

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Dilvana Maria Fiorini de Aguiar Moreira

**A EVOLUÇÃO TEMPORAL DA COMPETÊNCIA EM
FÍSICA ESCOLAR E HABILIDADES COGNITIVAS
AMPLAS NA PRIMEIRA SÉRIE DO ENSINO MÉDIO**

Belo Horizonte

2011

Dilvana Maria Fiorini de Aguiar Moreira

A EVOLUÇÃO TEMPORAL DA COMPETÊNCIA EM FÍSICA ESCOLAR E HABILIDADES COGNITIVAS AMPLAS NA PRIMEIRA SÉRIE DO ENSINO MÉDIO

Dissertação apresentada ao Curso de Mestrado da Faculdade de Educação da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Educação.

Linha de Pesquisa: Educação e Ciências

Orientador: Prof. Dr. Oto Néri Borges

Belo Horizonte

Faculdade de Educação da UFMG

2011

Universidade Federal de Minas Gerais

Faculdade de Educação

Programa de Pós-Graduação em Educação: Conhecimento e Inclusão Social

Dissertação intitulada “A evolução temporal da competência em física escolar e habilidades cognitivas amplas na primeira série do ensino médio”, de autoria da mestrande Dilvana Maria Fiorini de Aguiar Moreira, aprovada pela banca examinadora constituída pelos seguintes professores:

Prof. Dr. Oto Néri Borges – FAE/UFMG - Orientador

Prof. Dr. Geide Rosa Coelho - UFES

Profa. Dra. Ines Mendes Pinto – Rede Municipal de Ensino de Belo Horizonte

Profa. Dra. Carmen Maria de Caro Martins – FAE/UFMG

Profa. Dra. Marluce Alves Paraíso
Coordenadora do Programa de Pós-Graduação em Educação: Conhecimento e
Inclusão Social – FAE/UFMG

Belo Horizonte, 16 de agosto de 2011

Dedico este trabalho aos meus pais, Leny de Aguiar e Souza e Diva Fiorini de Aguiar, e à minha irmã, Divanir Fiorini de Aguiar, exemplos de coragem e perseverança, e que deram, de diferentes formas, apoio e suporte para que eu chegasse até aqui.

Agradecimentos

Agradeço ao professor Oto Borges, meu orientador, que me proporcionou a oportunidade e o acompanhamento na realização desta pesquisa.

Ao meu companheiro Adelson, que sempre esteve presente incentivando e apoiando as minhas escolhas.

Aos meus filhos Francisco e Lucas, que entenderam as minhas ausências e apoiaram a minha decisão de fazer o mestrado.

À amiga Vânia, que me ajudou, ouvindo em meus momentos de angústia e me transmitindo segurança e palavras de conforto.

Agradeço também aos meus colegas do grupo de orientação: Amanda, Inês, Geide, Marciana, Morgana, Maria Tereza, Terezinha e Rafael pelos momentos de discussão e trocas.

Agradeço, em especial, à Elizabeth, pela generosidade e disposição para as conversas sobre o meu trabalho.

Aos meus colegas da Espaço Escola, Valéria, Ainara, Daniel, Patrícia e Carla pelos momentos de muita conversa e apoio.

E, por fim, ao Sérgio, Wagner e Serginho que estiveram comigo apoiando e conversando desde o início de minha caminhada como mestranda.

Resumo

Esta dissertação relata uma investigação sobre a existência da relação entre a evolução temporal da aprendizagem em física escolar, através da sua representação como competência em física escolar, e seis das habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll (1993). Essa investigação se situa na área temática que investiga a relação entre a aprendizagem escolar e as habilidades cognitivas amplas. Investigaram-se duas questões: (1) a dependência da evolução temporal da competência em física escolar em relação às variáveis demográficas e contextuais e, após controlar para os efeitos de tais dependências, (2) a relação entre a evolução temporal da competência em física escolar e algumas das habilidades cognitivas amplas do estrato II do modelo de Carroll.

Para realizar esta investigação utilizaram-se os dados cognitivos dos estudantes obtidos de uma investigação anterior (Gomes, 2010) e dados das avaliações escolares ordinárias existentes nos registros acadêmicos da escola. A partir dos dados das avaliações escolares, de natureza ordinal, construiu-se uma escala intervalar para mensurar a competência em física escolar em três ocasiões distintas durante o ano letivo. Usando-se os resultados dos testes marcadores das habilidades cognitivas do estrato I do modelo de Carroll, construíram-se escalas intervalares (baseadas nos escores padronizados) para as seis habilidades amplas mensuradas pela bateria BaFaCaLo (Gomes, 2010). Construíram-se, também, indicadores dicotômicos ou categóricos para as variáveis demográficas e contextuais.

A evolução temporal da competência em física escolar foi analisada utilizando-se o modelo multinível de regressão múltipla, com um nível intra-individual e outro inter-individual. O nível intra-individual mostra que a competência em física escolar de cada estudante varia de forma sistemática e linear com o tempo, além de sofrer variações aleatórias entre as ocasiões. O nível inter-individual, como o intercepto e a inclinação de cada estudante, depende das variáveis demográficas (gênero, idade, e renda familiar), contextuais (frequência da presença em sala, turma, professor de física, escolaridade do pai e da mãe), e das seguintes habilidades cognitivas amplas:

a inteligência fluida (Gf), a inteligência cristalizada (Gc), a habilidade visuo-espacial (Gv), a fluência (Gr), a memória de curto prazo (Gy) e a rapidez cognitiva (Gs).

O resultado mostra que, depois de controlado para os efeitos das variáveis demográficas e contextuais, a inteligência fluida e a memória de curto prazo diferenciam os estudantes segundo a sua competência em física no início do ano letivo e que a inteligência fluida influencia a taxa de variação da competência em física escolar ao longo do ano letivo.

Abstract

This investigation report relates a research about the existence of relation between temporal evolution of learning in school physics, represented as school physics competency, and six human intellectual abilities, proposed by Carrol (1993). This investigation is situated in the thematic field in which is studied the relation between school learning and large cognitive abilities. Two questions oriented the research: (1) temporal evolution of school physics dependence with contextual and demographic variables, and after controlling of these dependency effects, (2) relation between temporal evolution of school physics competency and some large cognitive abilities detached from II extract of Carrol model.

Student cognitive data obtained from prior investigation (Gomes, 2010) and ordinary school tests results existing in academic school registries constituted the research data. It was constructed an interval scale from school tests results, these of ordinal nature, to measuring school physics competency in three different occasions during the school year. The cognitive ability marking test results from I extract of Carrol model provided the data to constructing interval scales (based in patterned scores) for six large abilities measured by BaFaCaLo battery (Gomes, 2010). It was also constructed dicotomic or categorical indexes for contextual and demographic variables.

The temporal evolution of school physics competency was analysed by multilayered model of multiple regression, with an intra-individual and other inter-individual layers. Intra-individual layer shows school physics competency of each student varying with time in a linear and systematic form, besides presenting aleatory variation between occasions durint the year. The inter-individual layer, as interception and inclination of each student, shows dependency with demographic variables (gender, age and familiar profits), contextual (presence frequency in class, school class, physics teacher, mother and father schooling) and with these large cognitive abilities: fluid intelligence (Gf), crystallized intelligence (Gc), visual-spatial ability (Gv), fluency (Gr), short-term memory (Gms) and cognitive speed (Gs).

After controlling of demographic and contextual variables effects, the result shows fluid intelligence and short-term memory differing students according to their school

physics competency in the beginning of school year, and fluid intelligence influencing rate of change of school physics competency during the school year.

Lista de Gráficos

GRÁFICO 1 – MÉDIA DA COMPETÊNCIA EM FÍSICA ESCOLAR DOS ESTUDANTES DA 1ª SÉRIE DO ENSINO MÉDIO	102
GRÁFICO 2 – COMPETÊNCIA EM FÍSICA ESCOLAR MÉDIA DA TURMA 1	109
GRÁFICO 3 – COMPETÊNCIA EM FÍSICA ESCOLAR MÉDIA DA TURMA 2	110
GRÁFICO 4 – COMPETÊNCIA EM FÍSICA ESCOLAR MÉDIA DA TURMA 3	110
GRÁFICO 5 – COMPETÊNCIA EM FÍSICA ESCOLAR MÉDIA DA TURMA 4	110
GRÁFICO 6 – COMPETÊNCIA EM FÍSICA ESCOLAR MÉDIA DA TURMA 5	111
GRÁFICO 7 - COMPETÊNCIA EM FÍSICA ESCOLAR MÉDIA DA TURMA 6	111
GRÁFICO 8 - COMPETÊNCIA EM FÍSICA ESCOLAR MÉDIA DA TURMA 7	111
GRÁFICO 9 - COMPETÊNCIA EM FÍSICA ESCOLAR MÉDIA DA TURMA 8	112
GRÁFICO 10 – EVOLUÇÃO TEMPORAL DA COMPETÊNCIA EM FÍSICA ESCOLAR NA CATEGORIA DE NGY (MB) PARA CATEGORIAS DE NGf(MB), NGf(B), NGf(A) E NGf (MA)	132
GRÁFICO 11 - EVOLUÇÃO TEMPORAL DA COMPETÊNCIA EM FÍSICA ESCOLAR (CFIS) NA CATEGORIA DE NGY(B) PARA CATEGORIAS DE NGf(MB), NGf(B), NGf(A) E NGf(MA)	133
GRÁFICO 12 - EVOLUÇÃO TEMPORAL DA COMPETÊNCIA EM FÍSICA ESCOLAR (CFIS) NA CATEGORIA DE NGY(A) PARA CATEGORIAS DE NGf (MB), NGf (B), NGf (A) E NGf (MA)	133
GRÁFICO 13 - EVOLUÇÃO TEMPORAL DA COMPETÊNCIA EM FÍSICA ESCOLAR (CFIS) NA CATEGORIA DE NGY (MA) PARA CATEGORIAS DE NGf(MB), NGf(B), NGf(A) E NGf(MA)	134

Lista de Quadros

QUADRO 1 HABILIDADES COGNITIVAS ESPECIALIZADAS DO ESTRATO I DO MODELO DE CARROLL MENSURADAS PELA BATERIA BAFACALO.....	40
QUADRO 2 RELAÇÃO ENTRE AS HABILIDADES COGNITIVAS DO ESTRATO II E DO ESTRATO I DO MODELO DE CARROLL, MENSURADAS PELA BATERIA BAFACALO	42
QUADRO 3 RELAÇÃO DE ESTUDANTES POR TURMA E GÊNERO NA 1ª SÉRIE	70
QUADRO 4 RELAÇÃO DO NÚMERO DE ESTUDANTES DO ENSINO MÉDIO E TÉCNICO, SEGUNDO O GÊNERO, TURMA E SÉRIE, FORNECIDA PELO PROFESSOR CRISTIANO.....	74
QUADRO 5 RELAÇÃO DE ESTUDANTES POR TURMA E GÊNERO NA PRIMEIRA SÉRIE DO ENSINO MÉDIO.....	76
QUADRO 6 TIPO E NATUREZA DAS VARIÁVEIS PREDITORAS DA PESQUISA.....	78
QUADRO 7 CONCEITOS UTILIZADOS NA UNIVERSIDADE SEGUNDO A FAIXA DE NOTAS.....	80
QUADRO 8 DUAS FORMAS DE ORGANIZAÇÃO DOS DADOS DE DUAS ONDAS DE MEDIDAS.....	82
QUADRO 9 ORGANIZAÇÃO DOS DADOS PARA EQUALIZAÇÃO VIRTUAL	83
QUADRO 10 SEIS HABILIDADES AMPLAS DO ESTRATO II DE CARROLL	84
QUADRO 11 CONCEITOS UTILIZADOS NA UNIVERSIDADE SEGUNDO A FAIXA DE NOTAS.....	91
QUADRO 12 CONCEITOS ATRIBUÍDOS AOS ESTUDANTES SEGUNDO A FAIXA DE NOTAS	92
QUADRO 13 DISTRIBUIÇÃO DE ESTUDANTES POR TURMA E CONCEITO, SEGUNDO A OCASIÃO	93
QUADRO 14 ESQUEMA DE TRANSFORMAÇÃO DOS CONCEITOS ATRIBUÍDOS AOS ESTUDANTES NA PRIMEIRA OCASIÃO EM INDICADORES DE UMA ESCALA DO TIPO GUTTMAN	97
QUADRO 15 RELAÇÃO DOS CONCEITOS ATRIBUÍDOS AOS ESTUDANTES POR OCASIÃO SEGUNDA UMA ESCALA GUTTMAN.....	98
QUADRO 16 CATEGORIZAÇÃO CONCEITUAL POR TURMA EM RELAÇÃO ÀS DIFICULDADES DOS ITENS	100
QUADRO 17 TRÊS SITUAÇÕES DE SUJEITOS PARA AS TRÊS OCASIÕES COM AS MEDIDAS CORRIGIDAS E NÃO CORRIGIDAS	101
QUADRO 18 HABILIDADES E SEU SCORE MÍNIMO E MÁXIMO, A MEDIDA DA HABILIDADE NORMALIZADA, SEUS VALORES MÍNIMO E MÁXIMO	105
QUADRO 19 RELAÇÃO DO NÚMERO DE MENINAS E MENINO POR TURMA	106
QUADRO 20 INCLINAÇÃO E INTERCEPTO SEGUNDO A CATEGORIA DE HABILIDADE DE MEMÓRIA DE CURTO PRAZO (NGY) E DE INTELIGÊNCIA FLUIDA (NGF)	134

