos utilizados na aplicação dos pré-testes foram novamente seguidos nos pós-testes.

“**Dia 29 de novembro:** Demonstração das Origens e Aplicações de Recursos (DOAR), conceitos, elaboração e formas de apresentação. Aplicação do **CCTDI** em ambas as turmas.”

Ações referentes ao quase-experimento: Nos primeiros 20 minutos da aula o monitor aplicou os pós-testes do CCTDI seguindo os mesmos procedimentos adotados nos pré-testes e ressaltando a importância de se manter o mesmo nível de seriedade empregado da primeira vez.

Posteriormente às aplicações, a professora apresentou os conceitos e os desenvolveu aplicando em exercícios; porém, ressaltou que apenas a parte conceitual cairia na Avaliação 2.

“**Dia 06 de dezembro:** Avaliação 2, data limite para entrega do artigo e aplicação da pesquisa demográfica.” Neste encontro ocorreu a Avaliação 2, que englobou todo o conteúdo do semestre. Grande parte dos artigos foi entregue neste dia.

Ações referentes ao quase-experimento: Nesta data aplicou-se a pesquisa demográfica, que foi entregue juntamente com as provas.

“Dia 13 de dezembro: Reavaliação.” Neste encontro houve a reavaliação para os alunos que não obtiveram 5,0 como média mínima e não ficaram abaixo de 3,0 pontos.

Ações referentes ao quase-experimento: Houve contato por *e-mail* com os alunos que perderam algum dos pós-testes para que participassem neste dia; mas nenhum aluno faltante se interessou.

No geral, a programação do curso seguiu fielmente o planejamento inicial, com um atraso na aplicação dos pré-testes do CCTDI e da primeira intervenção na turma experimento, conforme já foi explicado. As quatro intervenções consumiram em média 30 minutos e na turma controle este tempo foi sempre substituído por resoluções de exercícios e atendimento a dúvidas sobre a elaboração do artigo. Para os faltantes dos pré-testes foram agendadas sessões após o horário normal de aulas em duas datas, mas nenhum aluno compareceu. Os resultados dos testes e dados da pesquisa demográfica são analisados no Capítulo 4 desta pesquisa.

Na UNIFIG, considerando-se as classificações de Ennis (1993), a abordagem de instrução de raciocínio crítico foi a Mista, já que o autor deste trabalho foi o professor tanto da turma experimento quanto da controle. O modelo e os conceitos foram ensinados da mesma forma que na FEA-USP (intervenções explicando o modelo, a apostila e a utilização dos guias); porém, o modelo e seus conceitos também foram aplicados e lembrados durante todo o semestre em todas as atividades da turma experimento, enquanto que a turma controle possuiu um modelo mais tradicional de ensino da disciplina. É importante ressaltar que após a aplicação do quase-experimento piloto na FEA-USP houve maior compreensão do modelo por parte do autor desta pesquisa e constatou-se que seria necessário aumentar a intensidade das intervenções para avaliar a eficiência do modelo no desenvolvimento de habilidades e atitudes de raciocínio crítico nos alunos.

No início do primeiro semestre de 2011, o cronograma de aulas para ambas as turmas (com a diferença das intervenções e da metodologia de ensino) foi definido e disponibilizado aos alunos. Os textos entre aspas refletem exatamente o que foi apresentado aos discentes, em seguida o autor explica as atividades desenvolvidas em cada data.

“Dia 11 de fevereiro: Introdução ao Plano de Ensino: Objetivo da disciplina, Ementa, Cronograma de aulas, apresentação do Sistema de Avaliação, Parte Prática, Metodologia de Ensino e Bibliografia. Os alunos deverão formar grupos e trazer demonstrações de jornais para a próxima aula”. Neste dia foi apresentado o Plano de Ensino e explicada a pesquisa para ambas as turmas. Além disso, demonstrou-se todo o cronograma de atividades incentivando a participação na pesquisa com o acréscimo de 1,0 ponto na média geral daqueles alunos que atendessem todos os testes durante o semestre. Como a disciplina enfoca a análise de demonstrações financeiras, os alunos de ambas as turmas foram divididos em grupos de no máximo cinco integrantes para que pesquisassem em *sites* e jornais e trouxessem demonstrações reais para o próximo encontro.

“Dia 18 de fevereiro: Revisão das demonstrações financeiras, estrutura e finalidades. Análise e escolha das demonstrações coletadas pelos alunos. Participação na Pesquisa de Doutorado do Professor (**Teste EWCTET**).”

Ações referentes ao quase-experimento: O professor iniciou a sessão com a aplicação do pré-teste EWCTET seguindo os mesmos procedimentos detalhados no cronograma da FEA-USP.

Posteriormente continuou com uma aula expositiva projetando em *Datashow* uma revisão de todos os demonstrativos obrigatórios de acordo com a Lei 11.638/07, abordou as demonstrações e informações opcionais e depois recolheu as demonstrações de empresas reais trazidas pelos grupos.

“Dia 25 de fevereiro: Introdução à análise das demonstrações financeiras: objetivo, diferença entre dados e informações e usuários. Padronização das demonstrações contábeis: objetivo, prática e exercícios. Verificação da padronização do balanço patrimonial. Verificação da padronização da DRE. Participação na Pesquisa de Doutorado do Professor (**Teste CCTDI**).”

Ações referentes ao quase-experimento: O professor iniciou a sessão com a aplicação dos pré-testes CCTDI seguindo os mesmos procedimentos detalhados no cronograma da FEA-USP.

Posteriormente, devolveu os demonstrativos entregues na última sessão pelos grupos, colocou seu parecer sobre qual empresa deveria ser escolhida para ser analisada por todos os grupos e concedeu 15 minutos para que eles analisassem as empresas e votassem em apenas uma. Em ambas as turmas as empresas escolhidas foram as sugeridas pelo professor, uma vez que apresentou argumentos favoráveis à escolha, como: parecer da auditoria sem ressalvas, transparência nas práticas detalhas em notas explicativas, boa aparência dos demonstrativos, bom relatório da administração etc. Como cada sessão é de apenas duas horas-aula, não houve tempo suficiente para abordar os conceitos programados inicialmente.

“Dia 04 de março: Exercícios de fixação/Aplicação no trabalho Prático/Análise de artigos.” Nesta sessão o professor iniciou os conteúdos programados para o dia 25 de fevereiro, apresentou os conceitos em *Datashow* e explicou uma planilha de *Excel* contendo um modelo de sintetização e padronização dos demonstrativos da empresa escolhida pela turma, a qual foi analisada durante o semestre em grupo. Todos os conceitos vistos em sala de aula foram aplicados neste trabalho prático (entregue na data da última prova). O professor tinha enviado os materiais desta aula aos alunos da turma controle e experimento no *e-mail* das turmas; porém, nesta data passou uma lista na turma experimento coletando os *e-mails* particulares dos alunos (para enviar os materiais das intervenções e não correr o risco de enviar para alunos da turma controle).

“Dia 11 de março: Exercícios de fixação/Aplicação no trabalho Prático/Análise de artigos.” Este encontro foi disponibilizado para que os alunos em sala de aula resumissem os demonstrativos da empresa escolhida no modelo sugerido e disponibilizado pelo professor. Esta atividade foi desenvolvida em ambas as turmas com supervisão do professor.

Primeira Intervenção – Apresentação do modelo de desenvolvimento de raciocínio crítico de Richard Paul: Na turma experimento o autor desta pesquisa aproveitou os últimos 30 minutos para apresentar e explicar o modelo de Richard Paul. Da mesma forma que

realizou a primeira intervenção na FEA-USP, o professor introduziu o assunto explicando a atual necessidade do desenvolvimento de habilidades e atitudes de raciocínio crítico nos profissionais da área de negócios. Em seguida apresentou o modelo de Richard Paul projetando em *Datashow* a Figura 1 (Elementos de Raciocínio Crítico) e explicou cada um dos oito elementos necessários nas análises de problemas, discussões, artigos, livros etc. Depois apresentou a Figura 2 (Relacionamento entre os Elementos, Habilidades, Atitudes e Padrões) explicando que para desenvolver as habilidades é necessário identificar os elementos, possuir as atitudes e aplicar os padrões intelectuais universais. Por fim, o professor abriu espaço para as supostas dúvidas sobre o modelo e, após respondê-las, encerrou dizendo que tinha enviado aos *e-mails* pessoais passados na última aula uma apostila sobre raciocínio crítico (Apêndice A) que seria apresentada e explicada no próximo encontro.

“Dia 18 de março: Análise Vertical e Horizontal: Conceitos e exercícios. Os alunos aplicarão o conceito na parte prática.” O professor iniciou a sessão com uma aula expositiva sobre os conceitos programados para a data e depois disponibilizou o restante da aula para que os grupos os aplicassem no trabalho prático.

Segunda Intervenção – Apresentação da apostila de desenvolvimento de raciocínio crítico (Apêndice A): Na turma experimento, nos últimos 30 minutos do encontro o professor apresentou e explicou a apostila enviada por *e-mail* utilizando os mesmos *slides* da intervenção realizada na FEA-USP. Primeiramente o professor ressaltou a importância do desenvolvimento de habilidades e atitudes de raciocínio crítico expondo algumas citações de pesquisas sobre o tema. Depois projetou a apostila em *Datashow* e percorreu rapidamente as páginas explicando a essência de cada tópico e a definição de raciocínio crítico defendida pelo autor desta pesquisa. Enfatizou os guias de análise de artigos ou livros que identificam os elementos de raciocínio crítico e apresentou um exemplo de análise de artigo (Quadro 6). Falou sobre o livro *O Homem que Calculava*¹⁴, do autor Malba Tahan, e contou a história do capítulo “Trinta e cinco camelos”. Comentou sobre a importância de adquirir o hábito da leitura, elogiou o livro indicado como instrumento para melhorar o raciocínio e pediu que utilizassem o guia da apostila para análise de um capítulo a escolha (atividade que acrescentou 0,5 ponto na nota do primeiro bimestre). Finalizou respondendo algumas

¹⁴ TAHAN, Malba. *O Homem que Calculava*. 55. ed. São Paulo: Editora Record 2001.

questões sobre a apostila, o guia, a bibliografia indicada e avisou que enviaria por *e-mail* artigos para análises em sala de aula.

“Dia 25 de março: Exercícios de fixação/Aplicação no trabalho Prático/Análise de artigos.” Nesta data o professor fez uma revisão para ambas as turmas dos conceitos abordados até a data e que foram objetos de avaliação do primeiro bimestre.

Ações referentes ao quase-experimento: Quatro alunos da turma experimento tiraram dúvidas sobre a utilização do guia para análise de um capítulo do livro indicado.

“Dia 01 de abril: Período de Avaliações do 1º Bimestre”. São aproximadamente duas semanas de prova, com apenas uma avaliação por dia. Nesta data a turma experimento foi avaliada.

Ações referentes ao quase-experimento: Os alunos da turma experimento entregaram a análise do capítulo do livro utilizando o guia da apostila.

“Dia 08 de abril: Período de Avaliações do 1º Bimestre.” Nesta data a turma controle foi avaliada e entregaram um exercício de Análise Vertical e Horizontal.

“Dia 15 de abril: Índices de Liquidez: Conceito e Exercícios.” O professor iniciou o encontro com a correção da primeira avaliação. Seguiu com uma aula expositiva sobre Índices de Liquidez, com alguns exercícios para entrega e o restante da aula disponível para aplicação dos conceitos no trabalho prático em grupo com a supervisão do professor.

Terceira Intervenção – Apresentação do guia para análise de um artigo ou livro: Nos últimos 30 minutos da aula da turma experimento, o autor apresentou detalhadamente o guia de análise de um artigo ou livro dando exemplos de aplicações. Posteriormente comentou as análises dos capítulos entregues na data da primeira avaliação, abrindo para discussão e respondendo às questões. Enfatizou a objetividade na análise quando da utilização do modelo e cinco alunos concordaram com esta opinião. Os demais não mostraram interesse e não apresentaram dúvidas ou críticas.

“Dia 29 de Abril: Índices de Endividamento: Conceito e Exercícios.” No início da sessão o professor abordou os conceitos programados para a data, corrigiu os exercícios entregues e disponibilizou o restante do encontro para o desenvolvimento do trabalho prático em grupos.

Quarta Intervenção – Apresentação do guia para análise de um exercício: Na turma experimento durante os últimos 30 minutos da aula o professor apresentou detalhadamente o guia para análise de um exercício disponibilizado na apostila. Posteriormente corrigiu um exercício utilizando também o guia para identificação dos elementos de raciocínio (no mesmo molde do Quadro 7). Este exercício apresentava demonstrativos de duas empresas e perguntava, de acordo com diferentes visões (usuários: investidor, banco, cliente, fornecedor, concorrente e governo), qual empresa seria a escolhida e a razão desta decisão.

“Dia 06 de maio: Índices de Rentabilidade: Conceito e Exercícios.” O professor iniciou a sessão com uma aula expositiva dos conceitos programados para a data. Posteriormente disponibilizou o restante da aula para aplicação dos conceitos no trabalho prático em grupo.

Ações referentes ao quase-experimento: Nos últimos 20 minutos da aula da turma experimento projetou em *Datashow* o artigo “Lucro Gasoso: uma interpretação do velho EVA¹⁵”, do Prof. José Roberto Kassai, e explicou os principais elementos de raciocínio do artigo. O professor avisou os alunos de que o artigo já havia sido enviado por *e-mail* e pediu para que o analisassem utilizando o guia da apostila (esta atividade somou 0,5 ponto à avaliação final e a data limite para entrega foi dia 27 de maio).

“Dia 13 de maio: Encerramento do trabalho prático em grupo em sala de aula. Entrega do Exercício de análise ‘Verdadeiro ou Falso e Justifique’”. Este encontro foi disponibilizado para encerramento do trabalho prático em grupos com a supervisão do professor. Ocorreu também a correção de um exercício com demonstrativos contábeis e índices já calculados e com 15 afirmações, as quais deveriam ser assinaladas como verdadeiras ou falsas e depois justificadas.

¹⁵ KASSAI, José Roberto. Lucro gasoso: uma interpretação do velho EVA. In: ENCONTRO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, 28., 2004, Curitiba. **Anais...** Curitiba, 2004.

Ações referentes ao quase-experimento: Na turma experimento, durante a correção do exercício, o professor relembrou os oito elementos de raciocínio crítico, principalmente na elaboração das justificativas.

“Dia 20 de maio: Avaliação Institucional.” Nesta data houve a aplicação da avaliação institucional, que se trata de uma prova com questões de múltipla escolha de todas as disciplinas que o aluno estiver cursando no semestre.

“Dia 27 de maio: Revisão para a Avaliação Final. Participação na Pesquisa de Doutorado do Professor (Teste EWCTET).”

Ações referentes ao quase-experimento: No início da sessão o professor aplicou os pós-testes do EWCTET, seguindo os mesmos procedimentos descritos anteriormente e ressaltando a importância em responder o teste com a mesma seriedade empregada pela primeira vez.

Após a aplicação, iniciou uma revisão dos conceitos para a avaliação final e abriu uma sessão para tirar dúvidas.

Ações referentes ao quase-experimento: Na turma experimento, nos últimos 20 minutos da aula o professor apresentou a análise do artigo “Lucro Gasoso” (entregue pelos alunos nesta data) nos mesmos moldes do Quadro 6 e abriu a questionamentos. Apenas um aluno mostrou dúvidas na apresentação e teceu comentários positivos à utilização do guia para a análise. Os demais alunos não criticaram e acompanharam a apresentação até o final.

“Dia 03 de junho: Revisão para a Avaliação Final. Participação na Pesquisa de Doutorado do Professor (Teste CCTDI).”

Ações referentes ao quase-experimento: No início da aula o autor aplicou os pós-testes do CCTDI, seguindo os mesmos procedimentos descritos anteriormente e ressaltando a importância em manter o mesmo nível de seriedade empregado nos pré-testes.

Em seguida, continuou a revisão dos conceitos para a avaliação final e encerrou com uma sessão para tirar dúvidas.

“Dia 10 de junho: Período de Avaliações, Pesquisa Demográfica e Data Limite para entrega do Relatório em Grupo.” Nesta data foi aplicada a avaliação final para a turma experimento e os grupos entregaram o relatório final do trabalho prático desenvolvido durante o semestre.

Ações referentes ao quase-experimento: A pesquisa demográfica foi entregue juntamente com a prova.

“Dia 17 de junho: Período de Avaliações, Pesquisa Demográfica e Data Limite para entrega do Relatório em Grupo.” Nesta data foi aplicada a avaliação final para a turma controle e os grupos entregaram o relatório final do trabalho prático desenvolvido durante o semestre.

Ações referentes ao quase-experimento: A pesquisa demográfica foi entregue juntamente com a prova.

“Dia 24 de junho: Correção da Avaliação Final.”

Ações referentes ao quase-experimento: Além da correção, esta data foi disponibilizada para os alunos que perderam algum dos testes de raciocínio crítico; porém, nenhum mostrou interesse.

A execução dos dois quase-experimentos (o piloto na FEA-USP e o da UNIFIG) durou cerca de um ano, e, como todas as intervenções, aulas e aplicações dos testes tinham de ser previamente planejadas, estima-se um total de 90 horas de planejamento despendidas para cada experimento.

No geral, como no quase-experimento piloto na FEA-USP, a programação do curso na UNIFIG seguiu fielmente o planejamento inicial, com apenas um atraso de conteúdo, que foi compensado na aula seguinte. Como orientado nos manuais dos testes, as aplicações do EWCTET consumiram em média 40 minutos e as do CCTDI, 20 minutos. Foram quatro intervenções apenas sobre o modelo (no mesmo molde de Reed (1998)), porém as ações e atividades utilizando os conceitos do modelo de Richard Paul e os guias da apostila foram frequentes na turma experimento, já que o professor procurava sempre que possível ligá-los com os conceitos, exemplos, exercícios e atividades desenvolvidas em sala e extraclasse. Para

os faltantes dos pré-testes foram agendadas sessões antes do horário normal de aulas em duas datas; porém, nenhum aluno compareceu. Os resultados dos testes e dados da pesquisa demográfica são apresentados e analisados no Capítulo 4.

3.5 Método do Trabalho

Segundo Lakatos (1985, p. 81),

Todas as ciências caracterizam-se pela utilização de métodos científicos; em contrapartida, nem todos os ramos de estudo que empregam estes métodos são ciências. Dessas afirmações podemos concluir que a utilização de métodos científicos não é da alçada exclusiva da ciência, mas não há ciência sem o emprego de métodos científicos.

Como o presente trabalho é científico, faz-se necessário seguir uma metodologia científica, ou seja, um caminho científico que direcione aos objetivos da investigação. De acordo com Godoy (1995, p. 61), existem pesquisas de cunho quantitativo e qualitativo:

[...] em linhas gerais, num estudo quantitativo o pesquisador conduz seu trabalho a partir de um plano estabelecido a priori, com hipóteses claramente especificadas e variáveis operacionalmente definidas. Preocupa-se com a medição objetiva e a quantificação dos resultados. Busca a precisão, evitando distorções na etapa de análise e interpretação dos dados, garantindo assim uma margem de segurança em relação às inferências obtidas.

E, sobre a pesquisa qualitativa, Godoy (1995, p. 61) disserta:

[...] de maneira diversa, a pesquisa qualitativa não procura enumerar e / ou medir os eventos estudados, nem emprega instrumental estatístico na análise dos dados. Parte de questões ou focos de interesses amplos, que vão se definindo a medida que o estudo se desenvolve. Envolve a obtenção de dados descritivos sobre pessoas, lugares e processos interativos pelo contato direto do pesquisador com a situação estudada, procurando compreender os fenômenos segundo a perspectiva dos sujeitos, ou seja, dos participantes da situação em estudo.

Como esta pesquisa envolve obtenção de dados descritivos sobre pessoas, lugares e processos com o contato direto do pesquisador, pode-se dizer que possui traços qualitativos na interpretação e compreensão do fenômeno em questão. Porém, como se trata de um quase-experimento com a utilização de instrumental estatístico, a pesquisa é classificada como quantitativa.

A metodologia da pesquisa empírica seguiu as seguintes etapas:

- 1) Aplicação do modelo de Richard Paul em duas classes de disciplinas contábeis dos cursos de graduação em Ciências Contábeis e Administração, uma de IES pública (piloto) e outra privada.
- 2) Aplicação de um método tradicional de ensino em duas classes dos mesmos semestres e disciplinas citadas anteriormente.
- 3) Aplicação de um pré-teste e um pós-teste (EWCTET) que avaliou as habilidades de raciocínio crítico de ambas as classes (experimento e controle).
- 4) Aplicação de um pré-teste e um pós-teste (CCTDI) que avaliou as atitudes de raciocínio crítico de ambas as classes (experimento e controle).
- 5) Análises estatísticas dos resultados com comparação das turmas controle e experimento, considerando também as seguintes variáveis: gênero e idade (menor de 22 anos ou 22 anos ou mais velho).
- 6) Comparação dos resultados com a pesquisa de Reed (1998).

3.5.1 Modelo de Pesquisa

Neste tópico são abordados os aspectos teóricos do modelo de pesquisa, as definições dos termos, as preocupações com aspectos de validade e confiabilidade que servirão de guia na aplicação empírica da presente pesquisa e na análise dos dados e resultados.

Muitos experimentos em educação empregam alguma forma de modelo de variável. Experimentos de uma variável envolvem a manipulação de uma simples variável de tratamento seguida pela observação dos efeitos desta manipulação sobre uma ou mais variáveis dependentes. A variável a ser manipulada pode receber as seguintes denominações: variável de tratamento experimental, independente, experimental, de tratamento ou de intervenção. Muitos experimentos em educação envolvem um grupo que recebe o tratamento experimental e um grupo de comparação que não o recebe ou recebe um tratamento

alternativo. O grupo que recebe o tratamento é normalmente chamado de “grupo experimento” e o de comparação, de “grupo controle” (CAMPBELL; STANLEY, 1981).

Em experimentos e quase-experimentos são feitas comparações entre o que acontece quando um tratamento está presente e o que acontece quando nenhum tratamento (ou um tratamento alternativo) está presente. Uma das diferenças entre esses dois tipos de modelo é que em um experimento, um processo aleatório determina quem vai receber ou não o tratamento ou o tratamento alternativo, enquanto que em um quase-experimento, a escolha dos participantes do tratamento não é aleatória (BICKMAN; ROG, 1998).

Em um quase-experimento, em que o grupo como um todo foi escolhido aleatoriamente e não cada um dos participantes, o pré-teste é necessário para ser considerado como covariável nos ajustes das médias dos pós-testes na procura da diminuição e controle das diferenças iniciais (antes do tratamento).

Bickman e Rog (1998, p. 211-212) listam pelo menos quatro vantagens em se ter um pré-teste em modelos de pesquisa experimentais ou quase-experimentais:

1. Pré-testes podem aumentar o poder e a precisão das análises, com mais benefício, maior o valor absoluto da correlação entre o pré-teste e pós-teste.
2. Escolha aleatória não é sempre fácil de implementar e algumas vezes é corrompida (Boruch & Wothke, 1985; Braucht & Reichardt, 1993; Conner, 1977) e pré-testes podem ajudar o pesquisador a avaliar a probabilidade do procedimento de seleção aleatória ter sido comprometido e o dano causado.
3. Diferentes desgastes entre grupos de tratamento podem introduzir viés na seleção aleatória do experimento e pré-testes permitem que o pesquisador verifique estes vieses.
4. O efeito do tratamento pode variar entre os indivíduos com diferentes características e a inclusão de um ou mais pré-testes permite que o pesquisador estime o tamanho destas interações do tratamento com características individuais.

Os passos envolvidos no modelo quase-experimento desta pesquisa são os mesmos de um “modelo experimental com pré-teste, pós-teste e grupo controle” (CAMPBELL; STANLEY, 1981).

A Figura 3 a seguir demonstra o modelo quase-experimento desta pesquisa:

Figura 3 - Modelo da pesquisa.

Um dos cuidados que se deve ter para desenhar um experimento é com as validades internas. Trata-se de um dos fatores que pode enfraquecer o poder de um experimento. A discussão de validade interna fundamenta-se na extensão de quais variáveis, além da variável tratamento, proveem explicações plausíveis dos resultados de um experimento. Existe também a validade externa de um experimento, que trata da extensão em que os achados podem ser generalizados a uma determinada população (BICKMAN; ROG, 1998; CAMPBELL; STANLEY, 1981).

Campbell e Stanley (1981) listam diversos fatores que afetam a validade interna e externa de um experimento, cujas variáveis devem ser mantidas constantes ou eliminadas para que não afetem as mensurações do pós-teste.

Quadro 8 - Fatores que afetam a validade interna e externa de experimentos.

Validade Interna

- 1) História
- 2) Maturação
- 3) Testando
- 4) Instrumentação
- 5) Regressão Estatística

- 6) Seleção Diferencial
- 7) Mortalidade Experimental
- 8) Interação seleção-maturação
- 9) Difusão do tratamento experimental
- 10) Rivalidade compensatória pelo grupo controle
- 11) Equalização compensatória dos tratamentos
- 12) Desmoralização ressentida do grupo controle

Validade Externa

Validade da População

- 1) A extensão que alguém pode generalizar de uma amostra experimental para uma população definida
- 2) A extensão que variáveis pessoais da amostra interagem com os efeitos do tratamento

Validade Ecológica

- 1) Descrição explícita do tratamento experimental
- 2) Interferência do tratamento múltiplo
- 3) Efeito *Hawthorne*
- 4) Originalidade e efeitos transtornados
- 5) Efeito experimentador
- 6) Sensibilização do pré-teste
- 7) Sensibilização do pós-teste
- 8) Interação da história e dos efeitos do tratamento
- 9) Mensuração da variável dependente
- 10) Interação do tempo de mensuração e dos efeitos do tratamento

Fonte: CAMPBELL; STANLEY, 1981, p. 368.

Segue a explicação de cada uma das validades reportadas anteriormente e de como foram tratadas na aplicação desta pesquisa:

Validade Interna

- 1) **História:** ligada ao ambiente e ao tempo a que os grupos ficaram expostos, deve-se garantir que todas as características são iguais, somente a intervenção é diferente. Esta pesquisa possuiu o modelo de um quase-experimento com grupos tratamento e experimento em cada uma das Instituições de Ensino Superior com as mesmas

disciplinas e semestre; logo, tanto o tempo quanto o ambiente são semelhantes dentro de cada quase-experimento.

- 2) **Maturação:** em relação à idade dos participantes ou ao tempo do experimento, torna-se necessário possuir um grupo controle para garantir que o efeito medido não se deve ao tempo do experimento nem às diferenças de idades do grupo controle e experimento. Esta pesquisa possuiu um grupo controle e um experimento dentro de cada um dos dois quase-experimentos (piloto na FEA-USP e quase-experimento na UNIFIG), e as diferenças de idade e gênero foram consideradas e analisadas e eventuais diferenças são reportadas como limitações nas conclusões (Capítulo 5).
- 3) **Testando:** devido ao problema de se ter o pré-teste igual ao pós-teste, houve a necessidade de possuir um grupo controle. Esta pesquisa possuiu um grupo controle dentro de cada um dos dois quase-experimentos.
- 4) **Instrumentação:** ligada à tendência dos avaliadores em mensurar melhor o pós-teste. Ambos os instrumentos de avaliação utilizados foram padronizados e possuíam critérios específicos de mensuração (Manuais). Um deles foi corrigido pela própria entidade que o comercializou (*Insight Assessment*) e o outro foi corrigido por dois avaliadores, tendo sua confiabilidade analisada. Limitações referentes à possíveis vieses dos avaliadores são analisadas no item limitações dentro das conclusões (Capítulo 5).
- 5) **Regressão Estatística:** em relação ao problema dos que obtiveram menores escores nos pré-testes, a tendência é a de que tivessem maiores ganhos nos pós-testes. Nesta pesquisa, esta suposta tendência foi analisada por meio de regressão estatística nos dois quase-experimentos e, aparentemente, não ocorreu este problema.
- 6) **Seleção Diferencial:** devido às diferenças iniciais, houve a necessidade de se aplicar pré-testes. Nesta pesquisa se aplicou pré-testes nos dois quase-experimentos.
- 7) **Mortalidade Experimental:** refere-se ao problema de a intervenção causar um número maior de desistentes. Logo, deve-se verificar e comparar as desistências entre o grupo controle e experimento. Nesta pesquisa, no quase-experimento piloto na

instituição pública, o percentual de desistentes (não completaram todos os requisitos da pesquisa) entre as turmas experimento e controle foi muito semelhante, 42% e 43%, respectivamente. No quase-experimento na instituição particular, a turma controle foi a que mais teve desistentes, 49%, enquanto a turma experimento teve 38%. Possíveis limitações referentes às desistências são analisadas nas conclusões (Capítulo 5).

- 8) **Interação seleção-maturação:** refere-se ao problema de um dos grupos possuir maior faixa etária. Esta pesquisa considerou e analisou as diferenças entre as idades dos grupos controle e experimento dentro de cada quase-experimento e as possíveis limitações são analisadas nas conclusões (Capítulo 5).
- 9) **Difusão do tratamento experimental:** refere-se ao risco de o pessoal do grupo controle procurar saber quais são as intervenções no grupo experimento. No quase-experimento piloto da FEA-USP esta possibilidade era pequena, já que uma das turmas era do período matutino e a outra, do noturno. Além disso, não houve transferência de alunos entre elas durante o semestre. Na UNIFIG a possibilidade era maior, pois ambas eram do período noturno. No entanto, pela disposição física das classes (uma ficava no 4º andar e a outra no 2º andar do edifício da faculdade), pelo professor ter tido o cuidado de mandar o material das intervenções para os *e-mails* pessoais, por não ter ocorrido transferências entre as turmas e nenhum comentário sobre metodologia ou matérias diferentes entre as turmas, considera-se pequeno o risco a esta validade. Porém, as possíveis limitações são analisadas nas conclusões (Capítulo 5).
- 10) **Rivalidade compensatória pelo grupo controle:** devido à possibilidade de tomarem conhecimento de que estão sendo comparados, o grupo controle se esforçar para obter melhores resultados. Nesta pesquisa, em ambos os quase-experimentos o autor não notou e nem obteve evidências de que os grupos sabiam que estavam sendo comparados. Eles apenas tinham conhecimento de que faziam parte de uma pesquisa. Porém, possíveis limitações são analisadas nas conclusões (Capítulo 5).
- 11) **Equalização compensatória dos tratamentos:** refere-se ao problema de se procurar compensar o tempo da intervenção gasto com o grupo experimento com outras

atividades no grupo controle, sendo que estas podem interferir nos resultados. Nesta pesquisa, em ambos os quase-experimentos, as atividades compensatórias utilizadas nos grupos controle durante o tempo em que os grupos experimentos recebiam o tratamento foram comuns entre os grupos. Os grupos controle não receberam tipo algum de atividade diferenciada. Receberam apenas uma quantidade maior de atividades tradicionais das disciplinas. Porém, como o tempo para analisar os exercícios clássicos da disciplina foram maiores para a turma controle, possíveis limitações são analisadas nas conclusões (Capítulo 5).

- 12) **Desmoralização ressentida do grupo controle:** ligada ao risco de o grupo controle se sentir desmotivado por saber que não recebe a mesma atenção na pesquisa (e o seu desempenho cair por esta razão). Nesta pesquisa, em ambos os quase-experimentos o autor não notou nem obteve evidências de que os grupos obtiveram conhecimento sobre quaisquer diferenças de tratamento, atenção ou metodologia de ensino entre eles. Porém, possíveis limitações são analisadas nas conclusões (Capítulo 5).

Validade Externa

Validade da População

- 1) **A extensão que alguém pode generalizar de uma amostra experimental para uma população definida:** verificar o erro padrão é uma maneira, já que se pode analisar o desvio padrão das médias amostrais em relação à média da população e analisar a representatividade da amostra. Nesta pesquisa são calculados e analisados os erros padrão¹⁶ das amostras para as principais descobertas (Capítulo 5).
- 2) **A extensão que variáveis pessoais da amostra interagem com os efeitos do tratamento:** neste caso, a análise das características da amostra que podem interferir na generalização dos resultados. Nesta pesquisa foi aplicado um questionário (Pesquisa Demográfica – Apêndice C) que levantou dados primários sobre aspectos socioeconômicos dos participantes da pesquisa que foram analisados e considerados nos

¹⁶ “O erro padrão é o desvio padrão das médias das amostras. Como tal, ele é uma medida de quão representativa a amostra poderá ser da população.” FIELD, Andy. **Descobrendo a estatística usando o SPSS**. Trad. Lorí Viali. 2. ed. Porto Alegre: Artmed, 2009.

resultados finais. Portanto, possíveis limitações são analisadas nas conclusões (Capítulo 5).

Validade Ecológica

- 1) **Descrição explícita do tratamento experimental:** detalhar os procedimentos para que outros pesquisadores possam replicá-lo. Esta pesquisa relatou detalhadamente cada aula das turmas controle e experimento e a maneira como foram aplicadas as intervenções sobre o modelo e as ações relativas a cada um dos quase-experimentos.
- 2) **Interferência do tratamento múltiplo:** o problema relacionado à aplicação de vários tratamentos ao mesmo tempo. Nesta pesquisa houve a aplicação de apenas um modelo de desenvolvimento de raciocínio crítico (Modelo de Richard Paul) em diferentes atividades (análise de artigos, livros, exercícios e debates). Porém, para aplicar o modelo os alunos da turma experimento foram frequentemente incentivados a lerem livros, artigos e resolver exercícios, logo, as possíveis limitações referentes a estas ações são analisadas nas conclusões (Capítulo 5).
- 3) **Efeito *Hawthorne*:** Basicamente o efeito *Hawthorne* ocorre por causa de outras variáveis (que não são as variáveis independentes do estudo) que podem explicar as variações nos resultados do que se quer avaliar. De acordo com Adair (1984), as variáveis *Hawthorne* que normalmente são controladas em estudos são: atenção especial ao grupo experimento, consciência de ser um participante e a característica de novidade da intervenção. Segundo o autor, um exemplo de controle seria: para que o grupo controle não sentisse que o grupo experimento estivesse tendo uma atenção especial por parte do professor, poder-se-ia aplicar uma atividade (placebo) que não interferisse no que se quisesse avaliar, logo, eles se sentiriam como se fossem da turma experimento. Portanto, as variáveis *Hawthorne* e seus efeitos são basicamente todas as possíveis variáveis que podem vir a interferir nos resultados e que não foram consideradas como as principais variáveis independentes da pesquisa (as intervenções). Muitos são os estudos que não compreendem claramente seu conceito (ADAIR, 1984). Nas conclusões da presente pesquisa listam-se as possíveis limitações dos dois quase-experimentos realizados, dentre elas as ligadas aos efeitos *Hawthorne*.