Lista de Tabelas

TABELA 1 MÉDIA E DESVIO PADRÃO DA COMPETÊNCIA EM FÍSICA ESCOLAR DOS ESTUDANTES DA 1ª SÉRIE DO ENSINO MÉDIO POR OCASIÃO	102
TABELA 2 COMPETÊNCIA EM FÍSICA ESCOLAR SEGUNDO GÊNERO E OCASIÃO DE MEDIDA	107
TABELA 3 MÉDIA, DESVIO PADRÃO DA COMPETÊNCIA EM FÍSICA ESCOLAR E NUMERO DE ESTUDANTES DA 1ª SÉRIE DO ENSINO MÉDIO DE 2007 POR CATEGORIA DE IDADE EM ANOS.	108
TABELA 4 COMPETÊNCIA EM FÍSICA ESCOLAR SEGUNDO O PROFESSOR E A OCASIÃO	113
TABELA 5 COMPETÊNCIA EM FÍSICA ESCOLAR DOS ESTUDANTES SEGUNDO A ESCOLARIDADE DO PAI E A OCASIÃO	114
TABELA 6 COMPETÊNCIA EM FÍSICA ESCOLAR SEGUNDO A ESCOLARIDADE DA MÃE EM CADA OCASIÃO	116
TABELA 7 SEQUÊNCIA DE MODELOS PARA A EVOLUÇÃO TEMPORAL DA COMPETÊNCIA EM FÍSICA ESCOLAR	121
TABELA 8 ESTATÍSTICAS DESCRITIVAS SEGUNDO AS CATEGORIAS DE INTELIGÊNCIA FLUIDA PARA AS AMOSTRAS BAFACALO E Eq. 7	130
TABELA 9 ESTATÍSTICAS DESCRITIVAS SEGUNDO AS CATEGORIAS DE MEMÓRIA DE CURTO PRAZO PARA AS AMOSTRAS BAFACALO E Eq. 7	131

Sumário

1. INTRODUÇÃO.....	15
2. REFERENCIAIS TEÓRICOS.....	25
2.1 – INTRODUÇÃO.....	25
2.2 – OS CONCEITOS DE COMPETÊNCIA E A COMPETÊNCIA EM FÍSICA ESCOLAR.....	26
2.3 – O MODELO DA ESTRUTURA DAS HABILIDADES INTELECTUAIS HUMANAS, PROPOSTO POR CARROLL	33
2.4 – IMPLICAÇÕES EDUCACIONAIS DE UM MODELO DE HABILIDADES INTELECTUAIS HUMANAS	43
2.5 – AVALIAÇÕES EDUCACIONAIS COMO DADO DE MEDIDA EDUCACIONAL	47
2.6 - MODELO RASCH	54
2.7 – MODELO DE REGRESSÃO MÚLTIPLA MULTINÍVEL PARA ESTUDOS LONGITUDINAIS	60
3. ASPECTOS METODOLÓGICOS DA PESQUISA.....	68
3.1 - CONTEXTO DA PESQUISA.....	68
3.2 - A DISCIPLINA FÍSICA NA 1ª SÉRIE DO ENSINO MÉDIO EM 2007.....	71
3.3 - FONTES DE DADOS.....	72
3.3.1 - <i>Preparação dos dados</i>	75
3.4 - AS VARIÁVEIS DA PESQUISA	77
3.4.1 - <i>Competência em física escolar</i>	78
3.4.2 <i>Habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll</i>	84
3.4.3 - <i>Variáveis demográficas</i>	85
3.4.4 - <i>Variáveis de Contextos</i>	86
3.5 - ANÁLISE DOS DADOS.....	88
4. RESULTADOS.....	90
4.1 - A COMPETÊNCIA EM FÍSICA ESCOLAR	90
4.1.1 - <i>Construção da medida da competência em física escolar</i>	90
4.1.2 – <i>Uma análise exploratória da competência em física escolar</i>	101
4.2 - HABILIDADES COGNITIVAS AMPLAS DO ESTRATO II DO MODELO DE ESTRUTURA DAS HABILIDADES INTELECTUAIS HUMANAS, PROPOSTO POR CARROLL.....	102
4.3 - ANÁLISE EXPLORATÓRIA DO EFEITO DAS VARIÁVEIS DEMOGRÁFICAS E CONTEXTUAIS.....	105
4.3.1 - <i>Gênero</i>	105
4.3.2 - <i>Idade</i>	107
4.3.3 - <i>Renda familiar</i>	108

4.3.4 - Turma.....	109
4.3.5 - Professor.....	112
4.3.6 - Escolaridade do pai.....	114
4.3.7 - Escolaridade da mãe.....	115
4.4 - MODELO MULTINÍVEL DE REGRESSÃO MÚLTIPLA DE MUDANÇAS	116
4.5 - ANÁLISE DO RESULTADO.....	128
4.5 - CONCLUSÃO DO CAPÍTULO	135
5. CONCLUSÃO.....	137
5.1 - RESULTADOS E CONCLUSÕES	138
5.2 - LIMITAÇÕES DA PESQUISA	142
5.3 - IMPLICAÇÕES DA PESQUISA	143
6. REFERÊNCIAS BIBLIOGRÁFICAS	146
ANEXO A - PLANEJAMENTO DO CURSO DE FÍSICA – 1º ANO – 2007	158

1. INTRODUÇÃO

Ao longo de minha experiência firmei uma convicção de que não existe em uma sala estudantes iguais, pois cada um deles vem de uma experiência que é única, e esta história já vivida interfere na sua maneira de ver e de se relacionar com o conhecimento. Sempre que se inicia um trabalho em sala de aula, esta se apresenta com várias novidades. Dentre elas estão os estudantes novatos que vieram de outras experiências e que trouxeram seus conhecimentos e suas dificuldades. Existem aqueles que, mesmo não tendo vindo de outras escolas, estiveram em outras turmas envolvidos com outros colegas, professores e outras dinâmicas de aprendizagem. E por fim, aqueles que foram reprovados e que apesar de permanecerem nas mesmas séries estiveram envolvidos em turmas e experiências diferentes. Desta forma, os estudantes têm trajetórias de aprendizagem distintas. Essas trajetórias, ainda que pessoais e enraizadas em histórias de aprendizagens anteriores, possuem semelhanças e diferenças. Ainda que elas possam ser percebidas nos comportamentos, ações e discursos dos estudantes nas atividades cotidianas em sala de aula, a investigação sobre trajetórias de aprendizagem ganhou impulso com métodos quantitativos, principalmente de modelamento de variáveis latentes (HOX, 2002; STOEL et al., 2006).

Em um trabalho inspirador, Yan (2000) mostrou que mesmo em disciplinas de cursos de pós-graduação de uma universidade americana de grande renome, as histórias de aprendizagem vistas através das histórias de desempenho, são singulares e únicas. Ainda assim, é possível identificar alguns tipos característicos de trajetórias na sala de aula. Empregando uma sofisticada metodologia qualitativa e quantitativa em um estudo longitudinal de média duração, ele identificou quatro grupos distintos de estudantes cujas trajetórias de desempenho apresentavam características comuns se comparadas dentro do grupo e muito diferentes se comparadas entre os grupos. No entanto, empregando a metodologia quantitativa tradicional ele não foi capaz de distinguir grupos de estudantes com base em suas trajetórias de aprendizagem.

Este resultado é inspirador por mostrar que o uso de metodologias mais avançadas permite entender melhor o que se passa no espaço da sala de aula, especialmente permite identificar diferentes tipos de trajetórias de aprendizagem.

As diferenças de desempenho escolar dos estudantes têm sido estudadas há décadas, principalmente usando métodos quantitativos e estudos transversais. Essas pesquisas mostram que, em praticamente todos os conteúdos escolares, a regra é a diversidade de desempenhos dos estudantes. Há duas abordagens tradicionais para entender essas diferenças entre os desempenhos dos estudantes. A primeira delas, de inspiração sociológica, atribui as diferenças individuais nos desempenhos às diferenças dos contextos sociais e econômicos de origem dos estudantes. A segunda, de inspiração psicológica, atribui as diferenças de desempenho entre os estudantes às diferenças de suas habilidades cognitivas. Do ponto de vista de pesquisa empírica a primeira abordagem busca identificar, dentre as variáveis que descrevem o contexto social e econômico do qual se origina o estudante, aquelas que são boas variáveis preditoras do desempenho escolar dos estudantes. Já a segunda tenta associar o desempenho escolar às habilidades intelectuais dos estudantes, aferidas usando testes e instrumentos especialmente desenvolvidos para mensurar estas habilidades cognitivas.

Esta dissertação relata uma investigação sobre a existência de uma possível relação entre a evolução temporal da aprendizagem em física escolar e seis das habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll (1993). O modelo da estrutura das habilidades intelectuais humanas, proposto por Carroll (1993) organiza as dezenas de habilidades cognitivas especializadas, identificadas na literatura psicológica educacional, em um primeiro nível hierárquico, denominado estrato I. Organiza um conjunto menor de habilidades cognitivas amplas em um segundo nível hierárquico, denominado estrato II, e coloca no terceiro nível hierárquico, denominado estrato III uma habilidade cognitiva geral. A ideia básica é que as habilidades cognitivas de um estrato controlam o funcionamento das habilidades cognitivas do estrato hierarquicamente abaixo. Essa investigação se situa, de forma ampla, na área temática que contempla a relação entre a aprendizagem e as habilidades cognitivas amplas.

A relação entre a habilidade cognitiva geral ou inteligência e o desempenho dos estudantes na escola tem sido objeto de investigação e debate há mais de um século, tanto nos Estados Unidos quanto na Europa e mesmo no Brasil. No Brasil, há registro de pesquisas sobre a relação entre as habilidades cognitivas e o desempenho escolar, realizadas por Helena Antipoff, já na década de 30 (CAMPOS, 2003).

Na França, o próprio desenvolvimento de um instrumento para se medir a inteligência, por Binet em 1904, é um marco inicial do estudo desta relação. De fato, para construir o seu instrumento, Binet indagou junto aos professores sobre o tipo de tarefas que eram mais difíceis para os estudantes que tinham dificuldades de aprendizagem (BINET & SIMON, 1916). O interesse de Binet era desenvolver um procedimento que permitisse separar os estudantes considerados normais daqueles com dificuldades de aprendizagem, classificados por ele, como mentalmente retardados (STERNBERG, 2000).

Nos Estados Unidos, Terman, baseou-se no trabalho de Binet para desenvolver as conhecidas escalas de inteligência Stanford-Binet, (TERMAN & MERRILL, 1937), amplamente utilizadas no ambiente escolar americano. Desde a década de 20, os estudantes americanos são intensamente avaliados por testes psicológicos. Talvez o mais influente resultado das pesquisas americanas sobre a relação entre inteligência e desempenho escolar seja de que os escores nos testes de inteligência são bons preditores do desempenho escolar (correlação de 0,55, ANASTASI, 1968)¹, medido pelo grau² alcançado pelo aluno.

¹ Citado por, S.J.Ceci, (1991).

² Estou traduzindo a palavra *Grade*, como grau. No sistema educacional americano, *Grade* é uma avaliação ordinal do desempenho escolar. Normalmente os graus são atribuídos na forma de letras A, B, C, D, E ou F, sendo A o maior grau e F o menor grau. Eventualmente, é necessário calcular o grau médio do aluno (grade point average) na série ou no curso. Para fazer isto, substituem-se os conceitos nominais por numéricos, por exemplo, A=5, B=4... e F=0, e obtém-se a média aritmética dos diversos graus alcançados pelo aluno nas disciplinas da série ou do curso, dependendo do caso. Como a atribuição de grau é usada universalmente nas escolas americanas, esse sistema de registro do desempenho acadêmico é considerado uma medida padronizada dos diferentes níveis de entendimento em cada disciplina ou área de conteúdo.

Em geral, os interesses que mobilizam este tipo de pesquisa decorrem de uma das seguintes posições: (a) uma visão de que a inteligência ou as habilidades cognitivas antecedem e são causas do desempenho escolar, e é exatamente por isto que os resultados dos testes de habilidades são bons preditores de desempenho escolar em qualquer disciplina; (b) uma visão de que o desempenho escolar antecede e favorece a obtenção de bons escores nos testes de habilidades cognitivas; (c) uma visão interacionista em que a habilidade cognitiva e desempenho escolar evoluem conjuntamente de forma entrelaçada em que uma antecede e é antecedita pela outra, sendo ambas causa e consequência uma da outra. Nesta última visão, as habilidades cognitivas e o desempenho escolar têm forte interação com os sistemas motivacionais, afetivos, atitudinais, volitivos e com o contexto social e cultural.