- 4) **Originalidade e efeitos transtornados:** em relação ao problema de algumas intervenções darem certo pelo fato de serem pioneiras ou novas; porém, depois de conhecidas, não conseguem obter o mesmo efeito desejado. O autor desta pesquisa notou que apresentar explicitamente um modelo de raciocínio crítico é algo novo para os alunos; no entanto, observou pouco entusiasmo e dedicação da maioria dos participantes. Portanto, possíveis limitações ligadas ao fato de ser uma novidade ou à desmotivação dos alunos são analisadas nas conclusões (Capítulo 5).
- 5) **Efeito experimentador:** devido ao risco de os resultados se deverem exclusivamente ao público ou instrutor e não ao instrumento de intervenção. Nesta pesquisa, os participantes são alunos aparentemente comuns aos cursos e períodos que cursam. O instrutor foi monitor no quase-experimento piloto e professor quase-experimento que é comparado ao estudo de Reed (1998). Nas avaliações docentes das instituições em que atua, o autor fica em torno da média do corpo docente, sem grandes elogios ou reclamações por parte dos discentes. Porém, possíveis limitações são analisadas nas conclusões (Capítulo 5).
- 6) **Sensibilização do pré-teste:** considerando-se que o pré-teste está relacionado ao instrumento de intervenção, o problema de quanto ele pode afetar os resultados do grupo controle. Nesta pesquisa, após a aplicação dos pré-testes, não houve correção, comentário ou debate sobre seus conteúdos, diminuindo, assim, a interferência nos resultados dos pós-testes. Porém, possíveis limitações são analisadas nas conclusões (Capítulo 5).
- 7) **Sensibilização do pós-teste:** em relação à influência de alguma experiência que o grupo controle pode ter tido que afetou os resultados do pós-teste. Nesta pesquisa, as turmas controle dos dois quase-experimentos tiveram os conteúdos ministrados de forma tradicional (aula expositiva seguida de exercícios) e não realizaram atividade que os grupos experimento não tivessem realizado. Portanto, caso alguma atividade ou conteúdo tradicional das disciplinas tenham influenciado os pós-testes, isto ocorreu para ambas às turmas. Porém, possíveis limitações ligadas ao fato de algumas atividades terem tido maior tempo para a turma controle, já que este foi utilizado na turma experimento para as intervenções, são analisadas nas conclusões (Capítulo 5).

- 8) **Interação da história e dos efeitos do tratamento:** o problema de o método ter dado resultado positivo, pois está atualmente sendo muito defendido e incentivado. Nesta pesquisa, o autor notou que ensinar explicitamente um modelo de raciocínio crítico foi novidade para os alunos. Na revisão de literatura deste trabalho constatou-se que o raciocínio crítico é internacionalmente discutido, defendido e incentivado há anos, mas no Brasil ainda não o é.
- 9) **Mensuração da variável dependente:** o problema de possuir diferentes formas de aplicar um teste e possuir diferentes resultados por isso (por exemplo: um grupo respondeu o teste no computador e o outro no papel). Nesta pesquisa, os instrumentos utilizados foram padronizados, seguiram os mesmos procedimentos de aplicação detalhados em seus Manuais e foram todos respondidos em papel.
- 10) **Interação do tempo de mensuração e dos efeitos do tratamento:** o risco de se perder o efeito da intervenção por causa de uma demora na aplicação dos pós-testes. Nesta pesquisa, os pós-testes foram aplicados uma sessão antes das avaliações finais e logo após as últimas intervenções.

3.5.2 Procedimentos para Análise dos Dados

O problema de pesquisa definido no Capítulo 1 deste trabalho é:

A aplicação do modelo instrucional de raciocínio crítico de Richard Paul melhora o nível de habilidades e atitudes de raciocínio crítico dos alunos de disciplinas da área contábil?

No Capítulo 1 também foi exposto o objetivo geral do trabalho, que é conhecer, analisar e avaliar empiricamente a efetividade em ensinar o modelo de Richard Paul (1995) para raciocínio crítico em alunos de disciplinas da área contábil no desenvolvimento de habilidades para pensar criticamente sobre contabilidade e atitudes de raciocínio crítico em geral.

Com base no problema e objetivo definidos acima e no trabalho de Reed (1998), esta pesquisa possui as seguintes hipóteses:

- 1) Uma classe que receber explicitamente treinamento sobre o modelo de desenvolvimento de raciocínio crítico de Richard Paul obterá melhor resultado em um teste (EWCTET) que avalia habilidade em análise e resposta de argumentos sobre um assunto contemporâneo do que uma classe que não receber instruções explícitas sobre o modelo.

Para esta hipótese, realizaram-se análises descritivas das amostras em relação ao desempenho no EWCTET e método de instrução, checkou-se a possível existência de diferenças significativas entre as médias das turmas em cada quase-experimento e se as turmas experimento apresentaram melhor desempenho. Logo, esperou-se que:

$$\mu_{\text{experimento}} > \mu_{\text{controle}}$$

- 2) Uma classe que receber explicitamente treinamento sobre o modelo de desenvolvimento de raciocínio crítico de Richard Paul obterá melhor resultado em um teste (CCTDI) que avalia atitudes de raciocínio crítico do que uma classe que não receber instruções explícitas sobre o modelo.

Neste caso, assim como no anterior, realizaram-se análises descritivas das amostras em relação ao desempenho no CCTDI e método de instrução, checkou-se a possível existência de diferenças significativas entre as médias das turmas em cada quase-experimento e se as turmas experimento apresentaram melhor desempenho. Logo, esperou-se que:

$$\mu_{\text{experimento}} > \mu_{\text{controle}}$$

- 3) Alunos mais velhos (22 anos e acima) obterão melhor desempenho nos testes (EWCTET e CCTDI) de habilidades e atitudes de raciocínio crítico do que alunos mais novos (abaixo de 22 anos) de acordo com o método instrucional.

Para esta hipótese, realizaram-se análises descritivas de acordo com a idade e método de instrução, checkaram-se possíveis diferenças significativas entre as médias dos desempenhos entre os grupos experimento e controle considerando a idade e se os participantes com maior faixa etária obtiveram melhor desempenho. Logo, no geral, esperou-se que:

$$\mu_{\text{maisvelhos}} > \mu_{\text{maisnovos}}$$

- 4) Esperam-se diferenças significativas nos testes (EWCTET e CCTDI) de habilidades e atitudes de raciocínio crítico dos alunos de acordo com o método instrucional e gênero (feminino/masculino).

Para esta hipótese, realizaram-se análises descritivas de acordo com o gênero e método de instrução e checaram-se possíveis diferenças significativas entre os desempenhos por método de instrução e gênero. Logo, no geral, esperou-se que:

$$\mu_{\text{masculino}} \neq \mu_{\text{feminino}}$$

Para todas as hipóteses definidas foram realizadas análises descritivas para resumir os desempenhos das notas no começo (pré-teste) e final (pós-teste) do curso por método de instrução, gênero e idade. Para cada situação foi analisado se a amostra demonstrou aspectos de uma distribuição normal. Foram examinados padrões de interações entre variáveis para descrever relacionamentos entre o método de instrução e desempenho de acordo com idade e gênero (REED, 1998).

Como já discutido, não é possível sortear cada aluno para o grupo experimento, e sim a classe como um todo. Portanto, esta é uma das razões pela qual o modelo da pesquisa é classificado como um quase-experimento. De acordo com Martins e Theóphilo (2009) outra característica de um quase-experimento que é típico das pesquisas na área das ciências sociais é que não há total controle sobre as variáveis do estudo, faltando assim pleno controle da aplicação dos estímulos experimentais.

Uma das ameaças às validades internas deste modelo de pesquisa é a possibilidade de que diferenças entre os grupos nos pós-testes sejam devidas às diferenças que já existiam antes mesmo dos efeitos do tratamento. A análise de covariância (ANCOVA) é frequentemente utilizada para lidar com este problema. A ANCOVA estatisticamente reduz os efeitos das diferenças iniciais dos grupos compensando com ajustes nas médias dos pós-testes dos dois grupos. A utilização desta técnica estatística exige que os dados satisfaçam seus requisitos (BICKMAN; ROG, 1998; CAMPBELL; STANLEY, 1981; FIELD, 2009).

A pesquisa de Reed (1998) também utilizou a ANCOVA, já que aplicou pré-testes e pós-testes. Como já comentado anteriormente, a intenção é reduzir os efeitos das diferenças iniciais entre os grupos utilizando os pré-testes para ajustar os pós-testes. O resultado final, a média ajustada, possibilita analisar o desempenho durante o experimento, já que o seu cálculo considera os escores iniciais e finais nos testes.

A ANCOVA seguiu o modelo: 2 (turma) x 2 (idade) x 2 (gênero), ou seja, analisou os dois grupos (classe experimento e controle), idade (menor de 22 anos e 22 anos ou mais velho) e gênero (feminino e masculino) considerando os pré-testes como covariáveis. As análises apropriadas dos requisitos para a utilização da ANCOVA foram realizadas (REED, 1998; FIELD, 2009).

Para a execução dos tratamentos estatísticos dos dados primários e secundários foram utilizados os *softwares* SPSS for *Windows* versão 13.0 e o *Excel* 2010.

4 ANÁLISE DOS DADOS E RESULTADOS

Na análise dos dados e resultados desta pesquisa foram consideradas três variáveis independentes: método de instrução (instrução que inclui o modelo de Richard Paul e instrução que não o inclui), idade (alunos com menos de 22 anos e com 22 anos ou mais velho) e gênero (masculino e feminino). As variáveis dependentes ou de saída foram obtidas por meio de dois instrumentos: o que avaliou as habilidades de raciocínio crítico (EWCTET) e o que avaliou as atitudes de raciocínio crítico (CCTDI). Adicionalmente foram obtidos dados socioeconômicos dos participantes por meio de um questionário (Pesquisa Demográfica).

Este capítulo relata os resultados deste estudo e como eles se relacionam com as hipóteses da pesquisa, realiza as descrições das amostras e relata os procedimentos das análises.

4.1 Análises Descritivas das Amostras

Como se tratam de dois quase-experimentos (um piloto e o quase-experimento comparável ao estudo de Reed (1998)), os itens são divididos entre FEA-USP (piloto) e UNIFIG.

4.1.1 FEA-USP

As médias e os desvios padrão das amostras em cada um dos dois pré-testes são apresentados na Tabela 1. Os dados referem-se ao total de alunos que iniciaram a pesquisa (EWCTET, n = 85 e CCTDI, n = 90) e aos que completaram todos os instrumentos (n = 52).

Tabela 1 - Distribuição dos pré-testes na FEA-USP.

<u>Instrumentos</u>	<u>Alunos que iniciaram a pesquisa (n = 85 e 90)</u>		<u>Alunos que completaram a pesquisa (n = 52)</u>	
	<u>Média</u>	<u>Desv. Padrão</u>	<u>Média</u>	<u>Desv. Padrão</u>
EWCTET	13,74	6,58	14,13	6,70
CCTDI	300,64	26,28	297,92	26,59

É possível notar que os alunos que terminaram a pesquisa obtiveram média um pouco superior no EWCTET do que os alunos que iniciaram e um pouco inferior no CCTDI. Porém, devido à semelhança das estatísticas descritivas e dos histogramas analisados, pode-se dizer que ambas as amostras aparentaram ser normais, da mesma população e nenhum valor atípico (*outlier*) foi encontrado. Quando realizado um teste t para as médias do EWCTET entre os alunos que iniciaram e completaram a pesquisa, não houve diferença significativa, $t(51) = -1,42$, $p > 0,05$ ($p = 0,16$). Também não houve diferença significativa para o CCTDI, $t(51) = 0,16$, $p > 0,05$ ($p = 0,88$).

As medidas de curtose e assimetria ficaram próximas de zero, e para se ter certeza de que eram normais, realizou-se o teste de Kolmogorov-Smirnov (K-S) indicando os seguintes resultados:

- Alunos que iniciaram a pesquisa:
 - EWCTET_{pré}: $D(85) = 0,07$, $p > 0,05$
 - CCTDI_{pré}: $D(90) = 0,06$, $p > 0,05$
- Alunos que completaram a pesquisa:
 - EWCTET_{pré}: $D(52) = 0,10$, $p > 0,05$
 - CCTDI_{pré}: $D(52) = 0,08$, $p > 0,05$

Portanto, como o teste K-S não foi significativo ($p > 0,05$), pode-se dizer que as amostras se assemelham a uma distribuição normal. Para a amostra que completou todos os requisitos da pesquisa, a Tabela 2 apresenta as estatísticas descritivas para os pré-testes e pós-testes.

Tabela 2 - Distribuição dos pré-testes e pós-testes na FEA-USP.

<u>Instrumentos</u>	<u>(n = 52)</u>			
	<u>Pré-teste</u>		<u>Pós-teste</u>	
	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>
EWCTET	14,13	6,70	12,50	6,21
CCTDI	297,92	26,59	294,15	30,90

Nota-se que, no geral, as médias dos pós-testes foram menores que as dos pré-testes. Os gráficos de caixa e bigode não demonstraram valores atípicos. Analisaram-se os histogramas dos pós-testes (os pré-testes foram analisados anteriormente) e o CCTDI aparentou normalidade; porém, o mesmo não ocorreu com o EWCTET, o qual pareceu bimodal. Logo, realizaram-se os testes K-S para os pós-testes e os resultados foram:

- $EWCTET_{pós}: D(52) = 0,14, p < 0,05$
- $CCTDI_{pós}: D(52) = 0,05, p > 0,05$

Portanto, no geral, a amostra é normal para os pós-testes do CCTDI, porém foram significativos para os EWCTET, ou seja, a distribuição difere de uma normal ao nível de significância de 5%. No entanto, antes de proceder a uma transformação dos dados para procurar normalizá-los, realizou-se uma análise de suas distribuições de acordo com as turmas (experimento e controle) e obtiveram-se os seguintes resultados:

– Turma Experimento:

- $EWCTET_{pré}: D(25) = 0,12, p > 0,05$
- $EWCTET_{pós}: D(25) = 0,16, p > 0,05$
- $CCTDI_{pré}: D(25) = 0,10, p > 0,05$
- $CCTDI_{pós}: D(25) = 0,10, p > 0,05$

– Turma Controle:

- $EWCTET_{pré}: D(27) = 0,11, p > 0,05$
- $EWCTET_{pós}: D(27) = 0,16, p > 0,05$
- $CCTDI_{pré}: D(25) = 0,14, p > 0,05$
- $CCTDI_{pós}: D(25) = 0,07, p > 0,05$

Logo, como a maioria das hipóteses se baseia na comparação entre os grupos controle e experimento, decidiu-se por não transformar os dados dos pós-testes do EWCTET, já que eles não foram significativos ($p > 0,05$) em cada turma (experimento e controle).

Para as futuras considerações realizadas a partir das análises das amostras, torna-se importante calcular o erro padrão. De acordo com Field (2009, p. 44),

[...] o erro padrão é o desvio padrão das médias das amostras. Como tal, ele é uma medida de quão representativa a amostra poderá ser da população. Um erro padrão grande (comparado à média da amostra) informa que existe muita variabilidade entre as médias das diferentes amostras e, dessa forma, a amostra que temos pode não ser representativa da população. Um erro padrão pequeno indica que muitas médias amostrais são similares (estão próximas) à média da população e, assim, a nossa amostra será provavelmente uma boa representação da população.

Para calcular uma estimativa do erro padrão, pode-se dividir o desvio padrão da amostra pela raiz quadrada do tamanho da amostra. A Tabela 3 a seguir apresenta as médias e erros padrão dos pré-testes e pós-testes da instituição pública.

Tabela 3 - Distribuição dos erros padrão dos pré-testes e pós-testes na FEA-USP.

<u>Instrumentos</u>	<u>(n = 52)</u>			
	<u>Pré-teste</u>		<u>Pós-teste</u>	
	<u>Média</u>	<u>Erro padrão</u>	<u>Média</u>	<u>Erro padrão</u>
EWCTET	14,13	0,93	12,50	0,86
CCTDI	297,92	3,69	294,15	4,29

Ao analisar os erros padrão verifica-se que eles são pequenos em relação às médias; logo, indicam que a amostra é provavelmente uma boa representação da população.

Como o estudo mede os efeitos da instrução do modelo de Richard Paul considerando gênero e idade para cada tipo de instituição (pública ou privada), estatísticas descritivas dos dois instrumentos foram obtidas por gênero e idade.

A distribuição dos escores nos pré-testes e pós-testes por idade é apresentada na Tabela 4 a seguir.

Tabela 4 - Distribuição dos pré-testes e pós-testes por idade na FEA-USP.

<u>Instrumentos</u>	<u>Abaixo de 22 anos (n = 42)</u>				<u>22 anos ou acima (n = 10)</u>			
	<u>Pré-teste</u>		<u>Pós-teste</u>		<u>Pré-teste</u>		<u>Pós-teste</u>	
	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>
EWCTET	13,98	6,90	12,25	6,67	14,75	6,06	13,55	3,77
CCTDI	295,48	25,22	290,64	31,19	308,20	31,03	308,90	26,07

No geral, os alunos mais velhos obtiveram escores maiores do que os alunos mais novos, tanto nos pré-testes quanto nos pós-testes. A maioria dos pós-testes foi menor que os pré-testes, com exceção do CCTDI para os alunos mais velhos, que na média permaneceu praticamente estável. A maior queda foi do EWCTET para os alunos mais novos (-1,73 ou 12%). As análises estatísticas que verificaram se as médias são significativamente diferentes considerando método de instrução e idade são realizadas no tópico 4.4.

Na Tabela 5 tem-se a distribuição dos pré-testes e pós-testes de acordo com o gênero.

Tabela 5 - Distribuição dos pré-testes e pós-testes por gênero na FEA-USP.

<u>Instrumentos</u>	<u>Feminino (n = 19)</u>				<u>Masculino (n = 33)</u>			
	<u>Pré-teste</u>		<u>Pós-teste</u>		<u>Pré-teste</u>		<u>Pós-teste</u>	
	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>
EWCTET	11,26	7,51	11,89	7,07	15,77	5,67	12,85	5,74
CCTDI	288,53	24,09	293,47	21,77	303,33	26,79	294,55	35,42

Verifica-se que tanto os pré-testes quanto os pós-testes dos homens foram maiores que os das mulheres. As mulheres tiveram um crescimento médio em ambos os instrumentos, enquanto os valores dos homens caíram. As análises estatísticas que verificaram se as médias são significativamente diferentes considerando método de instrução e gênero são realizadas no tópico 4.5.

4.1.2 UNIFIG

As médias e os desvios padrão das amostras em cada um dos dois pré-testes são demonstrados na Tabela 6. Os dados referem-se ao total de alunos que iniciaram a pesquisa (EWCTET, $n = 73$ e CCTDI, $n = 56$) e aos que completaram todos os instrumentos ($n = 40$).

Tabela 6 - Distribuição dos pré-testes na UNIFIG.

<u>Instrumentos</u>	<u>Alunos que iniciaram a pesquisa</u> <u>(n = 73 e 56)</u>		<u>Alunos que completaram a pesquisa</u> <u>(n = 40)</u>	
	<u>Média</u>	<u>Desv. Padrão</u>	<u>Média</u>	<u>Desv. Padrão</u>
EWCTET	4,42	4,81	3,60	4,51
CCTDI	294,57	28,78	297,18	30,22

Nota-se que os alunos que terminaram a pesquisa obtiveram média um pouco inferior no EWCTET do que os alunos que iniciaram e um pouco superior no CCTDI. Porém, devido à semelhança das estatísticas descritivas e dos histogramas analisados, pode-se dizer que ambas as amostras aparentaram ser normais, da mesma população e nenhum valor atípico da amostra dos alunos que completaram a pesquisa foi encontrado. Quando realizado um teste t para as médias do EWCTET entre os alunos que iniciaram e completaram a pesquisa, não houve diferença significativa, $t(39) = 1,62$, $p > 0,05$ ($p = 0,11$). Também não houve diferença significativa para o CCTDI, $t(39) = -0,53$, $p > 0,05$ ($p = 0,60$).

Como as medidas de curtose e assimetria ficaram próximas de zero, para se ter certeza de que eram aparentemente normais realizou-se o teste de Kolmogorov-Smirnov (K-S), que indicou os seguintes resultados:

- Alunos que iniciaram a pesquisa:
 - EWCTET_{pré}: $D(73) = 0,08$, $p > 0,05$
 - CCTDI_{pré}: $D(56) = 0,06$, $p > 0,05$
- Alunos que completaram a pesquisa:
 - EWCTET_{pré}: $D(40) = 0,08$, $p > 0,05$
 - CCTDI_{pré}: $D(40) = 0,10$, $p > 0,05$

Portanto, como o teste K-S não foi significativo ($p > 0,05$), pode-se dizer que as amostras se assemelham à uma distribuição normal. Para a amostra que completou todos os requisitos da pesquisa a Tabela 7 apresenta as estatísticas descritivas para os pré-testes e pós-testes.

Tabela 7 - Distribuição dos pré-testes e pós-testes na UNIFIG.

(n = 40)				
<u>Instrumentos</u>	<u>Pré-teste</u>		<u>Pós-teste</u>	
	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>
EWCTET	3,60	4,51	4,91	4,74
CCTDI	297,18	30,22	294,63	32,97

Os gráficos de caixa e bigode apresentaram um valor atípico para os pós-testes do EWCTET; portanto, ele foi excluído. A Tabela 8, com a distribuição dos pré-testes e pós-testes sem valores atípicos, pode ser analisada a seguir.

Tabela 8 - Distribuição dos pré-testes e pós-testes na UNIFIG sem valores atípicos.

(n = 39)				
<u>Instrumentos</u>	<u>Pré-teste</u>		<u>Pós-teste</u>	
	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>
EWCTET	3,40	4,38	4,58	4,30
CCTDI	297,08	30,60	294,38	33,37

Nota-se que, no geral, as médias dos pós-testes do EWCTET foram maiores que as dos pré-testes e que o contrário ocorreu com os valores do CCTDI. Os gráficos de caixa e bigode não apresentaram valores atípicos. Analisaram-se os histogramas dos pré-testes e pós-testes para ambos os instrumentos, os quais aparentaram ser normais. Realizaram-se os testes K-S e os resultados foram os seguintes:

- $EWCTET_{pré}: D(39) = 0,11, p > 0,05$
- $EWCTET_{pós}: D(39) = 0,11, p > 0,05$
- $CCTDI_{pré}: D(39) = 0,08, p > 0,05$
- $CCTDI_{pós}: D(39) = 0,10, p > 0,05$

Portanto, no geral, as amostras são significativamente normais. Como nesta pesquisa há comparações entre as turmas controle e experimento, testou-se a normalidade para os pré-testes e pós-testes de cada turma:

- Turma Experimento:
 - EWCTET_{pré}: $D(16) = 0,11$, $p > 0,05$
 - EWCTET_{pós}: $D(16) = 0,15$, $p > 0,05$
 - CCTDI_{pré}: $D(16) = 0,17$, $p > 0,05$
 - CCTDI_{pós}: $D(16) = 0,14$, $p > 0,05$
- Turma Controle:
 - EWCTET_{pré}: $D(23) = 0,14$, $p > 0,05$
 - EWCTET_{pós}: $D(23) = 0,10$, $p > 0,05$
 - CCTDI_{pré}: $D(23) = 0,09$, $p > 0,05$
 - CCTDI_{pós}: $D(23) = 0,15$, $p > 0,05$

Logo, o pressuposto de normalidade não é violado para ambos os instrumentos e turmas.

Para as futuras considerações realizadas a partir das análises das amostras, tem-se a Tabela 9, que apresenta as médias e os erros padrão dos pré-testes e pós-testes da instituição particular.

Tabela 9 - Distribuição dos erros padrão dos pré-testes e pós-testes na UNIFIG.

<u>(n = 39)</u>				
<u>Instrumentos</u>	<u>Pré-teste</u>		<u>Pós-teste</u>	
	<u>Média</u>	<u>Erro padrão</u>	<u>Média</u>	<u>Erro padrão</u>
EWCTET	3,40	0,70	4,58	0,69
CCTDI	297,08	4,90	294,38	5,34

Ao analisar os erros padrão verifica-se que eles são pequenos em relação às médias; logo, indicam que a amostra é provavelmente uma boa representação da população.

Como o estudo mede os efeitos da instrução do modelo de Richard Paul considerando gênero e idade, estatísticas descritivas dos dois instrumentos foram obtidas considerando estas variáveis.

A distribuição dos escores nos pré-testes e pós-testes por idade é apresentada na Tabela 10.

Tabela 10 - Distribuição dos pré-testes e pós-testes por idade na UNIFIG.

<u>Instrumentos</u>	<u>Abaixo de 22 anos (n = 21)</u>				<u>22 anos ou acima (n = 18)</u>			
	<u>Pré-teste</u>		<u>Pós-teste</u>		<u>Pré-teste</u>		<u>Pós-teste</u>	
	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>
EWCTET	3,21	5,02	3,43	4,24	3,61	3,62	5,92	4,08
CCTDI	293,81	30,18	290,52	33,29	300,89	31,52	298,89	33,84

No geral, os alunos mais velhos obtiveram escores maiores do que os alunos mais novos, tanto nos pré-testes quanto nos pós-testes. Os pós-testes do EWCTET foram maiores para ambos os grupos, enquanto os pós-testes do CCTDI foram menores. As análises estatísticas que verificaram se as médias são significativamente diferentes considerando método de instrução e idade são realizadas no tópico 4.4.

Na Tabela 11 tem-se a distribuição dos pré-testes e pós-testes de acordo com o gênero.

Tabela 11 - Distribuição dos pré-testes e pós-testes por gênero na UNIFIG.

<u>Instrumentos</u>	<u>Feminino (n = 18)</u>				<u>Masculino (n = 21)</u>			
	<u>Pré-teste</u>		<u>Pós-teste</u>		<u>Pré-teste</u>		<u>Pós-teste</u>	
	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>
EWCTET	3,19	3,99	4,19	4,40	3,57	4,78	4,90	4,29
CCTDI	290,28	28,35	282,17	31,27	302,90	31,92	304,86	32,17

Verifica-se que tanto os pré-testes quanto os pós-testes dos homens foram maiores que os das mulheres. Os homens tiveram um crescimento médio em ambos os instrumentos, enquanto os pós-testes do EWCTET das mulheres aumentaram em relação aos pré-testes e os CCTDI caíram. As análises estatísticas que verificaram se as médias são significativamente diferentes considerando método de instrução e gênero são realizadas no tópico 4.5.

4.2 Hipótese 1 – Habilidades de Raciocínio Crítico

Relembrando, a Hipótese 1 definida para esta pesquisa foi:

Uma classe que receber explicitamente treinamento sobre o modelo de desenvolvimento de raciocínio crítico de Richard Paul obterá melhor resultado em um teste (EWCTET) que avalia

habilidade em análise e resposta de argumentos sobre um assunto contemporâneo do que uma classe que não receber instruções explícitas sobre o modelo.

Para esta hipótese, realizaram-se análises descritivas das amostras em relação ao desempenho no EWCTET e método de instrução, checkou-se a possível existência de diferenças significativas entre as médias das turmas em cada quase-experimento e se as turmas experimento apresentaram melhor desempenho.

4.2.1 FEA - USP

A distribuição dos escores no EWCTET dos grupos experimento e controle são apresentados na Tabela 12.

Tabela 12 - Distribuição do EWCTET por método de instrução na FEA-USP.

Medidas	Grupo Experimento (n = 25)		Grupo Controle (n = 27)	
	Pré-teste	Pós-teste	Pré-teste	Pós-teste
Média	11,28	12,36	16,76	12,63
DP	6,84	7,33	5,46	5,09

Analísaram-se os histogramas e os gráficos de caixa e bigode e as distribuições aparentaram ser normais (com assimetrias e curtoses próximas de zero) e nenhum valor atípico foi encontrado. No item 4.1.1 os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras são normais. Os testes de homogeneidade das variâncias foram realizados utilizando a estatística de Levene e apresentaram os seguintes resultados:

- $EWCTET_{pré}: F(1, 50) = 1,72, p > 0,05$.
- $EWCTET_{pós}: F(1, 50) = 3,32, p > 0,05$.

Portanto, como ambos os testes de Levene não foram significativos a um nível de 5%, pode-se dizer que existe homogeneidade das variâncias e este requisito para as análises estatísticas não foi violado. Os pré-testes foram em média bem maiores para o grupo controle do que para o grupo experimento; no entanto, o grupo experimento obteve aumento (1,08 ou 10%) no pós-teste em relação ao pré-teste, enquanto o grupo controle obteve grande queda (4,13 ou 25%). No final, as médias obtidas nos pós-testes foram bem similares entre os grupos.

Os pré-requisitos para a ANCOVA foram examinados para os escores do EWCTET. O pressuposto de independência dos escores de cada indivíduo não foi violado, já que cada aluno realizou os testes individualmente e as turmas foram ensinadas pela mesma professora e da mesma forma (com a única diferença do método de instrução). O segundo pré-requisito é o da normalidade e já se verificou anteriormente que as amostras assemelham-se à uma distribuição normal. O terceiro pré-requisito é o da homogeneidade da variância, que também já foi testado e tido como não violado. O quarto pré-requisito refere-se à mensuração das variáveis dependentes em uma escala de intervalo, que também não foi violado para esta pesquisa, já que tanto o EWCTET quanto o CCTDI são intervalares. O quinto pré-requisito é o da confiabilidade da correção dos testes pelos avaliadores, o qual não foi violado (discutido no item 3.3.1). O sexto pré-requisito refere-se à homogeneidade dos parâmetros da regressão, que, de acordo com Field (2009, p. 358),

Quando uma ANCOVA é realizada, olhamos para o relacionamento total entre a saída (variável dependente) e a covariável: ajustamos uma linha de regressão para todo o conjunto de dados, ignorando o grupo que a pessoa pertence. Ajustando esse modelo nós, portanto, assumimos que esse relacionamento é verdadeiro para todos os grupos de participantes. Por exemplo, se existe um relacionamento positivo entre a covariável e a saída em um grupo, assumimos que existe um relacionamento positivo em todos os grupos também. Se, entretanto, o relacionamento entre a saída (variável dependente) e a covariável difere entre os grupos, todo o modelo de regressão não é preciso (ele não representa todos os grupos). Essa suposição é muito importante e é chamada de hipótese da homogeneidade dos parâmetros da regressão.

A seguir tem-se a Tabela 13 que se refere à saída do SPSS (Anexo C – Saída 1) que testa a homogeneidade dos parâmetros da regressão considerando uma interação entre a covariável e as turmas (FIELD, 2009).

Tabela 13 - Interação entre a covariável e as turmas.

Fonte	Soma dos Quadrados		gl	F	Sig
	Tipo III				
Modelo Corrigido	509,951		3	5,604	0,002
Intercepto	267,475		1	8,818	0,005
Turma	2,017		1	0,067	0,798
EWCTETpre	424,744		1	14,002	0,000
Turma*EWCTETpre	27,339		1	0,901	0,347
Erro	1.456,049		48		
Total	10.091,000		52		
Total Corrigido	1.966.000		51		

Analisada a interação “Turma*EWCTETpre”, verifica-se que o efeito não é significativo e que o pré-requisito da homogeneidade dos parâmetros da regressão não foi violado, $F(1, 48) = 0,90$, $p > 0,05$.

Na Tabela 14 a seguir pode-se analisar a saída do SPSS (Anexo C – Saída 2) da Análise de Variância (ANOVA) quando a covariável (pré-teste) não está incluída.

Tabela 14 - ANOVA para habilidades de raciocínio crítico – FEA-USP.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig
Modelo Corrigido	0,944	1	0,024	0,877
Intercepto	8.106,251	1	206,260	0,000
Turma	0,944	1	0,024	0,877
Erro	1.965,056	50		
Total	10.091,000	52		
Total Corrigido	1.966,000	51		

Nota-se pelo valor da significância (0,88) maior que 5% que não há diferenças significativas nos pós-testes entre as turmas experimento e controle, $F(1, 50) = 0,02$, $p > 0,05$. Logo, pode-se dizer que o modelo de desenvolvimento de raciocínio crítico não teve efeito significativo para desenvolver as habilidades no grupo experimento. Verifica-se que o total da variação a ser explicada foi de 1966,00 unidades, sendo que a manipulação experimental foi responsável por apenas 0,94 unidades, enquanto 1965,06 unidades não foram explicadas.

A Tabela 15 apresenta uma saída do SPSS (Anexo C – Saída 3) que representa a ANOVA quando o pré-teste é incluído como uma covariável, ou seja, a ANCOVA.

Tabela 15 - ANCOVA para habilidades de raciocínio crítico – FEA-USP.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig
Modelo Corrigido	482,612	2	7,971	0,001
Intercepto	240,199	1	7,934	0,007
EWCTETpre	481,669	1	15,911	0,000
Turma	66,814	1	2,207	0,144
Erro	1.483,388	49		
Total	10.091,000	52		
Total Corrigido	1.966,000	51		

Analisando-se os dados é possível verificar que a covariável prevê significativamente a variável dependente, $F(1, 49) = 15,91$, $p < 0,05$. A ANCOVA é interessante porque remove o

efeito da covariável, e, comparando com os dados da ANOVA, pode-se verificar que o efeito na variável dependente devido ao método de instrução (Turma) aumentou, porém ainda não se tornou significativo, $F(1, 49) = 2,21$, $p > 0,05$ (FIELD, 2009). Nota-se que a quantidade de variação de responsabilidade do modelo aumentou para 482,61 unidades, da qual o método de instrução (Turma) é responsável por 66,81 unidades. Também é importante analisar que a grande variação, que é de responsabilidade da covariável (481,67 unidades), fez com que a variação não explicada fosse reduzida para 1.483,39 unidades.

Das saídas do SPSS, a que apresenta as médias ajustadas dos grupos, que devem ser utilizadas para interpretação, mostra-se bem interessante (Anexo C – Saída 4) (FIELD, 2009).

Nota-se que apesar de as médias das turmas não serem significativamente diferentes, após os ajustes dos pós-testes por meio dos pré-testes, a média da turma experimento (13,79) é superior ao da turma controle (11,30).

Outra importante saída do SPSS (Anexo C – Saída 5) que realiza testes t é a tabela das Estimativas dos Parâmetros.