Sabe-se que muitas das habilidades do estrato I do modelo de Carroll são educáveis, pois podem ser aprendidas ou treinadas principalmente no ambiente escolar. Mas, se este fato, por si só, é educacionalmente relevante, boa parte dos educadores mantêm esperança de descobrir algumas habilidades cognitivas amplas que sejam educáveis (ADEY, 1997), visando a sua transferência e aplicação para situações e contextos inéditos. Estas duas implicações educacionais permitem afirmar a relevância educacional sobre a relação entre a evolução da aprendizagem em física e as habilidades cognitivas amplas.

A maioria dos trabalhos nesta área normalmente relata e discute os resultados de pesquisas quantitativas em que se emprega com muita frequência algum instrumento para aferir o QI do pesquisado como uma medida do seu índice de inteligência geral. As discussões são muito polarizadas, confusas e complicadas, e quase sempre dominadas pelos posicionamentos ideológicos dos pesquisadores. Mesmo nos trabalhos que seguem uma linha de investigação mais qualitativa também houve forte polarização, mas agora entre os pólos Piaget e Vygotsky. Esse cenário conflituoso levou a que a literatura educacional adotasse uma linguagem mais vaga, mas que também é menos comprometida ideologicamente, para expressar os variados graus de concordância com uma das mais antigas metas educacionais, aquela que postula ser meta da educação propiciar o desenvolvimento das habilidades intelectuais humanas, e que já aparece nos escritos de Aristóteles. De fato, Aristóteles afirmava que

Não se deve deixar no esquecimento qual deve ser a educação e como se há de educar. Nos tempos modernos, as opiniões sobre este tema diferem. Não há acordo sobre o que os jovens devem aprender, nem no que se refere à virtude nem quanto ao necessário para uma vida melhor. Tampouco está claro se a educação deveria preocupar-se mais com a formação do intelecto ou do caráter. Do ponto de vista do sistema educativo atual, a investigação é confusa e não há certeza alguma sobre se devem ser praticadas as disciplinas úteis para a vida ou as que tendem à virtude, ou as que se sobressaem do ordinário (pois todas elas têm seus partidários). (Aristóteles, s.d.)³

No cenário educacional brasileiro, ao introduzir o Exame Nacional do Ensino Médio (ENEM), na década de 90, o Instituto Nacional de Estudos e Pesquisa (INEP) utilizou o conceito de competências para se referir às metas curriculares. Competências eram definidas como

Competências são as modalidades estruturais da inteligência, ou melhor, ações e operações que utilizamos para estabelecer relações com e entre objetos, situações, fenômenos e pessoas que desejamos conhecer. As habilidades decorrem das competências adquiridas e referem-se ao plano imediato do “saber fazer”. Por meio das ações e operações, as habilidades aperfeiçoam-se e articulam-se, possibilitando nova reorganização das competências. (INEP, s.d.).

Uma linguagem similar foi adotada pela Organização para a Cooperação e Desenvolvimento Econômico (Organization for Economic Co-operation and development OECD) ao estabelecer as competências-chave como base para a definição do Programa de Avaliação de Sistemas Educacionais PISA (Programme for International Student Assessment).

Anteriormente, os conceitos de letramento⁴ (SOARES, 2004), alfabetização matemática⁵ (UK. DES, 1999) e alfabetização científica⁶ (NELSON, 1999) foram amplamente utilizadas para se referir às metas curriculares desejáveis. Todas estas

³ Aristóteles, *Política*, VIII, 1 e 2, citado no Parecer relatado por Guiomar Namó de Melo, para o conselho nacional de Educação quando da análise e aprovação das diretrizes Curriculares Nacionais para o ensino Médio (BRASIL, 1998).

⁴ Em inglês literacy.

⁵ Em inglês numeracy

⁶ Em inglês Scientific Literacy

expressões utilizadas, de alguma forma, se referem ao desenvolvimento das habilidades intelectuais humanas, principalmente aquelas habilidades que são mais facilmente utilizadas em contextos diferentes daqueles em que elas foram aprendidas ou desenvolvidas. Assim, (SOARES 2004, p.19) refere-se ao letramento “como o desenvolvimento de comportamentos e habilidades de uso competente da leitura e da escrita em práticas sociais [...]”.

Alfabetização matemática é uma proficiência que é desenvolvida principalmente em matemática, mas também em outras disciplinas. Ela é mais que uma habilidade de fazer aritmética básica. Ela envolve desenvolver confiança e competência com números e medidas. Ela requer entendimento do sistema de números, um repertório de técnicas matemáticas e uma inclinação e habilidade para resolver problemas quantitativos ou espaciais em diversos contextos. Alfabetização matemática também demanda entendimento das formas pelas quais dados podem ser coletados via contagem e mensuração e apresentados em gráficos, diagramas, quadros e tabelas. (UK. DES, 1999)

O conceito de Alfabetização científica tornou-se proeminente no cenário educacional com o lançamento do Projeto 2061, da Associação Americana para o Progresso da Ciência (AAAS), em 1986. Em 1996, a AAAS, publicou uma declaração de princípios com que se comprometia relativamente à alfabetização científica. Nos documentos constam, dentre outros, os seguintes princípios:

(1) alfabetização científica consiste no conhecimento de certos e importantes fatos, conceitos e teorias científicas, o exercício de hábitos da mente e um entendimento da natureza das ciências, sua conexão com a matemática e tecnologia, seu impacto sobre os indivíduos e seu papel na sociedade. (2) Educação efetiva para a alfabetização científica requer que cada estudante seja frequente e ativamente envolvido em explorar a natureza de forma que se assemelhe a como os cientistas trabalham (NELSON, 1999, p.2).

Talvez enunciado com uma linguagem mais vaga e sutil do que no conceito de alfabetização matemática, ainda assim o conceito de alfabetização científica envolve uma competência genérica. Isto fica mais explícito na definição operacional adotada pela OECD, na definição do referencial teórico que norteou a elaboração do PISA,(2003):

Alfabetização científica é a capacidade de usar o conhecimento científico, de identificar questões e tirar conclusões baseadas em evidência de forma a compreender e ajudar na tomada de decisão sobre o mundo natural e as mudanças feitas nele pela atividade humana (OECD, 2003).

As mudanças educacionais ocorridas nas sociedades modernas, como as adotadas anteriormente, podem ser entendidas como resultando de uma busca de *“algo estável na mente que auxilie adaptação ao longo de toda a vida”* (ADEY et al., 2007). Segundo Adey et al., (2007), este algo buscado é uma habilidade cognitiva geral, mas eles notam certa aversão no meio educacional ao uso da palavra inteligência para designar esta noção. Entretanto, segundo eles, termos tais como habilidade geral, habilidade mental geral, habilidade cognitiva geral parecem ser bem aceitos entre os educadores já que são muito comuns na literatura educacional. Eles especulam que a rejeição ao termo inteligência parece estar associada a uma concepção, que eles situam na década de 70, de inteligência como algo fixo, imutável, inata e que foi base para muitas manifestações acadêmicas de racismo. Ao defender a revisão e superação desse preconceito conceitual, (ADEY et al., 2007) argumentam que *“sem a reintrodução da habilidade cognitiva geral no discurso educacional”*, todo o trabalho de teorização sobre metas educacionais de articular referenciais teóricos e estabelecer padrões para a avaliação sistêmica da educação fracassará.

Juntamente com a reformulação de metas educacionais houve uma crescente adoção de avaliações sistêmicas da educação baseadas na idéia de mensurar as habilidades cognitivas gerais dos estudantes ao longo do processo educacional. No Brasil, por exemplo, o Sistema Nacional de Avaliação Educacional do Brasil (SAEB), mensura as competências desenvolvidas pelos estudantes nas áreas de português e matemática em três momentos distintos da escolarização compulsória: na quarta e na oitava⁷ séries do ensino fundamental e na terceira série do ensino médio. Ainda no Brasil, no final da década de 90, foi introduzido o Exame Nacional do Ensino

⁷ Estas séries estão designadas segundo o sistema de oito séries no ensino fundamental. No sistema de nove séries, usado no Estado de Minas Gerais, o SAEB seria aplicado na quinta e na nona séries do ensino fundamental.

Médio (ENEM), que visava mensurar um conjunto de 21 habilidades específicas e 5 competências gerais que se aplicavam em todas as áreas de conhecimento e que deveriam ter sido desenvolvidas pelos estudantes até o final do ensino médio. Este ENEM era multidisciplinar e contrastava-se com a forma tradicional dos exames de vestibular aplicados aos estudantes, e que focavam no conhecimento disciplinar. O programa de avaliação sistêmica PISA desenvolvido pela Organização para a cooperação e desenvolvimento econômico (Organization for Economic Co-operation and Development, OECD) procurava comparar os resultados da educação mensurando as habilidades cognitivas gerais em estudantes de 15 anos de idade nas áreas da língua materna, matemática e ciências. Inicialmente, o PISA adotava um referencial para avaliação baseada nos conceitos de alfabetização, alfabetização científica e alfabetização matemática, e mais recentemente passou a adotar a idéia de competências-chave.

Essas avaliações sistêmicas baseadas na mensuração de habilidades cognitivas gerais caminhavam na mesma direção do conhecido teste americano SAT (Scholastic Aptitude Test) que existe desde a década de 20, e que era, até recentemente, o principal teste para seleção de candidatos para o sistema universitário americano. O SAT é um teste que mensura habilidades gerais independentes de conhecimentos escolares específicos. A partir de 2001, foi desenvolvido o SAT II, um exame mais fortemente apoiado nos conhecimentos adquiridos na escola, e que a partir de 2005 começou a substituir o SAT como teste de seleção de candidatos para o ensino universitário. No Brasil, em 2008, o ENEM sofreu uma grande mudança passando a ser um exame que mensura conhecimentos escolares específicos no esforço de se conformar como um exame nacional de seleção de candidatos para as universidades brasileiras. Estes movimentos na avaliação sistêmica refletem a tensão existente no mundo acadêmico sobre o que melhor prediz o desempenho futuro dos estudantes: o seu nível de habilidade geral ou o seu nível de conhecimentos escolares adquiridos. No entanto, ambas as vertentes deixam de considerar a possibilidade de se usar as avaliações escolares ordinárias feitas nos ambientes de sala de aula como um bom preditor do desempenho futuro dos estudantes.

Essa dissertação se alinha no esforço daqueles que defendem que as avaliações escolares ordinárias devem ser tomadas mais a sério, principalmente, como dados de pesquisa educacional e em decisões sobre políticas públicas. Ao tomar as avaliações escolares ordinárias, realizadas na disciplina física ao longo do ano letivo, como indicadores da aprendizagem em física escolar dos estudantes, as avaliações serão interpretadas como indicadores de uma competência em física escolar, um conceito teórico que será introduzido para designar aquilo que os estudantes aprendem e desenvolvem nesta disciplina. Assim, analisa-se a evolução temporal da competência em física escolar durante o ano letivo como uma representação da história de aprendizagem em física escolar. A questão é relevante pelas implicações curriculares e pelo fato de ser uma investigação inédita sobre a evolução temporal de uma competência e as habilidades cognitivas amplas.

O objeto da dissertação pode ser mais bem especificado tendo em vista as duas questões norteadoras da investigação. A primeira questão investiga quais variáveis organísmicas e contextuais afetam a evolução temporal da competência em física escolar e como afetam esta evolução. Variáveis organísmicas são aquelas específicas de cada sujeito, tais como, gênero, idade, habilidades cognitivas, etc. Variáveis contextuais caracterizam o ambiente social, econômico, escolar e cultural do estudante. A segunda questão investiga se há uma relação estável e significativa entre a evolução temporal da competência em física escolar e algumas das habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll (1993), após controlar para os efeitos das variáveis organísmicas e contextuais. Esta questão é o principal objetivo da pesquisa aqui relatada.

A primeira questão só é importante instrumentalmente, por permitir controlar os efeitos das variáveis organísmicas e contextuais na segunda questão. Este controle é importante porque o efeito destas variáveis pode obscurecer a relação entre a evolução da competência em física escolar, tomado como uma representação da evolução temporal da aprendizagem em física escolar, e as habilidades cognitivas amplas.

Essa dissertação está organizada em cinco capítulos, sendo o primeiro esta introdução. O capítulo dois é o capítulo em que se apresentam os referenciais teóricos necessários utilizados para desenhar e interpretar os resultados teóricos desta investigação. O capítulo três aborda os aspectos metodológicos da pesquisa. O capítulo quatro apresenta e analisa os resultados da investigação. O último capítulo sintetiza as conclusões e discute as implicações educacionais deste trabalho.