Verifica-se que $t = 1,49$ é maior que 5% (0,14), demonstrando mais uma vez que as médias ajustadas não diferem significativamente. Embora o teste t desta pesquisa não seja significativo, pode-se calcular o tamanho do efeito. De acordo com Field (2009, p. 56): “O tamanho do efeito é simplesmente uma medida de magnitude padronizada do efeito observado”. As mais conhecidas são o d de Cohen e coeficiente de correlação r de Pearson (FIELD, 2009). Neste estudo é utilizado o r de Pearson para medir o tamanho do efeito de todos os resultados e para os casos em que as hipóteses foram confirmadas calcula-se também o d de Cohen. De acordo com Rosenthal (1991 *apud* FIELD, 2009, p. 361), para a ANCOVA pode-se utilizar a seguinte fórmula para converter um valor t em r e assim calcular o tamanho do efeito:

$$r = \sqrt{\frac{t^2}{t^2 + gl}}$$

Utilizando $t = 1,486$ e $N-2$ graus de liberdade ($52 - 2 = 50$), chega-se a $r = 0,21$ ou 21%. De acordo com Cohen (1988; 1992 *apud* FIELD, 2009, p. 57):

Cohen (1988, 1992) estipulou o que é um efeito pequeno ou grande:

- $r = 0,10$ (efeito pequeno): nesse caso, o efeito explica 1% da variância total.
- $r = 0,30$ (efeito médio): o efeito é responsável por 9% da variância total.
- $r = 0,50$ (efeito grande): o efeito é responsável por 25% da variância total.

Logo, de acordo com a classificação anterior, pode-se dizer que o efeito medido neste experimento foi de pequeno a médio.

De acordo com Field (2009, p. 57),

O tamanho do efeito em uma população é intrinsecamente ligado a três outras propriedades estatísticas: (1) o tamanho da amostra no qual o tamanho do efeito da amostra é baseado; (2) o nível da probabilidade no qual aceitaremos que um efeito é estatisticamente significativo (nível α); (3) a habilidade de um teste detectar um efeito daquele tamanho (conhecido como o poder estatístico).

O poder estatístico de um teste é a probabilidade de encontrar um efeito assumindo que um já exista na população. Para Field (2009, p. 58): “Nosso objetivo deve ser alcançar um poder de 0,8 ou 80% de probabilidade de detectar um efeito se ele genuinamente existe”.

Utilizando-se o *software* G*Power¹⁷ é possível calcular o tamanho da amostra que seria necessária para com um tamanho de efeito de 21% e um nível de significância de 5% se obter um poder estatístico de 80%. Após o cálculo encontrou-se o resultado de 564 participantes, o que é bem superior à amostra deste estudo, que conta com 52 alunos. Após o cálculo do poder estatístico levando-se em consideração as variáveis citadas e a amostra utilizada, encontrou-se o resultado de 18,46%.

Apesar de as médias das turmas não serem significativamente diferentes, após os ajustes dos pós-testes por meio dos pré-testes a média da turma experimento foi superior ao da turma controle. Porém, a conclusão final é a de que na instituição pública não foram encontradas diferenças significativas entre as médias dos testes que avaliam habilidade de raciocínio crítico entre as turmas experimento e controle; logo, não se pode dizer que utilizando o

¹⁷ *Software* gratuito, disponível em: <www.artmed.com.br>. Acesso em: 1 ago. 2011.

modelo em uma abordagem geral se consegue desenvolver habilidades de raciocínio crítico nos alunos.

Este resultado foi contra o de Reed (1998), a qual constatou em sua tese que os alunos de graduação do curso de História que foram expostos ao modelo de raciocínio crítico de Richard Paul obtiveram melhor resultado no EWCTET do que os alunos de uma turma controle (REED, 1998). Porém, a autora utilizou uma abordagem mista, na qual as intervenções são frequentes, pois o modelo faz parte da metodologia das atividades da disciplina como um todo. No caso da abordagem geral utilizada na FEA-USP, o modelo foi apresentado como um curso separado (em quatro intervenções de 30 minutos), já que o autor desta pesquisa não era o professor. Logo, pode-se dizer que a intensidade do tratamento na abordagem geral é menor que na mista.

Comparando as médias obtidas nos pós-testes das turmas experimento (12,36) e controle (12,63) da instituição pública com dados de pesquisas anteriores do Suplemento do Manual do EWCTET (Universidade A = 13,1 e Universidade B = 14,3) (ENNIS, 1998) e com o resultado da pesquisa de Cornachione, Duncan e Johnson (2007) (Grupo USP = 15,1), pode-se dizer que o desempenho da amostra da FEA-USP ficou um pouco abaixo das outras pesquisas que utilizaram o mesmo instrumento de avaliação de raciocínio crítico.

Além do estudo de Reed, este resultado foi contra as teorias ou conclusões das seguintes pesquisas listadas na revisão de literatura deste trabalho: Abrami *et al.* (2008); Allen *et al.* (1999); Bangert-Drowns e Bankert, (1990); Quitadamo *et al.* (2008); Pithers e Soden (2000), Riesenmy *et al.* (1991); Cornachione, Duncan e Johnson (2007); Baldwin e Ingram (1991); Somunen e Chalupa (1994); Kimmel (1995); Stone e Shelley (1997); Bonk e Smith (1998); Baril *et al.* (1998); Wilson (2002); Braun (2004); Springer e Borthick (2007). No entanto, corroborou os resultados dos seguintes estudos: McMillan (1987) e Adams (1999).

4.2.2 UNIFIG

A Tabela 16 a seguir apresenta a distribuição dos escores no EWCTET dos grupos experimento e controle.

Tabela 16 - Distribuição do EWCTET por método de instrução na UNIFIG.

<u>Medidas</u>	<u>Grupo Experimento (n = 16)</u>		<u>Grupo Controle (n = 23)</u>	
	<u>Pré-teste</u>	<u>Pós-teste</u>	<u>Pré-teste</u>	<u>Pós-teste</u>
Média	4,06	5,91	2,93	3,65
DP	4,18	3,96	4,54	4,37

Analisaram-se os histogramas e os gráficos de caixa e bigode e as distribuições aparentaram ser normais (com assimetrias e curtoses próximas de zero). Nos pós-testes foi encontrado um valor atípico na turma controle. Após a exclusão desse valor, chegou-se à distribuição referente à Tabela 17.

Tabela 17- Distribuição do EWCTET por método de instrução na UNIFIG sem valores atípicos.

<u>Medidas</u>	<u>Grupo Experimento (n = 16)</u>		<u>Grupo Controle (n = 22)</u>	
	<u>Pré-teste</u>	<u>Pós-teste</u>	<u>Pré-teste</u>	<u>Pós-teste</u>
Média	4,06	5,91	3,05	3,11
DP	4,18	3,96	4,62	3,60

Os gráficos de caixa e bigode foram novamente analisados e nenhum valor atípico foi encontrado. Os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras são normais. Os testes de homogeneidade das variâncias foram realizados utilizando a estatística de Levene e apresentaram os seguintes resultados:

- $EWCTET_{pré}$: $F(1, 36) = 0,18$, $p > 0,05$.
- $EWCTET_{pós}$: $F(1, 36) = 0,15$, $p > 0,05$.

Portanto, o pressuposto de homogeneidade das variâncias não foi violado. Tanto os pré-testes quanto os pós-testes foram maiores para a turma experimento. A turma controle apresentou um rendimento no pós-teste muito semelhante ao pré-teste, enquanto a turma experimento obteve um bom aumento (1,85 ou 46%).

Os pré-requisitos para a ANCOVA foram examinados para os escores do EWCTET. O pressuposto de independência dos escores de cada indivíduo não foi violado. O segundo pré-requisito é o da normalidade e já se verificou anteriormente que as amostras são normais. O terceiro pré-requisito é o da homogeneidade da variância, e este também já foi testado e não

violado. O quarto pré-requisito é a mensuração das variáveis dependentes em uma escala de intervalo e o EWCTET é intervalar. O quinto refere-se à confiabilidade da correção dos testes pelos avaliadores, que não foi violado (discutido no item 3.3.1). O sexto pré-requisito é o da homogeneidade dos parâmetros da regressão, que não foi violado, $F(1, 34) = 0,10$, $p > 0,05$.

A Tabela 18 apresenta a ANCOVA (Anexo C – Saída 6) com os pré-testes incluídos como covariáveis.

Tabela 18 - ANCOVA para habilidades de raciocínio crítico – UNIFIG.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig
Modelo Corrigido	174,465	2	7,532	0,002
Intercepto	223,300	1	19,281	0,000
EWCTETpre	102,225	1	8,827	0,005
Turma	52,945	1	4,572	0,040
Erro	405,351	35		
Total	1.279,000	38		
Total Corrigido	579,816	37		

Analisando-se a linha “Turma” é possível verificar que o efeito na variável dependente causado pelo método de instrução é significativo, $F(1, 35) = 4,57$, $p < 0,05$. Portanto, existe diferença significativa entre as médias ajustadas do grupo experimento e controle. Pode-se dizer também que o modelo de desenvolvimento de raciocínio crítico teve efeito significativo para desenvolver as habilidades no grupo experimento. Nota-se que a quantidade total de variação a ser explicada foi de 579,82 unidades, das quais o modelo consegue explicar 174,47 unidades e o método de instrução (Turma), 52,95 unidades.

Analisadas as médias ajustadas dos grupos (Anexo C – Saída 7), nota-se que a média ajustada dos pós-testes da turma experimento (5,68) é superior ao da turma controle (3,28).

Verifica-se que $t = 2,14$ é menor que 5% (0,04) e demonstra novamente que as médias ajustadas diferem de maneira significativa (Anexo C – Saída 8). Utilizando-se $t = 2,14$ e N-2 graus de liberdade ($38 - 2 = 36$), chega-se a $r = 0,34$ ou 34%. Portanto, pode-se dizer que o tamanho do efeito medido neste experimento foi médio. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 34% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 216 participantes. Ao

calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 38$) utilizada, encontrou-se 26,45 como resultado.

De acordo com Bickman e Rog (1998), existe outra medida de tamanho de efeito, o d de Cohen, o qual pode ser calculado pela divisão da diferença entre as médias da turma controle e experimento pelo desvio padrão comum. Utilizando as médias ajustadas pelas covariáveis e o desvio comum a todos os dados (pré-testes e pós-testes), chega-se a $d = 0,56$ ou 56%, o que é considerado um tamanho de efeito grande. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 56% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 82 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 38$) utilizada, encontrou-se 51,09% como resultado.

Portanto, na instituição particular encontrou-se diferença significativa entre as médias dos testes que avaliam habilidade de raciocínio crítico entre as turmas experimento e controle, e as médias ajustadas da turma experimento foram maiores. Logo, pode-se dizer que utilizando o modelo em uma abordagem mista foi possível desenvolver habilidades de raciocínio crítico nos alunos.

Este resultado corrobora o de Reed (1998), a qual em seu quase-experimento aplicou o modelo de desenvolvimento de raciocínio crítico de Richard Paul em uma abordagem mista (ressalta-se que o quase-experimento desta pesquisa seguiu o mesmo protocolo da pesquisa de Reed (1998), com a mesma quantidade e tempo de intervenções explicando o modelo e posteriormente aplicando-o em atividades em sala de aula). Este resultado demonstra a importância da utilização desta abordagem, já que utilizando a abordagem geral na instituição pública não se constatou maior desempenho dos alunos da turma experimento. Analisados os gráficos demográficos, notou-se que a turma experimento possui duas grandes diferenças em relação à controle: a turma experimento possui maior hábito de leitura e escrita. No entanto, sabendo que a pesquisa demográfica foi aplicada ao final do semestre, o modelo pode ter influenciado o desenvolvimento destas características na turma experimento.

Analisou-se no capítulo anterior que o erro padrão da amostra é pequeno, e isto é positivo, já que indica que provavelmente a amostra representa bem sua população. O tamanho do efeito foi alto (56%), e como a amostra não foi muito grande ($n = 38$), o poder estatístico não foi o

esperado (51%). Sendo assim, a generalização dos resultados deve ser realizada com cautela, e esta limitação deve ser levada em consideração.

Comparando as médias obtidas nos pós-testes das turmas experimento (5,91) e controle (3,11) da instituição particular com dados de pesquisas anteriores do Suplemento do Manual do EWCTET (Universidade A = 13,1 e Universidade B = 14,3) (ENNIS, 1998) e com o resultado da pesquisa de Cornachione, Duncan e Johnson (2007) (Grupo USP = 15,1), pode-se dizer que o desempenho da amostra da UNIFIG ficou bem abaixo das outras pesquisas que utilizaram o mesmo instrumento de avaliação de raciocínio crítico. As análises dos gráficos da pesquisa demográfica podem nos ajudar a procurar possíveis explicações para um desempenho tão inferior, já que, comparando as duas instituições analisadas, nota-se que a particular possui menor renda familiar, menor fluência em língua estrangeira, trabalha mais em período integral e estudou maior parte do ensino fundamental e médio em escola pública.

Além do estudo de Reed (1998), este resultado corroborou as teorias e conclusões das seguintes pesquisas listadas na revisão de literatura deste trabalho: Abrami *et al.* (2008); Allen *et al.* (1999); Bangert-Drowns e Bankert (1990); Quitadamo *et al.* (2008); Pithers e Soden (2000); Riesenmy *et al.* (1991); Cornachione, Duncan e Johnson (2007); Baldwin e Ingram (1991); Somunen e Chalupa (1994); Kimmel (1995); Stone e Shelley (1997); Bonk e Smith (1998); Baril *et al.* (1998); Wilson (2002); Braun (2004); Springer e Borthick (2007). No entanto, esta pesquisa foi contra os resultados dos seguintes estudos: McMillan (1987) e Adams (1999).

Portanto, a Hipótese 1 foi confirmada para a instituição particular, na qual a abordagem de aplicação do modelo instrucional foi a mista. Na instituição pública, na qual a abordagem utilizada foi a geral, apesar de a turma experimento ter melhorado seu desempenho nos pós-testes em relação aos pré-testes, enquanto a turma controle obteve um pior desempenho nos pós-teses, não houve diferença significativa entre as médias das turmas, ou seja, não houve efeito significativo do modelo instrucional no desenvolvimento de habilidades de raciocínio crítico.

Para o principal resultado deste item, apresenta-se a seguir o Gráfico 33.

Gráfico 33 - Pré-testes e pós-testes do EWCTET da UNIFIG.

Nota-se que a turma experimento obteve um desempenho levemente melhor nos pré-testes e altamente superior nos pós-testes em relação à turma controle, indicando o efeito significativo do modelo instrucional no desenvolvimento das habilidades de raciocínio crítico.

4.3 Hipótese 2 – Atitudes de Raciocínio Crítico

Relembrando, a Hipótese 2 definida para esta pesquisa foi:

Uma classe que receber explicitamente treinamento sobre o modelo de desenvolvimento de raciocínio crítico de Richard Paul obterá melhor resultado em um teste (CCTDI) que avalia atitudes de raciocínio crítico do que uma classe que não receber instruções explícitas sobre o modelo.

Realizaram-se análises descritivas das amostras em relação ao desempenho no CCTDI e método de instrução, checkou-se a possível existência de diferenças significativas entre as médias das turmas em cada quase-experimento e se as turmas experimento apresentaram melhor desempenho.

4.3.1 FEA - USP

A Tabela 19 apresenta a distribuição dos escores no CCTDI dos grupos experimento e controle.

Tabela 19 - Distribuição do CCTDI por método de instrução na FEA-USP.

<u>Medidas</u>	<u>Grupo Experimento (n = 25)</u>		<u>Grupo Controle (n = 27)</u>	
	<u>Pré-teste</u>	<u>Pós-teste</u>	<u>Pré-teste</u>	<u>Pós-teste</u>
Média	297,92	292,08	297,93	296,07
DP	28,42	31,31	25,32	30,98

Os histogramas e as distribuições aparentaram ser normais (com assimetrias e curtoses próximas de zero) e os gráficos de caixa e bigode não apresentaram valores atípicos. Os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras assemelham-se à uma distribuição normal. Ambos os testes de Levene não foram significativos a um nível de 5% (CCTDI_{pré}: $F(1, 50) = 0,20$, $p > 0,05$ e CCTDI_{pós}: $F(1, 50) = 0,01$, $p > 0,05$), indicando que existe homogeneidade das variâncias e este requisito para as análises estatísticas não foi violado. Os pré-testes foram similares entre os grupos controle e experimento. O grupo experimento obteve uma leve queda no pós-teste (5,84 ou 2%), enquanto o grupo controle se manteve praticamente estável. No final, as médias obtidas nos pós-testes pelo grupo controle foram um pouco maiores do que as do grupo experimento.

Os pré-requisitos para a ANCOVA foram examinados para os escores do CCTDI. O pressuposto de independência dos escores de cada indivíduo não foi violado. O segundo pré-requisito é o da normalidade, e já se verificou que as amostras são normais. O terceiro pré-requisito é o da homogeneidade da variância, o qual também já foi testado e não violado. O quarto pré-requisito refere-se à mensuração das variáveis dependentes em uma escala de intervalo, e o CCTDI é intervalar. O quinto pré-requisito refere-se à confiabilidade da correção dos testes pelos avaliadores, e este não foi violado (discutido no item 3.3.2). O sexto, da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(1, 48) = 0,76$, $p > 0,05$.

A Tabela 20 apresenta os dados da ANCOVA (Anexo C – Saída 9) com os pré-testes como covariáveis.

Tabela 20 - ANCOVA para atitudes de raciocínio crítico – FEA-USP.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig
Modelo Corrigido	18.499,046	2	15,010	0,000
Intercepto	2.702,064	1	4,385	0,041
CCTDIpre	18.291,969	1	29,683	0,000
Turma	206,640	1	0,335	0,565
Erro	30.195,723	49		
Total	4.548.072,000	52		
Total Corrigido	48.694,769	51		

Analísada a linha “Turma”, verifica-se que não há diferenças significativas entre as médias ajustadas dos grupos experimento e controle, $F(1, 49) = 0,34$, $p > 0,05$. Portanto, pode-se dizer que o modelo de desenvolvimento de raciocínio crítico não teve efeito significativo para desenvolver as atitudes no grupo experimento. Nota-se que a quantidade total de variação a ser explicada foi de 48.694,77, das quais o modelo consegue explicar 18.499,05 e o método de instrução (Turma), apenas 206,64 unidades.

Analísadas as médias ajustadas dos grupos (Anexo C – Saída 10), nota-se que, mesmo após os ajustes, a média dos pós-testes da turma controle (296,07) é superior ao da turma experimento (292,08).

Verifica-se que $t = -0,58$ é maior que 5% (0,57), e demonstra novamente que as médias ajustadas não diferem significativamente (Anexo C – Saída 11). Utilizando $t = -0,579$ e N-2 graus de liberdade ($52 - 2 = 50$), chega-se a $r = 0,08$ ou 8%. Portanto, pode-se dizer que o tamanho do efeito medido neste experimento foi pequeno. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 8% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 3.866 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra utilizada, encontrou-se 8,68% como resultado.

As análises acima foram das médias das atitudes. Tem-se na Tabela 21 a seguir a análise descritiva das sete atitudes de acordo com o método de instrução (Turma).

Tabela 21 - Distribuição de cada atitude do CCTDI por método de instrução na FEA-USP.

Atitudes	Turma Experimento (n = 25)				Turma Controle (n = 27)			
	Pré-teste		Pós-teste		Pré-teste		Pós-teste	
	Média	DP	Média	DP	Média	DP	Média	DP
Buscando a Verdade	35,20	4,85	33,96	7,08	38,70	6,80	36,56	7,42
Mente Aberta	44,96	5,35	41,92	5,37	43,81	5,07	41,96	6,80
Analiticidade	43,76	6,48	42,32	5,95	42,30	6,95	43,96	4,81
Sistematicidade	41,00	8,40	39,76	7,27	38,85	7,82	39,11	7,26
Autoconfiança	43,64	5,62	43,08	7,35	41,70	7,65	45,22	6,14
Curiosidade	48,44	6,12	48,24	5,95	49,30	5,75	48,30	6,34
Maturidade	40,92	7,10	42,80	6,63	43,26	5,84	40,96	7,18

Analisando-se a Tabela 21 é possível notar que a única atitude da turma experimento que apresentou uma melhora de desempenho médio no pós-teste em relação ao pré-teste foi a da maturidade. Já na turma controle foram três as atitudes que apresentaram melhora nos pós-testes: analiticidade, sistematicidade e autoconfiança.

Para analisar uma possível diferença relativa ao método de instrução e à atitude maturidade, realizou-se uma ANCOVA. Os pré-requisitos para a ANCOVA foram examinados para os escores relativos à atitude maturidade. O pressuposto de independência dos escores de cada indivíduo não foi violado. O segundo pré-requisito é o da normalidade, e o mesmo não foi violado. O terceiro pré-requisito é o da homogeneidade da variância, e este também foi testado e não violado, $F(1, 50) = 0,30, p > 0,05$. O quarto pré-requisito é o da mensuração das variáveis dependentes em uma escala de intervalo e a atitude maturidade é intervalar. O quinto pré-requisito refere-se à confiabilidade da correção dos testes pelos avaliadores, e este não foi violado (discutido no item 3.3.2). O sexto, da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(1, 48) = 0,01, p > 0,05$.

A Tabela 22 apresenta a ANCOVA com os pré-testes como covariáveis (Anexo C – Saída 12).

Tabela 22 - ANCOVA para a atitude maturidade – FEA-USP.

Fonte	Soma dos Quadrados		gl	F	Sig
	Tipo III				
Modelo Corrigido	677,601		2	9,416	0,000
Intercepto	406,546		1	11,298	0,002
Maturidade.pre	633,795		1	17,614	0,000
Turma	122,460		1	3,403	0,071
Erro	1.763,168		49		
Total	93.498,000		52		
Total Corrigido	2.440,769		51		

Analisada a linha “Turma” verifica-se que não há diferenças significativas entre as médias ajustadas referente à atitude maturidade dos grupos experimento e controle, $F(1, 49) = 3,40$, $p > 0,05$. Nota-se que a quantidade total de variação a ser explicada foi de 2.440,77 unidades, das quais o modelo consegue explicar 677,60, e o método de instrução, 122,46 unidades.

Analisando-se as médias ajustadas dos grupos (Anexo C – Saída 13) é possível notar que apesar da diferença não ter sido significativa, a média ajustada da turma experimento (43,47) foi maior do que a da turma controle (40,35).

Verificando as saídas do SPSS, nota-se que $t = 1,85$ é maior que 5%, e demonstra novamente que as médias ajustadas não diferem significativamente (Anexo C – Saída 14). Utilizando $t = 1,85$ e N-2 graus de liberdade ($52 - 2 = 50$), chega-se a $r = 0,25$ ou 25,31%. Portanto, pode-se dizer que o tamanho do efeito medido neste experimento foi de pequeno a médio. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 25,31% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 388 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 52$) utilizada, encontrou-se 22,81% como resultado.

Além de as médias das turmas não serem significativamente diferentes, após os ajustes dos pós-testes por meio dos pré-testes a média da turma controle foi superior ao da turma experimento. Portanto, a conclusão final é a de que na instituição pública não se encontrou diferenças significativas entre as médias dos testes que avaliam atitudes de raciocínio crítico entre as turmas experimento e controle. Logo, não se pode dizer que utilizando o modelo em uma abordagem geral se consegue desenvolver atitudes de raciocínio crítico nos alunos.

No geral, este resultado corroborou os de Reed (1998); porém, analisando-se a evolução dos desempenhos dos pós-testes em relação aos pré-testes da turma experimento para cada uma das sete atitudes, constatou-se aumento na atitude maturidade, porém também não significativa. O tamanho do efeito para esta atitude foi de pequeno a médio (25%) e o poder estatístico, de 23%, o que nos levaria novamente a tratar as generalizações do resultado com cautela.

Comparando as médias obtidas nos pós-testes das turmas experimento (292,08) e controle (296,07) da instituição pública com os critérios de interpretação dos desempenhos no CCTDI descritos no item 3.3.2, pode-se constatar que existe fraca presença de atitudes de raciocínio crítico na amostra da FEA-USP, já que para ser considerado forte o desempenho precisa ser superior a 350 pontos.

Comparando a média ajustada (43,47) da atitude maturidade da turma experimento da instituição pública com os critérios de interpretação dos desempenhos no CCTDI descritos no item 3.3.2, pode-se constatar que existe uma média presença em relação à atitude maturidade, já que para ser considerada forte a pontuação precisa ser superior a 50 pontos na atitude (abaixo de 40 é considerada fraca).

4.3.2 UNIFIG

A Tabela 23 apresenta a distribuição dos escores no CCTDI dos grupos experimento e controle.

Tabela 23 - Distribuição do CCTDI por método de instrução na UNIFIG.

Medidas	Grupo Experimento (n = 16)		Grupo Controle (n = 23)	
	Pré-teste	Pós-teste	Pré-teste	Pós-teste
Média	311,81	312,81	286,83	281,57
DP	23,00	26,05	31,45	32,27

Os gráficos de caixa e bigode apresentaram um valor atípico para os pré-testes da turma experimento. Portanto, o valor foi excluído e a distribuição sem valores atípicos pode ser analisada na Tabela 24 a seguir.

Tabela 24 - Distribuição do CCTDI por método de instrução na UNIFIG sem valores atípicos.

<u>Medidas</u>	<u>Grupo Experimento (n = 15)</u>		<u>Grupo Controle (n = 23)</u>	
	<u>Pré-teste</u>	<u>Pós-teste</u>	<u>Pré-teste</u>	<u>Pós-teste</u>
Média	316,20	316,27	286,83	281,57
DP	15,39	22,86	31,45	32,27

Os histogramas e as distribuições aparentaram ser normais e os gráficos de caixa e bigode não apresentaram valores atípicos. Os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras assemelham-se à uma distribuição normal. Os testes de Levene para os pré-testes foram significativos e os dados foram transformados (quando os dados foram elevados ao quadrado, passaram a atender aos requisitos) para a aplicação da ANCOVA ($\text{CCTDI}_{\text{pré}}$: $F(1, 36) = 4,81$, $p < 0,05$ e $\text{CCTDI}_{\text{pós}}$: $F(1, 36) = 3,15$, $p > 0,05$).

Tanto os pré-testes quanto os pós-testes foram maiores para a turma experimento. O grupo experimento obteve um rendimento muito similar nos pós-testes em relação aos pré-testes, enquanto a turma controle apresentou uma leve queda nos pós-testes (5,26 ou 2%).

Os pré-requisitos para a ANCOVA foram examinados para os escores transformados (elevados ao quadrado) do CCTDI. O pressuposto de independência dos escores de cada indivíduo não foi violado, já que cada aluno realizou os testes individualmente. O segundo pré-requisito é o da normalidade, e os dados transformados não violaram o pressuposto da normalidade:

- Turma Experimento:
 - $\text{CCTDI}_{\text{présquare}}$: $D(15) = 0,13$, $p > 0,05$
 - $\text{CCTDI}_{\text{póssquare}}$: $D(15) = 0,12$, $p > 0,05$
- Turma Controle:
 - $\text{CCTDI}_{\text{présquare}}$: $D(23) = 0,10$, $p > 0,05$
 - $\text{CCTDI}_{\text{póssquare}}$: $D(23) = 0,15$, $p > 0,05$

O terceiro pré-requisito é o da homogeneidade da variância, que não foi violado ($\text{CCTDI}_{\text{présquare}}$: $F(1, 36) = 1,72$, $p > 0,05$ e $\text{CCTDI}_{\text{póssquare}}$: $F(1, 36) = 0,09$, $p > 0,05$). O quarto pré-requisito é o da mensuração das variáveis dependentes em uma escala de intervalo,

e o CCTDI é intervalar. O quinto pré-requisito refere-se à confiabilidade da correção dos testes pelos avaliadores, que não foi violado (discutido no item 3.3.2). O sexto pré-requisito, da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(1, 34) = 0,12$, $p > 0,05$.

A Tabela 25 apresenta a ANCOVA com os pré-testes como covariáveis (Anexo C – Saída 15).

Tabela 25 - ANCOVA para atitudes de raciocínio crítico – UNIFIG.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig
Modelo Corrigido	9.647.959.410,632	2	36,565	0,000
Intercepto	256.871.749,344	1	1,947	0,172
CCTDIprequadrado	5.929.654.035,711	1	44,946	0,000
Turma	288.803.487,306	1	2,189	0,148
Erro	4.617.460.826,632	35		
Total	310.302.876.244,000	38		
Total Corrigido	14.265.420.237,264	37		

Analizada a linha “Turma”, verifica-se que não há diferenças significativas entre as médias ajustadas dos grupos experimento e controle, $F(1, 35) = 2,19$, $p > 0,05$. Portanto, pode-se dizer que o modelo de desenvolvimento de raciocínio crítico não teve efeito significativo para desenvolver as atitudes no grupo experimento. Nota-se que a quantidade total de variação a ser explicada foi de 14.265.420.237,26 unidades, das quais o modelo consegue explicar 9.647.959.410,63 unidades, e o método de instrução (Turma), 288.803.487,31 unidades.

Considerando-se as médias ajustadas (dados elevados ao quadrado) dos grupos é possível notar que, apesar de não haver diferenças significativas entre os grupos, a média ajustada dos pós-testes da turma experimento (92.144,78) é superior ao da turma controle (85.732,36) (Anexo C – Saída 16).

De acordo com as saídas do SPSS, nota-se que $t = 1,48$ é maior que 5% (0,15) e demonstra novamente que as médias ajustadas não diferem significativamente (Anexo C – Saída 17). Utilizando $t = 1,48$ e N-2 graus de liberdade ($38 - 2 = 36$), chega-se a $r = 0,24$ ou 23,95%. Portanto, pode-se dizer que o tamanho do efeito medido neste experimento é de pequeno a médio. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 23,95% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se

ao número de 434 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 38$) utilizada, encontrou-se 17,63% como resultado.

As análises acima foram das médias das atitudes. Tem-se a seguir a Tabela 26 com a análise descritiva das sete atitudes de acordo com o método de instrução (Turma).

Tabela 26 - Distribuição de cada atitude do CCTDI por método de instrução na UNIFIG.

Atitudes	Turma Experimento (n = 16)				Turma Controle (n = 23)			
	Pré-teste		Pós-teste		Pré-teste		Pós-teste	
	Média	DP	Média	DP	Média	DP	Média	DP
Buscando a Verdade	37,50	5,92	36,69	5,90	32,78	7,22	32,30	7,71
Mente Aberta	39,94	5,20	41,94	5,90	38,26	7,26	38,48	6,48
Analiticidade	45,06	5,82	44,81	6,30	42,74	7,14	41,26	6,96
Sistematicidade	44,94	6,37	44,69	5,44	40,87	6,28	39,13	6,36
Autoconfiança	46,75	6,08	44,75	8,35	45,30	6,25	42,52	7,77
Curiosidade	53,31	3,30	54,50	4,43	50,22	5,17	49,78	5,36
Maturidade	44,31	4,81	45,44	5,40	36,65	9,34	38,09	10,25

Analisada a Tabela 26, nota-se que a turma experimento apresentou melhora de desempenho médio nos pós-testes em relação aos pré-testes de três atitudes: mente aberta, curiosidade e maturidade. Já na turma controle foram duas as atitudes que apresentaram melhora nos pós-testes: mente aberta e maturidade.

Para analisar uma possível diferença relativa ao método de instrução e à atitude mente aberta, realizou-se uma ANCOVA. Os pré-requisitos para a ANCOVA foram examinados para os escores relativos à atitude mente aberta. O pressuposto de independência dos escores de cada indivíduo não foi violado. O segundo pré-requisito é o da normalidade, o qual não foi violado. O terceiro pré-requisito é o da homogeneidade da variância, que também foi testado e não violado, $F(1, 37) = 0,60$, $p > 0,05$. O quarto pré-requisito é a mensuração das variáveis dependentes em uma escala de intervalo, e a atitude mente aberta é intervalar. O quinto pré-requisito refere-se à confiabilidade da correção dos testes pelos avaliadores, e este não foi violado (discutido no item 3.3.2). O sexto pré-requisito, da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(1, 35) = 2,27$, $p > 0,05$.

A Tabela 27 a seguir apresenta a ANCOVA com os pré-testes como covariáveis (Anexo C – Saída 18).

Tabela 27 - ANCOVA para a atitude mente aberta – UNIFIG.

Fonte	Soma dos Quadrados Tipo III	Gl	F	Sig
Modelo Corrigido	579,985	2	10,657	0,000
Intercepto	354,239	1	13,018	0,001
Mente.aberta.pre	467,072	1	17,165	0,000
Turma	60,023	1	2,206	0,146
Erro	979,605	36		
Total	63.640,000	39		
Total Corrigido	1.559,590	38		

Analizada a linha “Turma”, verifica-se que não há diferenças significativas entre as médias ajustadas referentes à atitude mente aberta dos grupos experimento e controle, $F(1, 36) = 2,21$, $p > 0,05$. Nota-se que a quantidade total de variação a ser explicada foi de 1.559,59 unidades, das quais o modelo consegue explicar 579,99 unidades, e o método de instrução, apenas 60,02 unidades.

De acordo com as médias ajustadas dos grupos, nota-se que, apesar de a diferença não ter sido significativa, a média ajustada da turma experimento (41,10) foi maior do que a da turma controle (38,85) (Anexo C – Saída 19).

Analizadas as saídas do SPSS, verifica-se que $t = 1,49$ é maior que 5%, e demonstra novamente que as médias ajustadas não diferem significativamente (Anexo C – Saída 20). Utilizando $t = 1,49$ e $N-2$ graus de liberdade ($39 - 2 = 37$), chega-se a $r = 0,24$ ou 23,79%. Portanto, pode-se dizer que o tamanho do efeito medido neste experimento foi de pequeno a médio. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 23,79% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 440 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 39$) utilizada, encontrou-se 17,69% como resultado.

Para analisar uma possível diferença relativa ao método de instrução e a atitude curiosidade, realizou-se uma ANCOVA. Os pré-requisitos para a ANCOVA foram examinados para os escores relativos à atitude curiosidade. O pressuposto de independência dos escores de cada indivíduo não foi violado. O segundo pré-requisito, que se refere à normalidade, não foi violado. O terceiro pré-requisito é o da homogeneidade da variância, o qual também foi testado e não violado, $F(1, 37) = 0,27$, $p > 0,05$. O quarto pré-requisito é a mensuração das

variáveis dependentes em uma escala de intervalo, e a atitude mente aberta é intervalar. O quinto pré-requisito refere-se à confiabilidade da correção dos testes pelos avaliadores, que não foi violado (discutido no item 3.3.2). O sexto pré-requisito, da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(1, 35) = 0,30$, $p > 0,05$.

A Tabela 28 apresenta a tabela da ANCOVA com os pré-testes como covariáveis (Anexo C – Saída 21).

Tabela 28 - ANCOVA para a atitude curiosidade – UNIFIG.

Fonte	Soma dos Quadrados Tipo III	Gl	F	Sig
Modelo Corrigido	377,326	2	8,953	0,001
Intercepto	213,739	1	10,144	0,003
Curiosidade.pre	167,341	1	7,942	0,008
Turma	89,331	1	4,239	0,047
Erro	758,572	36		
Total	105.451,000	39		
Total Corrigido	1.135,897	38		

Considerando-se a linha “Turma”, é possível notar que há diferenças significativas entre as médias ajustadas referentes à atitude curiosidade dos grupos experimento e controle, $F(1, 36) = 4,24$, $p < 0,05$. Nota-se que a quantidade total de variação a ser explicada foi de 1.135,90 unidades, das quais o modelo consegue explicar 377,33 unidades, e o método de instrução, 89,33 unidades.

Analizadas as médias ajustadas dos grupos, nota-se que, além da diferença ter sido significativa, a média ajustada da turma experimento (53,64) foi maior do que a da turma controle (50,38) (Anexo C – Saída 22). Portanto, pode-se dizer que o método de instrução teve influência significativa no desenvolvimento da atitude curiosidade.

De acordo com as saídas do SPSS, verifica-se que $t = 2,06$ é menor que 5%, e demonstra novamente que as médias ajustadas diferem significativamente (Anexo C – Saída 23). Utilizando $t = 2,06$ e $N-2$ graus de liberdade ($39 - 2 = 37$), chega-se a $r = 0,32$ ou 32,08%. Portanto, pode-se dizer que o tamanho do efeito medido neste experimento foi de médio a forte. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 32,08% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se

ao número de 242 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 39$) utilizada, encontrou-se 24,91% como resultado.

Calculando o d de Cohen, chega-se a $d = 0,65$ ou 65%, o que é considerado um tamanho de efeito grande. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 65% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 60 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 39$) utilizada, encontrou-se 62,37% como resultado.

Para analisar uma possível diferença relativa ao método de instrução e a atitude maturidade, realizou-se uma ANCOVA. Os pré-requisitos para a ANCOVA foram examinados para os escores relativos à atitude maturidade. O pressuposto de independência dos escores de cada indivíduo não foi violado. O segundo pré-requisito, referente à normalidade, não foi violado. O terceiro pré-requisito é o da homogeneidade da variância, e este também foi testado e não violado, $F(1, 37) = 0,04$, $p > 0,05$. O quarto pré-requisito é a mensuração das variáveis dependentes em uma escala de intervalo, e a atitude maturidade é intervalar. O quinto pré-requisito refere-se à confiabilidade da correção dos testes pelos avaliadores, o qual não foi violado (discutido no item 3.3.2). O sexto, que se refere à homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(1, 35) = 3,72$, $p > 0,05$.

A Tabela 29 apresenta a ANCOVA com os pré-testes como covariáveis (Anexo C – Saída 24).

Tabela 29 - ANCOVA para a atitude maturidade – UNIFIG.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig
Modelo Corrigido	1.922,630	2	25,885	0,000
Intercepto	128,182	1	3,452	0,071
Maturidade.pre	1.412,804	1	38,042	0,000
Turma	12,871	1	0,347	0,560
Erro	1.336,959	36		
Total	69.147,000	39		
Total Corrigido	3.259,590	38		

Analísada a linha “Turma”, verifica-se que não há diferenças significativas entre as médias ajustadas referentes à atitude maturidade dos grupos experimento e controle, $F(1, 36) = 0,35$,

$p > 0,05$. Nota-se que a quantidade total de variação a ser explicada foi de 3.259,60 unidades, das quais o modelo consegue explicar 1.922,63 unidades, e o método de instrução apenas 12,87 unidades.

Com base nas médias ajustadas dos grupos, nota-se que, apesar de a diferença não ter sido significativa, a média ajustada da turma experimento (41,87) foi maior do que a da turma controle (40,57) (Anexo C – Saída 25).

Considerando-se a saída do SPSS, verifica-se que $t = 0,59$ é maior que 5%, e demonstra novamente que as médias ajustadas não diferem significativamente (Anexo C – Saída 26). Utilizando $t = 0,59$ e N-2 graus de liberdade ($39 - 2 = 37$), chega-se a $r = 0,10$ ou 9,65%. Portanto, pode-se dizer que o tamanho do efeito medido neste experimento foi pequeno. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 9,65% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 2.658 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 39$) utilizada, encontrou-se 8,60% como resultado.

Portanto, apesar de as médias das turmas não serem significativamente diferentes, após os ajustes dos pós-testes por meio dos pré-testes a média da turma experimento foi superior ao da turma controle. Porém, a conclusão final é a de que na instituição particular não se encontrou diferenças significativas entre as médias dos testes que avaliam atitudes de raciocínio crítico entre as turmas experimento e controle; logo, não se pode dizer que utilizando o modelo em uma abordagem mista se consegue desenvolver atitudes de raciocínio crítico nos alunos.

No geral, este resultado corroborou os de Reed (1998); porém, analisando-se a evolução dos desempenhos dos pós-testes em relação aos pré-testes da turma experimento para cada uma das sete atitudes, constatou-se aumento nas atitudes mente aberta, curiosidade e maturidade. Analisadas as diferenças entre as médias das turmas experimento e controle destas três atitudes, identificou-se diferença significativa e maior média ajustada para a turma experimento para a atitude curiosidade; logo, pode-se dizer que o modelo instrucional teve efeito no desenvolvimento desta atitude de raciocínio crítico. O tamanho do efeito para esta atitude foi alto (65%), porém, o poder estatístico foi de 62%, o que nos leva novamente a tratar as generalizações do resultado com cautela.

Esta pesquisa também corrobora a de Abrami *et al.* (2008), a qual apresentou evidências de desenvolvimento de habilidades e atitudes de raciocínio crítico de acordo com o modelo instrucional.

A teoria de Richard Paul (1995) defende que as atitudes de raciocínio crítico devem ser explicitadas para conscientizar os alunos de que não basta possuir as habilidades de raciocínio crítico, é preciso ter intenção ou atitude para utilizá-las. Na abordagem mista, o professor procurou sempre conscientizar a turma experimento sobre a importância de possuir as atitudes de raciocínio crítico, e isto pode ter sido importante no desenvolvimento significativo da atitude curiosidade e na maior média nos pós-testes da turma experimento em relação à controle.

Comparadas as médias das atitudes obtidas nos pós-testes das turmas experimento (312,81) e controle (281,57) da instituição particular com os critérios de interpretação dos desempenhos no CCTDI descritos no item 3.3.2, pode-se constatar que existe fraca presença de atitudes de raciocínio crítico na amostra da UNIFIG, já que para ser considerado forte o desempenho precisa ser superior a 350 pontos.

Comparando a média ajustada (53,64) da atitude curiosidade da turma experimento da instituição particular com os critérios de interpretação dos desempenhos no CCTDI descritos no item 3.3.2, pode-se constatar que existe forte presença em relação a esta atitude, a qual sofreu efeito significativo do modelo instrucional para seu desenvolvimento.

Para o principal resultado deste item, tem-se a seguir o Gráfico 34.

Gráfico 34 - Pré-testes e pós-testes da atitude curiosidade da UNIFIG.

Nota-se que a turma experimento obteve um melhor desempenho tanto nos pré-testes quanto nos pós-testes em relação à turma controle. Como houve efeito significativo do modelo instrucional no desenvolvimento da atitude curiosidade, a distância entre as linhas aumentou nos pós-testes, principalmente porque a turma experimento melhorou seu desempenho.

4.4 Hipótese 3 - Idade

A Hipótese 3 definida nesta pesquisa foi:

Alunos mais velhos (22 anos e acima) obterão melhor desempenho nos testes (EWCTET e CCTDI) de habilidades e atitudes de raciocínio crítico do que alunos mais novos (abaixo de 22 anos) de acordo com o método instrucional.

Realizaram-se análises descritivas de acordo com a idade e método de instrução, checaram-se possíveis diferenças significativas entre as médias dos desempenhos entre os grupos experimento e controle considerando a idade e se os participantes com maior faixa etária obtiveram melhor desempenho.

4.4.1 FEA - USP

Neste tópico é analisada a Hipótese 3 para as duas variáveis dependentes: EWCTET e CCTDI.

4.4.1.1 EWCTET

A distribuição dos escores no EWCTET dos grupos experimento e controle considerando a idade são apresentados na Tabela 30.

Tabela 30 - Distribuição do EWCTET por método de instrução e idade na FEA-USP.

<u>Grupo Idade</u>	<u>Turma Experimento (n = 25)</u>						<u>Turma Controle (n = 27)</u>				
	<u>Pré-teste</u>			<u>Pós-teste</u>			<u>Pré-teste</u>		<u>Pós-teste</u>		
	<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>
<u>Abaixo de 22 anos</u>	11,41	7,06	23	12,39	7,54		17,08	5,38	19	12,08	5,65
<u>22 anos ou acima</u>	9,75	4,60	2	12,00	6,36		16,00	5,94	8	13,94	3,41

Analisado o gráfico de caixa e bigode, encontrou-se um valor atípico para os pré-testes da turma controle do grupo 22 anos ou acima. Os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras assemelham-se à uma distribuição normal, com exceção do grupo 22 anos ou acima da turma experimento, que, devido ao pequeno número de participantes (2), não foi possível realizar o teste.

Retirando o valor atípico do pré-teste do EWCTET da turma controle, rodou-se novamente a análise e outro valor atípico foi encontrado e excluído. Após estas duas exclusões, não foram encontrados mais valores atípicos e o requisito da normalidade não foi violado.

A Tabela 31 apresenta a distribuição dos escores no EWCTET dos grupos experimento e controle considerando a idade e sem os valores atípicos.

Tabela 31 - Distribuição do EWCTET por método de instrução e idade na FEA-USP sem valores atípicos.

<u>Grupo Idade</u>	<u>Turma Experimento (n = 25)</u>						<u>Turma Controle (n = 25)</u>				
	<u>Pré-teste</u>			<u>Pós-teste</u>			<u>Pré-teste</u>		<u>Pós-teste</u>		
	<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>
<u>Abaixo de 22 anos</u>	11,41	7,06	23	12,39	7,54		17,08	5,38	19	12,08	5,65
<u>22 anos ou acima</u>	9,75	4,60	2	12,00	6,36		18,25	4,23	6	15,33	2,50

Comparados os dois grupos (abaixo de 22 anos e 22 anos ou acima) da turma experimento, nota-se que seus pós-testes foram superiores aos pré-testes, sendo que os mais velhos tiveram aumento maior (2,25 ou 23%). Ainda sobre a turma experimento, verifica-se que os mais novos obtiveram maiores médias tanto nos pré-testes quanto nos pós-testes.

Analisando-se os dados da turma controle, é possível constatar que tanto os alunos mais novos quanto os mais velhos tiveram menor rendimento nos pós-testes em relação aos pré-testes, sendo que a maior queda foi dos alunos mais novos (5,00 ou 29%).

Comparadas as turmas experimento e controle, nota-se que nos pré-testes a controle obteve rendimento médio superior em ambos os grupos. Nos pós-testes os alunos mais novos apresentaram valores médios bem semelhantes, enquanto os alunos mais velhos da turma controle se saíram em média melhor no pós-teste.

Os pré-requisitos para a ANCOVA foram examinados para os escores do EWCTET considerando-se turma e idade. Como já visto em tópicos anteriores, o pressuposto de independência não foi violado. Verificou-se anteriormente que as amostras assemelham-se à uma distribuição normal, porém existe uma limitação em relação aos alunos mais velhos da turma experimento. A homogeneidade das variâncias não foi violada, $F(3, 46) = 1,34$, $p > 0,05$ (Anexo C – Saída 27).

O quarto pré-requisito é a mensuração das variáveis dependentes em uma escala de intervalo, e o EWCTET é intervalar. Existiu confiabilidade da correção dos testes pelos avaliadores (discutido no item 3.3.1), e quanto ao sexto pré-requisito, da homogeneidade dos parâmetros da regressão, este foi testado e não foi violado, $F(3, 42) = 1,41$, $p > 0,05$.

A Tabela 32 a seguir apresenta a ANCOVA com os pré-testes como covariáveis e com a interação Turma e Grupo.Idade (abaixo de 22 anos e 22 anos ou acima) (Anexo C – Saída 28).

Tabela 32 - ANCOVA para habilidades de raciocínio crítico considerando idade – FEA-USP.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig	η^2
Modelo Corrigido	521,802	4	4,118	0,006	0,268
Intercepto	191,319	1	6,040	0,018	0,118
Turma	21,696	1	0,685	0,412	0,015
Grupo.Idade	12,757	1	0,403	0,529	0,009
EWCTETpre	470,100	1	14,842	0,000	0,248
Turma*Grupo.Idade	6,123	1	0,193	0,662	0,004
Erro	1.425,343	45			
Total	9.897,750	50			
Total Corrigido	1.947,145	49			

Analizada a linha “Turma*Grupo.Idade”, verifica-se que não há diferenças significativas entre as médias ajustadas dos grupos experimento e controle considerando a idade, $F(1, 45) = 0,19$, $p > 0,05$. Também foram analisadas as saídas do SPSS que realizam as comparações aos pares e não se encontrou diferenças significativas entre as idades e turmas. Portanto, pode-se dizer que não houve diferenças significativas no desenvolvimento de habilidades de raciocínio crítico entre os alunos mais velhos e mais novos das turmas controle e experimento. Nota-se que a quantidade total de variação a ser explicada foi de 1.947,15 unidades, das quais o modelo consegue explicar 521,80 unidades, e o método de instrução e idade (Turma*Grupo.Idade) explica apenas 6,12 unidades.

Observando a última coluna da saída da ANCOVA, verifica-se uma estimativa do tamanho do efeito denominada eta quadrado (η^2), que é semelhante ao coeficiente de correlação ao quadrado R^2 (conhecido também como coeficiente de determinação) (FIELD, 2009).

Logo, tirando a raiz quadrada do valor (0,004) da linha da interação da turma com a idade, chega-se a $r = 0,06$ ou 6,32 %, o que é considerado um efeito pequeno. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 6,32% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 6.194 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra utilizada, encontrou-se 7,71% como resultado.

Analizadas as médias ajustadas dos grupos, nota-se que, apesar de não haver diferenças significativas entre as turmas e idade, a média ajustada dos alunos mais velhos da turma

experimento foi a mais alta (14,38), seguida dos mais novos da turma experimento (13,91) e por último ficaram os mais novos da turma controle (10,64) (Anexo C – Saída 29).

A conclusão final é a de que na instituição pública não foram encontradas diferenças significativas entre as médias dos testes que avaliam habilidade de raciocínio crítico entre as turmas experimento e controle considerando a idade. Logo, não se pode dizer que utilizando o modelo em uma abordagem geral houve maior desenvolvimento de habilidades de raciocínio crítico nos alunos mais velhos.

Este resultado corrobora o de Reed (1998), a qual não constatou em sua tese diferenças significativas entre as turmas experimento e controle de acordo com a idade. O resultado vai contra o modelo de desenvolvimento intelectual e moral de Perry (1970), que foi posteriormente modificado por King e Kitchener (1994), e que estabelece que as concepções das pessoas sobre a natureza do conhecimento e seus entendimentos sobre elas mesmas como conhecedoras, pensadoras e pessoas racionais geralmente se desenvolvem com o tempo. De acordo com King e Kitchener (1994), pesquisadores têm demonstrado que estudantes mais velhos diferem de estudantes mais novos ou de idade tradicional para determinado curso em uma variedade de formas, incluindo abordagens de estudo, atitudes em relação à escola e desempenho.

Uma limitação desta constatação é que apenas dois alunos da turma experimento tinham 22 anos ou mais. Analisando-se os gráficos da pesquisa demográfica, verifica-se que a turma controle apresentou uma faixa etária levemente maior.

Comparando as médias obtidas nos pós-testes das turmas experimento dos mais novos (12,39) e mais velhos (12,00) e as da turma controle (mais novos 12,08 e mais velhos 15,33) da instituição pública com dados de pesquisas anteriores do Suplemento do Manual do EWCTET (Universidade A = 13,1 e Universidade B = 14,3) (ENNIS, 1998) e com o resultado da pesquisa de Cornachione, Duncan e Johnson (2007) (Grupo USP = 15,1), pode-se dizer que, no geral, o desempenho da amostra da FEA-USP ficou um pouco abaixo das outras pesquisas que utilizaram o mesmo instrumento de avaliação de raciocínio crítico, com exceção dos alunos mais velhos da turma controle que tiveram alto desempenho.

4.4.1.2 CCTDI

A Tabela 33 apresenta a distribuição dos escores no CCTDI dos grupos experimento e controle considerando a idade.

Tabela 33 - Distribuição do CCTDI por método de instrução e idade na FEA-USP.

<u>Grupo Idade</u>	<u>Turma Experimento (n = 25)</u>					<u>Turma Controle (n = 27)</u>					<u>D P</u>
	<u>Pré-teste</u>		<u>n</u>	<u>Pós-teste</u>		<u>Pré-teste</u>		<u>n</u>	<u>Pós-teste</u>		
	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	
<u>Abaixo de 22 anos</u>	294,22	26,50	23	291,52	32,46	297,00	24,21	19	289,58	30,43	
<u>22 anos ou acima</u>	340,50	0,71	2	298,50	16,26	300,13	29,42	8	311,50	28,24	

Analisado o gráfico de caixa e bigode, encontraram-se dois valores atípicos, um para os pré-testes da turma controle do grupo 22 anos ou acima e outro para os pós-testes da turma experimento do grupo abaixo de 22 anos. Os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras assemelham-se à uma distribuição normal, com exceção do grupo 22 anos ou acima da turma experimento, que, devido ao pequeno número de participantes (2), não foi possível realizar o teste.

Retirando os valores atípicos, rodou-se novamente a análise e não foram encontrados mais valores atípicos e o requisito da normalidade não foi violado.

A Tabela 34 apresenta a distribuição dos escores no CCTDI dos grupos experimento e controle considerando a idade e sem os valores atípicos.

Tabela 34 - Distribuição do CCTDI por método de instrução e idade na FEA-USP sem valores atípicos.

<u>Grupo Idade</u>	<u>Turma Experimento (n = 24)</u>					<u>Turma Controle (n = 26)</u>				
	<u>Pré-teste</u>		<u>n</u>	<u>Pós-teste</u>		<u>Pré-teste</u>		<u>n</u>	<u>Pós-teste</u>	
	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>
<u>Abaixo de 22 anos</u>	291,50	23,62	22	287,59	27,04	297,00	24,21	19	289,58	30,43
<u>22 anos ou acima</u>	340,50	0,71	2	298,50	26,26	291,57	18,08	7	304,14	20,62

Comparando-se os dois grupos da turma experimento, nota-se que seus pós-testes ficaram abaixo dos pré-testes, sendo que os mais velhos tiveram a maior queda (42,00 ou 12%). Ainda

sobre a turma experimento, verifica-se que os mais velhos obtiveram maiores médias tanto nos pré-testes quanto nos pós-testes.

Analizados os dados da turma controle, constata-se uma pequena queda nos pós-testes em relação aos pré-testes dos alunos mais novos e um pequeno aumento nos pós-testes dos mais velhos (12,57 ou 4%). Verifica-se também que enquanto os mais velhos se saíram melhor nos pós-testes, os mais novos ficaram em média acima nos pré-testes.

Comparadas as turmas experimento e controle, nota-se que os mais novos da turma controle obtiveram melhor desempenho tanto nos pré-testes quanto nos pós-testes. Observando os mais velhos, nota-se que enquanto os da turma experimento apresentaram maior resultado médio nos pré-testes, os alunos da turma controle apresentaram melhor desempenho médio nos pós-testes.

Os pré-requisitos para a ANCOVA foram examinados para os escores do CCTDI. Como já visto em tópicos anteriores, o pressuposto de independência não foi violado. Verificou-se anteriormente que as amostras assemelham-se à uma distribuição normal, porém existe uma limitação em relação aos alunos mais velhos da turma experimento. A homogeneidade das variâncias não foi violada, $F(3, 46) = 0,30, p > 0,05$. O quarto pré-requisito é a mensuração das variáveis dependentes em uma escala de intervalo, e o CCTDI é intervalar. Existiu confiabilidade da correção dos testes pelos avaliadores (discutido no item 3.3.2), e o sexto pré-requisito, que é o da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(3, 42) = 0,49, p > 0,05$.

A Tabela 35 apresenta a ANCOVA com os pré-testes como covariáveis e com a interação Turma e Grupo.Idade (abaixo de 22 anos e 22 anos ou acima) (Anexo C – Saída 30).

Tabela 35 - ANCOVA para atitudes de raciocínio crítico considerando idade – FEA-USP.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig	η^2
Modelo Corrigido	11.884,013	4	5,440	0,001	0,326
Intercepto	2.383,411	1	4,364	0,042	0,088
Turma	1.574,518	1	2,883	0,096	0,060
Grupo.Idade	10,227	1	0,019	0,892	0,000
CCTDIpre	10.268,820	1	18,804	0,000	0,295
Turma*Grupo.Idade	1.770,235	1	3,242	0,078	0,067
Erro	24.574,487	45			
Total	4.273.419,000	50			
Total Corrigido	36.458,500	49			

Analísada a linha “Turma*Grupo.Idade”, verifica-se que não há diferenças significativas entre as médias ajustadas dos grupos experimento e controle considerando a idade, $F(1, 45) = 3,24$, $p > 0,05$. Analisaram-se também as saídas do SPSS que realizam as comparações aos pares e não foram encontradas diferenças significativas entre as idades e turmas. Portanto, pode-se dizer que não houve diferenças significativas no desenvolvimento de atitudes de raciocínio crítico entre os alunos mais velhos e mais novos das turmas controle e experimento. Nota-se que a quantidade total de variação a ser explicada foi de 36.458,50 unidades, das quais o modelo consegue explicar 11.884,01 unidades, e o método de instrução e idade (Turma*Grupo.Idade), 1.770,24 unidades.

Tirando a raiz quadrada do valor do η^2 (0,067) da linha da interação da turma com a idade, chega-se a $r = 0,26$ ou 25,88 %, o que é considerado um efeito de pequeno a médio. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 25,88% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 372 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra utilizada, encontrou-se 22,88% como resultado.

Analísadas as médias ajustadas dos grupos, nota-se que, além de não haver diferenças significativas entre as turmas e idades, os alunos mais velhos da turma experimento apresentaram a menor média ajustada (269,24), e os alunos mais novos de ambas as turmas apresentaram médias bem semelhantes (experimento = 290,23 e controle = 288,64) (Anexo C – Saída 31).

A conclusão final é que na instituição pública não foram encontradas diferenças significativas entre as médias dos testes que avaliam atitude de raciocínio crítico entre as turmas experimento e controle considerando a idade. Logo, não se pode dizer que utilizando o modelo em uma abordagem geral houve maior desenvolvimento de atitudes de raciocínio crítico nos alunos mais velhos.

Este resultado corrobora o de Reed (1998) e vai contra as teorias e modelos de Perry (1970) e King e Kitchener (1994). Porém, deve-se lembrar da limitação de possuir apenas dois alunos da turma experimento que tinham 22 anos ou mais e do fato de a turma controle apresentar uma faixa etária levemente maior.

Comparando as médias obtidas nos pós-testes das turmas experimento dos mais novos (287,59) e mais velhos (298,50) e as da turma controle (mais novos 289,58 e mais velhos 304,14) da instituição pública com os critérios de interpretação dos desempenhos no CCTDI descritos no item 3.3.2, pode-se constatar que existe fraca presença de atitudes de raciocínio crítico na amostra da FEA-USP, já que para ser considerado forte o desempenho precisa ser superior a 350 pontos.

4.4.2 UNIFIG

Neste tópico é analisada a Hipótese 3 para as duas variáveis dependentes: EWCTET e CCTDI.

4.4.2.1 EWCTET

A Tabela 36 apresenta a distribuição dos escores no EWCTET dos grupos experimento e controle considerando a idade.

Tabela 36 - Distribuição do EWCTET por método de instrução e idade na UNIFIG.

Grupo Idade	Turma Experimento (n = 16)					Turma Controle (n = 22)				
	Pré-teste			Pós-teste		Pré-teste		Pós-teste		
	Média	DP	N	Média	DP	Média	DP	n	Média	DP
Abaixo de 22 anos	2,44	5,07	8	2,75	2,55	3,96	5,27	12	2,88	3,82
22 anos ou acima	5,69	2,40	8	9,06	2,08	1,95	3,66	10	3,40	3,50

Analisando-se o gráfico de caixa e bigode, não foram encontrados valores atípicos. Os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras assemelham-se à uma distribuição normal e os testes de Levene não violaram a homogeneidade das variâncias.

Comparados os dois grupos (abaixo de 22 anos e 22 anos ou acima) da turma experimento, nota-se que seus pós-testes foram superiores aos pré-testes, sendo que os mais novos tiveram um leve aumento e os mais velhos, um grande aumento (3,37 ou 59%). Ainda sobre a turma experimento, verifica-se que os mais velhos obtiveram maiores médias, tanto nos pré-testes quanto nos pós-testes.

Analisados os dados da turma controle, constata-se que os alunos mais novos tiveram um menor rendimento nos pós-testes em relação aos pré-testes, enquanto os mais velhos apresentaram uma melhora nos pós-testes. Enquanto os mais novos se saíram melhor do que os mais velhos nos pré-testes, o contrário aconteceu em relação aos pós-testes.

Comparadas as turmas experimento e controle, nota-se que tanto nos pré-testes quanto nos pós-testes a turma controle obteve rendimento médio superior no grupo dos mais novos e inferior no dos mais velhos.

Os pré-requisitos para a ANCOVA foram examinados para os escores do EWCTET considerando turma e idade. Como já visto em tópicos anteriores, o pressuposto de independência não foi violado. Verificou-se anteriormente que as amostras assemelham-se à uma distribuição normal. A homogeneidade das variâncias não foi violada, $F(3, 34) = 0,26$, $p > 0,05$. O quarto pré-requisito é a mensuração das variáveis dependentes em uma escala de intervalo, e o EWCTET é intervalar. Existiu confiabilidade da correção dos testes pelos avaliadores (discutido no item 3.3.1) e o sexto pré-requisito, da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(3, 30) = 0,69$, $p > 0,05$.

A Tabela 37 apresenta a ANCOVA com os pré-testes como covariáveis e com a interação Turma e Grupo.Idade (abaixo de 22 anos e 22 anos ou acima) (Anexo C – Saída 32).

Tabela 37 - ANCOVA para habilidades de raciocínio crítico considerando idade – UNIFIG.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig	η^2
Modelo Corrigido	290,884	4	8,306	0,000	0,502
Intercepto	261,464	1	29,863	0,000	0,475
Turma	53,894	1	6,155	0,018	0,157
Grupo.Idade	95,963	1	10,960	0,002	0,249
Turma*Grupo.Idade	37,323	1	4,263	0,047	0,114
EWCTETpre	57,749	1	6,596	0,015	0,167
Erro	288,932	33			
Total	1.279,000	38			
Total Corrigido	579,816	37			

Analizada a linha “Turma*Grupo.Idade”, verifica-se que há diferenças significativas entre as médias ajustadas dos grupos experimento e controle considerando-se a idade, $F(1, 33) = 4,26$, $p < 0,05$. Outra linha que apresenta diferença significativa é a do “Grupo.Idade”, $F(1, 33) = 10,96$, $p < 0,05$. Analisaram-se também as saídas do SPSS que realizam as comparações aos pares e foram encontradas diferenças significativas entre as idades e as turmas. Portanto, pode-se dizer que houve diferenças significativas no desenvolvimento de habilidades de raciocínio crítico entre os alunos mais velhos e mais novos das turmas controle e experimento. Nota-se que a quantidade total de variação a ser explicada foi de 579,82 unidades, das quais o modelo consegue explicar 290,88 unidades, idade (Grupo.Idade) explica 95,96 unidades e o método de instrução e idade (Turma*Grupo.Idade) explica 37,32 unidades.

Como houve efeito significativo da variável idade considerada isoladamente, tirando a raiz quadrada do valor do η^2 (0,249) da linha idade (Grupo.Idade) chega-se a $r = 0,50$ ou 49,90 %, o que é considerado um efeito grande. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 49,90% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 102 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 38$) utilizada, encontrou-se 43,85% como resultado.

Calculando o d de Cohen para a variável idade, chega-se a $d = 0,77$ ou 77%, o que é considerado um tamanho de efeito grande. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 77% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 44 participantes. Ao calcular o poder

estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 38$) utilizada, encontrou-se 75,07% como resultado.

Tirando a raiz quadrada do valor do η^2 (0,114) da linha da interação da turma com a idade, chega-se a $r = 0,34$ ou 33,76 %, o que é considerado um efeito médio. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 33,76% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 220 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 38$) utilizada, encontrou-se 26,22% como resultado.

Calculando o d de Cohen para a interação turma e idade, chega-se a $d = 0,58$ ou 58%, o que é considerado um tamanho de efeito grande. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 58% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 76 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 38$) utilizada, encontrou-se 54,14% como resultado.

Consideradas as médias ajustadas dos grupos, nota-se que a maior média ajustada é a dos alunos mais velhos da turma experimento (8,4), seguida dos mais velhos da turma controle (3,9), e a média mais baixa é a dos mais novos da turma controle (2,7). Vale lembrar que houve diferenças significativas entre os grupos e as idades (Anexo C – Saída 33).

A conclusão é que pode-se dizer que utilizando o modelo em uma abordagem mista houve maior desenvolvimento de habilidades de raciocínio crítico nos alunos mais velhos. Houve também diferença significativa entre as idades sem considerar o método de instrução. Este resultado vai contra o de Reed (1998) e a favor das teorias e modelos de Perry (1970) e King e Kitchener (1994). Refletindo sobre isto, pode-se pensar em diversas ações didáticas que considerem a idade um fator importante no processo de ensino-aprendizagem, como, por exemplo: classes divididas por idade; diferentes métodos de ensino, que levem em consideração a heterogeneidade da classe em relação à idade; diferentes currículos para um mesmo curso, levando-se em consideração a idade etc.

Comparando as médias obtidas nos pós-testes das turmas experimento dos mais novos (2,75) e mais velhos (9,06) e as da turma controle (mais novos 2,88 e mais velhos 3,40) da

instituição particular com dados de pesquisas anteriores do Suplemento do Manual do EWCTET (Universidade A = 13,1 e Universidade B = 14,3) (ENNIS, 1998) e com o resultado da pesquisa de Cornachione, Duncan e Johnson (2007) (Grupo USP = 15,1), pode-se dizer que, no geral, o desempenho da amostra da UNIFIG ficou bem abaixo das outras pesquisas que utilizaram o mesmo instrumento de avaliação de raciocínio crítico.

4.4.2.2 CCTDI

A Tabela 38 apresenta a distribuição dos escores no CCTDI dos grupos experimento e controle considerando a idade.

Tabela 38 - Distribuição do CCTDI por método de instrução e idade na UNIFIG.

<u>Grupo Idade</u>	<u>Turma Experimento (n = 15)</u>						<u>Turma Controle (n = 23)</u>				
	<u>Pré-teste</u>			<u>Pós-teste</u>			<u>Pré-teste</u>		<u>Pós-teste</u>		
	<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>N</u>	<u>Média</u>	<u>DP</u>
<u>Abaixo de 22 anos</u>	322,43	10,20	7	313,43	25,47		282,08	25,48	13	280,46	32,16
<u>22 anos ou acima</u>	310,75	17,66	8	318,75	21,78		293,00	38,44	10	283,00	34,09

Analisando-se o gráfico de caixa e bigode, não foram encontrados valores atípicos. Os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras assemelham-se à uma distribuição normal e a homogeneidade das variâncias não foi violada.

Comparados os dois grupos da turma experimento, nota-se que os pós-testes dos alunos mais novos caíram em relação aos pré-testes, enquanto os mais velhos obtiveram melhora de rendimento nos pós-testes (8,00 ou 3%). Ainda sobre a turma experimento, verifica-se que os mais velhos obtiveram maiores médias nos pós-testes em comparação aos mais novos e o contrário ocorreu com os pré-testes.

Considerados os dados da turma controle, verifica-se que ambos os grupos caíram de rendimento médio nos pós-testes; porém, enquanto os mais novos tiveram uma leve queda, os mais velhos caíram mais (10,00 ou 3%). Verifica-se também que os mais velhos apresentaram melhor rendimento do que os mais novos em ambos os testes.

Comparando as turmas experimento e controle, nota-se que os mais novos da turma experimento obtiveram melhor desempenho do que os mais novos da turma controle, tanto

nos pré-testes quanto nos pós-testes. Comparando os mais velhos das turmas, nota-se o mesmo fenômeno.

Os pré-requisitos para a ANCOVA foram examinados para os escores do CCTDI. Como já visto em tópicos anteriores, o pressuposto de independência não foi violado. Verificou-se anteriormente que as amostras assemelham-se à uma distribuição normal. A homogeneidade das variâncias não foi violada, $F(3, 34) = 1,60$, $p > 0,05$. O quarto pré-requisito é a mensuração das variáveis dependentes em uma escala de intervalo, e o CCTDI é intervalar. Existiu confiabilidade da correção dos testes pelos avaliadores (discutido no item 3.3.2), e o sexto pré-requisito, da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(3, 30) = 1,02$, $p > 0,05$.

A Tabela 39 apresenta a ANCOVA com os pré-testes como covariáveis e com a interação Turma e Grupo.Idade (abaixo de 22 anos e 22 anos ou acima) (Anexo C – Saída 34).

Tabela 39 - ANCOVA para atitudes de raciocínio crítico considerando idade – UNIFIG.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig	η^2
Modelo Corrigido	29.703,972	4	21,385	0,000	0,722
Intercepto	269,885	1	0,777	0,384	0,023
Turma	517,743	1	1,491	0,231	0,043
Grupo.Idade	163,462	1	0,471	0,497	0,014
CCTDIpre	18.629,049	1	53,647	0,000	0,619
Turma*Grupo.Idade	1.109,112	1	3,194	0,083	0,088
Erro	11.459,396	33			
Total	3.354.016,000	38			
Total Corrigido	41.163,368	37			

Analísada a linha “Turma*Grupo.Idade”, verifica-se que não há diferenças significativas entre as médias ajustadas dos grupos experimento e controle considerando a idade, $F(1, 33) = 3,19$, $p > 0,05$. Analisaram-se também as saídas do SPSS que realizam as comparações aos pares e não foram encontradas diferenças significativas entre as idades e turmas. Portanto, pode-se dizer que não houve diferenças significativas no desenvolvimento de atitudes de raciocínio crítico entre os alunos mais velhos e mais novos das turmas controle e experimento. Nota-se que a quantidade total de variação a ser explicada foi de 41.163,37 unidades, das quais o modelo consegue explicar 29.703,97 unidades, e o método de instrução e idade (Turma*Grupo.Idade), 1.109,11 unidades.

Tirando a raiz quadrada do valor do η^2 (0,088) da linha da interação da turma com a idade, chega-se a $r = 0,30$ ou 29,67 %, o que é considerado um efeito médio. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 29,67% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 284 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra utilizada, encontrou-se 22,14% como resultado.