2. REFERENCIAIS TEÓRICOS

2.1 – Introdução

Esta pesquisa foi desenvolvida no campo temática da relação entre a aprendizagem escolar e as habilidades cognitivas amplas. Para a sua realização utiliza as avaliações escolares ordinárias como fonte de dados para construir uma escala que mensura a competência em física escolar, que é tomada como uma representação da aprendizagem em física escolar desenvolvida pelos estudantes. Desta forma, pretende mostrar o valor dessas avaliações escolares como dados para construção de medidas educacionais que deveriam ser mais valorizadas e utilizadas nas pesquisas e políticas educacionais.

Especificamente esta pesquisa focaliza na construção de uma medida da competência em física escolar a partir das avaliações trimestrais ordinárias realizadas nessa disciplina, como base para a investigação da possível relação entre a evolução temporal da competência em física escolar ao longo do ano letivo e as habilidades cognitivas amplas do estrato II do modelo de Carroll. A próxima seção revê como o conceito de competência é utilizado na literatura e introduz a noção de competência em física escolar. Na terceira seção será apresentado o modelo da estrutura das habilidades intelectuais humanas desenvolvido por Carroll (1993), e a bateria de testes BaFaCaLo, desenvolvida por Gomes (2010) para mensurar as habilidades desse modelo. Na quarta seção será discutida a importância educacional das habilidades cognitivas e de sua mensuração. Na quinta seção será discutido o uso de avaliações escolares ordinárias como dados de medida educacional. Na sexta seção será discutido o modelo Rasch, que é usado nesta pesquisa para modelar as avaliações escolares ordinárias e construir a escala de competência em física escolar. Na sétima seção será apresentado o modelo de regressão múltipla multinível que é usado nesta pesquisa para modelar a evolução temporal da competência em física escolar e investigar a sua relação com as habilidades cognitivas amplas.

2.2 – Os conceitos de competência e a competência em física escolar

A entrada do termo Competência no meio educacional brasileiro veio principalmente através do parecer que fundamenta e recomenda a aprovação das Diretrizes Curriculares Nacionais para Ensino Médio, Brasil (1998). Neste longo documento, o conceito de competência é usado de forma que permite múltiplas significações. Ora aparece como Competências Gerais e Amplas, ora aparece como Competências Cognitivas e Sociais e ora como Competências Específicas por área de conteúdos. Como historia um dos protagonistas da reforma Curricular do Ensino Médio,

Quando da elaboração da proposta de Diretrizes Curriculares Nacionais a ser encaminhada ao Conselho Nacional de Educação e dos Parâmetros Curriculares Nacionais para o Ensino Médio e dos Referenciais Curriculares para a Educação Profissional, fizemos a opção de trabalhar a partir do conceito de competências. Entretanto, reconstruímos o conceito que vinha sendo utilizado, quer pela recente tradição anglo-saxônica, quer pela tradição francesa, aproximando-nos mais desta última. Tínhamos como referências básicas a epistemologia genética de Jean Piaget e a lingüística de Noam Chomsky (RUY BERGER, 1998).

No cenário internacional, as Competências passam a dominar a linguagem curricular principalmente devido à atuação da Organização para a Cooperação e Desenvolvimento Econômico (OECD), que desenvolveu, entre 1997 e 2003, um longo esforço para esclarecer conceitualmente e identificar as chamadas Competências-chave. O esforço desenvolvido pela OECD para delimitar e operacionalizar um conceito de Competência envolveu a organização de dois simpósios internacionais, um em 2001 e outro em 2003. Além da contratação de vários estudos específicos. Um destes foi o estudo feito por Weinert (1999) que reviu uma ampla literatura sobre as bases teóricas do conceito de competência como usado nas ciências do desenvolvimento, na psicologia, na lingüística, na sociologia, nas ciências políticas e na economia.

Em sua revisão, Weinert identificou nove diferentes tradições teóricas nas quais o conceito de competência era definido, descrito ou interpretado teoricamente. Essas nove diferentes tradições teóricas são: (a) Competência enquanto habilidade cognitiva geral; (b) Competência enquanto habilidades cognitivas especializadas; (c) O modelo de competência performance de Chomsky; (d) Modificações do modelo

competência performance; (e) Competência como tendências de ação motivada; (f) Competência como auto conceitos objetivas e subjetivas; (g) Competências de ação; (h) Competências-chave e (i) Metacompetências.

A primeira das tradições teóricas identificadas por Weinert concebe competência como habilidade cognitiva geral, e é aquela ligada ao estudo da inteligência. Ele identifica três abordagens teóricas dentro desta tradição, a abordagem psicométrica (CARROLL, 1993), a abordagem de processamento de informação e abordagem piagetiana. A primeira entende inteligência, que Weinert identifica como competência, como um sistema de habilidades e atitudes mais ou menos independente do contexto e do conteúdo. Na abordagem de processamento de informação, a inteligência ou competência geral, segundo Weinert, é entendida como uma máquina de processamento de informação cujos componentes (velocidade de processamento, capacidade de memória de trabalho e capacidade de processamento) permitem a aquisição de uma enorme variedade de conhecimentos específicos e habilidades. A abordagem piagetiana vê os processos ontológicos e psicológicos de adaptação organizada produzindo uma sequência universal de estágios de desenvolvimento que “leva a conhecimentos cada vez mais flexíveis e abstratos e as competências de ação que também são adaptações às condições do ambiente concreto” (WEINERT, 1999, p.6).

A segunda tradição teórica identificada por Weinert concebe a existência de um conjunto de competências cognitivas especializadas que são

conglomerados de pré-requisitos cognitivos que devem estar disponíveis para um indivíduo desempenhar-se bem em uma área de conteúdo particular (por exemplo, jogar xadrez, tocar piano, dirigir automóvel, resolver problemas matemáticos, etc.) (WEINERT, 2009, p.7).

Essas competências cognitivas especializadas são aprendidas ao longo da vida, requerendo um longo processo de aprendizagem, uma ampla base experiencial, o desenvolvimento de um profundo entendimento da área temática da competência e o desenvolvimento de rotinas de ação automatizadas que possam ser conscientemente controladas.

A terceira tradição teórica cresce em torno da distinção feita por Noan Chomsky, nas décadas de 1950 e 1960, entre competência e performance. Segundo Weinert,

Chomsky entende a competência lingüística como uma habilidade universal, hereditária e modularizada para adquirir a língua materna. A competência lingüística - um conjunto básico e inato de elementos cognitivos, princípios e regras lingüísticos abstratas - distingue-se da performance que inclui a habilidade de criar e entender uma variedade infindável de sentenças novas e gramaticalmente corretas. Nessa mesma tradição Weinert inclui um conjunto de modelos que incorporam uma ou mais das seguintes idéias: (i) especificidade do domínio; (ii) sistema inato de princípios e regras modularizadas (para fins específicos); (iii) aprendizado governado por regras; (iv) a performance cuja qualidade depende não somente de princípios universais, mas também das experiências de aprendizagens e do contexto situacional atual do aprendente. Assim nesta mesma tradição aparecem autores como Hymes, Habermas, Weber & Westmeyer.

A quarta tradição, é aquela que modifica a idéia básica do modelo Chomskyano, seja pela introdução de um mecanismo cognitivo que faz o papel de moderador entre a competência e a performance, ou pela diferenciação da competência em componentes. Um exemplo da primeira abordagem é o modelo de Overton (1985), no qual são citados como exemplos de moderadores, o estilo cognitivo, a capacidade de memória, a familiaridade com a situação da tarefa e outras variáveis relativas às diferenças individuais. Um exemplo da segunda abordagem é um modelo desenvolvido por Gelman & Geeno, (1989), em que a competência é diferenciada em três componentes: (1) a competência conceitual, (2) a competência procedimental e (3) as competências de performance.

A quinta tradição é aquela ligada à concepção de competência como um conceito motivacional. Esta abordagem foi desenvolvida por White, (1959), ele argumentava que *“é necessário tornar competência um conceito motivacional, isto é, há uma competência motivacional bem como uma competência no seu sentido mais familiar de capacidade realizada”*. Weinert engloba nesta tradição os trabalhos ligados às teorias de motivação baseadas na idéia de auto-eficácia. Nestes modelos, um sistema de conhecimentos e crenças é formado a partir das experiências pessoais com sua própria competência em situações concretas de realização. E este sistema de conhecimentos e crenças influencia as realizações e performance do sujeito por meio de suas expectativas, atitudes e esquemas interpretativos.

A sexta tradição é aquela que distingue competência objetiva, performance e disposições para performance que podem ser medidas por escalas e testes padronizados, da competência subjetiva, uma avaliação subjetiva das capacidades relevantes para a performance e das habilidades necessárias para executar tarefas e executar problemas (STERNBERG & KOLLIGINIAN, 1990).

A sétima tradição é aquela que concebe a competência como competência de ação “que inclui todos aqueles pré-requisitos cognitivos, motivacionais e sociais necessários e/ou disponíveis para aprendizagem e ação bem sucedidas”. (WEINERT, 1999, p.10). Ainda segundo Weinert,

o construto teórico de competência de ação combina compreensivelmente, em um sistema complexo, todas as habilidades intelectuais, os conhecimentos específicos por conteúdo, as habilidades cognitivas, as estratégias específicas por domínios, as rotinas e sub rotinas, as tendências motivacionais, os sistemas de controle volitivo, as inclinações baseadas em valores pessoais e os comportamentos sociais.

A oitava tradição é aquela que introduz o conceito de competências-chave que são competências independentes do contexto e que são equivalentes em seus usos e em suas efetividades em diferentes instituições, diferentes tarefas e sob diferentes condições de demanda. São exemplos de competências-chave: competências basais (por exemplo, cálculo mental aritmético, alfabetização e educação geral); competências metodológicas (por exemplo, capacidade de planejar a solução de problemas, habilidades ligadas a computadores, capacidade de usar diversas mídias); competências comunicativas (por exemplo, habilidades com línguas estrangeiras, retórica, habilidades de exposição escrita e oral) e competências de julgamento (por exemplo, habilidades de pensamento crítico, avaliação sobre a sua performance e sobre a de outros).

Finalmente, a nona tradição identificada por Weinert, focaliza os conceitos de metacompetências, uma generalização conceitual a partir da idéia de metaconhecimento. Metaconhecimento captura as habilidades que uma pessoa tem em julgar sobre o que sabe e o que não sabe, em julgar sobre seus pontos fortes e fracos intelectualmente falando, em julgar sobre o que sabe fazer para usar as habilidades e conhecimentos de que dispõe para resolver uma variedade de tarefas e adquirir competências que não tem, para julgar desde o início se tem ou não

chance de sucesso na solução de uma tarefa. A pessoa não é apenas capaz de avaliar as suas possibilidades na performance e seus pré-requisitos, mas também é capaz de utilizar esses julgamentos subjetivos para dirigir suas ações.

Metacompetências inclui conhecimento e experiência sobre as dificuldades das diferentes tarefas, conhecimentos sobre suas próprias capacidades, talentos, conhecimentos, habilidades e deficiências cognitivas; conhecimento sobre a aprendizagem, resolução de problemas, estratégias efetivas para aprender, lembrar, resolver problemas e tirar dúvidas; conhecimento sobre técnicas para dominar diversas tarefas com as competências cognitivas disponíveis para compensar a falta de conhecimento e para se propor metas realísticas. Todas essas características das metacompetências são conhecidas como metacompetências declarativas. As metacompetências procedimentais incluem as estratégias para organizar tarefas e problemas tornando-os fáceis de resolver e o uso efetivo de ferramentas e dispositivos cognitivos (gráficos, pinturas, títulos, analogias), aplicação dos recursos cognitivos de forma relevante na tarefa e no nível apropriado de dificuldade, o registro e a avaliação contínua do progresso da performance.

A partir da análise dessas nove tradições teóricas, Wienert propõe sete diferentes acepções teóricas para o conceito de competência. A primeira é competência enquanto uma disposição psicológica geral, os recursos cognitivos gerais que uma pessoa possui para realizar tarefas desafiadoras em diferentes conteúdos, para adquirir os conhecimentos necessários e para alcançar uma alta performance. Nesta acepção teórica, competência é essencialmente de natureza cognitiva, porém, ampla, aplicando-se a todos os domínios de conteúdo.

A segunda acepção teórica é competência enquanto uma disposição específica para a performance. Ela inclui todas as capacidades e habilidades cognitivas, conhecimentos, estratégias e rotinas necessárias para realizar os diferentes tipos de tarefas. Nessa acepção teórica, competência é de natureza cognitiva e específica por domínio.

Na terceira acepção teórica, a competência é concebida em termos de motivação, e não, de cognição. A competência é vista como uma avaliação subjetiva de seus

próprios recursos de performance (auto-conceito e auto-eficácia) e de suas tendências motivacionais para a ação.