De acordo com as médias ajustadas dos grupos, nota-se que, apesar de não haver diferenças significativas entre as turmas e idades, os alunos mais velhos da turma experimento apresentaram a maior média ajustada (307,87) (Anexo C – Saída 35).

A conclusão para esta hipótese é que na instituição particular não foram encontradas diferenças significativas entre as médias dos testes que avaliam atitude de raciocínio crítico entre as turmas experimento e controle considerando a idade. Logo, não se pode dizer que utilizando o modelo em uma abordagem mista houve maior desenvolvimento de atitudes de raciocínio crítico nos alunos mais velhos. Este resultado corrobora o de Reed (1998) e vai contra as teorias e modelos de Perry (1970) e King e Kitchener (1994).

Comparando as médias obtidas nos pós-testes das turmas experimento dos mais novos (313,43) e mais velhos (318,75) e as da turma controle (mais novos 280,46 e mais velhos 283,00) da instituição particular com os critérios de interpretação dos desempenhos no CCTDI descritos no item 3.3.2, pode-se constatar que existe fraca presença de atitudes de raciocínio crítico na amostra da UNIFIG, já que para ser considerado forte o desempenho precisa ser superior a 350 pontos.

Portanto, pode-se dizer que duas diferenças significativas foram encontradas para a instituição particular (abordagem mista). Houve efeito significativo do método instrucional no desenvolvimento de habilidades de raciocínio crítico de acordo com a idade e diferenças significativas foram encontradas para as habilidades considerando o efeito isolado da idade.

Para os principais resultados deste item, tem-se a seguir o Gráfico 35.

Gráfico 35 - Pré-testes e pós-testes do EWCTET considerando idade da UNIFIG.

Nota-se que os mais velhos da turma experimento possuem maior ângulo de inclinação positivo de sua linha do que os mais novos da mesma turma, corroborando o efeito significativo do modelo instrucional no desenvolvimento de habilidades de raciocínio crítico de acordo com a idade. Nota-se, também, que o desempenho nos pós-testes dos mais velhos foi maior em ambas as turmas, o que pode explicar a diferença significativa do efeito isolado da idade.

4.5 Hipótese 4 - Gênero

A Hipótese 4 definida nesta pesquisa foi:

Esperam-se diferenças significativas nos testes (EWCTET e CCTDI) de habilidades e atitudes de raciocínio crítico dos alunos de acordo com o método instrucional e gênero (feminino/masculino).

Para a hipótese acima, realizaram-se análises descritivas de acordo com o gênero e método de instrução e checaram-se possíveis diferenças significativas entre os desempenhos por método de instrução e gênero.

4.5.1 FEA - USP

Neste tópico é analisada a Hipótese 4 para as duas variáveis dependentes: EWCTET e CCTDI.

4.5.1.1 EWCTET

A Tabela 40 apresenta a distribuição dos escores no EWCTET dos grupos experimento e controle considerando o gênero.

Tabela 40 - Distribuição do EWCTET por método de instrução e gênero na FEA-USP.

<u>Gênero</u>	<u>Turma Experimento (n = 25)</u>						<u>Turma Controle (n = 27)</u>				
	<u>Pré-teste</u>			<u>Pós-teste</u>			<u>Pré-teste</u>		<u>Pós-teste</u>		
	<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>
<u>Feminino</u>	9,21	7,29	12	11,21	7,75		14,79	6,97	7	13,07	6,13
<u>Masculino</u>	13,19	6,05	13	13,42	7,07		17,45	4,84	20	12,48	4,85

Os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras assemelham-se à uma distribuição normal, com exceção dos pós-testes das mulheres da turma experimento, $D(12) = 0,28$, $p < 0,05$. Logo, para as análises de covariância, os dados foram elevados ao quadrado e passaram a não mais violar os pressupostos da normalidade e homogeneidade das variâncias. Analisaram-se os gráficos de caixa e bigode e não foram encontrados valores atípicos.

Comparados os gêneros da turma experimento, nota-se que os homens tiveram em média rendimento superior ao das mulheres nos pré-testes e pós-testes, sendo que tanto as mulheres quanto os homens foram melhores nos pós-testes; no entanto, as mulheres tiveram o maior aumento (2,00 ou 22%).

Analisando-se os dados da turma controle, constata-se que tanto os homens quanto as mulheres obtiveram um menor rendimento nos pós-testes em relação aos pré-testes, e que a maior queda ocorreu para os homens (4,97 ou 29%). Verifica-se também que enquanto os homens se saíram melhor do que as mulheres nos pré-testes, as mulheres foram um pouco melhor do que os homens nos pós-testes.

Comparadas as turmas experimento e controle, nota-se que as mulheres da turma controle obtiveram melhor desempenho tanto nos pré-testes quanto nos pós-testes. Observando os homens, nota-se que enquanto os da turma controle apresentaram maior resultado médio nos pré-testes, os homens da turma experimento apresentaram melhor desempenho médio nos pós-testes.

Os pré-requisitos para a ANCOVA foram examinados para os escores do EWCTET. Como já visto em tópicos anteriores, o pressuposto de independência não foi violado. Verificou-se anteriormente que os valores foram elevados ao quadrado para não violar a normalidade. A homogeneidade das variâncias não foi violada, $F(3, 48) = 0,81$, $p > 0,05$. O quarto pré-requisito refere-se à mensuração das variáveis dependentes em uma escala de intervalo, e o EWCTET é intervalar. Existiu confiabilidade da correção dos testes pelos avaliadores (discutido no item 3.3.1), e o sexto pré-requisito, que é o da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(3, 44) = 1,10$, $p > 0,05$.

A Tabela 41 a seguir apresenta a ANCOVA com os pré-testes como covariáveis e com a interação Turma e Gênero (Anexo C – Saída 36).

Tabela 41 - ANCOVA para habilidade de raciocínio crítico considerando gênero – FEA-USP.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig.	η^2
Modelo Corrigido	143.455,053	4	2,309	0,072	0,164
Intercepto	248.235,193	1	15,979	0,000	0,254
Turma	25.708,935	1	1,655	0,205	0,034
Gênero	1.386,428	1	0,089	0,766	0,002
EWCTETprequadrado	122.030,320	1	7,855	0,007	0,143
Turma*Gênero	13.232,532	1	0,852	0,361	0,018
Erro	730.142,773	47			
Total	2.831.834,000	52			
Total Corrigido	873.597,827	51			

Analizada a linha “Turma*Gênero”, verifica-se que não há diferenças significativas entre as médias ajustadas dos grupos experimento e controle considerando o gênero, $F(1, 47) = 0,85$, $p > 0,05$. Analisaram-se também as saídas do SPSS que realizam as comparações aos pares e não se encontrou diferenças significativas entre os gêneros e as turmas. Portanto, pode-se dizer que não houve diferenças significativas no desenvolvimento de habilidades de raciocínio crítico entre as turmas controle e experimento considerando o gênero. Nota-se que

a quantidade total de variação a ser explicada foi de 873.597,83 unidades, das quais o modelo consegue explicar 143.455,05 unidades, e o método de instrução e gênero explica 13.232,53 unidades.

Tirando a raiz quadrada do valor do η^2 (0,018) da linha da interação da turma com o gênero chega-se a $r = 0,13$ ou 13,42 %, o que é considerado um efeito pequeno. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 13.42% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 1.376 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra utilizada, encontrou-se 12,14% como resultado.

Analizadas as médias ajustadas dos grupos, nota-se que, apesar de não haver diferenças significativas entre as turmas e gêneros, os homens da turma experimento apresentaram a maior média ajustada de todos os grupos (237,28). Em contrapartida os homens da turma controle apresentaram a menor média ajustada (152,41) (Anexo C – Saída 37).

A conclusão final para esta hipótese é que na instituição pública não foram encontradas diferenças significativas entre as médias dos testes que avaliam habilidade de raciocínio crítico entre as turmas experimento e controle considerando o gênero. Logo, não se pode dizer que utilizando o modelo em uma abordagem geral houve desenvolvimento de habilidades de raciocínio crítico de acordo com gênero.

Este resultado corrobora o de Reed (1998), a qual não constatou em sua tese diferenças significativas entre as turmas experimento e controle de acordo com o gênero. O resultado vai contra estudos que têm demonstrado que as “formas de aprendizagem” das mulheres são diferentes das dos homens (MILLER; FINLEY; MCKINLEY, 1990). Analisando-se os gráficos da pesquisa demográfica é possível verificar que o percentual de homens na turma controle é bem maior do que na turma experimento, e esta é uma limitação que deve ser levada em consideração para este resultado.

Comparando as médias obtidas nos pós-testes da turma experimento das mulheres (11,21) e dos homens (13,42) e as da turma controle (mulheres 13,07 e homens 12,48) da instituição pública com dados de pesquisas anteriores do Suplemento do Manual do EWCTET (Universidade A = 13,1 e Universidade B = 14,3) (ENNIS, 1998) e com o resultado da

pesquisa de Cornachione, Duncan e Johnson (2007) (Grupo USP = 15,1), pode-se dizer que, no geral, o desempenho da amostra da FEA-USP ficou um pouco abaixo das outras pesquisas que utilizaram o mesmo instrumento de avaliação de raciocínio crítico, com exceção dos homens da turma experimento, que ficaram um pouco acima da média de uma das Universidades relatadas no Suplemento do Manual do EWCTET.

4.5.1.2 CCTDI

A Tabela 42 apresenta a distribuição dos escores no CCTDI dos grupos experimento e controle considerando o gênero.

Tabela 42 - Distribuição do CCTDI por método de instrução e gênero na FEA-USP.

<u>Gênero</u>	<u>Turma Experimento (n = 25)</u>						<u>Turma Controle (n = 27)</u>				
	<u>Pré-teste</u>			<u>Pós-teste</u>			<u>Pré-teste</u>		<u>Pós-teste</u>		
	<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>
<u>Feminino</u>	284,92	27,53	12	288,42	23,48		294,71	16,76	7	302,14	16,50
<u>Masculino</u>	309,92	24,41	13	295,46	37,81		299,05	27,99	20	293,95	34,77

Os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras assemelham-se à uma distribuição normal. Encontrou-se um valor atípico nos pré-testes dos homens na turma experimento, o qual foi excluído. Após a exclusão mantiveram-se não violados os pressupostos da normalidade e homogeneidade das variâncias.

A Tabela 43 a seguir apresenta a distribuição dos escores no CCTDI dos grupos experimento e controle considerando o gênero e sem o valor atípico.

Tabela 43 - Distribuição do CCTDI por método de instrução e gênero na FEA-USP sem valores atípicos.

<u>Gênero</u>	<u>Turma Experimento (n = 24)</u>						<u>Turma Controle (n = 27)</u>				
	<u>Pré-teste</u>			<u>Pós-teste</u>			<u>Pré-teste</u>		<u>Pós-teste</u>		
	<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>N</u>	<u>Média</u>	<u>DP</u>
<u>Feminino</u>	284,92	27,53	12	288,42	23,48		294,71	16,76	7	302,14	16,50
<u>Masculino</u>	306,25	21,42	12	288,58	29,82		299,05	27,99	20	293,95	34,77

Comparando os gêneros da turma experimento, nota-se que, em média, os homens tiveram rendimento superior às mulheres nos pré-testes e que os pós-testes foram bem semelhantes.

As mulheres melhoraram o rendimento médio nos pós-testes, enquanto os homens pioraram (17,67 ou 6%).

Analisados os dados da turma controle, constata-se também que enquanto as mulheres apresentaram uma melhora média no desempenho dos pós-testes em relação aos pré-testes (7,43 ou 3%), os homens pioraram. Os homens tiveram melhor rendimento do que as mulheres nos pré-testes, enquanto as mulheres apresentaram maior valor médio nos pós-testes.

Comparadas as turmas experimento e controle, nota-se que as mulheres da turma controle obtiveram melhor desempenho tanto nos pré-testes quanto nos pós-testes. Observando os homens, nota-se que, enquanto os da turma experimento apresentaram maior resultado médio nos pré-testes, os homens da turma controle apresentaram melhor desempenho médio nos pós-testes.

Os pré-requisitos para a ANCOVA foram examinados para os escores do CCTDI. Como já visto em tópicos anteriores, o pressuposto de independência não foi violado. Verificou-se anteriormente que as amostras assemelham-se à uma distribuição normal. A homogeneidade das variâncias não foi violada, $F(3, 47) = 1,30$, $p > 0,05$. O quarto pré-requisito é a mensuração das variáveis dependentes em uma escala de intervalo, e o CCTDI é intervalar. Existiu confiabilidade da correção dos testes pelos avaliadores (discutido no item 3.3.2), e o sexto pré-requisito, que é o da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(3, 43) = 0,67$, $p > 0,05$.

A Tabela 44 seguir apresenta a ANCOVA com os pré-testes como covariáveis e com a interação Turma e Gênero (Anexo C – Saída 38).

Tabela 44 - ANCOVA para atitude de raciocínio crítico considerando gênero – FEA-USP.

Fonte	Soma dos Quadrados Tipo		F	Sig	η^2
	III	GI			
Modelo Corrigido	15.556,515	4	6,889	0,000	0,375
Intercepto	2.581,782	1	4,573	0,038	0,090
Turma	830,668	1	1,471	0,231	0,031
Gênero	1.754,664	1	3,108	0,085	0,063
Turma*Gênero	32,405	1	0,057	0,812	0,001
CCTDIpre	14.479,411	1	25,647	0,000	0,358
Erro	25.970,230	46			
Total	4.405.188,000	51			
Total Corrigido	41.526,745	50			

Analisada a linha “Turma*Gênero”, é possível verificar que não há diferenças significativas entre as médias ajustadas dos grupos experimento e controle considerando o gênero, $F(1, 46) = 0,06$, $p > 0,05$. Analisaram-se também as saídas do SPSS que realizam as comparações aos pares e não foram encontradas diferenças significativas entre os gêneros e turmas. Portanto, pode-se dizer que não houve diferenças significativas no desenvolvimento de atitudes de raciocínio crítico entre as turmas controle e experimento considerando o gênero. Nota-se que a quantidade total de variação a ser explicada foi de 41.526,75 unidades, das quais o modelo consegue explicar 15.556,52 unidades, e o método de instrução e gênero, apenas 32,41 unidades.

Tirando a raiz quadrada do valor do η^2 (0,001) da linha da interação da turma com o gênero, chega-se a $r = 0,03$ ou 3,16 %, o que é considerado um efeito pequeno. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 3,16% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 24.768 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra utilizada, encontrou-se 6,25% como resultado.

Analisadas as médias ajustadas dos grupos, nota-se que, apesar de não haver diferenças significativas entre as turmas e gêneros, as mulheres (experimento = 296,70 e controle = 303,61) apresentaram maiores médias ajustadas do que os homens (experimento = 282,03 e controle = 292,40) nas duas turmas (Anexo C – Saída 39).

A conclusão final para esta hipótese é que na instituição pública não foram encontradas diferenças significativas entre as médias dos testes que avaliam atitude de raciocínio crítico entre as turmas experimento e controle considerando o gênero. Logo, não se pode dizer que utilizando o modelo em uma abordagem geral houve desenvolvimento de atitudes de raciocínio crítico de acordo com o gênero. Este resultado corrobora o de Reed (1998) e vai contra os estudos de Miller, Finley e Mckinley (1990). Não se pode esquecer a limitação referente ao fato do percentual de homens na turma controle ter sido bem maior do que na turma experimento.

Comparando as médias obtidas nos pós-testes da turma experimento das mulheres (288,42) e dos homens (288,58) e as da turma controle (mulheres 302,14 e homens 293,95) da instituição pública com os critérios de interpretação dos desempenhos no CCTDI descritos no item 3.3.2, pode-se constatar que existe fraca presença de atitudes de raciocínio crítico na amostra da FEA-USP, já que para ser considerado forte o desempenho precisa ser superior a 350 pontos.

4.5.2 UNIFIG

Neste tópico é analisada a Hipótese 4 para as duas variáveis dependentes: EWCTET e CCTDI.

4.5.2.1 EWCTET

A Tabela 45 apresenta a distribuição dos escores no EWCTET dos grupos experimento e controle considerando o gênero.

Tabela 45 - Distribuição do EWCTET por método de instrução e gênero na UNIFIG.

<u>Gênero</u>	<u>Turma Experimento (n = 16)</u>						<u>Turma Controle (n = 22)</u>			
	<u>Pré-teste</u>			<u>Pós-teste</u>			<u>Pré-teste</u>		<u>Pós-teste</u>	
	<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>N</u>	<u>Média</u> <u>DP</u>
<u>Feminino</u>	2,11	4,13	9	3,83	3,06		4,75	3,73	8	3,19 4,09
<u>Masculino</u>	6,57	2,82	7	8,57	3,46		2,07	4,91	14	3,07 3,46

Encontrou-se um valor atípico nos pós-testes dos homens na turma experimento, o qual foi excluído.

A Tabela 46 apresenta a distribuição do EWCTET por método de instrução e gênero sem os valores atípicos.

Tabela 46 - Distribuição do EWCTET por método de instrução e gênero na UNIFIG sem valores atípicos.

<u>Gênero</u>	<u>Turma Experimento (n = 15)</u>			<u>Turma Controle (n = 22)</u>						
	<u>Pré-teste</u>		<u>n</u>	<u>Pós-teste</u>		<u>Pré-teste</u>		<u>N</u>	<u>Pós-teste</u>	
	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>
<u>Feminino</u>	2,11	4,13	9	3,83	3,06	4,75	3,73	8	3,19	4,09
<u>Masculino</u>	5,92	2,44	6	9,67	2,07	2,07	4,91	14	3,07	3,46

Os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras assemelham-se à uma distribuição normal. Nenhum valor atípico foi encontrado e há homogeneidade das variâncias.

Comparados os gêneros da turma experimento, nota-se que os homens tiveram, em média, rendimento superior às mulheres nos pré-testes e pós-testes, sendo que tanto as mulheres quanto os homens melhoraram o rendimento nos pós-testes.

Analisando-se os dados da turma controle, constata-se que enquanto as mulheres tiveram pós-testes menores que os pré-testes os homens melhoraram nos pós-testes. Verifica-se também que as mulheres se saíram melhor que os homens tanto nos pré-testes quanto nos pós-testes.

Comparadas as turmas experimento e controle, nota-se que as mulheres da turma controle obtiveram melhor desempenho nos pré-testes e pior nos pós-testes. Observando os homens, nota-se que os da turma experimento ficaram com médias superiores em ambos os testes.

Os pré-requisitos para a ANCOVA foram examinados para os escores do EWCTET. Como já visto em tópicos anteriores, o pressuposto de independência não foi violado. Verificou-se anteriormente que as amostras não violaram a normalidade. A homogeneidade das variâncias também não foi violada, $F(3, 33) = 0,42$, $p > 0,05$. O quarto pré-requisito refere-se à mensuração das variáveis dependentes em uma escala de intervalo, e o EWCTET é intervalar. Existiu confiabilidade da correção dos testes pelos avaliadores (discutido no item 3.3.1), e o sexto pré-requisito, que é o da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(3, 29) = 1,62$, $p > 0,05$.

A Tabela 47 apresenta a ANCOVA com os pré-testes como covariáveis e com a interação Turma e Gênero (Anexo C – Saída 40).

Tabela 47 - ANCOVA para habilidade de raciocínio crítico considerando gênero – UNIFIG.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig	η^2
Modelo Corrigido	285,595	4	7,910	0,000	0,497
Intercepto	236,753	1	26,230	0,000	0,450
Turma	96,778	1	10,722	0,003	0,251
Gênero	58,920	1	6,528	0,016	0,169
EWCTETpre	79,894	1	8,851	0,006	0,217
Turma*Gênero	22,880	1	2,535	0,121	0,073
Erro	288,837	32			
Total	1.275,000	37			
Total Corrigido	574,432	36			

Analizada a linha “Turma*Gênero”, verifica-se que não há diferenças significativas entre as médias ajustadas dos grupos experimento e controle considerando o gênero, $F(1, 32) = 2,54$, $p > 0,05$. Porém, nota-se que houve diferença significativa considerando o efeito isolado do gênero (sem considerar a interação com o método de instrução), $F(1, 32) = 6,53$, $p < 0,05$, e da turma (já analisada na Hipótese 1). Analisaram-se também as saídas do SPSS que realizam as comparações aos pares, e, como já confirmado anteriormente, encontraram-se diferenças significativas entre os efeitos isolados do método de instrução (Turma) e também entre os gêneros, mas não da interação das turmas e gêneros. Portanto, pode-se dizer que não houve diferenças significativas no desenvolvimento de habilidades de raciocínio crítico entre as turmas controle e experimento considerando o gênero.

Nota-se que a quantidade total de variação a ser explicada foi de 574,43 unidades, das quais o modelo consegue explicar 285,60 unidades, o efeito isolado do método de instrução explica 96,78, o efeito isolado do gênero explica 58,92 e o método de instrução e gênero explica apenas 22,88 unidades.

Como houve efeito significativo da variável gênero, tirando a raiz quadrada do valor do η^2 (0,169) da linha gênero chega-se a $r = 0,41$ ou 41,11 %, o que é considerado um efeito de médio a grande. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 41,11% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 148 participantes. Ao calcular o poder estatístico levando-se em

consideração as variáveis citadas e a amostra ($n = 37$) utilizada, encontrou-se 32,97% como resultado.

Calculando o d de Cohen para a variável gênero, chega-se a $d = 0,64$ ou 64%, o que é considerado um tamanho de efeito grande. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 64% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 62 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 37$) utilizada, encontrou-se 59,73% como resultado.

Tirando a raiz quadrada do valor do η^2 (0,073) da linha da interação da turma com o gênero, chega-se a $r = 0,27$ ou 27,02 %, o que é considerado um efeito de pequeno a médio. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 27,02% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 342 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 37$) utilizada, encontrou-se 19,67% como resultado.

Analizadas as médias ajustadas dos grupos, nota-se que, apesar de não haver diferenças significativas entre as turmas e gêneros, as diferenças entre as médias ajustadas considerando apenas turma ou gênero foram significativas, sendo que a turma experimento (6,48) apresentou maior média que a controle (3,08) e os homens (6,11) apresentaram maior média que as mulheres (3,46). Considerada a interação, verifica-se que a maior média foi a dos homens da turma experimento (8,69), seguidos das mulheres da mesma turma (4,27) e por último ficaram as mulheres da turma controle (2,64) (Anexo C – Saída 41).

Apesar das diferenças significativas para os efeitos isolados de método de instrução (Turma) e gênero, a conclusão final para esta hipótese é que utilizando o modelo em uma abordagem mista não houve o desenvolvimento de habilidades de raciocínio crítico de acordo com o gênero.

Este resultado corrobora o de Reed (1998) e, de certa maneira, os estudos de Miller, Finley e Mckinley (1990), pois apesar de a interação do método e gênero não ter apresentado influência significativa no desenvolvimento de habilidades, houve diferença significativa no desempenho de acordo com gênero. Portanto, pode-se pensar em ações didáticas que

considerem o gênero como fator importante no processo de ensino-aprendizagem, como, por exemplo: classes divididas por gênero; métodos de ensino diferentes, que levem em consideração os percentuais de cada gênero; diferentes currículos para um mesmo curso, levando-se em consideração o gênero etc.

O tamanho do efeito medido isoladamente para o gênero foi alto (64%), e, em função do tamanho da amostra ($n = 37$), o poder estatístico foi de apenas 60%. Logo, novamente se deve considerar o resultado com cautela, sendo aconselháveis futuras investigações com amostras maiores para se chegar ao poder estatístico desejado de 80%.

Comparando as médias obtidas nos pós-testes da turma experimento das mulheres (3,83) e dos homens (9,67) e as da turma controle (mulheres 3,19 e homens 3,07) da instituição particular com dados de pesquisas anteriores do Suplemento do Manual do EWCTET (Universidade A = 13,1 e Universidade B = 14,3) (ENNIS, 1998) e com o resultado da pesquisa de Cornachione, Duncan e Johnson (2007) (Grupo USP = 15,1), pode-se dizer que, no geral, o desempenho da amostra da UNIFIG ficou bem abaixo das outras pesquisas que utilizaram o mesmo instrumento de avaliação de raciocínio crítico.

4.5.2.2 CCTDI

A Tabela 48 apresenta a distribuição dos escores no CCTDI dos grupos experimento e controle considerando o gênero.

Tabela 48 - Distribuição do CCTDI por método de instrução e gênero na UNIFIG.

<u>Gênero</u>	<u>Turma Experimento (n = 15)</u>						<u>Turma Controle (n = 23)</u>				
	<u>Pré-teste</u>			<u>Pós-teste</u>			<u>Pré-teste</u>		<u>Pós-teste</u>		
	<u>Média</u>	<u>DP</u>	<u>N</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>
<u>Feminino</u>	309,63	14,74	8	305,50	24,72		278,00	27,00	9	263,78	23,71
<u>Masculino</u>	323,71	13,25	7	328,57	13,25		292,50	33,72	14	293,00	32,51

Os testes K-S não foram significativos ($p > 0,05$), indicando que as amostras assemelham-se à uma distribuição normal. Encontraram-se dois valores atípicos, um nos pré-testes e outro nos pós-testes dos homens na turma experimento, os quais foram excluídos. Após a exclusão mantiveram-se não violados os pressupostos da normalidade e homogeneidade das variâncias.

A Tabela 49 apresenta a distribuição dos escores no CCTDI dos grupos experimento e controle considerando o gênero sem os valores atípicos.

Tabela 49 - Distribuição do CCTDI por método de instrução e gênero na UNIFIG sem valores atípicos.

<u>Gênero</u>	<u>Turma Experimento (n = 13)</u>						<u>Turma Controle (n = 23)</u>				
	<u>Pré-teste</u>			<u>Pós-teste</u>			<u>Pré-teste</u>		<u>Pós-teste</u>		
	<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>		<u>Média</u>	<u>DP</u>	<u>n</u>	<u>Média</u>	<u>DP</u>
<u>Feminino</u>	309,63	14,74	8	305,50	24,72		278,00	27,00	9	263,78	23,71
<u>Masculino</u>	330,00	6,35	5	332,40	8,44		292,50	33,72	14	293,00	32,51

Comparados os gêneros da turma experimento, nota-se que, em média, os homens tiveram rendimento superior às mulheres nos pré-testes e pós-testes. Os homens melhoraram levemente o rendimento médio nos pós-testes, enquanto as mulheres pioraram (4,13 ou 1%).

Analisando-se os dados da turma controle, constata-se o mesmo padrão da turma experimento. Os homens se saíram melhor que as mulheres em ambos os testes, e, enquanto os homens tiveram um rendimento muito semelhante nos pós-testes em relação ao pré-testes, as mulheres pioraram (14,22 ou 5%).

Comparadas as turmas experimento e controle, nota-se que as mulheres da turma experimento obtiveram melhor desempenho médio tanto nos pré-testes quanto nos pós-testes. Observando os homens, nota-se o mesmo fenômeno, ou seja, os homens da turma experimento tiveram rendimento maior do que os da turma controle em ambos os testes.

Os pré-requisitos para a ANCOVA foram examinados para os escores do CCTDI. Como já visto em tópicos anteriores, o pressuposto de independência não foi violado. Verificou-se acima que as amostras assemelham-se à uma distribuição normal. A homogeneidade das variâncias não foi violada, $F(3, 32) = 1,89$, $p > 0,05$. O quarto pré-requisito, referente à mensuração das variáveis dependentes em uma escala de intervalo e o CCTDI é intervalar. Existiu confiabilidade da correção dos testes pelos avaliadores (discutido no item 3.3.2), e o sexto pré-requisito, que é o da homogeneidade dos parâmetros da regressão, foi testado e não foi violado, $F(3, 28) = 0,35$, $p > 0,05$.

A Tabela 50 apresenta a ANCOVA com os pré-testes como covariáveis e com a interação Turma e Gênero (Anexo C – Saída 42).

Tabela 50 - ANCOVA para atitude de raciocínio crítico considerando gênero – UNIFIG.

Fonte	Soma dos Quadrados Tipo III	gl	F	Sig	η^2
Modelo Corrigido	30.225,471	4	25,361	0,000	0,766
Intercepto	951,668	1	3,194	0,084	0,093
Turma	1.026,104	1	3,444	0,073	0,100
Gênero	1.488,493	1	4,996	0,033	0,139
CCTDIpre	13.560,338	1	45,512	0,000	0,595
Turma*Gênero	93,624	1	0,314	0,579	0,010
Erro	9.236,418	31			
Total	3.149.982,000	36			
Total Corrigido	39.461,889	35			

Analizada a linha “Turma*Gênero”, verifica-se que não há diferenças significativas entre as médias ajustadas dos grupos experimento e controle considerando o gênero, $F(1, 31) = 0,31$, $p > 0,05$. Porém, nota-se que o efeito isolado do gênero foi significativo, $F(1, 31) = 5,00$, $p < 0,05$. Analisaram-se também as saídas do SPSS que realizam as comparações aos pares, e novamente não foram encontradas diferenças significativas entre a interação gênero e turma; porém, encontrou-se diferença significativa isoladamente em relação ao gênero. Portanto, pode-se dizer que não houve diferenças significativas no desenvolvimento de atitudes de raciocínio crítico entre as turmas controle e experimento considerando o gênero. Nota-se que a quantidade total de variação a ser explicada foi de 39.461,89 unidades, das quais o modelo consegue explicar 30.225,47 unidades, o efeito isolado do gênero explica 1.488,49 unidades e o método de instrução e gênero explica apenas 93,62 unidades.

Como houve efeito significativo da variável gênero, tirando a raiz quadrada do valor do η^2 (0,139) da linha gênero chega-se a $r = 0,37$ ou 37,28 %, o que é considerado um efeito de médio a grande. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 37,28% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 180 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 36$) utilizada, encontrou-se 27,70% como resultado.

Calculando o d de Cohen para a variável gênero, chega-se a $d = 0,46$ ou 46%, o que é considerado um tamanho de efeito de médio a grande. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 46% e um nível de significância de 5% se

obter um poder estatístico de 80%, chegou-se ao número de 120 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 36$) utilizada, encontrou-se 36,48% como resultado.

Tirando a raiz quadrada do valor do η^2 (0,010) da linha da interação da turma com o gênero, chega-se a $r = 0,10$ ou 10,00 %, o que é considerado um efeito pequeno. Calculando o tamanho da amostra que seria necessária para com um tamanho de efeito de 10,00% e um nível de significância de 5% se obter um poder estatístico de 80%, chegou-se ao número de 2.476 participantes. Ao calcular o poder estatístico levando-se em consideração as variáveis citadas e a amostra ($n = 36$) utilizada, encontrou-se 8,65% como resultado.

Analizadas as médias ajustadas dos grupos, nota-se que, no geral, os homens (302,28) apresentaram maior média ajustada do que as mulheres (287,82), e esta diferença foi significativa. Apesar de não haver diferença significativa entre as médias considerando método de instrução e gênero, verifica-se que os homens da turma experimento (307,36) apresentaram as maiores médias ajustadas, seguidos pelos homens da turma controle (297,20) e por último ficaram as mulheres da turma controle (279,29) (Anexo C – Saída 43).

Apesar de ter sido constatada diferença significativa para o efeito isolado do gênero, a conclusão final é que na instituição particular não foram encontradas diferenças significativas entre as médias dos testes que avaliam atitude de raciocínio crítico entre as turmas experimento e controle considerando o gênero. Logo, não se pode dizer que utilizando o modelo em uma abordagem mista houve desenvolvimento de atitudes de raciocínio crítico de acordo com gênero. Este resultado corrobora o de Reed (1998) e, de certa maneira, os estudos de Miller, Finley e Mckinley (1990), pois houve diferença significativa entre os gêneros.

O tamanho do efeito medido isoladamente para o gênero foi de médio a alto (46%), e, em função do tamanho da amostra ($n = 36$), o poder estatístico foi de 36%. Logo, novamente se deve considerar o resultado com cautela, sendo aconselháveis futuras investigações com amostras maiores.

Comparando as médias obtidas nos pós-testes da turma experimento das mulheres (305,50) e dos homens (332,40) e as da turma controle (mulheres 263,78 e homens 293,00) da instituição particular com os critérios de interpretação dos desempenhos no CCTDI descritos no item

3.3.2, pode-se constatar que existe fraca presença de atitudes de raciocínio crítico na amostra da UNIFIG, já que para ser considerado forte o desempenho precisa ser superior a 350 pontos. Vale ressaltar que os homens da turma experimento quase alcançaram os 350 pontos.

Resumindo, para a Hipótese 4 houve diferenças significativas para habilidades e atitudes de raciocínio crítico no efeito isolado do gênero na instituição particular (abordagem mista). Não houve efeito significativo do modelo instrucional para o desenvolvimento de habilidades ou atitudes de raciocínio crítico de acordo com gênero em ambas as IES.

Para os principais resultados deste item, tem-se a seguir os Gráficos 36 e 37.

Gráfico 36 - Pré-testes e pós-testes do EWCTET considerando gênero da UNIFIG.

Nota-se que a diferença significativa no efeito isolado do gênero para habilidades de raciocínio crítico pode ter como principal explicação o alto desempenho dos homens da turma experimento em relação aos outros grupos.

Gráfico 37 - Pré-testes e pós-testes do CCTDI considerando gênero da UNIFIG.

Nota-se novamente que a diferença significativa no efeito isolado do gênero para atitudes de raciocínio crítico pode ter como principal explicação o alto desempenho dos homens da turma experimento em relação aos outros grupos.

5 CONCLUSÕES

Neste capítulo é resgatado o problema da pesquisa, seu objetivo geral, a literatura revisada, os dados demográficos e os resultados para analisar as principais conclusões e contribuições do trabalho, tecendo comentários sobre suas limitações e sugestões para futuras pesquisas.

5.1 Principais Conclusões

O problema de pesquisa contextualizado e definido no Capítulo 1 foi: A aplicação do modelo instrucional de raciocínio crítico de Richard Paul melhora o nível de habilidades e atitudes de raciocínio crítico dos alunos de disciplinas da área contábil?

O objetivo geral do trabalho foi conhecer, analisar e avaliar empiricamente a efetividade em ensinar o modelo de Richard Paul (1995) para raciocínio crítico em alunos de disciplinas da área contábil no desenvolvimento de habilidades para pensar criticamente sobre contabilidade e atitudes de raciocínio crítico em geral.

O conhecimento do modelo foi adquirido principalmente pela sua utilização no quase-experimento piloto desenvolvido na instituição pública (FEA-USP), sendo que após a análise dos seus resultados, decidiu-se aumentar a intensidade das intervenções no quase-experimento na instituição particular (UNIFIG) e seguir o mesmo protocolo do estudo de Reed (1998) para fins de comparabilidade dos resultados. A avaliação das variáveis dependentes foi realizada por meio de dois instrumentos padronizados (EWCTET e CCTDI). Seus resultados foram expostos e analisados no capítulo anterior. Neste capítulo, estes resultados são analisados com base nos dados demográficos e na literatura revisada e as eventuais limitações e sugestões para futuras pesquisas são tecidas.

Antes de iniciar a discussão sobre os principais resultados, segue o Quadro 9, que os resume por cada quase-experimento de acordo com cada hipótese da pesquisa.

Quadro 9 - Resumo dos resultados das IES por hipótese.

Hipóteses		FEA-USP (Abordagem Geral)	UNIFIG (Abordagem Mista)
1 – Habilidade		Não Confirmada	Confirmada
2 – Atitude		Não Confirmada	Não Confirmada ¹
3 – Idade	Habilidade	Não Confirmada	Confirmada ²
	Atitude	Não Confirmada	Não Confirmada
4 – Gênero	Habilidade	Não Confirmada	Não Confirmada ²
	Atitude	Não Confirmada	Não Confirmada ²

¹ Confirma-se para a atitude curiosidade.

² Diferenças significativas para efeitos isolados (desconsiderando o método de instrução).

Ao analisar o Quadro 9, verifica-se que apenas duas hipóteses foram confirmadas para a instituição particular (que possuiu uma abordagem mista de aplicação do método instrucional):

- 1) Os alunos da turma experimento se saíram melhor do que os alunos da turma controle em um teste que avalia habilidades de raciocínio crítico. Portanto, pode-se dizer que o modelo instrucional (abordagem mista) teve efeito significativo no desenvolvimento de habilidades de raciocínio crítico.
- 2) Os alunos mais velhos (22 anos ou acima) se saíram melhor do que os alunos mais novos (abaixo de 22 anos) em um teste que avalia habilidades de raciocínio crítico. Portanto, pode-se dizer que o modelo instrucional (abordagem mista) teve efeito significativo no desenvolvimento de habilidades de raciocínio crítico considerando a idade.

Além destas duas confirmações, houve também para a instituição particular (abordagem mista) uma diferença significativa no desenvolvimento da atitude curiosidade. Diante disso pode-se dizer que o modelo instrucional teve efeito significativo no desenvolvimento desta atitude de raciocínio crítico.

Encontraram-se também na instituição particular duas diferenças significativas no desenvolvimento de habilidades de raciocínio crítico considerando os efeitos isolados de

idade e gênero e uma diferença significativa no desenvolvimento das atitudes de raciocínio crítico considerando o efeito isolado do gênero.

Analisando uma das contribuições desta pesquisa que foi o resultado positivo na utilização da abordagem mista no desenvolvimento das habilidades de raciocínio crítico nos alunos da instituição particular, pode-se procurar possíveis explicações para o sucesso na utilização do modelo de desenvolvimento de raciocínio crítico de Richard Paul. O resultado em questão vai de encontro com o pensamento de Siegel (1988) que afirma que um pensador crítico deve ter bom entendimento de princípios de assuntos específicos e de assuntos neutros (lógicos) que governam a avaliação das razões e que o deixam apto a utilizá-los. Logo, para Siegel, o debate sobre a generalização ou não de habilidades de raciocínio crítico não é necessário. No experimento na UNIFIG as intervenções foram realizadas utilizando conceitos da disciplina em questão e também abordando apenas o modelo, os alunos da turma experimento utilizaram para a análise de artigos, capítulos de livro e exercícios, os guias disponíveis na apostila de raciocínio crítico. Porém, de acordo com os resultados e com as impressões do autor desta pesquisa, os alunos da turma experimento da UNIFIG aprenderam a identificar os elementos de raciocínio crítico em um diálogo, artigo ou livro, tendo compreendido o modelo sem necessariamente os materiais analisados serem específicos da disciplina (por exemplo: análise de um capítulo do livro “O Homem que calculava” de Malba Tahan), logo, concordando com a visão de Siegel.

Os resultados da pesquisa também sugerem estar em harmonia com o pensamento de Lipman (2003), já que para ele pensar significa a posse de um conceito, uma opinião ou uma crença, e isso não implica o uso de bom julgamento. Qualquer pensamento que envolve critério, comparação e se baseie em padrões é raciocinar. O modelo de raciocínio crítico aplicado na turma experimento fez com que os alunos utilizassem critérios e padrões para pensar sobre determinado objeto de estudo (artigo, livro, problema etc.), logo como os resultados apontaram um desenvolvimento significativo das habilidades de raciocínio crítico da turma experimento da UNIFIG, pode-se dizer que eles foram capazes de raciocinar criticamente sobre o teste de argumentação de uma situação cotidiana abordada no EWCTET (*Ennis Weir Critical Thinking Essay Test*).

Outra importante conclusão e contribuição do trabalho se refere à abordagem que deve ser empregada na aplicação do modelo, de acordo com Ennis (1993) existem quatro:

- Geral: independentemente do conteúdo, as habilidades e atitudes de raciocínio crítico são ensinadas aos alunos.
- Infusão: requer profunda reflexão sobre um assunto e, adicionalmente, princípios gerais de raciocínio crítico são ensinados e aplicados nas atividades desenvolvidas.
- Imersão: os alunos são incentivados a pensar criticamente sobre determinado assunto; porém, os princípios gerais de raciocínio crítico não são apresentados.
- Misto: uma mistura do Geral com Infusão ou Imersão, os alunos estão envolvidos em um curso em que analisam criticamente um assunto com ou sem os conceitos gerais de raciocínio crítico explícitos; porém, há uma parte separada do curso que tem como objetivo ensinar princípios gerais de raciocínio crítico.

No quase-experimento piloto na FEA-USP o modelo de desenvolvimento de raciocínio crítico de Richard Paul foi aplicado na abordagem Geral e não foram constatadas diferenças significativas entre as turmas experimento e controle, logo não se pode afirmar que houve efeito do modelo. Porém, no experimento na UNIFIG, a abordagem foi a Mista, mesma utilizada na pesquisa de Reed (1998) e encontraram-se os resultados que estão sendo discutidos e analisados neste tópico. Portanto, a abordagem Mista que também já foi mencionada na meta-análise de Abrami *et al.* (2008) como a mais eficaz para o desenvolvimento de habilidades e atitudes de raciocínio crítico, também obteve resultados positivos nesta pesquisa e entende-se que isto ocorreu não necessariamente por ser necessário vincular o modelo à determinado assunto, mas sim por causa da maior intensidade das intervenções e aplicações em atividades em sala de aula e extra-classe.

Os resultados também corroboram as ideias de Paul, já que ele disserta que raciocínio crítico requer uma integração de campos cognitivos e afetivos. Para ele, o conteúdo em qualquer disciplina deveria ser visto e ensinado como um modelo de pensamento (por exemplo, história como pensamento histórico, biologia como pensamento biológico) e seu modelo instrucional defende que se deve ensinar os alunos a avaliar seus próprios pensamentos, seja por meio da leitura, escrita, audição ou fala, já que alguém incapaz de avaliar seu próprio pensamento não pode ser considerado um pensador crítico (PAUL, 1995). No quase-experimento da UNIFIG

houve grande incentivo à leitura, discussão e análise crítica de artigos, capítulos de livros e exercícios com a utilização dos guias de análise disponibilizados na apostila sobre o modelo de desenvolvimento de raciocínio crítico de Richard Paul. Logo, o objetivo da disciplina não foi de apenas transmitir um conhecimento pronto, mas também desenvolver nos alunos as habilidades e atitudes de raciocínio crítico tendo como suporte o modelo de Paul.

Pensando no modelo de desenvolvimento de raciocínio crítico de Richard Paul como um suporte para ajudar os alunos a pensarem criticamente e assim melhorar seus resultados em um teste que avalia habilidades de raciocínio crítico em geral (EWCTET), pode-se fazer uma analogia com o conceito metafórico do andaime¹⁸ de Vygotsky, no qual ele compara o processo de aprendizagem à um andaime que suporta a construção de um prédio. No começo da construção o andaime é essencial, pois sem o mesmo os trabalhadores não conseguem alcançar os andares seguintes, porém conforme o prédio vai se erguendo, sua própria estrutura começa a servir de apoio aos trabalhadores, sem tanta dependência do andaime (VYGOTSKY, 1962). Logo, no começo é essencial a ajuda do professor e de outros colegas para dar suporte ao aluno no processo de aprendizagem, sendo que após a compreensão dos pilares de uma disciplina, teoria, conceito ou modelo, o aluno conseguirá sozinho compreender assuntos ou problemas que anteriormente precisaria de ajuda.

Refletindo sobre as teorias de Vygotsky e sua relação com os resultados deste trabalho pode-se citar sua teoria da zona de desenvolvimento proximal que basicamente se trata do que as pessoas conseguem fazer inicialmente apenas com a ajuda de terceiros e que na sequência conseguirão fazer de forma independente (LEFRANÇOIS, 2009). De acordo com Lefrançois (2009, p. 269):

A maneira mais simples de explicar o conceito de zona de desenvolvimento proximal é dizer que ela é uma espécie de potencial para o desenvolvimento. Esclarecendo: Leonardo e Daniel são dois garotos de 7 anos que podem, ambos, em circunstâncias normais, responder adequadamente as mesmas perguntas respondidas pela média das crianças de 7 anos e realizar tarefas mais ou menos no mesmo tempo. A inteligência dessas duas crianças é média. Suponha que, quando estimulado e ajudado por um adulto competente ou por uma criança mais velha, Leonardo consiga, com sucesso, realizar tarefas e responder perguntas mais características de crianças de 9 anos, e Daniel não. Nesse caso, seria apropriado afirmar que a zona de desenvolvimento proximal de Leonardo é maior do que a de Daniel; ou seja, abrange uma esfera maior de funções mentais.”

¹⁸ Tradução do termo *scaffolding* na língua inglesa.

Portanto, na explicação e aplicação do modelo de desenvolvimento de Richard Paul buscou-se ajudar os alunos a desenvolverem seu raciocínio crítico primeiramente com o suporte do professor e colegas, para posteriormente eles conseguirem por si próprios analisarem argumentos em um carta, um documento, analisarem um artigo, capítulo de livro ou resolução de um exercício. Logo, resgatando a definição de raciocínio crítico do autor desta obra, buscou-se ajudar os alunos a desenvolver um pensamento que facilita julgamento, pois identifica, analisa e julga os elementos de um raciocínio. Os elementos de um raciocínio são: objetivo, problema, informações, interpretações, conceitos, suposições, implicações e ponto de vista.

Uma outra contribuição desta pesquisa é que na instituição particular encontrou-se diferença significativa entre as médias ajustadas das turmas experimento e controle considerando idade. Portanto, pode-se dizer que utilizando o modelo em uma abordagem mista houve maior desenvolvimento de habilidades de raciocínio crítico nos alunos mais velhos. Houve também diferença significativa entre as idades sem considerar o método de instrução. Este resultado vai contra o de Reed (1998) e a favor das teorias e modelos de Perry (1970) e King e Kitchener (1994) que defendem que alunos mais velhos possuem maior facilidade para o desenvolvimento do raciocínio crítico. Refletindo sobre este resultado, pode-se pensar em ações didáticas que considerem a idade um fator importante no processo de ensino-aprendizagem, como, por exemplo: classes divididas por idade; diferentes métodos de ensino, que levem em consideração a heterogeneidade da classe em relação à idade; diferentes currículos para um mesmo curso, levando-se em consideração a idade etc. A impressão do autor da pesquisa é que dos poucos interessados na utilização do modelo, como já comentado em tópicos anteriores, a maioria era composta por alunos mais velhos. Logo, pode-se refletir que os alunos mais velhos não só possuem maior facilidade para o desenvolvimento do raciocínio crítico, como também possuem maior interesse nestes assuntos.

Apesar dos resultados positivos na utilização do modelo para desenvolvimento de habilidades de raciocínio crítico na instituição particular, comparando com outros estudos que utilizaram o EWCTET, o desempenho da amostra da UNIFIG ficou bem abaixo. Possíveis explicações para o baixo desempenho podem ser analisadas por alguns dados demográficos que trazem indícios de uma formação básica dos alunos e um nível socioeconômico mais baixo do que os das amostras que servem como comparação dos resultados do EWCTET.

Uma outra contribuição do presente estudo foi a constatação no quase-experimento da instituição particular de diferença significativa e maior média ajustada da turma experimento para a atitude curiosidade. Portanto, pode-se dizer que o modelo instrucional teve efeito no desenvolvimento desta atitude de raciocínio crítico. Este resultado corrobora o de Abrami *et al.* (2008), o qual apresentou evidências de desenvolvimento de habilidades e atitudes de raciocínio crítico de acordo com o modelo instrucional. A teoria de Richard Paul (1995) defende que as atitudes de raciocínio crítico devem ser explicitadas para conscientizar os alunos de que não basta possuir as habilidades de raciocínio crítico, é preciso ter intenção ou atitude para utilizá-las. Na abordagem mista, o professor procurou sempre conscientizar a turma experimento sobre a importância de possuir as atitudes de raciocínio crítico, e isto pode ter sido importante no desenvolvimento significativo da atitude curiosidade e na maior média nos pós-testes da turma experimento em relação à controle. Outra importante constatação foi que ao se comparar a média ajustada (53,64) da atitude curiosidade da turma experimento da instituição particular com os critérios de interpretação dos desempenhos no CCTDI descritos no item 3.3.2, pode-se constatar que existe forte presença em relação a esta atitude.

Apesar de a interação do método e gênero não ter apresentado influência significativa no desenvolvimento de habilidades e atitudes, houve diferença significativa no desempenho de acordo com o efeito isolado do gênero. Portanto, pode-se pensar em ações didáticas que considerem o gênero como fator importante no processo de ensino-aprendizagem, como, por exemplo: classes divididas por gênero; métodos de ensino diferentes, que levem em consideração os percentuais de cada gênero; diferentes currículos para um mesmo curso, levando-se em consideração o gênero etc. Este resultado corrobora o de Reed (1998) e, de certa maneira, os estudos de Miller, Finley e Mckinley (1990) que defendem que as formas de raciocínio são diferentes entre os gêneros.

Estas foram as principais conclusões e contribuições do estudo no que diz respeito aos resultados positivos encontrados na utilização do modelo de desenvolvimento de raciocínio crítico de Richard Paul no quase-experimento da UNIFIG. Devido às possíveis limitações que são tratadas no tópico a seguir, as conclusões e resultados devem ser analisados com cautela. De maneira geral, pode-se concluir que nenhum efeito significativo foi evidenciado no estudo piloto na FEA-USP devido à baixa intensidade de intervenções utilizando a abordagem Geral de aplicação do modelo. Por se tratar de um estudo pioneiro no Brasil em relação ao assunto aplicado à área contábil e também à metodologia do quase-experimento aplicada na sua parte

empírica, a pesquisa se justifica principalmente por contribuir com a abertura do debate sobre o assunto na área contábil brasileira, cobrir uma lacuna na literatura internacional e servir de protocolo para futuros estudos experimentais. No tópico a seguir, listam-se possíveis limitações que devem ser consideradas nas análises dos resultados e devem procurar ser controladas em futuras investigações.

5.2 Limitações

Abaixo são listadas e analisadas possíveis limitações da pesquisa referentes, principalmente, à diferenças entre as turmas experimento e controle e à variáveis que não foram objeto de controle dos quase-experimentos e que podem ter interferido de alguma maneira nos resultados encontrados.

Analisados os gráficos da pesquisa demográfica da instituição pública, notou-se que duas características se apresentaram fortemente diferentes entre as turmas controle e experimento: gênero e emprego. A turma controle apresentou um número maior de homens em relação ao total se comparada com a turma experimento; além disso, possui um bom percentual que trabalha em período integral, enquanto ninguém trabalha o dia todo na turma experimento.

Com base na análise destes mesmos gráficos, também é possível notar que existe uma leve diferença entre as turmas experimento e controle em relação a outras três características: idade, renda e fluência em língua estrangeira. A turma experimento aparenta possuir maior renda familiar em relação à turma controle, menor faixa etária e menor fluência em língua estrangeira.

Considerando os gráficos da pesquisa demográfica da instituição particular, notou-se que duas características se apresentaram levemente diferentes entre as turmas controle e experimento: gênero e renda. Comparada à turma experimento, a turma controle apresentou maior percentual de homens e maior renda familiar.

Em sua maioria, os tamanhos de efeito dos resultados foram altos, exceto o do gênero em relação às atitudes de raciocínio crítico da instituição particular, que foi médio. Outra

limitação se refere aos poderes estatísticos, já que nenhum atingiu o nível desejável de 80%. Portanto, outra limitação que tem ligação com os tamanhos de efeito e os poderes estatísticos foram os tamanhos das amostras.

Analisados os gráficos demográficos, notou-se que a turma experimento da UNIFIG apresentou duas grandes diferenças em relação à controle: a turma experimento possuiu maior hábito de leitura e escrita. No entanto, sabendo que a pesquisa demográfica foi aplicada ao final do semestre, o modelo pode ter influenciado o desenvolvimento destas características na turma experimento. Porém, o professor incentivou durante todo o semestre a turma experimento a ler artigos e capítulos de livros aplicando o guia de análise crítica disponibilizado na apostila de raciocínio crítico, logo, o incentivo a pesquisa pode ser considerado uma variável *Hawthorne*, já que o maior hábito de leitura pode ter influenciado os resultados dos testes, os quais exigem interpretação e análise crítica.

Outro fato que pode ser considerado uma limitação dos resultados encontrados se refere à utilização de apenas dois métodos estatísticos de avaliação, a ANCOVA e os testes t. Portanto, futuras pesquisas podem pensar na utilização, também, de outras ferramentas estatísticas, como por exemplo, testes não paramétricos.

O tempo entre a primeira intervenção e os pós-testes em ambos os quase-experimentos foi de aproximadamente dois meses e meio. Alguns autores citados na revisão de literatura defendem que para o desenvolvimento do raciocínio crítico as intervenções devem ser realizadas durante um longo período. Portanto, o período de análise pode ser considerado como uma limitação do presente estudo.

Apesar das desistências entre as turmas controle e experimento serem semelhantes entre as turmas de cada quase-experimento, o fato de se ter utilizado apenas os alunos que completaram todos os testes pode ser considerado uma limitação da pesquisa, já que eles podem possuir uma tendência à utilização do modelo e isto pode ter interferido de alguma forma os resultados. Portanto, dentro da definição exposta no trabalho, isto pode ser considerado um possível efeito *Hawthorne*.

Os testes utilizados para avaliar habilidades (EWCTET) e atitudes (CCTDI) de raciocínio crítico são padronizados e estrangeiros. Apesar de se referirem à situações do cotidiano, o fato

de não serem nacionais pode ter influenciado de alguma forma os resultados, já que as amostras pertencem à uma realidade brasileira e este fato pode ser considerado uma possível limitação do presente trabalho.

As avaliações do EWCTET foram realizadas por duas pessoas que primeiramente analisaram o manual e o seu formulário de critério de avaliação. Apesar das altas correlações entre as avaliações e dos cuidados tomados para que os avaliadores não soubessem a que turma determinada redação se referia, o fato de um dos avaliadores também ser o autor desta pesquisa pode ter acarretado certo viés aos resultados e isto deve, também, ser considerado como uma possível limitação.

Apesar do professor não ter notado comunicação entre os alunos das turmas experimento e controle; e ter tomado todos os cuidados para que as intervenções e a apostila não viessem ao conhecimento dos alunos da turma controle, a possibilidade de os alunos terem conversado entre eles sobre o modelo e as diferenças de metodologia de ensino pode ter influenciado de alguma forma os resultados. Portanto, isto pode ser considerado um possível efeito *Hawthorne*.

Houve um incentivo de um ponto na média para os alunos que participassem de todas as etapas da pesquisa. Portanto, refletindo sobre uma hipótese de alunos que já tivessem um alto desempenho não considerarem isto atrativo, ou sobre alunos que se empenharam não por causa do modelo ou do aprendizado, mas apenas por causa do ponto na média, o incentivo do ponto pode ter influenciado de alguma forma os resultados, portanto, pode também ter sido um efeito *Hawthorne* da pesquisa.

No quase-experimento piloto na FEA-USP existia a obrigatoriedade da elaboração de um artigo e isto pode ser considerado uma variável *Hawthorne*, já que pode ter influenciado positivamente ou negativamente os resultados desta pesquisa.

De acordo com a Teoria de Orne é importante realizar entrevistas com os participantes de um experimento para buscar captar suas percepções (ADAIR, 1984). Esta pesquisa não entrevistou os participantes e isto pode ser considerado como uma limitação.

O quase-experimento na UNIFIG foi realizado com alunos do curso de graduação em Administração, devido, principalmente, a falta de duas turmas do mesmo semestre no curso de Ciências Contábeis. Apesar dos perfis dos alunos da área de negócios serem mais semelhantes entre seus cursos do que com outras áreas, os perfis entre os dois cursos (Administração e Ciências Contábeis) não são idênticos. Portanto, esta é uma limitação do presente trabalho.

Embora todas estas limitações listadas, pode-se concluir que durante o quase-experimento na UNIFIG, os alunos da turma experimento apresentaram um maior desenvolvimento das habilidades de raciocínio crítico e da atitude curiosidade e que os alunos mais velhos também apresentaram um maior desenvolvimento das habilidades do que os mais novos. Porém, as limitações listadas apresentam variáveis que podem também, além do modelo de Richard Paul, ter sua parcela de explicação nos resultados. Portanto, as limitações relatadas devem ser consideradas nas generalizações e análises dos resultados encontrados, e, em alguns casos, poderão ajudar as futuras pesquisas na tentativa de controlá-las ou de diminuir seus efeitos.

5.3 Sugestões para Futuras Pesquisas

De acordo com os principais resultados, sugere-se que futuras pesquisas investiguem a adoção da abordagem mista do modelo instrucional para desenvolver habilidades de raciocínio crítico utilizado neste trabalho em outros cursos de graduação, pós-graduação, em cursos de graduação em Ciências Contábeis de outras regiões do país e até mesmo em outros países, comparando seus resultados com os encontrados nesta pesquisa.

Como se constatou que o método de instrução teve maior efeito no desenvolvimento de habilidades de raciocínio crítico em alunos com 22 anos ou mais velhos, sugerem-se futuras investigações que aprofundem mais o tema, que enfoquem uma revisão de literatura sobre modelos de raciocínio de acordo com a idade e criem experimentos que analisem a idade como variável principal. Refletindo sobre este resultado, pode-se pensar em ações didáticas que considerem a idade um fator importante no processo de ensino-aprendizagem, como, por exemplo: classes divididas por idade; diferentes métodos de ensino, que levem em

consideração a heterogeneidade da classe em relação à idade; diferentes currículos para um mesmo curso, levando-se em consideração a idade etc.

Houve diferença significativa entre os gêneros para habilidades e atitudes de raciocínio crítico sem considerar o método de instrução. Portanto, sugere-se que futuras pesquisas estudem a literatura específica sobre formas de pensamento e comportamento de acordo com gênero e desenhem experimentos que aprofundem a discussão sobre o tema.

Existem outros testes padronizados para avaliar habilidades e atitudes de raciocínio crítico, como, por exemplo: um teste comercializado pela *Insight Assessment* exclusivo para avaliar atitudes de raciocínio crítico de profissionais da área de negócios. Portanto, sugere-se que futuras pesquisas investiguem e utilizem outros testes para fins de comparação de resultados.

Por fim, sugere-se que futuros experimentos trabalhem com amostras maiores para que seja possível aumentar o poder estatístico dos resultados e observar as limitações listadas neste trabalho.

REFERÊNCIAS

ABRAMI, Phillip C. *et al.* Instructional Interventions Affecting Critical Thinking Skills and Dispositions: A Stage 1 Meta-Analysis. **Review of Educational Research**. [S.l.], v. 78, n. 4, p. 1102-1134, 2008.

ABRÃO, Bernadette Siqueira. **História da Filosofia**. São Paulo: Editora Nova Cultural, 1999. (Os Pensadores).

ADAIR, John G. The Hawthorne Effect: A Reconsideration of the Methodological Artifact. **Journal of Applied Psychology**, v. 69, n. 2, pp. 334-345, 1984.

ADAMS, Barbara L. Nursing education for critical thinking: An integrative review. **Journal of Nursing Education**. [S.l.], v. 38, p. 111-119, 1999.

AICPA. American Institute of Certified Public Accountants. **The CPA Vision Project: 2011 and Beyond**. Disponível em: < <http://www.aicpa.org> >. Acesso em: 11/07/2010.

ALBERT, R. T. *et al.* Relationships among bilingualism, critical thinking ability, and critical thinking disposition. **Journal of Professional Nursing**. [S.l.], v. 18, p. 220-229, 2002.

ALBRECHT, W. S.; SACK, R. J. **Accounting Education: Charting the Course through a Perilous Future**. Sarasota, FL: American Accounting Association, 2000.

ALLEGRETTI, Christine; FREDERICK, J. Norris. A model for thinking critically about ethical issues. **Teaching of Psychology**, v. 22, n. 1, pp. 46-58, 1995.

ALLEN, Mike *et al.* A meta-analysis of the impact of forensics and communication education on critical thinking. **Communication Education**. [S.l.], v. 48, issue 1, 1999.

ANDERSON, Tony; HOWE, Christine; SODEN, Rebecca; HALLIDAY, John; LOW, Jennifer. Peer interaction and the learning of critical thinking skills in further education students. **Instructional Science**, v. 29, n. 1, pp. 1-32, 2001.

ANNIS, Linda F.; ANNIS, David B. The impact of philosophy on student's critical thinking ability. **Contemporary Educational Psychology**, v. 4, Issue 3, pp. 219-226, 1979.

ARTMED. Disponível em: <<http://www.atmed.com.br>>. Acesso em: 03/07/2011.

ATHANASSIOU, Nicholas; MCNETT, Jeanne M.; HARVEY, Carol. Critical Thinking in the Management Classroom: Bloom's Taxonomy as a Learning Tool. **Journal of Management Education**, v.27, n5, p.533-555, 2003.

BALDWIN, B.; INGRAM, R. Rethinking the objectives and content of elementary accounting. **Journal of Accounting Education**. [S.l.], v. 9, p. 1-14, 1991.

BANGERT-DROWNS, R. L.; BANKERT, E. Meta-analysis of effects of explicit instruction for critical thinking. In: ENCONTRO ANUAL DA ASSOCIAÇÃO DE PESQUISA EDUCACIONAL AMERICANA, 1990, Boston. **Anais...** Boston, 1990. Disponível em: <<http://eric.ed.gov>> (Nº ED 328614). Acesso em: 25/06/2010.

BARIL, Charles P. *et al.* Critical Thinking in the Public Accounting Profession: Aptitudes and Attitudes. **Journal of Accounting Education**. [S.l.], v. 16, n. 3, p. 381-406, 1998.

BARTLETT, D. J.; COX, P. D. Measuring change in students' critical thinking ability: implications for health care education. **Journal of Allied Health**. [S.l.], v. 31, p. 64-69, 2002.

BICKMAN, Leonard; ROG, Debra J. **Handbook of applied social research methods**. CA, USA: SAGE, 1998.

BLEDSON, Joseph C. A comparative study of values and critical thinking skills of a group of educational workers. **Journal of Educational Psychology**, v. 46, n. 7, pp. 408-417, 1955.

BONK, Jay Curtis; SMITH, G. Stevenson. Alternative Instructional Strategies for Creative and Critical Thinking in the Accounting Curriculum. **Journal of Accounting Education**. [S.l.], v. 16, n. 2, p. 261-293, 1998.

BRAUN, N. M. Critical thinking in the business curriculum. **Journal of Education for Business**. [S.l.], v. 79, p. 232-236, 2004.

BROOKE, Stephanie L. Using the Case Method to Teach Online Classes: Promoting Socratic Dialogue and Critical Thinking Skills. **International Journal of Teaching and Learning in Higher Education**, v. 18, n2, pp. 142-149, 2006.

CAMPBELL, D. T.; STANLEY, J. C. **Experimental and quasi-experimental designs for research**. Boston: Houghton Mifflin, 1981.

CARRITHERS, David; BEAN, John. Using a Client Memo to Assess Critical Thinking of Finance Majors. **Business Communication Quarterly**, v. 71, n1, pp. 10-26, 2008.

CHAU, J.P.C; CHANG, A.M; LEE, I.F.K.; LEE, D.T.F; WOOTTON, Y. Effects of using videotaped vignettes on enhancing student's critical thinking ability in a baccalaureate nursing programme. **Journal of Advanced Nursing**, v. 36, Issue 1, pp. 112-119, 2001.

COLBERT, Kent R. The effects of debate participation on argumentativeness and verbal aggression. **Communication Education**, v. 42, Issue 3, 1993.

CORNACHIONE JR., Edgard Bruno *et al.* Critical Thinking Skills of Accounting Students: Habilidade de Raciocínio Crítico de Alunos de Contabilidade. **Revista de Educação e Pesquisa em Contabilidade – REPEC**. Brasília, v. 1, n. 2, p. 43-70, maio/ago., 2007.

CRUMBLEY, D. Larry; SMITH, L. Murphy. Using short stories to teach critical thinking and communication skills to tax students. **Accounting Education**, v. 9, Issue 3, 2000.

DETMERING, Robert; JOHNSON, Anna Marie. Focusing on the Thinking, Not the Tools: Incorporating Critical Thinking Into an Information Literacy Module for an Introduction to Business Course. **Journal of Business & Finance Librarianship**, v. 16, Issue 2, 2011.

DURAND, T. Forms of incompetence. In: **INTERNATIONAL CONFERENCE ON COMPETENCE BASED MANAGEMENT**, IV, 1998, Oslo. Anais... Oslo: Norwegian School Management. 1998.

ENNIS, ROBERT H. Critical Thinking Assessment. **Theory into Practice**, College of Education, The Ohio State University, v. 32, n. 3, 1993.

_____. **Critical Thinking**. New Jersey: Prentice Hall, 1996.

_____. **Manual supplement for The Ennis-Weir Critical Thinking Essay Test**, 1998.

ENNIS, R. H.; MILLMAN J.; TOMKO, T. N. **Cornell Critical Thinking Test level X and level Z manual**: The Critical Thinking Co. [S.l.], 2005.

ENNIS, ROBERT H.; WEIR, Eric. **The Ennis-Weir Critical Thinking Essay Test**. Pacific Grove, CA: Midwest Publications, 1985.

FACIONE, Peter A. **Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction “The Delphi Report”**. Millbrae, CA: Santa Clara University, California Academic Press, 1990.

FACIONE, N.; FACIONE, Peter A. **Manual of the California Critical Thinking Disposition Inventory (CCTDI)**. Millbrae, CA: California Academic Press, 2010.

FIELD, Andy. **Descobrendo a estatística usando o SPSS**. Trad. Lorí Viali. 2. ed., Porto Alegre: Artmed, 2009.

FIPECAFI. **Manual de Contabilidade Societária**: Aplicável a todas as Sociedades de acordo com as Normas Internacionais e do CPC. São Paulo: Atlas, 2010.

FRANCIS, Marlene C. *et al.* **Intentional Learning: A Process for Learning to Learn in the Accounting Curriculum**. Sarasota, Florida: Accounting Education Change Commission (AECC) e American Accounting Association (AAA), 1995. Disponível em: <<http://aaahq.org/aecc/intent/index.htm>>. Acesso em: 05/07/2011.

GARVEY, John; BUCKLEY, Patrick. Using Technology to Encourage Critical Thinking and Optimal Decision Making in Risk Management Education. **Risk Management and Insurance Review**, v. 14, Issue 2, pp. 299-309, 2011.

GELDER, Tim van. Teaching Critical Thinking: Some Lessons from Cognitive Science. **College Teaching**, v.53, n1, 2005.

GODOY, Arilda Schmidt. Introdução à Pesquisa Qualitativa e suas Possibilidades. **Revista de Administração de Empresas**. São Paulo, v. 35, n. 2, p. 58, 1995.

GOLD, Jeff; HOLMAN, David; THORPE, Richard. The Role of Argument Analysis and Story Telling in Facilitating Critical Thinking. **Management Learning**, v. 33, n.3, pp. 371-388, 2002.

HAPER, Diane F. Teaching critical thinking for transfer across domains: Disposition, skills, structure training and metacognitive monitoring. **American Psychologist**, v. 53, n. 4, pp. 449-455, 1998.

HARREL, Maralee. Using Argument Diagramming Software to Teach Critical Thinking Skills. Carnegie Mellon University. **5th Conference on Education and Information Systems**, 2007.

HENNINGER, Edward A.; HURLBERT, Janet McNeil. Critical Thinking and Information across the Undergraduate Business Curriculum. **Journal of Business & Finance Librarianship**, v. 2, n. 1, pp. 29-40, 1996.

IUDÍCIBUS, Sérgio de. Ensaio sobre algumas raízes profundas da Contabilidade, em apoio aos Princípios Fundamentais. **Revista de Contabilidade e Organizações - RCO** (FEARP/USP). Ribeirão Preto, v. 1, n. 1, p. 8-15, set./dez. 2007.

IBGE. Instituto Brasileiro de Geografia e Estatística. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 08/09/2010.

JEONG, Allan C. The Sequential Analysis of Group Interaction and Critical Thinking in Online Threaded Discussions. **The American Journal of Distance Education**, v.17, n1, pp. 25-43, 2003.

KASSAI, José Roberto. Lucro gasoso: uma interpretação do velho EVA. In: ENCONTRO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, 28., 2004, Curitiba. **Anais...** Curitiba, 2004.

KERMIS, George; KERMIS, Marguerite. Professional Presence and Soft Skills: A Role for Accounting Education. **Journal of Instructional Pedagogies**. [S.l.], v. 2, marc. 2010.

KIMEL, Paul. A Framework for Incorporating Critical Thinking into Accounting Education. **Journal of Accounting Education**. [S.l.], v. 13, n. 3, p. 299-318, 1995.

KING, Patricia; KITCHENER, Karen Strohm. **Developing reflective judgment**. San Francisco, CA: Jossey-Bass, 1994.

KUHN, Deanna. A Developmental Model of Critical Thinking. **Educational Researcher**, v.28, n2, p.16-46, 1999.

LADYSHEWSKY, Richard. Peer Coaching: a constructivist methodology for enhancing critical thinking in postgraduate business education. **Higher Education Research & Development**, v.25, n1, pp. 67-84, 2006.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Fundamentos de Metodologia Científica**. São Paulo: Atlas, 1985.

LEDERER, Naomi. New form(at): using the Web to teach research and critical thinking skills. **Reference Services Review**, v.28, Issue 2, pp. 130-153, 2000.

LEDERER, J. M. Disposition toward critical thinking among occupational therapy students. **American Journal of Occupational Therapy**. [S.l.], v. 61, p. 519-526, 2007.