Na quarta acepção teórica, competência é concebida como competência de ação, um conceito que combina aspectos cognitivos e motivacionais relacionados às metas, demandas e tarefas de um contexto de ação particular, por exemplo, uma profissão.

A quinta acepção entende a competência como competências-chave, que podem ser usadas para alcançar boa performance em uma ampla variedade de situações distintas.

Na sexta acepção teórica, a competência é entendida como metacompetências, que se referem a conhecimentos atributos, motivacionais, habilidades volitivas que permitem que os recursos sejam usados eficientemente em diferentes tarefas, em diferentes áreas de conteúdos e para diferentes propósitos.

Na sétima acepção teórica, prevalente em algumas áreas das Ciências Sociais, tal como a economia, as competências individuais são consideradas apenas mais um dos recursos humanos necessários para o desenvolvimento de uma instituição, de uma economia, de ou de uma sociedade.

Por mais abrangente que tenha sido a revisão feita por Weinert sobre a literatura ligada ao conceito de competência, não se pode negar que ao final todas as suas acepções são mais relevantes para o campo psicológico do que para os demais. Neste sentido, seu trabalho deixa de contemplar de forma mais incisiva o mundo do trabalho e suas demandas. Não se deve esquecer que a força motriz que levou a linguagem de competência para a linguagem curricular vem das pressões do campo do trabalho e do campo econômico sobre o campo educacional, como já identificou Casemiro Lopes (2001) ao afirmar que

No momento atual, em diferentes países no mundo ocidental, o conceito de competências tem configurado as reformas curriculares. Não que exista um discurso homogêneo em todas essas reformas. Há sempre recontextualizações locais nos diferentes países, produzidas pelas interseções entre diretrizes de órgãos de fomento internacionais, dos órgãos de governo locais e de países com os quais estabelecem relações de intercâmbio, em busca de legitimação e de acordos, bem como por interseções com os campos de controle simbólico e de produção.

A OECD, por sua vez, interessava em atender a estas demandas externas. Em um artigo de discussão, proposto como uma preparação para o segundo Simpósio Internacional no âmbito do projeto *DeSeCo (Defining and Selecting Key Competencies)*, Rychen & Salganik (2002), se propõem a consolidar um entendimento sobre o conceito de competência. Propõem uma concepção para atender tanto aos aspectos externos quanto aos aspectos internos. Do ponto de vista externo, competências devem ser vistas como habilidades para proporcionar resultados externos, desempenhos e comportamentos facilmente observáveis e detectáveis e de interesse do mundo econômico e do trabalho. Sinteticamente propõem que *“uma competência é a habilidade de satisfazer com sucesso uma demanda complexa ou realizar uma atividade ou tarefa complexa.”*

No entanto, Rychen & Salganik (2002) reconhecem que essa definição precisa ser complementada para incluir uma concepção de competência enquanto competência interna do indivíduo. Neste ponto, acatam o argumento de Witt & Lehman (2001) de que sem uma visão de competência como uma estrutura interna ao indivíduo seria fácil identificar uma nova competência bastando para isto descrever uma demanda ou uma tarefa do mundo do trabalho antecedida da expressão *“uma capacidade para”*. Desta forma, reconhecem que uma competência inclui os componentes cognitivo, emocional, motivacional, social e comportamental, aproximando a conceituação proposta do modelo de Competência de Ação descrito por Weinert (1999). Desta forma esclarecem que

é a demanda, a tarefa ou atividade que define a estrutura interna de uma competência, incluindo as atitudes, valores, conhecimentos e habilidades inter relacionadas que juntas tornam possível a ação efetiva. Competências não podem ser reduzidas às suas componentes cognitivas.

Por mais atraente que esta definição possa ser, ela traz consigo o risco de uma regressão infinita de competências: todo ser humano realiza uma infinidade de tarefas e atividades ao longo de sua vida, significando que ele deveria ser caracterizado por igual número de competências. Uma visão mais parcimoniosa consiste em admitir que a estrutura interna de uma competência é determinada por um domínio temático com as suas demandas, tarefas e atividades. Tal restrição atenderia ao princípio de Humboldt, identificado por Witt & Lehman (2001, p.5) que consiste em realizar infinitamente usando meios finitos.

Nesta investigação, é introduzido o conceito de competência em física escolar, que entendemos como uma competência para a ação especializada para o domínio de conteúdo da física escolar do ensino médio. Neste sentido, competência em física escolar é uma amálgama de capacidades, habilidades, conhecimentos, valores, atitudes e crenças relativas ao domínio da física escolar de ensino médio, sendo inseparáveis os aspectos cognitivos, afetivos, motivacionais, volitivos, sociais, atitudinais e comportamentais.

Seria possível cair na tentação de fragmentar a competência em física em diversas componentes, tais como competência em mecânica, competência em termodinâmica, competência em física moderna, ou ainda, em componentes mais especializadas como, por exemplo, competência em interpretar gráficos, competência em realizar observações experimentais, em realizar medidas, etc. Se adotada tal postura, correr-se-ia o risco de termos de caracterizar as competências do sujeito no domínio da física escolar por um conjunto inumerável de competências, uma para cada tarefa ou demanda que conseguíssemos identificar. A opção por trabalhar com um único conceito de competência em física escolar decorre da crença de que o princípio de Humboldt vale tanto para o ser humano quanto para as ciências.

2.3 – O modelo da estrutura das habilidades intelectuais humanas, proposto por Carroll

Existem várias tradições de pesquisa em torno do termo inteligência. Dentre elas destacam-se a tradição psicométrica, a tradição da psicologia do desenvolvimento e a das ciências cognitivas. A tradição psicométrica contenta-se em criar modelos descritivos das relações detectadas através da matriz de correlação entre os escores de testes desenvolvidos para mensurar as habilidades cognitivas. Essa tradição se inicia com Francis Galton no século XIX, mas ganha muito impulso a partir dos trabalhos de Charlis Spearman nas primeiras décadas do século XX. Os expoentes desta tradição são Thurstone, Guilford, Cattell, Horn, Carroll.

A tradição da psicologia do desenvolvimento em geral adota uma metodologia qualitativa e um paradigma genético. Os principais expoentes dessa tradição são

Piaget, Vygotsky e Fischer. Essa tradição se preocupa essencialmente em traçar como um ser humano partindo de um conjunto limitado de recursos cognitivos inatos desenvolve-se e adquire na fase de jovem adulto um conjunto exuberante de habilidades intelectuais amplas, e também, especializadas.

A tradição das ciências cognitivas é um campo interdisciplinar da inteligência e da mente encampando pesquisas nas áreas de psicologia, filosofia, inteligência artificial, neurociências e antropologia. Esta tradição se estruturou em meados da década de 70, mas as suas origens remontam a década de 50 com o pioneiro trabalho de George Miller, que mostrou a limitada capacidade de processamento de informação do ser humano. Desta época, também, surge a rejeição de Chomsky das hipóteses behavioristas sobre a linguagem como um hábito aprendido. A partir daí, se desenvolve o paradigma dominante nesta tradição de pesquisa que concebe a mente como um sistema de processamento simbólico. Como expoentes desta tradição destacam-se Herbert Simon, Allen Newel, Johnson-Laird, Anderson e Demetriou. Nesta tradição as pesquisas em geral buscam propor mecanismos modeláveis, via programas de computadores, que explique o funcionamento da mente.

Nesta pesquisa, investiga-se a relação entre as habilidades cognitivas do estrato II do modelo de Carroll com a evolução da competência em física escolar. Por esta razão, a seguir se detalha um pouco mais sobre a tradição psicométrica e o modelo de Carroll.

A pesquisa em psicometria começa, via de regra, hipotetizando-se uma habilidade cognitiva específica. Em seguida busca-se identificar conjuntos de tarefas distintas e específicas cuja resolução esteja essencialmente relacionada à ativação desta habilidade cognitiva específica. Com este conjunto de tarefas compõem-se testes que contenham tarefas da mesma natureza, por exemplo, um teste pode conter trinta itens e cada item consiste de uma multiplicação de um número de dois dígitos por um número de um dígito. Em seguida, deve-se construir uma escala para mensurar esta habilidade. Há duas estratégias possíveis para isto: (1) mensurar a habilidade usando três ou mais testes ou (2) mensurar a habilidade usando apenas um teste marcador da habilidade.

No primeiro caso, a análise fatorial (um procedimento de análise estatística multivariada que não será explicada aqui) da matriz de correlação entre os escores dos testes permite definir se os testes de fato mensuram a habilidade pretendida. Se for este o caso, a matriz de correlação dos testes pode ser reproduzida supondo a existência de uma variável latente, portanto, não observável diretamente, que é a habilidade identificada. O escore de cada sujeito na habilidade pode ser determinado somando-se seus escores nos diversos testes. A carga fatorial de um teste representa a correlação entre os escores na habilidade (escores fatoriais) e os escores no teste. Se um teste tem uma carga fatorial alta (acima de 0,7) ele pode ser considerado um teste marcador da habilidade.

A segunda estratégia para mensurar a habilidade consiste em igualar o escore na habilidade ao escore no teste marcador. Neste caso, a habilidade é uma variável manifesta, portanto, diretamente observável, não uma variável latente, como é na estratégia anterior.

Uma vez identificada em diversos domínios cognitivos um conjunto de habilidades cognitivas específicas e tendo determinado os escores nas habilidades de uma amostra da população, prossegue-se calculando a matriz de correlação entre os testes. Essa matriz de correlação contém um conjunto de relações lineares que os testes guardam entre si. A hipótese psicométrica básica é que este conjunto de relações lineares pode ser mais parcimoniosamente explicado se suposta a existência de certas variáveis latentes, linearmente associadas às habilidades específicas. A técnica para identificar as variáveis latentes é análise fatorial exploratória ou a análise fatorial confirmatória. A primeira técnica é adequada quando não se tem um modelo hipotético que mostre como as variáveis latentes se relacionam com as variáveis manifestas e a segunda técnica é empregada quando este modelo já foi identificado em análises fatoriais exploratórias prévias ou deduzido de alguma teoria psicológica ou cognitiva.

Em ambas as técnicas há a possibilidade de admitir-se, usando o procedimento de rotação oblíqua que as variáveis latentes identificadas sejam correlacionadas entre si. Neste caso é possível partir-se da matriz de correlação das variáveis latentes

identificadas para se determinar variáveis latentes de ordem mais alta que podem explicar as relações entre as variáveis latentes da ordem inferior.

A FIG.1 mostra um modelo com dois estratos de variáveis latentes e um estrato de variáveis manifestas. As dez variáveis, nomeadas de X1 a X10, são manifestas, isto é, diretamente observadas: o escore de cada pessoa é determinado pela aplicação de um instrumento de medida, usualmente um questionário ou um teste. Se o instrumento é um teste indicador de uma habilidade cognitiva, as variáveis manifestas correspondem às variáveis cognitivas específicas do estrato I. As correlações entre os escores dessas variáveis manifestas podem ser descritas de maneira mais sucinta admitindo-se que em cada sujeito há uma estrutura cognitiva latente, não diretamente observável, composta de três habilidades amplas do estrato II (F1, F2 e F3). Finalmente, as relações entre as habilidades amplas do estrato II podem ser explicadas se admitida a existência de uma habilidade ainda mais ampla em um estrato III (g).

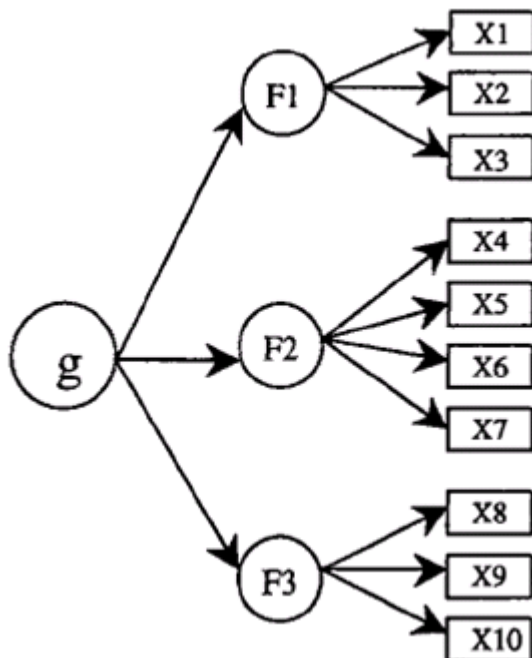


Figura 1- Exemplo hipotético de um modelo hierárquico com três estratos de habilidades

John Carroll realizou, no final da década de 80, uma extensa reanálise de um conjunto de 461 matrizes de correlação entre testes psicológicos identificadas pela literatura da área no período entre 1927 e 1987. A partir dessa reanálise, Carroll

desenvolveu um modelo hierárquico com três estratos de habilidades para representar a estrutura geral das habilidades intelectuais humana.