LEFRANÇOIS, Guy R. **Teorias da Aprendizagem**. 5. ed. São Paulo: Cengage Learning, 2009.

LIMA, Marcos; KOEHLER, Matthew J.; SPIRO, Rand. Cognitive Flexibility Hypertexts and the development of creative and critical thinking in business education: The Panteon Project. **FACEF Pesquisa**, v.7, n.3, pp. 109-127, 2004.

LIPMAN, Matthew. Critical Thinking: what can it be? **Analytic Teaching**. [S.l.], v. 8, p. 5-12, 1988.

_____. **Thinking in Education**. 2nd ed. New York: Cambridge University Press, 2003.

MACPHERSON, Karen. The Development of Critical Thinking Skills in Undergraduate Supervisory Management Units: efficacy of student peer assessment. **Assessment & Evaluation in Higher Education**, v. 24, Issue 3, 1999.

MARTINS, Gilberto A.; THEÓFILO, Carlos Renato. **Metodologia da Investigação Científica para Ciências Sociais Aplicadas**. 2. ed. São Paulo: Atlas, 2009.

MAUDSLEY, Gillian; STRIVENS, Janet. Promoting professional knowledge, experiential learning and critical thinking for medical students. **Medical Education**, v.34, Issue 7, pp. 535-544, 2000.

MCEWEN, Beryl C. Teaching Critical Thinking Skills in Business Education. **Journal of Education for Business**, v70, n2, p.99-103, 1994.

MCMILLAN, James H. Enhancing college students' critical thinking: A review of studies. **Research in Higher Education**. [S.l.], v. 26, p. 3-29, 1987.

MCPECK, John. **Critical Thinking and Education**. New York: St. Martin's Press, 1981.

MILLER, C.D. *et al.* Learning approaches and motives: Male and female differences and implications for learning assistance programs. **Journal of College Student Development**. [S.l.], v. 31, p. 147-154, 1990.

MINISTÉRIO DA EDUCAÇÃO. Disponível em: <<http://portal.mec.gov.br/index.html>>. Acesso em: 11/09/2010.

MORIN, Edgar. Os sentidos do trabalho. In: WOOD, T. (Ed.). **Gestão empresarial: o fator humano**. São Paulo: Atlas, 2002.

MUIR, Clive. Using Consulting Projects to Teach Critical Thinking Skills in Business Communication. **Business Communication Quarterly**, v. 59, n.4, pp. 77-87, 1996.

NENTL, Nancy; ZIETLOW, Ruth. Using Bloom's Taxonomy to Teach Critical Thinking Skills to Business Students. **College Undergraduate Libraries**. V. 15, Issue: 1, p. 159-172, 2008.

PAISEY, C.; PAISEY, N. An analysis of accounting education research in accounting education: an international journal – 1992–2001. **Accounting Education: An International Journal**. [S.l.], v. 13, issue 1, p. 69-99, 2004.

PARKINSON, Michael G.; EKACHAI, Daradirek. The Socratic method in the introductory PR course: an alternative pedagogy. *Public Relations Review*, v.28, Issue 2, pp. 167-174, 2002.

PAUL, Richard W. **Critical Thinking: How to Prepare Students for a Rapidly Changing World**. Santa Rosa, CA: Foundation for Critical Thinking, 1995.

PAUL, Richard W. **Critical Thinking: Basic Theory & Instructional Structures**. Dillon Beach, California: Foundation for Critical Thinking, 1999.

PETER, G.; WILSON, Carolyn R. **Helping Students Make More Informed Judgments**. Boston: Accounting Education News, 2010.

PITHERS, R.T.; SODEN, Rebecca. Critical Thinking in education: a review. **Educational Research**. [S.l.], v. 42, n. 3, p. 237-249, 2000.

PONTE, Vera Maria Rodrigues *et al.* Análise do grau de cumprimento das práticas de divulgação definidas pelo Pronunciamento Técnico CPC 13 no âmbito das companhias abertas. In: CONGRESSO USP DE CONTROLADORIA E CONTABILIDADE, 10., 2010, São Paulo. **Resumo...** São Paulo: USP, 2010.

QUITADAMO, Ian J. *et al.* Community-Based Inquiry Improves Critical Thinking in General Education Biology. **CBE – Life Sciences Education**. [S.l.], v. 7, 2008.

RAUEN, C.A. Using simulation to teach critical thinking skills: You can't just throw the book at them. **Critical Care Nursing Clinics of North America**, v.13, Issue 1, p.93-103, 2001.

REED, Jennifer H. **Effect of a Model for Critical Thinking on Student Achievement in Primary Source Document Analysis and Interpretation, Argumentative Reasoning, Critical Thinking Dispositions, and History Content in a Community College History Course**. 1998. Tese (Doutorado em Filosofia) – Faculdade de Educação, Universidade do Sul da Florida, Florida, 1998.

RIESENMY, M.R. *et al.* Retention and transfer of children's self-directed critical thinking skills. **Journal of Educational Research**. [S.l.], v. 85, p. 14-25, 1991.

ROEVER, Carol. Using "The Wall Street Journal" to Stimulate Critical Thinking. **Business Communication Quarterly**, v. 61, n.3, pp. 66-70, 1998.

ROBBINS, S. Comportamento organizacional. São Paulo: Prentice Hall, 2002.

RUDD, Rick; BAKER, Matt; HOOVER, Tracy. Undergraduate Agriculture Student Learning Styles and Critical Thinking Abilities: Is There a Relationship? **Journal of Agricultural Education**. V. 41, n.3, pp. 2-12, 2000.

SCHEFFER, B.K.; RUBENFELD, M.G. A consensus statement on critical thinking in nursing. Eastern Michigan University. **Journal of Nursing Education**. 39 (8), 352-9, 2000.

SIEGEL, Harvey. **Educating Reason: Rationality, Critical Thinking and Education**. New York: Routledge, 1988.

SMITH, Gerald F. Beyond Thinking and Decision Making: Teaching Business Students How to Think. **Journal of Management Education**, v.27, n1, p.24-51, 2003.

SOLON, Tom. Improving Critical Thinking in an Introductory Psychology Course. **Michigan Community College Journal: Research & Practice**, v.7, n.2, pp. 73-80, 2001.

SORMUNEN, C.; CHALUPA, M. Critical thinking skills research: Developing evaluation techniques. **Journal of Education for Business**. [S.l.], v. 69, p. 172-177, 1994.

STOIBER, Karen Callan. The Effect of Technical and Reflective Preservice Instruction on Pedagogical Reasoning and Problem Solving. **Journal of Teacher Education**, v. 42, nº. 2, pp. 131-139, 1991.

STONE, D. N.; SHELLEY, M. K. Education for accounting expertise: A field study. **Journal of Accounting Research**. [S.l.], v. 35, p. 35-61, 1997.

TAHAN, Malba. **O Homem que Calculava**. 55. ed. São Paulo: Editora Record, 2001.

TEIXEIRA, Marco Antônio Pereira; GOMES, William Barbosa. Estou me formando... e agora?: Reflexões e perspectivas de jovens formandos universitários. **Rev. Bras. Orientação Profissional**. [S.l.], v. 5, n. 1, p. 47-62, 2004.

TERENZINI, Patrick T.; SPRINGER, Leonard; PASCARELLA, Ernest T.; NORA, Amaury. Influences affecting the development of student's critical thinking skills. **Research in Higher Education**, v.36, n1, pp. 23-39, 1995.

THAYER- BACON, Barbara J. Transforming Critical Thinking – Thinking Constructively. New York: Teachers College, Columbia University, 2000.

TIWARI, Agnes; LAI, Patrick; SO, Mike; YUEN, Kwan. A comparison of the effects of problem-based learning and lecturing on the development of student's critical thinking. **Medical Education**, v. 40, Issue 6, pp. 547-554, 2006.

TURNER, Yvonne. Students from mainland China and critical thinking in postgraduate Business and Management degrees: teasing out tensions of culture, style and substance. **International Journal of Management Education**, v.5, n.1, pp. 3-11, 2006.

TSUI, Lisa. Courses and Instruction Affecting Critical Thinking. **Research in Higher Education**, v.40, n2, pp. 185-200, 1999.

UNRAU, Norman J. The Effects of Explicit Instruction on Critical Reading and Argumentative Writing: The TASK of Reading and Writing. Paper presented at the Annual Meeting of the **American Educational Research Association**, 72nd, Chicago, IL, 1991.

VOCKELL, Edward; RIVERS, Robert. Computerized Science Simulations Stimulus to Generalized Problem Solving Capabilities. **Science Education Research**. Paper presented at the Annual Convention of the American Education Research Association, 68th, New Orleans, LA, 1984.

VYGOTSKY, Lev Semenovich. Thought and Language. Cambridge, MA, MIT Press, 1962.

WALSH, Catherine; HARDY, Robert C. Dispositional Differences in Critical Thinking Related to Gender and Academic Major. **Journal of Nursing Education**, v.38, n.4, pp.149-155, 1999.

WILSON, Richard M. S. Accounting education research: a retrospective over ten years with some pointers to the future. **Accounting Education: An International Journal**. [S.l.], v. 11, issue 4, p. 295-310, 2002.

ZABIT, Mohd Nazir Md. Problem-Based Learning on Student's Critical Thinking Skills in Teaching Business Education in Malaysia: A Literature Review. **American Journal of Business Education**, v. 3, n6, 2010.

BIBLIOGRAFIA

BALL, Deborah Loewenberg; FORZANI, Francesca M. What Makes Education Research “Educational”? Wallace Foundation Distinguished Lecture. **Educational Researcher**. [S.l.], v. 36, n. 9, p. 529-540, 2007.

CAPRA, Fritjof. **O ponto de mutação: A Ciência, a Sociedade e a Cultura Emergente**. 24. ed. São Paulo: Editora Cultrix, 2003.

CORNACHIONE JR., Edgard Bruno. **Tecnologia da educação e cursos de ciências contábeis: Modelos Colaborativos Virtuais**. 2004. 383f. Tese (Livre-Docência) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2004.

EIDE, Bárbara J. Integrating Learning Strategies in Accounting Courses. **Advances in Accounting Education**. [S.l.], v. 2, p. 37-55, 1989.

FREIRE, Paulo. **Pedagogia do oprimido**. Rio de Janeiro: Editora Paz e Terra, 2005.

LOMBARDI, Marily. M. Authentic Learning for the 21st Century: An Overview. **Educause Learning Initiative**, 2007. Disponível em:
<<http://net.educause.edu/ir/library/pdf/ELI3009.pdf>>. Acesso em: 23/02/2010.

MARTINS, Gilberto de Andrade. **Estatística Geral e Aplicada**. 3. ed. São Paulo: Atlas, 2001.

MARZANO, Robert J. *et al.* **What Works in Classroom Instruction**. Aurora, Colorado: Mid-continent Research for Education and Learning, 2000.

MINGERS, J. What is to be critical? Teaching a critical approach to management undergraduates. **Management Learning**. [S.l.], v. 31, p. 219-237, 2000.

PASSOS, Ivan Carlin. **A interdisciplinaridade no ensino e na pesquisa contábil: um estudo do município de São Paulo**. 2004, 165f. Dissertação (Mestrado) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2004

REEVES, Thomas C. *et al.* Authentic activities and online learning. **HERDSA**, 2002. p. 562-567.

SHIN, Sharon. Ready or Not, Here Comes IFRS. **BizEd**, July/August, 2009.

SOARES, Mara Alves. **Aplicação do método de ensino Problem Based Learning (PBL) no curso de Ciências Contábeis: um estudo empírico**. 2008. 214f. Dissertação (Mestrado) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, Ribeirão Preto, 2008.

SPRINGER, Carol W.; BORTHICK, A. Faye. Improving performance in accounting: Evidence for insisting on cognitive conflicts. **Issues in Accounting Education**. [S.l.], v. 22, p. 1-19, 2007.

TOMPKINS, G. E. Review of the Ennis-Weir Critical Thinking Essay Test. In: CONOLEY, Jane Close; KRAMER, Jack J. **Tenth Mental Measurements Yearbook**. Lincoln, Nebraska: Buross Institute of Mental Measurement, University of Nebraska Press, 1989.

UMAR, Irfan Naufal; MASWAN, Sajap. **A Guided Inquiry Learning Approach in a Web Environment: Theory and Application**. School of Educational Studies, University Sains Malaysia, 2004.

WERNER, P. H. The Ennis-Weir Critical Thinking Essay Test: An instrument for testing and teaching. **Journal of Reading**. [S.l.], v. 34, p. 494-495, 1991.

APÊNDICES

APÊNDICE A - Apostila sobre Raciocínio Crítico

APÊNDICE B - Modelo de Consentimento Explícito

APÊNDICE C - Pesquisa Demográfica

APÊNDICE D - Exemplo do Modelo de Controle dos Participantes por instrumento e data

APÊNDICE A

Apostila sobre Raciocínio Crítico

RACIOCÍNIO CRÍTICO E CONTABILIDADE

Esta disciplina da área Contábil/Financeira coloca à sua disposição esta apostila¹⁹ com um modelo para desenvolvimento do raciocínio crítico. O modelo foi desenvolvido por Richard Paul (1995) e pelo Centro de Raciocínio Crítico e Crítica Moral da Universidade Estadual de Sonoma, Califórnia, EUA.

O modelo possui elementos e padrões intelectuais sobre raciocínio crítico, os quais são detalhados e explicados nesta apostila. Conforme você aprende os elementos e padrões de raciocínio, poderá utilizá-los em cada aspecto desta disciplina, bem como em outras situações acadêmicas e do cotidiano que requeiram bom raciocínio. Este modelo pode melhorar suas habilidades e atitudes de pensar criticamente durante a leitura e escrita de artigos, livros, redações, exames e debates. Esta apostila contém:

- 1) definição de raciocínio crítico;**
- 2) uma figura que mostra os elementos de raciocínio e padrões intelectuais universais;**
- 3) definições dos elementos de raciocínio;**
- 4) a sessão “Ajudando os Alunos a Mensurar Seus Pensamentos”, que dá dicas para você guiar seu próprio raciocínio e avaliar o raciocínio de outros;**
- 5) explanação dos padrões intelectuais por meio de questões que você pode fazer a si mesmo sobre seu próprio pensamento ou sobre o de outros;**
- 6) uma figura que mostra o relacionamento entre os elementos, padrões, atitudes e habilidades de raciocínio crítico;**
- 7) dois guias para identificar os principais elementos de raciocínio de um artigo ou de um livro-texto;**
- 8) algumas falácias comuns de raciocínio;**
- 9) bibliografia.**

¹⁹ Fonte: REED, Jennifer H. **Effect of a Model for Critical Thinking on Student Achievement in Primary Source Document Analysis and Interpretation, Argumentative Reasoning, Critical Thinking Dispositions, and History Content in a Community College History Course**. 1998. Tese (Doutorado em Filosofia) – Faculdade de Educação, Universidade do Sul da Florida, Florida, 1998.

COMO UTILIZAR ESTA APOSTILA:

Consulte a Figura 1 sobre os elementos e padrões de raciocínio sempre que você for avaliar o raciocínio de outros (por exemplo: na leitura de artigos, livros ou mesmo em debates) ou seu próprio raciocínio (por exemplo: na confecção de tarefas, exercícios, redações ou relatórios).

Utilize as explicações dos elementos e padrões sempre que você precisar ter certeza de que compreendeu bem os vários aspectos de raciocínio. Conforme os elementos e padrões se tornem familiares a você, comece a examinar como suas atitudes estão em relação às atitudes intelectuais ideais e cheque se está desenvolvendo as habilidades necessárias para se tornar um pensador crítico.

Quando estiver lendo um artigo ou um livro-texto, você pode utilizar os guias disponíveis para identificar os principais elementos de raciocínio. O informativo sobre as falácias comuns de raciocínio servirá para que você procure identificá-las nos argumentos dos outros e nos seus próprios.

Procure desenvolver suas habilidades de raciocínio crítico e utilizá-las sempre que possível em outras disciplinas do seu curso, assim como no seu ambiente de trabalho e na vida pessoal.

1 DEFINIÇÃO DE RACIOCÍNIO CRÍTICO

Raciocínio crítico é pensamento que facilita julgamento, pois identifica, analisa e julga os elementos de um raciocínio. Os elementos de um raciocínio são: objetivo, problema, informações, interpretações, conceitos, suposições, implicações e ponto de vista.

2 ELEMENTOS DE RACIOCÍNIO & PADRÕES INTELLECTUAIS UNIVERSAIS¹⁷

UM PENSADOR CRÍTICO

Considera os Elementos de Raciocínio

COM SENSITIVIDADE AOS PADRÕES INTELLECTUAIS UNIVERSAIS

¹⁷ Fonte: PAUL, Richard W. **Critical Thinking**: How to Prepare Students for a Rapidly Changing World. Santa Rosa, CA: Foundation for Critical Thinking, 1995.

3 DEFINIÇÕES: ELEMENTOS DE RACIOCÍNIO²⁰

PONTOS DE VISTA: pensamento humano é relacional e seletivo. É impossível entender qualquer pessoa, evento ou fenômeno de todos os pontos de vista simultaneamente. Raciocínio crítico requer que este fato seja levado em consideração quando se analisa ou avalia algo.

PROPÓSITO DO PENSAMENTO: a intenção, objetivo ou fim de um documento, discussão, atividade, relacionamento etc.

QUESTÃO SOBRE O ASSUNTO: um problema, situação ou pessoa que é difícil de compreender, lidar ou resolver. Problemas e questões podem ser divididos em muitos tipos, incluindo monológicos (problemas que podem ser resolvidos por meio do raciocínio exclusivamente dentro de uma disciplina ou ponto de vista) e multilógicos (problemas que podem ser analisados e abordados por mais de uma disciplina e, geralmente, conflitando pontos de vista).

INFORMAÇÃO: os dados (fatos, figuras ou informações) nos quais um julgamento ou conclusão pode estar baseado ou por qual uma prova ou probabilidade pode ser estabelecida. Pensadores críticos distinguem a evidência ou os dados primários sobre os quais eles baseiam suas interpretações ou conclusões das inferências ou suposições que conectaram os dados às conclusões. Pensadores não críticos tratam suas conclusões como alguma coisa gerada por suas experiências, como alguma coisa que eles diretamente observaram no mundo. Como resultado, eles encontram dificuldade em enxergar a razão pela qual qualquer pessoa poderá discordar com suas conclusões.

INTERPRETAÇÃO & INFERÊNCIA: uma inferência é um passo da mente, um ato intelectual pelo qual alguém conclui que alguma coisa é ou parece ser claramente algo. Se você vem em minha direção com uma faca na mão, eu provavelmente irei inferir que você pretende me ferir. Inferências podem ser fortes ou fracas, justificadas ou injustificadas. Inferências são baseadas em suposições.

CONCEITOS: ideias ou pensamentos, especialmente uma ideia generalizada de uma coisa ou de uma classe de coisas. Humanos pensam dentro de conceitos ou ideias. Nós nunca teremos controle sobre nossos pensamentos a não ser que nós aprendamos como comandar nossos conceitos ou ideias. Então, nós devemos aprender como identificar os conceitos ou

²⁰ Fonte: PAUL, Richard W. **Critical Thinking**: How to Prepare Students for a Rapidly Changing World. Santa Rosa, CA: Foundation for Critical Thinking, 1995.

ideias que nós estamos usando, contrastá-los com conceitos ou ideias alternativas e clarificar o que nós incluímos ou excluímos por meio deles. Por exemplo, a maioria das pessoas diz que acredita fortemente na democracia, mas poucas podem clarificar com exemplos o que esta palavra realmente significa.

SUPOSIÇÕES: afirmações aceitas ou supostas como verdade sem prova ou demonstração; uma premissa ou crença não confirmada. Todo pensamento e experiência humana são baseados em suposições. Nosso pensamento deve iniciar com alguma coisa que tomamos como verdade em um contexto particular. Muito do que está errado no pensamento humano pode ser encontrado nas suposições não críticas ou não examinadas que o baseiam. Por exemplo, nós geralmente experimentamos o mundo de tal maneira que assumimos que o que nós estamos observando é realmente daquele jeito; como esse pensamento, nós estamos vendo o mundo sem o filtro de um ponto de vista diferente. Há uma conexão, portanto, entre pensamento quando estamos conscientes das nossas suposições e quando estamos intelectualmente limitados.

IMPLICAÇÕES & CONSEQUÊNCIAS: uma alegação ou verdade que é seguida por outras alegações ou verdades. Uma das mais importantes habilidades de raciocínio crítico é a de distinguir entre o que é na verdade implicado por uma afirmação ou situação do que pode ser inferido descuidadamente pelas pessoas. Dizer que um ato é assassinato, por exemplo, implica dizer que é intencional e não justificado.

4 AJUDANDO OS ALUNOS A MENSURAR SEUS PENSAMENTOS²¹

1) Todo raciocínio tem um PROPÓSITO.

- Leve o tempo necessário para estabelecer seu propósito claramente.
- Distinga seu propósito de outros relacionados ao assunto.
- Durante a análise, cheque-o periodicamente para ter certeza de que você ainda está no foco.
- Procure escolher propósitos que sejam significantes e realísticos.

2) Todo raciocínio é uma tentativa de IMAGINAR algo, de formular alguma QUESTÃO, resolver algum PROBLEMA.

²¹ Fonte: PAUL, Richard W. **Critical Thinking: Basic Theory & Instructional Structures**. Dillon Beach, California: Foundation for Critical Thinking, 1999.

- Leve o tempo necessário para estabelecer clara e precisamente a questão sobre o assunto.
- Expresse a questão de diversas maneiras para clarificar seu significado e objetivo.
- Quebre a questão em subquestões.
- Identifique se a questão tem uma resposta correta, se é uma questão de mera opinião ou se requer raciocínio de mais de um ponto de vista.

3) Todo raciocínio é baseado em SUPOSIÇÕES.

- Identifique claramente suas suposições e determine se elas são justificáveis.
- Considere como suas suposições estão modelando seu ponto de vista.

4) Todo raciocínio é realizado de algum PONTO DE VISTA.

- Identifique seu ponto de vista.
- Procure outros pontos de vista e identifique suas forças, bem como suas fraquezas.
- Lute para ser justo no raciocínio em avaliar todos os pontos de vista.

5) Todo raciocínio é baseado em DADOS, INFORMAÇÃO e EVIDÊNCIA.

- Restrinja suas alegações para aquelas suportadas pelos dados que você possui.
- Procure toda informação que opõem sua posição, bem como as que a suportam.
- Tenha certeza de que toda informação utilizada é clara, precisa e relevante para o assunto em questão.
- Tenha certeza de que você levantou as informações suficientes.

6) Todo raciocínio é expresso e modelado por meio de CONCEITOS e IDEIAS.

- Identifique conceitos-chave e explique-os claramente.
- Considere conceitos alternativos ou definições alternativas para os conceitos.
- Tenha certeza de que você está utilizando os conceitos com cuidado e precisão.

7) Todo raciocínio contém INFERÊNCIAS ou INTERPRETAÇÕES pelas quais nós desenhamos CONCLUSÕES e damos significado aos dados.

- Infira apenas sobre o que a evidência implica.

- Cheque inferências por suas consistências com cada uma das outras.
- Identifique suposições que o guiam às suas inferências.

8) Todo raciocínio leva a algum lugar ou tem IMPLICAÇÕES e CONSEQUÊNCIAS.

- Trace as implicações e consequências que se seguem de seu raciocínio.
- Procure por implicações negativas e positivas.
- Considere todas as consequências possíveis.

5 PADRÕES INTELLECTUAIS UNIVERSAIS²²

Neste tópico são abordadas questões que, quando aplicadas, auxiliam a utilização dos padrões intelectuais universais.

Os padrões intelectuais devem ser aplicados nos momentos em que se quer checar à qualidade do raciocínio de alguém sobre um problema, assunto ou situação. Pensar criticamente envolve ter comando destes padrões. Para ajudar os alunos, os professores deveriam expor questões que os instigassem a investigar, questões que os fizessem pensar por um determinado tempo, Estas questões, consistentemente utilizadas pelos professores na sala de aula, são internalizadas pelos alunos, os quais sentem necessidade de reproduzi-las a eles próprios.

O principal objetivo deste tópico é que estas questões fiquem impressas em seus pensamentos, guiando-os a um melhor raciocínio. Abaixo os padrões e as questões norteadoras.

CLAREZA: Você poderia elaborar mais determinado ponto de vista? Você poderia expressar aquele ponto de outra maneira? Você poderia ilustrar esse ponto de vista? Você poderia dar um exemplo? Clareza é um portão padrão. Se uma afirmação não é clara, nós não podemos determinar se ela é precisa ou relevante. Em fato, nós não podemos dizer algo sobre isto porque ainda não sabemos o que isto está dizendo. Por exemplo, a questão “O que pode ser feito sobre o sistema educacional no Brasil?” não é clara. Para adequadamente endereçar a questão, nós necessitaríamos compreender mais claramente qual o problema considerado pela pessoa que formulou a questão. Uma questão mais clara seria: “O que os educadores podem

²² Fonte: PAUL, Richard W. **Critical Thinking**: Basic Theory & Instructional Structures. Dillon Beach, California: Foundation for Critical Thinking, 1999.

fazer para assegurar que seus alunos aprenderão habilidades que os ajudarão a atuar com sucesso em suas carreiras profissionais e nas decisões da vida cotidiana?”.

ACURÁCIA: Aquilo realmente é verdade? Como poderíamos checar aquilo? Como poderíamos descobrir se aquilo é verdade? Uma afirmação pode ser clara, mas não correta, como “a maioria dos cachorros pesa mais do que 135 kg”.

PRECISÃO: Você poderia dar mais detalhes? Você poderia ser mais específico? Uma afirmação pode ser ambas, clara e correta, mas não precisa, como em “Pedro está acima do peso” (nós não sabemos o quanto Pedro está acima do peso, um quilo ou 100 quilos).

RELEVÂNCIA: Como aquilo está conectado com a questão? Como aquilo confirma o assunto? Uma afirmação pode ser clara, correta e precisa, mas não relevante ao assunto em questão. Por exemplo, alunos geralmente pensam que a quantidade de esforço que eles colocaram em um curso deveria ser usada em sua nota. Geralmente, entretanto, “esforço” não mede a qualidade do aprendizado do aluno, e quando isto acontece, esforço é irrelevante para a nota apropriada.

PROFUNDIDADE: Como suas respostas tratam as complexidades inerentes à questão? Como você está levando em conta os problemas da questão? Está lidando com os fatores mais significativos? Uma afirmação pode ser clara, correta, precisa e relevante, mas superficial. Por exemplo, a frase “Apenas diga Não”, que geralmente é utilizada para desencorajar crianças e adolescentes a usarem drogas, é clara, correta, precisa e relevante. Porém, ela não é profunda, porque trata de forma superficial um assunto extremamente complexo.

ALCANCE: Nós precisamos considerar outro ponto de vista? Há outra maneira de olhar para esta questão? Como isso pareceria de um ponto de vista conservador? Como isso pareceria do ponto de vista de...? Uma linha de raciocínio pode ser clara, correta, precisa, relevante e profunda, mas com falta de alcance (como em um argumento tanto do ponto de vista conservador ou liberal que se aprofunda no assunto, mas reconhece apenas os discernimentos de um lado da questão).

LÓGICA: Isso realmente faz sentido? Aquilo segue o que você disse? Como é que aquilo segue? Mas antes você implicou aquilo e agora você está dizendo isto, eu não vejo como ambos podem ser verdadeiros. Quando pensamos, trazemos uma variedade de pensamentos juntos em alguma ordem. Quando as combinações não se suportam mutuamente, são contraditórias em certo sentido ou “não fazem sentido”, a combinação “não é lógica”.

6 RELACIONAMENTO ENTRE OS ELEMENTOS, PADRÕES, ATITUDES E HABILIDADES DE RACIOCÍNIO CRÍTICO

Fonte: PAUL, Richard W. **Critical Thinking**: Basic Theory & Instructional Structures. Dillon Beach, California: Foundation for Critical Thinking, 1999.

7 GUIAS PARA IDENTIFICAR OS PRINCIPAIS ELEMENTOS DE UM ARTIGO OU DE UM LIVRO²³

A LÓGICA DE UM ARTIGO

- 1) O principal propósito deste artigo é _____
- 2) A questão-chave é _____
- 3) A informação mais importante no artigo é _____
- 4) As principais conclusões do artigo são _____
- 5) O(s) conceito(s)-chave que nós precisamos para entender o artigo é (são) _____
- 6) A(s) principal(is) suposição(ões) suportando o pensamento do autor é (são) _____
- 7) Se as pessoas adotarem esta linha de raciocínio seriamente as implicações são _____

Se as pessoas falharem em adotar esta linha de raciocínio seriamente as implicações são _____
- 8) O(s) principal(is) ponto(s) de vista apresentado(s) no artigo é (são) _____

²³ Fonte: PAUL, Richard W. **Critical Thinking**: Basic Theory & Instructional Structures. Dillon Beach, California: Foundation for Critical Thinking, 1999.

A LÓGICA DE UM LIVRO

- 1) O principal propósito do livro é _____
- 2) A questão-chave que o autor está tentando responder é _____
- 3) A informação mais importante neste livro é _____
- 4) As principais conclusões do livro são _____
- 5) O(s) conceito(s)-chave que nós precisamos entender para compreender o raciocínio do autor é (são) _____
- 6) O(s) significado(s) deste(s) conceito(s) é (são) _____
- 7) A(s) principal(is) suposição(ões) suportando o pensamento do autor é (são) _____
- 8) Se as pessoas adotarem as ideias deste livro seriamente as implicações são _____
- 9) Se as pessoas falharem em adotar as idéias deste livro as implicações são _____
- 10) O(s) principal(is) ponto(s) de vista apresentado(s) neste livro é (são) _____

8 ALGUMAS FALÁCIAS COMUNS DE RACIOCÍNIO²⁴

MUDANDO O ASSUNTO: Em discussões geralmente uma pessoa pode argumentar por algo diferente do assunto em questão. De certa forma, às vezes as pessoas fazem isso tentando desacreditar a posição contrária atacando a pessoa que a está oferecendo (por exemplo: atacando sua inteligência, sua integridade e seu histórico familiar). Isso é chamado de argumento de ataque pessoal (*argumentum ad hominem*). Outra maneira de as pessoas

²⁴ Adaptado de ENNIS, Robert H. **Critical Thinking**. New Jersey: Prentice Hall, 1996.

mudarem um assunto é argumentando sobre uma posição que é suportável, mas não completamente relevante à questão em si. Por exemplo: uma pessoa que apoia a utilização de impostos sobre as receitas para uma nova estação policial pode dizer que as pessoas que são opostas não se importam com o crescimento das taxas de criminalidade (argumento de ataque pessoal) ou ela pode dar razões para a necessidade de mais oficiais de polícia, desviando a atenção do fato que eles não podem prover razões suficientes para a construção de uma nova estação policial. Em ambas as situações a pessoa que muda o assunto está evitando o real assunto.

ARGUMENTO CIRCULAR: Este argumento não realiza progresso algum, pois começa explicitamente assumindo o que ele quer provar ou concluir (isto se torna cíclico). Nenhuma razão ou evidência é fornecida para apoiar este argumento. O propositor assume que todos partilham das suas suposições e crenças básicas quando na verdade eles não o fazem. Fique atento quando alguém diz “todo mundo sabe”, “nós todos concordamos” ou “isto é óbvio”, pois muito frequentemente estas frases assumem o que eles precisam provar (por exemplo: “Todo mundo sabe que nós precisamos de uma nova estação policial. Precisamos angariar os fundos e começar o mais rápido possível”).

EQUÍVOCO OU LINGUAGEM AMBÍGUA: Às vezes um argumento pode parecer bom, mas não o é, pois depende da mudança de significado de um termo-chave durante o desenvolvimento da argumentação. Equivocar é mudar o significado da palavra no âmago do argumento para explorar esta mudança, chegando a uma conclusão. Linguagem ambígua permite ao autor definir o conceito de uma forma que preencha o seu propósito, mas que pode não ser aceita pela maioria das pessoas razoáveis (com bom senso). O autor define claramente a importância de conceitos como “liberdade”, “viés” ou “dúvida razoável”? A definição proposta é correta, racional e normalmente aceita? Faz uso de um termo consistentemente ao longo de um argumento?

SUPOSIÇÕES FALSAS OU DÚBIAS: Uma posição não é mais forte que suas suposições. Um exemplo é o argumento de uma pessoa de fraco caráter. Nesta falácia, suposições falsas são feitas sobre uma posição que você se opõe e, então, esta e suas fontes são apresentadas e recusadas. Por exemplo, os comerciantes de madeira podem alegar que os ambientalistas querem acabar com todo uso econômico das reservas naturais e, então, refutar aquela posição em editoriais e propagandas. Se esta não é realmente a posição dos ambientalistas, os comerciantes estão utilizando o argumento de uma pessoa de fraco caráter.

RAZÕES OU EVIDÊNCIAS INSUFICIENTES: Há muitas maneiras pelas quais razões e evidências oferecidas podem ser insuficientes.

- **Outras explicações ou possibilidades plausíveis:** algumas vezes uma hipótese explica a evidência, mas outras hipóteses ou possibilidades alternativas poderiam explicar a evidência tão bem quanto. Sempre olhe por causas rivais ou outras interpretações para a evidência. Analisando-se os assuntos educacionais, raramente existe apenas uma simples causa de um problema. Múltiplas causas e complexas explicações são mais prováveis.
- **Super generalização ou generalização descuidada:** geralmente pessoas chegam a conclusões gerais que aplicam para as coisas ou populações sobre as quais a evidência não é representativa ou é insuficiente. Afirmações como “escolas públicas estão decaindo” (todas elas? em tudo?) ou “políticos eleitos são desonestos” são supergeneralizações. Atente-se para termos absolutos, tais como: “Todos”, “Nenhum”, “Nunca” e “Sempre”.
- **Supersimplificação:** um tipo padrão de supersimplificação é a suposição de que há apenas duas alternativas (certo/errado, bom/mau) quando de fato há outras. Por exemplo: “Nós vamos construir uma nova estação policial nesta cidade ou vamos abandoná-la aos criminosos, gangues e traficantes?” Outro tipo de supersimplificação é adotar respostas simples para problemas complexos em vez de pensar de maneira profunda e abrangente.
- **Linguagem emocional sem suficiente substância:** geralmente as pessoas apelam para *slogans* ou utilizam palavras (incluindo “generalidades atrativas” [termo formal oferecido sem suporte] ou “chamar de um nome” [uso de um termo desrespeitoso sem suporte]) que despertam em muitos de nós reações emocionais, mas não oferecem razões substanciais. Chamar alguém de comunista ou racista porque ele não concorda com determinada posição é “chamar de um nome”. Chamá-lo de patriota porque ele apoia sua posição pode ser uma “generalidade atrativa”.
- **Analogias incorretas:** um argumento analógico ocorre mostrando que duas coisas são comuns em alguns aspectos e, então, conclui-se que elas também são similares em outro aspecto. Uma analogia pode ajudar leitores a entender ideias abstratas ou difíceis. Alguns argumentos analógicos são fortes, alguns

são incorretos. Suas potenciais fraquezas estão nos aspectos que as duas coisas são diferentes e que há sempre diferenças. Você precisa considerar se as similaridades são fortes o suficiente para garantir as conclusões. Por exemplo, alguém pode argumentar que todas as nações do mundo poderiam se unir, pois são similares aos treze Estados que se juntaram para formar os Estados Unidos da América. Um desafio à analogia poderia ser o argumento de que os casos são diferentes, já que os Estados que se uniram para formar os EUA o fizeram para se proteger de inimigos externos, causa esta que o mundo não poderia utilizar.