No estrato III do modelo de Carroll, o de nível hierárquico mais alto, há apenas uma habilidade ampla e geral, que ele identifica com o fator g, muito semelhante ao fator g de Spearman⁸. O estrato II, ou Intermediário, é composto por oito habilidades amplas, que ele identificou como inteligência fluida (Gf), inteligência cristalizada (Gc), memória geral e aprendizagem (Gy), percepção visual ampla (Gv), percepção auditória ampla (Gu), habilidade de recuperação ampla (Gr), rapidez cognitiva ampla (Gs) e velocidade de processamento (Gt). Finalmente no estrato I, hierarquicamente, o mais baixo, há um conjunto de 69 habilidades especializadas, tais como: conhecimento léxico, entendimento de leitura, atitude para a língua estrangeira, habilidade de escrita, habilidade auditiva, e etc. A figura 2 mostra uma representação gráfica do modelo proposto por Carroll. Desde a publicação do livro de Carroll, em 1993, em que ele expõe os resultados de sua reanálise e o seu modelo de estrutura das habilidades intelectuais humanas, os estudos sobre essa teoria e a sua aceitação passaram a ser um consenso entre os psicólogos (FLANAGRAN & MCGREW, 1997; MARQÑON & PUEYO, 1999; MCGREW, 1998).

Recentemente, McGrew (2009), propôs um modelo de habilidades intelectuais humanas, que compatibiliza o modelo de Cattell- Horn e o modelo de Carroll, e propôs que este modelo integrativo, denominado modelo Cattell-Horn-Carroll (CHC), seja considerado como um consenso na taxonomia das habilidades humanas. Uma tarefa importante para a psicometria, segundo McGrew, é validar baterias de testes cognitivos que mensurem as habilidades cognitivas do estrato II e III do modelo CHC.

⁸ John Spearman é o autor que forneceu as bases para psicometria a partir da primeira teoria da inteligência (teoria do fator g ; HOGAN, 2006) baseada na análise fatorial (ANASTASI & URBINA, 2000; PASQUALI, 2001). Segundo sua teoria a inteligência poderia ser definida através de um fator simples denominado fator g (ANASTASI & URBINA, 2000), além de existirem vários fatores específicos (fator S).

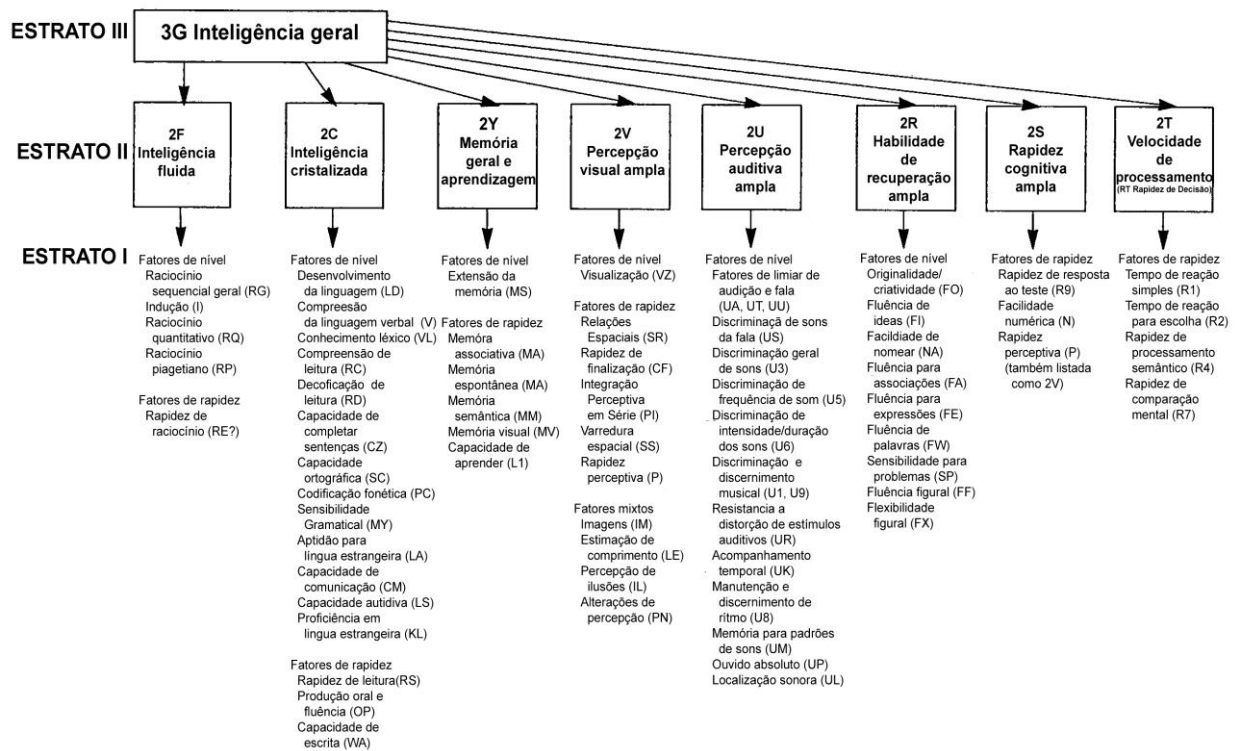


FIGURA 2 - Representação gráfica do modelo de habilidades cognitivas proposto por Carroll

A Bateria de Fatores Cognitivos de Alta Ordem (BaFaCaLo) é composta de 18 testes cognitivos do tipo lápis e papel, tendo sido desenvolvida por Gomes (2010), com o propósito de mensurar, em uma amostra brasileira, uma parte considerável de fatores cognitivos do segundo estrato de Carroll (1993). A bateria baseou-se em dois modelos: o modelo dos três estratos de Carroll (1993) e o modelo Cattell-Horn-Carroll (GOMES, 2010). Em ambos os modelos é argumentado que as habilidades cognitivas são organizadas em uma disposição hierárquica de níveis distintos em que cada um dos níveis correspondendo a diferentes graus de abrangências e especialização. Da bateria de testes, foram obtidas seis das oito habilidades cognitivas do segundo estrato de Carroll (GOMES, 2010). A mensuração das habilidades cognitivas especializadas do estrato I utiliza a estratégia de teste marcador: cada um dos testes da bateria foi cuidadosamente escolhido dentre os testes que carregam em cada habilidade especializada de modo a funcionar como um teste marcador da habilidade. Desta forma o escore em cada teste é o escore na habilidade cognitiva especializada do estrato I do modelo de Carroll.

O QUADRO 1, apresenta as dezoito habilidades cognitivas especializadas mensuradas pela bateria BaFaCaLo. Neste quadro a primeira coluna nomeia a habilidade, a segunda, descreve resumidamente o conteúdo da habilidade e a terceira apresenta o tipo de tarefa que compõe o teste marcador da habilidade. O detalhamento sobre a metodologia de aplicação e o tamanho de cada teste pode ser encontrado em Gomes (2010).

QUADRO 1
Habilidades cognitivas especializadas do estrato I do modelo de Carroll mensuradas pela bateria BaFaCaLo

Habilidade cognitiva do estrato I	Conceituação	Tipo de Tarefa usada no teste
Raciocínio Geral (RG)	Habilidade para raciocinar indutivamente ou dedutivamente com materiais baseados nas propriedades e relações matemáticas	Raciocínio Aritmético; Série de números; Aptidão matemática.
Raciocínio lógico (RL)	Habilidade para raciocinar e tirar conclusões a partir de condições dadas, ou premissas, com vários tipos de estímulos materiais- verbal, figural, etc.	Raciocínio verbal; Avaliação lógica de silogismos sem sentido
Indução (I)	Habilidade de induzir uma função ou um conceito ou outra característica comum subjacente a um dado conjunto de materiais de estímulo.	Matriz progressiva Série de letras Analogia verbal
Compreensão Verbal (VI) (VII) e (VIII)	Desenvolvimento geral da linguagem nativa em testes impressos, que exigem habilidade de leitura	Compreensão de leitura e testes de vocabulários
Memória associativa (MA1) e (MA2)	Habilidade de recordar uma parte de um par previamente aprendido, mas outras independentes dos itens quando a outra parte do par é apresentada.	Testes associativos em pares, por exemplo, número e palavra, primeiro e último nome
Memória Visual (MV)	Memória por imagens ou configurações visuais, não é capaz de ser codificada pela modalidade não visual, exemplo designer geométrico	Estudo de 40 figuras geométricas por 1 minuto e então, reconhecer essas figuras em um grupo de 80.
Flexibilidade de Fechamento (CF)	Velocidade de detecção e de desentrelaçamento a uma série de estímulos conhecidos a partir de uma matriz mais complexa	Copiando figuras ocultas
Vizualização (Vz)	Habilidade para aprender e manipular padrões espaciais ou visuais, geralmente envolvendo duas ou três dimensões	Rotação de cartas; Rotação de blocos; Tarefas de dobrar papéis;
Velocidade Perceptiva (P1), (P2) e (P3)	Velocidade de fazer comparações corretas de símbolos ou padrões em um campo visual, às vezes com estímulos dispersos	Localizar símbolos dentro de um pequeno tempo limite
Rapidez numérica	Mede a habilidade de reconhecer	Rapidez em resolver uma

(N)	números e suas ordens e para executar funções básicas operacionais tais como adição, subtração, multiplicação e divisão com velocidade e precisão em um curto limite de tempo	operação básica de aritmética
Fluência Figural (FF)	Habilidade de produzir uma variedade de desenhos simples ou esboços com um limite de tempo	Desenhar diferentes decorações em esboço de objetos comuns
Fluência Ideacional (F1) e (F2)	Velocidade de pensamento ou de recordação de idéias de experiências comuns em uma cultura	Pensar coisas que são brancas em um tempo limite

Cada grupo de três ou quatro testes, dentre os 18 testes, estão relacionados a uma das habilidades cognitivas do estrato II do modelo de Carroll. O QUADRO 2 mostra a relação entre as habilidades cognitivas do estrato II e do estrato I.

QUADRO 2
Relação entre as habilidades cognitivas do estrato II e do estrato I do modelo de Carroll, mensuradas pela bateria BaFaCaLo

Habilidades cognitivas do estrato II	Habilidades cognitivas do estrato I
Inteligência Fluida (Gf)	I; RL; RG e N
Inteligência Cristalizada (Gc)	V1; V2 e V3
Inteligência Visuo Espacial (Gv)	Vz; CF; I; MV; P2; P3; N e RL
Fluência (Gr)	FI1; FI2 e FF
Memória de Curto Prazo (Gy)	MV; MA1 e MA2
Rapidez Cognitiva (Gs)	P1; P2; P3 e N
Fator g	Gf, Gy, Gc, Gv, Gr, Gs

A inteligência fluida (Gf), segundo Gomes e Borges (2009), “é entendida como a capacidade das pessoas em aprender coisas novas que requerem relações lógicas e abstratas, assim como a capacidade em resolver problemas onde o conhecimento prévio é incipiente”.

A Inteligência cristalizada (Gc), “é entendida como a capacidade das pessoas em aprender coisas que requerem prioritariamente a utilização do seu conhecimento prévio, assim como capacidade em resolver problemas onde o conhecimento adquirido sustenta razoavelmente a resolução do problema” (Gomes, 2009).

Inteligência visuo espacial (Gv), “é genérica e ampla, correspondendo a um tipo de processamento geral capaz de produzir, registrar, relembrar transformar imagens e sensações visuais” (Gomes, 2009).

Fluência (Gr), “é a habilidade de armazenar e consolidar novas informações na memória de longo prazo e mais tarde recuperar a informação armazenada por associação” (FLANAGRAN & HARRISON, 2005).

Memória de curto prazo (Gy), “é habilidade de apreender e manter na consciência os elementos de informações das situações imediatas” (eventos que ocorreram na escala de segundos ou de unidades de minutos). (FLANAGRAN & HARRISON, 2005).

Rapidez cognitiva (Gs), “é a habilidade de executar automática e fluentemente tarefas relativamente fáceis ou tarefas cognitivas mais complexas, porém já aprendidas, especialmente quando é demandada alta eficiência mental” (atenção e concentração focal) (FLANAGRAN & HARRISON, 2005).

2.4 – Implicações educacionais de um modelo de habilidades intelectuais humanas

Como já foi dito na introdução, existem três tipos de interesse que mobilizam a pesquisa sobre habilidade cognitiva e desempenho acadêmico. O primeiro vê as habilidades cognitivas como antecedentes e causa do desempenho escolar e mais à frente do desempenho nos ambientes profissional, cultural e social. A dificuldade dessa visão é que ela tende a ver a inteligência como sendo um traço latente pouco maleável, que tem servido de base a muitas posições ideologicamente comprometidas, politicamente conservadoras e, até mesmo, segregacionistas do ponto de vista étnico, social e cultural. Os dois outros interesses convergem na idéia de que as habilidades cognitivas podem e são desenvolvidas ao longo da escolarização. Divergem por um considerar a escolarização como um antecedente e causa da inteligência e o outro por professar uma visão interacionista, em que a inteligência e a escolarização antecedem e são antecidas uma pela outra, influenciando-se reciprocamente.

É sabido que inúmeras das habilidades cognitivas especializadas do estrato I, de Carroll, são educáveis, visto que resultam de aprendizagem que ocorre, talvez principalmente, ainda que não de forma exclusiva, no ambiente escolar ou ambiente de aprendizagem específico, como por exemplo, aulas especializadas em casa. São exemplos de habilidades cognitivas especializadas educáveis, o raciocínio quantitativo, o raciocínio seqüencial geral, a compreensão de leitura, o conhecimento léxico, a sensibilidade gramatical, a memória associativa, a discriminação de fonemas, a manutenção e julgamento do ritmo musical, a facilidade numérica, dentre muitas outras. Carroll, ao realizar o estudo de revisão que levou ao desenvolvimento do seu modelo de estrutura hierárquica das habilidades cognitivas humanas não considerou uma série de habilidades especializadas e vinculadas aos conteúdos escolares, por considerar que as evidências sobre a existência e

estabilidade dessas habilidades eram muito inferiores às evidências das habilidades cognitivas incorporadas no estrato I do seu modelo. Spearritt (1996) ao discutir as implicações educacionais do modelo Carroll focaliza principalmente na possibilidade do desenvolvimento das habilidades cognitivas especializadas incorporadas ao estrato I do modelo de Carroll e mesmo de algumas ainda não incorporadas.

O fato de que muitas das habilidades cognitivas especializadas são educáveis é por si só educacionalmente relevante. Apesar disso, o maior interesse dos educadores é descobrir habilidades cognitivas mais amplas que sejam mais educáveis (Adey, 1997), na esperança de que habilidades cognitivas mais amplas possam ser aplicadas de forma mais geral em situações e contextos não experienciados anteriormente. Segundo Adey et al (2007), nas últimas décadas, diversos métodos foram propostos para aumentar as habilidades cognitivas gerais. Segundo eles, esses métodos têm em comum o propósito de *“(...) proporcionar aos estudantes estímulos específicos que sejam necessários para utilizar o desenvolvimento cognitivo, os quais de outra, faltariam em seus ambientes de aprendizagem”* (p.87).

Adey et al (2007) dividem os métodos propostos de aceleração cognitiva em dois grandes tipos: métodos diretos e métodos baseados em conteúdo. Por um lado, os métodos diretos de aceleração cognitiva não exigem o suporte de nenhum conteúdo de disciplina escolar para serem implementados. Por outro lado, anunciam resultados que mostram amplo progresso dos estudantes no desempenho escolar nas diversas disciplinas do currículo. Já os métodos baseados em conteúdo também pretendem propiciar um desenvolvimento otimizado das habilidades cognitivas amplas, porém pressupõem uma intervenção educacional centrada no conteúdo específico de uma disciplina escolar, por exemplo, ciências, matemática ou filosofia. Eles anunciam resultados que mostram uma melhoria substancial do desempenho dos estudantes, não apenas na disciplina em torno da qual se deu a intervenção educacional, mas estendendo-se para outras disciplinas do currículo.

O exemplo mais significativo de programas baseados nos métodos diretos de aceleração cognitiva é o conhecido Programa de Enriquecimento Instrumental (PIE) proposto por Feuerstein (FEUERSTEIN, RAND, HOFFMAN, & MILLER, 1980). Este programa foi desenvolvido para lidar com os problemas enfrentados pelas crianças

imigrantes para Israel que tiveram experiências traumáticas ou viveram em ambientes culturalmente desfavorecidos. O leitor interessado em conhecer mais sobre este programa pode consultar Gomes (2002). Arbitman-Smith, Haywood & Bransford, (1994)⁹ relatam tamanho de efeito de 0,6 DP no desempenho do teste de Raven e em outros testes de inteligência Fluida; e em média, tamanho de 0,77 DP em testes de inteligência cristalizada apurados em uma replicação do PIE nos Estados Unidos. Já na Inglaterra, o uso do PIE em uma turma de uma escola especial produziu tamanho de efeitos de 1,22 DP em uma bateria de testes piagetianos e 1,07 DP no teste de Raven (SHAYER & BEASLEY, 1987)¹⁰.

Outro exemplo de método direto é o Programa de Treinamento do Raciocínio Indutivo proposto por Klauer (KLAUER 1989, 1990, 1997; KLAUER & PHYE, 1994). Este método foi desenvolvido com uma série de programas de treinamento na descoberta de similaridades e dissimilaridades entre objetos e relações destinados a crianças na faixa do pré-escolar e início da escolarização formal. A meta-análise de inúmeras avaliações da aplicação desses programas indica o efeito de tamanho de 0,5DP para estudantes mais velhos e 0,63DP para crianças mais novas em medidas da inteligência geral, além de evidências de boa qualidade sobre a transferência do resultado para o desempenho acadêmico (ADEY et.al. 2007).

Os principais exemplos dos métodos baseados em conteúdos são os programas de aceleração cognitiva desenvolvidos por Adey e Shayer desde a década de 80. Estes programas fundamentam-se num referencial que combina o construtivismo piagetiano, construtivismo social de Vigotsky e a ideia de metacognição. O primeiro destes programas foi o CASE (Cognitive Acceleration Through Science Education), um programa desenvolvido em 1982 e que focalizava o ensino de ciências para crianças na faixa de 12 a 14 anos de idade (faixa da escolarização ginásial). Posteriormente, foram desenvolvidos o CAME (Cognitive Acceleration Through

⁹ Citado em Adey, Csapó, Demetriou, Hautamaki & Shayer (2007).

¹⁰ Citado em Adey, Csapó, Demetriou, Hautamaki & Shayer (2007).

Mathematical Education), baseado em matemática, destinado a estudantes da mesma faixa etária do CASE; o PCAME, também baseado em matemática, mas destinado a crianças de 9 a 11 anos; o “Let’s Think!”, baseado em ciências, e no raciocínio geral, para crianças de 5 a 6 anos de idade; o “Let’s Think through Science” baseado em ciências, e destinado a crianças de 7 a 9 anos de idade; o CATE, baseado em tecnologia, destinado a crianças na faixa de 12 a 14 anos de idade; e o ARTS, destinado à crianças na faixa de 12 a 14 anos de idade e baseado em atividades de música, teatro e artes visuais.

Crianças de 12 a 14 anos que usaram o CASE por dois anos consecutivos, logo após o término do programa, apresentaram níveis de desenvolvimento cognitivo superiores aos de estudantes que não utilizaram o CASE. Além disto, três anos mais tarde, os estudantes que utilizaram o CASE tiveram desempenho superior, situado na faixa de 0,35 a 1,0 DP ao dos estudantes que não utilizaram o CASE, e isto ocorreu não apenas em ciências, mas também em matemática e em inglês (ADEY et al. 2007). O programa de intervenção baseado em matemática teve seus efeitos transferidos para ciências e inglês (ADEY et al. 2007). O leitor interessado em saber mais sobre estes programas pode consultar os trabalhos de Adey e seus colaboradores (ADEY, & SHAYER, 1994; ADEY, 2001; ADEY, SHAYER & YATES, 2001; ADEY, ROBERTSON & VENVILLE, 2002; ADEY, 2005; ADEY, CSAPÓ, DEMETRIOU, HAUTAMAKI, SHAYER, 2007; LIN, HU, ADEY & SHEN, 2003)

Outro programa de intervenção educacional para aceleração cognitiva, que é baseado em conteúdo, é o programa Philosophy for Children proposto por Lipman na década de 70 (LIPMAN, SHARP, & OSCANYAN, 1980). Este programa apresenta uma série de pequenas histórias que contêm dilemas sobre a racionalidade, ética, moral, estética, raciocínio científico e valores cívicos, com um propósito de gerar um ambiente de discussão inteligente na sala de aula. Ele é normalmente usado com crianças de 9 a 11 anos e desenvolvido em torno das disciplinas inglês (língua materna) e estudos sociais. Segundo Adey et al. (2007), há evidências de que o programa promove ganhos significativos no raciocínio lógico, ganhos que se transferem para o desempenho escolar em leitura e matemática. Ainda segundo eles, há na Europa mais de 40 propostas de programas similares em uso.

Os que professam a visão interacionista, dentre os quais a presente autora, têm esperança de que os sistemas educacionais e os responsáveis por ele, acabem por prestar atenção nessa relação sinérgica entre inteligência e escolarização. E que, em virtude disto, patrocinem reformas curriculares enfatizando as disciplinas e as atividades que promovem o desenvolvimento das habilidades cognitivas amplas. Essa dissertação, ainda que limitada em seu escopo alinha-se no esforço de chamar a atenção para a importância das habilidades cognitivas amplas educáveis, pois sem a sua reintrodução *“no discurso educacional, esforços de teorizar propósitos, propor padrões e estabelecer referências teóricas para avaliação não podem ser bem sucedidos”* (ADEY et al. 2007). Esta dissertação limita-se a investigar a existência de uma relação bem definida entre habilidades cognitivas amplas e o desenvolvimento da competência em física escolar, dentro das possibilidades proporcionadas pela base de dados produzida ao longo da investigação.

2.5 – Avaliações educacionais como dado de medida educacional

A necessidade de comparar desempenho de pessoas diferentes, educadas em ambientes distintos por professores, também distintos, seja para fins de seleção escolar ou profissional, seja para informar a tomada de decisões de políticas públicas estão bem enraizadas na história das diversas nações. Já em 115 a.C. exames escritos eram usados na China para selecionar candidatos para o serviço público. Estes candidatos deveriam mostrar proficiência na arte da música, dos arqueiros, da cavalaria, da escrita e da aritmética (DUBOIS, 1965, citado por GROUWS, 1992)¹¹. Este princípio de selecionar servidores, baseado na proficiência em testes escrito, foi importado e usado desde o século XVIII na França, nos Estados Unidos e na Inglaterra, e permanece até hoje como base de seleção de servidores públicos no Brasil contemporâneo. Talvez o primeiro registro do uso de testes escritos para informar políticas públicas tenha sido a aplicação de testes de

¹¹ GROUWS, D.A. Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics. New York, NY, England: Macmillan Publishing Co, Inc. 1992. p 14

desempenho em matemática a todas as escolas de Boston em 1845 (KILPATRICK, 1992). Esta prática sobrevive hoje em programas do tipo PISA, SAEB, ENEN, etc.

A necessidade de comparar o desempenho, as habilidades, as proficiências, as competências, de pessoas diferentes, educadas de formas diferentes, por professores diferentes, em ambientes educacionais diferentes gera um problema metodológico complicado: o estabelecimento de uma métrica única para se avaliar essas pessoas. De fato, se o desempenho, as habilidades, a proficiência ou as competências das pessoas não forem mensuradas em uma mesma escala, elas não podem ser comparadas entre si de forma minimamente objetiva.

A solução tradicional para este problema metodológico, introduzida na China antiga e mantido até os dias atuais, consiste em submeter as pessoas que se quer comparar a um mesmo conjunto de testes aplicado na mesma ocasião. Há outras soluções técnicas para o mesmo problema e que pressupõem o aproveitamento da massa de dados educacionais gerados pelas avaliações escolares ordinárias.

Se a primeira prática, a de usar exames externos, é aceitável e desejável para fins de seleção profissional pública ou privada, no ambiente escolar ela tem levado a uma intensificação da prática de avaliação desvinculada dos processos educacionais que ocorrem nas escolas. Há muitos debates sobre os possíveis benefícios ou os malefícios que este movimento de avaliação externa provoca sobre as escolas (ATKINSON & GEISER, 2009; LINN, 2009). Por exemplo, entre 1999 e o final da década passada, os candidatos ao sistema universitários eram avaliados por exames de vestibular, quase sempre específicos por instituição. Os programas desses exames de vestibular acabaram tendo mais influência nos currículos, implementados nas escolas, do que os documentos curriculares oficiais publicados pelas autoridades educacionais públicas. Desta forma, as instituições de ensino superior capturaram, sem o devido mandato político, uma função essencial do estado, qual seja a regulação do currículo escolar.

O Brasil possui milhares de sistemas educacionais relativamente autônomos: um sistema federal, 26 sistemas estaduais, e milhares de sistemas municipais de educação. Cada um desses sistemas é composto por inúmeras escolas e salas de aula. Assim, em cada ano letivo algumas centenas de milhares de professores

avaliam milhões de estudantes em diversas ocasiões ao longo do ano letivo. Também nas instituições de ensino superior dezenas ou talvez centenas de milhares de professores avaliam alguns milhões de estudantes em cada uma das disciplinas em diversas ocasiões do semestre letivo. Todos esses dados são mantidos cuidadosamente arquivados na instituição em que foram gerados, segundo determina a legislação, e são fornecidos a cada um dos estudantes em um documento que atesta o seu desempenho escolar (histórico escolar). Essa situação ocorre em outros países: os sistemas educacionais são grandes geradores de massas incalculáveis de dados sobre desempenho escolar e acadêmico dos milhões de estudantes.