- **Negligência de um ponto de vista:** geralmente uma conclusão depende do ponto de vista do argumentador. Algumas vezes ajuda mostrar aquilo de outro ponto de vista, já que a evidência é insuficiente. A negligência de um ponto de vista pode ocorrer deliberadamente, por exemplo, se um argumentador apresenta apenas os elementos negativos (ou positivos) de uma instituição ou sistema, ele está provendo uma estranha ou incorreta figura da realidade.
- **Falha para seguir:** algumas vezes a conexão alegada entre razões ou evidências e conclusões está apenas lá. A conclusão falha em seguir seu suporte ou o suporte oferecido é irrelevante para a conclusão. Por exemplo: “Ele possui seu próprio negócio. Ele deve ser rico e não se importar com que acontece conosco, pobres e lutadores trabalhadores.” Estas conclusões não necessariamente seguem o fato de que ele possui um negócio.
- **Falsa causa:** argumentos assumem que porque um evento segue o outro, o primeiro evento causou o outro. Propagandas normalmente utilizam este tipo de falso raciocínio. Por exemplo: “se você comprar o carro certo sua vida social irá melhorar”. Aceitar uma causa falsa também é frequentemente utilizado para evitar pensamentos mais profundos sobre assuntos complexos. Alguém aceita uma causa que “faz sentido” ou se encaixa em suas ideias preconcebidas, evitando, portanto, o esforço de examinar outros pontos de vista ou outras possíveis causas.

ESTATÍSTICAS: Estatísticas provêm dados ou informações sobre grupos de coisas ou pessoas, mas elas precisam ser interpretadas. As estatísticas devem ser checadas de forma cuidadosa para verificar se elas realmente suportam as afirmações que foram feitas. Pergunte

a você mesmo se as estatísticas estão fundamentadas em uma amostra suficiente (suficiente pessoas, tempo, objetos etc., que permita generalizar) e cuidadosos controles (outras possíveis causas controladas). Cheque se as estatísticas providas não deixaram de fora informação relevante. A maneira que as questões foram organizadas em um questionário também pode afetar os resultados.

FONTE FRACA: Toda informação, argumento etc. tem fontes. Geralmente a credibilidade (o grau pelo qual uma fonte merece ser creditada) de uma fonte é suspeita. Para determinar se uma fonte é confiável, considere as seguintes qualificações: (1) o grau de especialidade no assunto; (2) o possível conflito de interesse; (3) se está de acordo com outras fontes; (4) o risco para reputação; (5) o uso de procedimentos estabelecidos; (6) a habilidade de prover razões; (7) os hábitos cuidadosos. “Apelo a uma autoridade” é um movimento em um argumento que em efeito diz que, por causa da autoridade, alguma coisa é, logo ela é. Geralmente apelar para um especialista ou autoridade é uma boa maneira de suportar uma posição (tal como citar fontes em um trabalho de pesquisa ou basear-se na testemunha de um especialista em uma corte). Em geral é considerado racional aceitar o julgamento de um especialista em determinado assunto, porque é um especialista e porque o que suporta seu julgamento está disponível para avaliação. Note que usar um produto (testemunho pessoal – uma tentativa de suportar uma posição fundamentando-se em uma testemunha a seu favor) ou participar em um testemunho de um evento não torna uma fonte confiável. É preciso considerar os “fatos” contraditórios tipicamente providos pela testemunha de um crime.

9 BIBLIOGRAFIA

ABRAMI, Phillip C. *et al.* Instructional Interventions Affecting Critical Thinking Skills and Dispositions: A Stage 1 Meta-Analysis. **Review of Educational Research**. [S.l.], v. 78, n. 4, p. 1102-1134, 2008.

BRAUN, N. M. Critical thinking in the business curriculum. **Journal of Education for Business**. [S.l.], v. 79, p. 232-236, 2004.

ENNIS, ROBERT H. Critical Thinking Assessment. **Theory into Practice**, College of Education, The Ohio State University, v. 32, n. 3, 1993.

_____. **Critical Thinking**. New Jersey: Prentice Hall, 1996.

FACIONE, Peter A. **Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction “The Delphi Report”**. Millbrae, CA: Santa Clara University, California Academic Press, 1990.

FREIRE, Paulo. **Pedagogia do oprimido**. Rio de Janeiro: Editora Paz e Terra, 2005.

LIPMAN, Matthew. Critical Thinking: what can it be? **Analytic Teaching**. [S.l.], v. 8, p. 5-12, 1988.

_____. **Thinking in Education**. 2nd ed. New York: Cambridge University Press, 2003.

MCPECK, John. **Critical Thinking and Education**. New York: St. Martin’s Press, 1981.

PAUL, Richard W. **Critical Thinking: How to Prepare Students for a Rapidly Changing World**. Santa Rosa, CA: Foundation for Critical Thinking, 1995.

PETER, G.; WILSON, Carolyn R. **Helping Students Make More Informed Judgments**. Boston: Accounting Education News, 2010.

REED, Jennifer H. **Effect of a Model for Critical Thinking on Student Achievement in Primary Source Document Analysis and Interpretation, Argumentative Reasoning, Critical Thinking Dispositions, and History Content in a Community College History Course**. 1998. Tese (Doutorado em Filosofia) – Faculdade de Educação, Universidade do Sul da Florida, Florida, 1998.

SIEGEL, Harvey. **Educating Reason: Rationality, Critical Thinking and Education**. New York: Routledge, 1988.

APÊNDICE B

Modelo de Consentimento Explícito

**Habilidades e Atitudes de Raciocínio Crítico
de alunos de Graduação em Ciências Contábeis**

Doutorando: Prof. Ms. Ivan Carlin Passos
Orientador: Prof. Dr. Edgard B. Cornachione Jr. (FEA/USP)

Consentimento Explícito

Esta pesquisa faz parte da tese de doutorado do autor identificado acima, a qual está desenhada para identificar habilidades e atitudes de raciocínio crítico de alunos do curso de Ciências Contábeis da FEA-USP e Administração da UNIFIG.

Participantes voluntários irão completar dois instrumentos de avaliação em sala de aula para avaliar suas habilidades e atitudes de raciocínio crítico. O primeiro instrumento é um teste que mensura as atitudes de Raciocínio Crítico denominado CCTDI (*California Critical Thinking Disposition Inventory – Insight Assessment – California Academic Press*) e levará 20 minutos para resposta. O segundo instrumento é o EWCTE (*Ennis-Weir Critical Thinking Essay – UIUC*), que consiste em escrever uma resposta para uma carta, devendo consumir no máximo 40 minutos (10 minutos para leitura e 30 minutos para a resposta). Após os testes, você receberá retorno sobre seu desempenho e informações adicionais sobre raciocínio crítico que irá lhe ajudar a aprimorar a forma usada por você para leitura e escrita, visando melhorar seu desempenho acadêmico.

Seu nome será registrado nos instrumentos, mas em hipótese alguma será divulgado, de forma que não há risco de dados individuais serem identificados como sendo seus. Os resultados serão tratados por um código numérico atribuído ao seu teste, de modo que apenas você poderá guardar e utilizar para identificar seus próprios resultados.

Caso haja quaisquer questões ou preocupações sobre esta pesquisa, você poderá contatar Ivan Carlin Passos (ivanep@usp.br / ivan.passos@fig.br) ou Edgard Bruno Cornachione Júnior (edgardbc@usp.br).

Você receberá uma cópia deste formulário de consentimento.

Data

Assinatura

Nome por extenso

APÊNDICE C

Pesquisa Demográfica

Pesquisa Demográfica

Doutorado: Prof. Ms. Ivan Carlin Passos
Orientador: Prof. Dr. Edgard B. Cornachione Jr.

Nome:.....nº:.....

Estas informações são apenas para propósitos estatísticos e não afetarão sua nota neste curso. Por favor, responda às questões e não se esqueça de preencher seu nome e número de registro acadêmico.

1. Gênero:

- a-) feminino
- b-) masculino

2. Cor/Raça (IBGE 2010):

- a-) Preta
- b-) Amarela
- c-) Parda
- d-) Indígena
- e-) Branca

3. Idade:anos.

4. Meu estado civil é:.....

5. Quantidade de filhos (se nenhum = 0):.....

6. Atualmente eu:

- a-) Não trabalho nem faço estágio
- b-) Trabalho ou faço estágio em meio período
- c-) Trabalho ou faço estágio em período integral
- d-) Trabalho informalmente sem uma carga horária determinada

7. Cursei o ensino básico, fundamental e médio:

- a-) Todo em escolas particulares
- b-) Maior parte em escolas particulares
- c-) Maior parte em escolas públicas
- d-) Todo em escolas públicas

8. Minha renda familiar é (salário mínimo paulista = R\$580,00):
- a-) menor ou igual a 3 salários mínimos
 - b-) entre 3,1 e 10 salários mínimos
 - c-) entre 10,1 e 15 salários mínimos
 - d-) entre 15,1 e 20 salários mínimos
 - e-) acima de 20 salários mínimos
9. Eu leio livros ou artigos:
- a-) Todos os dias
 - b-) Frequentemente
 - c-) Raramente
 - d-) Muito raramente
 - e-) Nunca
10. Eu escrevo redações ou textos:
- a-) Todos os dias
 - b-) Frequentemente
 - c-) Raramente
 - d-) Muito raramente
 - e-) Nunca
11. Eu faço cálculos ou estudo Matemática:
- a-) Todos os dias
 - b-) Frequentemente
 - c-) Raramente
 - d-) Muito raramente
 - e-) Nunca
12. Fala fluentemente outra língua?
- () Sim () Não
13. Já fez ou faz algum curso de leitura? (Não considere as disciplinas comuns de seu curso de graduação)
- () Sim () Não
14. Já fez ou faz algum curso de escrita? (Não considere as disciplinas comuns de seu curso de graduação)

() Sim () Não

15. Já fez ou faz algum curso de Matemática? (Não considere as disciplinas comuns de seu curso de graduação)

() Sim () Não

16. Já fez ou faz algum curso para desenvolver o raciocínio crítico? (Pode ter outro nome, como por exemplo: lógica, raciocínio reflexivo, técnicas de debate etc.)

() Sim () Não

Se sim, quando você fez?

De quanto tempo em horas foi o curso?.....

APÊNDICE D

Exemplo do Modelo de Controle dos Participantes
por instrumento e data

Quantidade	Registro Acadêmico	Nome	Data de Realização dos testes por cada participante						Pesquisa Demográfica	Entrevistas
			EWCTET - Pré-teste	CCTDI - Pré-teste	EWCTET - Pós-teste	CCTDI - Pós-teste	Prova Contábil 1	Prova Contábil 2		
1	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
2	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
3	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
4	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
5	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
6	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
7	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
8	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
9	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
10	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
11	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
12	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
13	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
14	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
15	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
16	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
17	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						
18	xxxxxxxxxxxx	xxxxxx	16/08/2010	13/09/2010						

ANEXOS

ANEXO A - Consentimento para a Pesquisa Cartas dos Coordenadores de Curso

ANEXO B - Atestado de Tradução Técnica de citações da Língua Inglesa para a Língua Portuguesa

ANEXO A

Consentimento para a Pesquisa

Cartas dos Coordenadores de Curso

UNIVERSIDADE DE SÃO PAULO
Faculdade de Economia, Administração e Contabilidade
Departamento de Contabilidade e Atuária
Graduação EAC

São Paulo, 17 de Setembro de 2010

Assunto: Autorização para pesquisa no Curso de Graduação

Caríssimo Ivan Carlin Passos,

Autorizo a sua pessoa a realizar pesquisa modelo instrucional para as turmas de graduação em Ciências Contábeis da FEA-USP sob tutoria da Profa. Tânia Relvas na disciplina Contabilidade Intermediária I.

Atenciosamente,

Prof. Dr. Gerlando Lima
Coordenador da CoC/EAC

FIG-UNIMESP
CENTRO UNIVERSITÁRIO METROPOLITANO DE SÃO PAULO

Guarulhos, 15 de março de 2011.

Venho por meio deste, autorizar o Prof. Ivan Carlin Passos a realizar sua pesquisa de doutoramento utilizando as turmas do 3º (terceiro) semestre do Curso de Administração no 1º semestre de 2011.

Atenciosamente

A handwritten signature in black ink, consisting of a large, stylized 'E' followed by a horizontal line and a small vertical stroke.

Prof. Eduardo Gimenes Romero
Coordenador do Curso de Administração

ANEXO B

Atestado de Tradução Técnica de citações
da Língua Inglesa para a Língua Portuguesa

Atestado de Verificação das Traduções da Língua Inglesa para a Língua Portuguesa

As citações desta tese ("Raciocínio Crítico de Alunos de Graduação em Ciências Contábeis: Aplicação do Modelo Instrucional de Richard Paul" do autor Ivan Carlin Passos) que foram traduzidas da Língua Inglesa para o Português foram todas revisadas e em alguns casos corrigidas. As traduções estão adequadas e transmitem bem os significados dos textos originais na Língua Inglesa.

Paulo César Passos

Tradutor e redator técnico em Inglês, por mais de 35 anos atuando na indústria. Participou da tradução de normas técnicas para fabricação do primeiro satélite brasileiro junto ao INPE. Trabalhou para a Monsanto, Petrobras, Indústrias Villares-Vibasa, Neiva e Tecnasa, indústria de equipamentos de navegação aérea e radares meteorológicos, onde teve o cargo de tradutor e redator técnico Sr.

Ribeirão Preto – SP, 10 de agosto de 2011.

ANEXO C
Saídas (*Outputs*) do SPSS

Saída 1**Tests of Between-Subjects Effects**

Dependent Variable: EWCTETpos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	509,951(a)	3	169,984	5,604	,002
Intercept	267,475	1	267,475	8,818	,005
Turma	2,017	1	2,017	,067	,798
EWCTETpre	424,744	1	424,744	14,002	,000
Turma * EWCTETpre	27,339	1	27,339	,901	,347
Error	1456,049	48	30,334		
Total	10091,000	52			
Corrected Total	1966,000	51			

a R Squared = ,259 (Adjusted R Squared = ,213)

Saída 2**Tests of Between-Subjects Effects**

Dependent Variable: EWCTETpos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	,944(a)	1	,944	,024	,877
Intercept	8106,251	1	8106,251	206,260	,000
Turma	,944	1	,944	,024	,877
Error	1965,056	50	39,301		
Total	10091,000	52			
Corrected Total	1966,000	51			

a R Squared = ,000 (Adjusted R Squared = -,020)

Saída 3**Tests of Between-Subjects Effects**

Dependent Variable: EWCTETpos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	482,612(a)	2	241,306	7,971	,001
Intercept	240,199	1	240,199	7,934	,007
EWCTETpre	481,669	1	481,669	15,911	,000
Turma	66,814	1	66,814	2,207	,144
Error	1483,388	49	30,273		
Total	10091,000	52			
Corrected Total	1966,000	51			

a R Squared = ,245 (Adjusted R Squared = ,215)

Saída 4**Estimates**

Dependent Variable: EWCTETpos

Turma	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimento	13,793(a)	1,158	11,467	16,120
Controle	11,302(a)	1,110	9,072	13,533

a Covariates appearing in the model are evaluated at the following values: EWCTETpre = 14,1250.

Saída 5**Parameter Estimates**

Dependent Variable: EWCTETpos

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	4,186	2,367	1,769	,083	-,570	8,943
EWCTETpre	,504	,126	3,989	,000	,250	,758
[Turma=1,000]	2,491	1,677	1,486	,144	-,878	5,860
[Turma=2,000]	0(a)	.	.	.	.	.

a This parameter is set to zero because it is redundant.

Saída 6**Tests of Between-Subjects Effects**

Dependent Variable: EWCTETpos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	174,465(a)	2	87,233	7,532	,002
Intercept	223,300	1	223,300	19,281	,000
EWCTETpre	102,225	1	102,225	8,827	,005
Turma	52,945	1	52,945	4,572	,040
Error	405,351	35	11,581		
Total	1279,000	38			
Corrected Total	579,816	37			

a R Squared = ,301 (Adjusted R Squared = ,261)

Saída 7**Estimates**

Dependent Variable: EWCTETpos

Turma	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimento	5,683(a)	,854	3,949	7,417
Controle	3,276(a)	,728	1,799	4,753

a Covariates appearing in the model are evaluated at the following values: EWCTETpre = 3,4737.

Saída 8**Parameter Estimates**

Dependent Variable: EWCTETpos

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	1,958	,823	2,379	,023	,287	3,630
EWCTETpre	,379	,128	2,971	,005	,120	,639
[Turma=1,00]	2,407	1,126	2,138	,040	,122	4,692
[Turma=2,00]	0(a)	.	.	.	.	.

a This parameter is set to zero because it is redundant.

Saída 9**Tests of Between-Subjects Effects**

Dependent Variable: CCTDIpos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	18499,046(a)	2	9249,523	15,010	,000
Intercept	2702,064	1	2702,064	4,385	,041
CCTDIpre	18291,969	1	18291,969	29,683	,000
Turma	206,640	1	206,640	,335	,565
Error	30195,723	49	616,239		
Total	4548072,000	52			
Corrected Total	48694,769	51			

a R Squared = ,380 (Adjusted R Squared = ,355)

Saída 10**Estimates**

Dependent Variable: CCTDIpos

Turma	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimento	292,082(a)	4,965	282,105	302,059
Controle	296,072(a)	4,777	286,471	305,673

a Covariates appearing in the model are evaluated at the following values: CCTDIpre = 297,9231.

Saída 11**Parameter Estimates**

Dependent Variable: CCTDIpos

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	83,871	39,241	2,137	,038	5,014	162,729
CCTDIpre	,712	,131	5,448	,000	,450	,975
[Turma=1,000]	-3,990	6,890	-,579	,565	-17,836	9,856
[Turma=2,000]	0(a)	.	.	.	.	.

a This parameter is set to zero because it is redundant.

Saída 12**Tests of Between-Subjects Effects**

Dependent Variable: Maturidade.pos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	677,601(a)	2	338,801	9,416	,000
Intercept	406,546	1	406,546	11,298	,002
Maturidade.pre	633,795	1	633,795	17,614	,000
Turma	122,460	1	122,460	3,403	,071
Error	1763,168	49	35,983		
Total	93498,000	52			
Corrected Total	2440,769	51			

a R Squared = ,278 (Adjusted R Squared = ,248)

Saída 13**Estimates**

Dependent Variable: Maturidade.pos

Turma	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimento	43,468(a)	1,210	41,036	45,900
Controle	40,345(a)	1,164	38,006	42,683

a Covariates appearing in the model are evaluated at the following values: Maturidade.pre = 42,1346.

Saída 14**Parameter Estimates**

Dependent Variable: Maturidade.pos

Parameter	B	Std. Error	T	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	17,181	5,783	2,971	,005	5,559	28,802
Maturidade.pre	,550	,131	4,197	,000	,287	,813
[Turma=1,00]	3,123	1,693	1,845	,071	-,279	6,525
[Turma=2,00]	0(a)	.	.	.	.	.

a This parameter is set to zero because it is redundant.

Saída 15**Tests of Between-Subjects Effects**

Dependent Variable: CCTDIpossquare

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9647959410,632(a)	2	4823979705,316	36,565	,000
Intercept	256871749,344	1	256871749,344	1,947	,172
CCTDIpresquare	5929654035,711	1	5929654035,711	44,946	,000
Turma	288803487,306	1	288803487,306	2,189	,148
Error	4617460826,632	35	131927452,189		
Total	310302876244,000	38			
Corrected Total	14265420237,264	37			

a R Squared = ,676 (Adjusted R Squared = ,658)

Saída 16**Estimates**

Dependent Variable: CCTDIpossquare

Turma	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimento	92144,784 (a)	3217,608	85612,692	98676,877
Controle	85732,358 (a)	2529,540	80597,118	90867,598

a Covariates appearing in the model are evaluated at the following values: CCTDIpresquare = 89921,3158.

Saída 17**Parameter Estimates**

Dependent Variable: CCTDIpossquare

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	12553,690	10381,382	1,209	,235	-8521,637	33629,016
CCTDIpresquare	,814	,121	6,704	,000	,567	1,060
[Turma=1,00]	6412,427	4333,999	1,480	,148	-2386,059	15210,912
[Turma=2,00]	0(a)	.	.	.	.	.

a This parameter is set to zero because it is redundant.

Saída 18**Tests of Between-Subjects Effects**

Dependent Variable: Mente.Aberta.pos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	579,985(a)	2	289,992	10,657	,000
Intercept	354,239	1	354,239	13,018	,001
Mente.Aberta.pre	467,072	1	467,072	17,165	,000
Turma	60,023	1	60,023	2,206	,146
Error	979,605	36	27,211		
Total	63640,000	39			
Corrected Total	1559,590	38			

a R Squared = ,372 (Adjusted R Squared = ,337)

Saída 19**Estimates**

Dependent Variable: Mente.Aberta.pos

Turma	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimento	41,397(a)	1,311	38,739	44,055
Controle	38,854(a)	1,091	36,640	41,068

a Covariates appearing in the model are evaluated at the following values: Mente.Aberta.pre = 38,9487.

Saída 20**Parameter Estimates**

Dependent Variable: Mente.Aberta.pos

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	17,579	5,160	3,406	,002	7,113	28,045
Mente.Aberta.pre	,546	,132	4,143	,000	,279	,814
[Turma=1,00]	2,543	1,713	1,485	,146	-,930	6,017
[Turma=2,00]	0(a)	.	.	.	.	.

a This parameter is set to zero because it is redundant.

Saída 21**Tests of Between-Subjects Effects**

Dependent Variable: Curiosidade.pos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	377,326(a)	2	188,663	8,953	,001
Intercept	213,739	1	213,739	10,144	,003
Curiosidade.pre	167,341	1	167,341	7,942	,008
Turma	89,331	1	89,331	4,239	,047
Error	758,572	36	21,071		
Total	105451,000	39			
Corrected Total	1135,897	38			

a R Squared = ,332 (Adjusted R Squared = ,295)

Saída 22**Estimates**

Dependent Variable: Curiosidade.pos

Turma	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimento	53,639(a)	1,188	51,230	56,047
Controle	50,382(a)	,980	48,393	52,370

a Covariates appearing in the model are evaluated at the following values: Curiosidade.pre = 51,4872.

Saída 23**Parameter Estimates**

Dependent Variable: Curiosidade.pos

Parameter	B	Std. Error	T	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	26,083	8,464	3,082	,004	8,918	43,249
Curiosidade.pre	,472	,167	2,818	,008	,132	,812
[Turma=1,00]	3,257	1,582	2,059	,047	,049	6,465
[Turma=2,00]	0(a)	.	.	.	.	.

a This parameter is set to zero because it is redundant.

Saída 24**Tests of Between-Subjects Effects**

Dependent Variable: Maturidade.pos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1922,630(a)	2	961,315	25,885	,000
Intercept	128,182	1	128,182	3,452	,071
Maturidade.pre	1412,804	1	1412,804	38,042	,000
Turma	12,871	1	12,871	,347	,560
Error	1336,959	36	37,138		
Total	69147,000	39			
Corrected Total	3259,590	38			

a R Squared = ,590 (Adjusted R Squared = ,567)

Saída 25**Estimates**

Dependent Variable: Maturidade.pos

Turma	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimento	41,871(a)	1,630	38,566	45,176
Controle	40,568(a)	1,333	37,865	43,271

a Covariates appearing in the model are evaluated at the following values: Maturidade.pre = 39,7949.

Saída 26**Parameter Estimates**

Dependent Variable: Maturidade.pos

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	9,150	4,861	1,883	,068	-,707	19,008
Maturidade.pre	,789	,128	6,168	,000	,530	1,049
[Turma=1,00]	1,303	2,213	,589	,560	-3,185	5,791
[Turma=2,00]	0(a)	.	.	.	.	.

a This parameter is set to zero because it is redundant.

Saída 27**Levene's Test of Equality of Error Variances(a)**

Dependent Variable: EWCTETpos

F	df1	df2	Sig.
1,335	3	46	,275

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a Design: Intercept+Turma+Grupo.Idade+EWCTETpre+Turma * Grupo.Idade

Saída 28**Tests of Between-Subjects Effects**

Dependent Variable: EWCTETpos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	521,802(a)	4	130,450	4,118	,006	,268
Intercept	191,319	1	191,319	6,040	,018	,118
Turma	21,696	1	21,696	,685	,412	,015
Grupo.Idade	12,757	1	12,757	,403	,529	,009
EWCTETpre	470,100	1	470,100	14,842	,000	,248
Turma * Grupo.Idade	6,123	1	6,123	,193	,662	,004
Error	1425,343	45	31,674			
Total	9897,750	50				
Corrected Total	1947,145	49				

a. R Squared = ,268 (Adjusted R Squared = ,203)

Saída 29**3. Grupo.Idade * Turma**

Dependent Variable: EWCTETpos

Grupo.Idade	Turma	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Abaixo de 22	Experimento	13,907 ^a	1,238	11,414	16,400
	Controle	10,640 ^a	1,344	7,933	13,347
22 ou acima	Experimento	14,383 ^a	4,027	6,272	22,495
	Controle	13,284 ^a	2,358	8,534	18,034

a. Covariates appearing in the model are evaluated at the following values:
EWCTETpre = 14,3200.

Saída 30**Tests of Between-Subjects Effects**

Dependent Variable: CCTDIpos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	11884,013(a)	4	2971,003	5,440	,001	,326
Intercept	2383,411	1	2383,411	4,364	,042	,088
Turma	1574,518	1	1574,518	2,883	,096	,060
Grupo.Idade	10,227	1	10,227	,019	,892	,000
CCTDIpre	10268,820	1	10268,820	18,804	,000	,295
Turma * Grupo.Idade	1770,235	1	1770,235	3,242	,078	,067
Error	24574,487	45	546,100			
Total	4273419,000	50				
Corrected Total	36458,500	49				

a. R Squared = ,326 (Adjusted R Squared = ,266)

Saída 31**3. Grupo.Idade * Turma**

Dependent Variable: CCTDIpos

Grupo.Idade	Turma	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Abaixo de 22	Experimento	290,234 ^a	5,019	280,125	300,344
	Controle	288,641 ^a	5,366	277,835	299,448
22 ou acima	Experimento	269,241 ^a	17,849	233,292	305,190
	Controle	306,740 ^a	8,853	288,909	324,570

a. Covariates appearing in the model are evaluated at the following values:
CCTDIpre = 295,5600.

Saída 32**Tests of Between-Subjects Effects**

Dependent Variable: EWCTETpos

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	290,884(a)	4	72,721	8,306	,000	,502
Intercept	261,464	1	261,464	29,863	,000	,475
Turma	53,894	1	53,894	6,155	,018	,157
Grupo.Idade	95,963	1	95,963	10,960	,002	,249
Turma * Grupo.Idade	37,323	1	37,323	4,263	,047	,114
EWCTETpre	57,749	1	57,749	6,596	,015	,167
Error	288,932	33	8,756			
Total	1279,000	38				
Corrected Total	579,816	37				

a. R Squared = ,502 (Adjusted R Squared = ,441)

Saída 33**Estimates**

Dependent Variable: EWCTETpos

Grupo.Idade	Turma	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Abaixo de 22 anos	Experimento	3,060 ^a	1,053	,917	5,202
	Controle	2,730 ^a	,856	,988	4,472
22 anos ou acima	Experimento	8,401 ^a	1,077	6,209	10,593
	Controle	3,856 ^a	,952	1,918	5,793

a. Covariates appearing in the model are evaluated at the following values: EWCTETpre = 3,4737.

Saída 34**Tests of Between-Subjects Effects**

Dependent Variable: CCTDIpos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	29703,972(a)	4	7425,993	21,385	,000	,722
Intercept	269,885	1	269,885	,777	,384	,023
Turma	517,743	1	517,743	1,491	,231	,043
Grupo.Idade	163,462	1	163,462	,471	,497	,014
CCTDIpre	18629,049	1	18629,049	53,647	,000	,619
Turma * Grupo.Idade	1109,112	1	1109,112	3,194	,083	,088
Error	11459,396	33	347,254			
Total	3354016,000	38				
Corrected Total	41163,368	37				

a. R Squared = ,722 (Adjusted R Squared = ,688)

Saída 35**Estimates**

Dependent Variable: CCTDIpos

Grupo.Idade	Turma	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Abaixo de 22 anos	Experimento	292,232 ^a	7,615	276,740	307,724
	Controle	294,892 ^a	5,531	283,639	306,145
22 anos ou acima	Experimento	307,865 ^a	6,754	294,124	321,606
	Controle	287,786 ^a	5,929	275,724	299,849

a. Covariates appearing in the model are evaluated at the following values: CCTDIpre = 298,4211.

Saída 36**Tests of Between-Subjects Effects**

Dependent Variable: EWCTETpossquare

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	143455,053(a)	4	35863,763	2,309	,072	,164
Intercept	248235,193	1	248235,193	15,979	,000	,254
Turma	25708,935	1	25708,935	1,655	,205	,034
Gênero	1386,428	1	1386,428	,089	,766	,002
EWCTETpresquare	122030,320	1	122030,320	7,855	,007	,143
Turma * Gênero	13232,532	1	13232,532	,852	,361	,018
Error	730142,773	47	15534,953			
Total	2831834,000	52				
Corrected Total	873597,827	51				

a. R Squared = ,164 (Adjusted R Squared = ,093)

Saída 37**Estimates**

Dependent Variable: EWCTETpossquare

Gênero	Turma	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
feminino	Experimento	214,401 ^a	37,942	138,071	290,732
	Controle	197,894 ^a	47,145	103,051	292,738
masculino	Experimento	237,281 ^a	34,789	167,295	307,267
	Controle	152,414 ^a	29,326	93,418	211,410

a. Covariates appearing in the model are evaluated at the following values:
EWCTETpresquare = 243,5048.

Saída 38**Tests of Between-Subjects Effects**

Dependent Variable: CCTDIpos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	15556,515(a)	4	3889,129	6,889	,000	,375
Intercept	2581,782	1	2581,782	4,573	,038	,090
Turma	830,668	1	830,668	1,471	,231	,031
Gênero	1754,664	1	1754,664	3,108	,085	,063
Turma * Gênero	32,405	1	32,405	,057	,812	,001
CCTDIpre	14479,411	1	14479,411	25,647	,000	,358
Error	25970,230	46	564,570			
Total	4405188,000	51				
Corrected Total	41526,745	50				

a R Squared = ,375 (Adjusted R Squared = ,320)

Saída 39**Estimates**

Dependent Variable: CCTDIpos

Gênero	Turma	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
feminino	Experimento	296,695 ^a	7,051	282,502	310,889
	Controle	303,609 ^a	8,985	285,523	321,696
masculino	Experimento	282,029 ^a	6,980	267,979	296,080
	Controle	292,402 ^a	5,322	281,690	303,114

a. Covariates appearing in the model are evaluated at the following values:
CCTDIpre = 296,8235.

Saída 40**Tests of Between-Subjects Effects**

Dependent Variable: EWCTETpos

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	285,595(a)	4	71,399	7,910	,000	,497
Intercept	236,753	1	236,753	26,230	,000	,450
Turma	96,778	1	96,778	10,722	,003	,251
Gênero	58,920	1	58,920	6,528	,016	,169
EWCTETpre	79,894	1	79,894	8,851	,006	,217
Turma * Gênero	22,880	1	22,880	2,535	,121	,073
Error	288,837	32	9,026			
Total	1275,000	37				
Corrected Total	574,432	36				

a R Squared = ,497 (Adjusted R Squared = ,434)

Saída 41**Estimates**

Dependent Variable: EWCTETpos

Turma	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Experimento	6,478(a)	,797	4,855	8,102
Controle	3,082(a)	,666	1,726	4,439

a Covariates appearing in the model are evaluated at the following values: EWCTETpre = 3,2838.

Estimates

Dependent Variable: EWCTETpos

Gênero	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Feminino	3,456(a)	,730	1,969	4,943
masculino	6,105(a)	,738	4,601	7,609

a Covariates appearing in the model are evaluated at the following values: EWCTETpre = 3,2838.

Estimates

Dependent Variable: EWCTETpos

Gênero	Turma	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
feminino	Experimento	4,269 ^a	1,012	2,208	6,331
	Controle	2,642 ^a	1,078	,447	4,838
masculino	Experimento	8,687 ^a	1,270	6,101	11,274
	Controle	3,522 ^a	,817	1,858	5,187

a. Covariates appearing in the model are evaluated at the following values:
EWCTETpre = 3,2838.

Saída 42**Tests of Between-Subjects Effects**

Dependent Variable: CCTDIpos

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	30225,471(a)	4	7556,368	25,361	,000	,766
Intercept	951,668	1	951,668	3,194	,084	,093
Turma	1026,104	1	1026,104	3,444	,073	,100
Gênero	1488,493	1	1488,493	4,996	,033	,139
CCTDIpre	13560,338	1	13560,338	45,512	,000	,595
Turma * Gênero	93,624	1	93,624	,314	,579	,010
Error	9236,418	31	297,949			
Total	3149982,000	36				
Corrected Total	39461,889	35				

a R Squared = ,766 (Adjusted R Squared = ,736)

Saída 43**Estimates**

Dependent Variable: CCTDIpos

Gênero	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
feminino	287,818(a)	4,220	279,211	296,425
masculino	302,280(a)	4,754	292,584	311,977

a Covariates appearing in the model are evaluated at the following values: CCTDIpre = 297,8889.

Estimates

Dependent Variable: CCTDIpos

Gênero	Turma	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
feminino	Experimento	296,348 ^a	6,252	283,597	309,098
	Controle	279,288 ^a	6,196	266,651	291,925
masculino	Experimento	307,358 ^a	8,566	289,889	324,828
	Controle	297,202 ^a	4,655	287,708	306,697

a. Covariates appearing in the model are evaluated at the following values:
CCTDIpre = 297,8889.