Nos Estados Unidos, esses dados escolares são muito utilizados, ao lado de exames externos, tanto na seleção para cursos universitários quanto para a seleção profissional. Assim o GPA (Grade Point Averages), da “*High School*” é utilizado ao lado do SAT (Scholastic Aptitude Test), do SAT II, e do ACT (American College Testing) na seleção de candidatos para o sistema universitário americano. O GPA relativo aos cursos de graduação é utilizado tanto na seleção de candidatos para cursos de pós-graduação quanto para a seleção profissional. Há uma ampla literatura nas áreas de educação e psicologia educacional comparando as vantagens de se usar o GPA ou testes externos como preditor do desempenho acadêmico na universidade, do desempenho profissional pós-universidade e mesmo do desenvolvimento de habilidades cognitivas (Atkinson & Geiser, 2009). No Brasil, não há registro de uso semelhante das avaliações escolares para a seleção de candidatos para a universidade, mas até o início da década passada o conceito global médio, análogo ao GPA, alcançado pelo estudante durante a graduação era usado na seleção para a pós- graduação. Também ao longo da década passada o histórico escolar da graduação foi perdendo força como elemento de seleção profissional.

É difícil entender a razão deste volume de dados ser gerado no funcionamento normal das instituições escolares e ao mesmo tempo ser desprezado quer nas pesquisas educacionais quer na seleção de estudantes para os níveis mais elevados de escolarização.

Pode-se argumentar que isto se deve ao fato de que a avaliação escolar tem menor fidedignidade e menor validade que os testes externos. Testes externos são especialmente desenhados para apresentarem boa fidedignidade e satisfazerem aos critérios usuais de validade. Fidedignidade é a propriedade de um instrumento ou de um procedimento atribuir o mesmo escore a sujeitos ou objetos mensurados sempre que nas mesmas condições. Os pesquisadores precisam estabelecer a fidedignidade de seus instrumentos de mensuração para reivindicarem o direito de afirmar que há ou não há uma relação sistemática e consistente entre as variáveis mensuradas.

Neste sentido a comparação do desempenho de estudantes diferentes usando o GPA é problemática, pois os diferentes professores usam critérios diferentes e com variados graus de rigor para atribuir os conceitos aos alunos. Mesmo quando se pretende comparar estudantes de uma mesma instituição universitária usando o GPA, observa-se que as comparações são enviesadas prejudicando estudantes envolvidos em cursos mais “duros”, em que prevalece uma tradição de maior exigência e maior rigor na atribuição de conceitos (JOHNSON, 1997). Na literatura da avaliação educacional e psicológica, há diversas propostas para tornar objetivas e fidedignas tais comparações, todas compartilham a idéia de construir uma escala estável em que os diferentes sujeitos possam ser mensurados, na mesma ocasião ou em ocasiões distintas. O ganho adicional advindo da adoção de uma estratégia dessa natureza é que ela também favorece o estabelecimento da validade do instrumento ou processo de mensuração.

A validade de um instrumento ou processo de mensuração informa se ele realmente mensura o que se propõe a mensurar, o quanto nossa mensuração é influenciada pelas circunstâncias em que é feita, e, se essa mensuração não inclui acidentalmente outros fatores que não pretendia mensurar. Por exemplo, se um pesquisador pretende mensurar uma habilidade cognitiva ampla ele precisa estabelecer que seu instrumento (um ou mais testes) mensura apenas esta habilidade cognitiva ampla, e não uma amálgama dela e de outros fatores não visados, como a motivação intrínseca, por exemplo. Estabelecer a validade de um instrumento é um processo complicado, que se baseia tanto em argumentos lógicos e teóricos quanto em evidências empíricas (YOUNG & KOBRIN, 2001).

Há muitos aspectos de validade a serem considerados, e que podem ser reunidos em quatro grandes categorias: (1) validade de conteúdo, (2) validade de face, (3) validade de critério e (4) validade de construto.

(1) A validade de conteúdo deve responder à questão: o conteúdo do que se mede está adequadamente refletido no instrumento ou processo de medida? Ela também está relacionada à aceitabilidade social do instrumento ou processo de mensuração pela comunidade que o utilizará.

(2) A validade de face é estabelecida quando um grupo de pesquisadores e especialistas no campo científico, em que o instrumento ou processo de mensuração será usado, chega a um consenso de que o instrumento de fato mensura o que ele se propõe a mensurar. Ela está, portanto, ligada à aceitabilidade social do instrumento ou processo de mensuração.

(3) A validade de critério indica se uma medida feita usando o instrumento ou processo de medida prediz a performance em outro instrumento ou processo de medida independente que mensura o mesmo conceito ou um conceito similar. A validade de critério pode ser considerada como validade preditiva ou, ainda, como validade concorrente. Ela é preditiva se a medida, que se quer validar, antecede no tempo a medida da variável de critério. Ela é validade concorrente se ambas, a variável que se valida e a de critério são medidas ao mesmo tempo.

(4) A validade de construto deve responder à questão: meu instrumento ou processo de mensuração mensura realmente o que pretendo que ele mesure? Ele indica o grau de precisão do vínculo entre o conceito mensurado e o instrumento ou processo de sua medida.

Segundo Young & Kobrin (2001), o consenso atual no campo educacional é o de que o processo de validação de um instrumento de mensuração é uma forma de validação do construto. Qualquer processo de mensuração baseado em avaliações escolares ordinárias facilmente satisfaz os critérios de face e de conteúdo. De fato, dificilmente se pode argumentar que as avaliações escolares ordinárias mensuram outra coisa que não seja o desempenho escolar, a proficiência na disciplina específica avaliada, alguma forma de habilidade cognitiva específica ligada ao

conteúdo da disciplina ou a competência na disciplina avaliada. A forma como o pesquisador conceituará aquilo que a avaliação mensura depende mais dos seus argumentos teóricos e metodológicos e menos do conteúdo específico da avaliação. Pode se argumentar que as avaliações escolares ordinárias têm aceitabilidade entre os seus maiores usuários, os professores, os alunos, os pais dos alunos e mesmo pesquisadores educacionais. No entanto, ela é menos aceita entre formuladores de políticas educacionais e equipes de seleção para níveis universitários de ensino. Entre esses dois últimos grupos, em geral, as ressalvas recaem sobre o caráter mais ou menos idiossincrático da avaliação escolar do que sobre a sua validade de face ou sua validade de conteúdo.

Quanto a validade de critério, há poucos estudos no Brasil que podem reivindicar a produção de evidências sobre a validade preditiva ou concorrentes das avaliações escolares ordinárias. Esses poucos estudos, em geral, fazem relação entre as habilidades cognitivas mais amplas e desempenho escolar. (BORGES, & GOMES, 2004; PRIMI et al, 2001). Entretanto, nos Estados Unidos, há um amplo conjunto de evidências empíricas sobre a validade preditiva e concorrente das avaliações escolares, por exemplo, segundo Atkinson & Geiser (2009). Ainda que os padrões utilizados para atribuir conceitos aos alunos variem de uma escola para outra, o grau médio cumulativo (GPA) nas disciplinas da “high school” (ensino médio) provaram ser o melhor preditor global de performance do estudante na graduação universitária. Este resultado estabelece a validade preditiva do GPA. Desta forma, uma escala baseada em avaliações escolares ordinárias precisa estabelecer apenas a sua validade de construto.

As diversas alternativas de construção de uma escala de mensuração, a partir de dados escolares via modelamento matemático de resposta ao item, fornecem diversas evidências para se julgar sobre a validade de construto do processo de mensuração construído. Uma das soluções, que se propõe a construir uma escala a partir de dados escolares ordinários, está na metodologia do índice de desempenho acadêmico proposta por Johnson (1997), como uma alternativa ao uso do GPA para representar o desempenho agregado. Johnson articula a área da moderna estatística aplicada Bayesiana e as teorias de resposta ao item para propor o seu modelo (JUNKER & BRADLOW, 1997). Comentando sobre o trabalho de Johnson,

Larkey (1997) diz que o problema com o uso do GPA se insere em uma classe mais geral de problemas, que é a necessidade de comparar sujeitos diferentes desempenhando em circunstâncias distintas, e que é surpreendentemente comum. Segundo ele, o mesmo problema ocorre no golfe profissional e na avaliação de performance das companhias aéreas.

Em seu modelo, Johnson mensura os sujeitos (os estudantes) e os estímulos (as turmas que eles freqüentam) em uma mesma escala, permitindo comparar o desempenho dos diversos estudantes, independente das disciplinas ou curso que eles realizaram. Segundo Junker & Bradlow (1997), outros modelos estatísticos de teoria de resposta ao item conseguem realizar o mesmo intento, de mensurar sujeitos e estímulos na mesma escala (FISCHER & MOLENAAR, 1995; VAN DER LINDEN & HAMBLETON, 1997). Segundo eles, Hemker et al. (1997) fazem uma revisão comparativa de diversos modelos que exploram a relação entre os ranqueamentos baseados no escore total (GPA não ajustado) e os baseados em variável latente (GPA ajustado). Ainda segundo Junker & Bradlow (1997), comparar estudantes diferentes usando GPA é um exemplo do problema chamado na literatura estatística da área educacional de problema da equalização (“Equating problem”) em torno do qual há uma extensa literatura, como por exemplo, o livro de Holland e Rubin (1982).

A formulação mais geral do problema da equalização é o desenvolvimento de um procedimento operacional que permita tornar comparáveis os resultados de avaliações educacionais feitas ou em ocasiões distintas ou usando instrumentos distintos. Na literatura ligada à avaliação sistêmica, há inúmeras propostas de procedimentos para calibrar e equalizar formas distintas do mesmo teste ou testes similares, aplicados em ocasiões distintas. Alguns destes procedimentos se baseiam na teoria clássica de testes e outros, nas teorias de respostas ao item. Nesta dissertação utiliza-se um procedimento baseado no modelamento Rasch, um paradigma no campo das teorias de resposta ao item, como uma técnica de equalização, que torna comparáveis os ranqueamentos dos estudantes baseados nas avaliações trimestrais da disciplina física, ao mesmo tempo em que constrói uma escala intervalar para mensurar a competência em física escolar.

2.6 - Modelo Rasch

Em situações de avaliação educacional ou psicológica aplica-se um teste, geralmente contendo um grande número de itens, e cada pessoa é avaliada a partir do seu escore total no teste. A hipótese básica que justifica somar-se os escores de cada sujeito nos diversos itens para compor um escore total é a de que a resposta dada pela pessoa ao item revela uma manifestação de uma mesma e única habilidade. Caso os itens não expressem a manifestação de uma mesma habilidade, somar os escores dos itens para compor o escore total corresponderia somar quantidades que se referem a habilidades de naturezas distintas. A situação é semelhante a somar três laranjas com dois cachorros. Sem dúvida resulta em uma quantidade igual a cinco, mas a que se refere essa quantidade? Essa hipótese básica é chamada de hipótese da unidimensionalidade do teste.

Um teste é perfeitamente unidimensional se cada um dos itens que o compõe expressa a manifestação de uma única e mesma habilidade. Em situações de avaliação psicológica cognitiva os testes, por princípio, não são unidimensionais, a menos que se parta de uma teoria sobre as estruturas das habilidades intelectuais humanas em que não pressupõe qualquer estrutura hierárquica dessas habilidades. Tais teorias existem, por exemplo, Aikens, Thorndike e Hubbell (1902, p.74), viam a mente como *“uma multiplicidade de capacidades particulares [...] que não podem ser muito independentes uma das outras”*.

Como esta pesquisa referencia-se na teoria de Carroll, que é uma teoria hierárquica da estrutura das habilidades intelectuais humanas, por princípio, admite-se que nenhum teste cognitivo é unidimensional. Cada item do teste é uma expressão não apenas da habilidade cognitiva especializada que mensura, mas é também a expressão das habilidades dos níveis hierárquicos mais altos. Desta forma, por um lado, a unidimensionalidade de um teste não pode ser avaliada de forma dicotômica, ela é uma questão de grau. Por outro lado, um teste só tem validade de conteúdo se todos os seus itens mensuram uma mesma habilidade. Assim, a tarefa do pesquisador consiste em julgar se o grau de unidimensionalidade alcançada pelo teste cognitivo é ou não suficiente para se estabelecer a sua validade de conteúdo.

Nos testes educacionais, a situação se repete. Nenhum teste que avalia a aprendizagem focaliza habilidades tão específicas quanto as habilidades especializadas do estrato I do modelo de Carroll. Em geral, testes que possuem valor educacional, mobilizam conhecimentos e um conjunto de habilidades que concorrem na conformação de um construto mais complexo que se refere às aprendizagens escolares que se quer mensurar. Por exemplo, um teste de física envolvendo o conteúdo de dinâmica mobiliza não só os conhecimentos sobre dinâmica, assim como aqueles que envolvem a cinemática. Mobilizam também habilidades no domínio verbal, habilidades matemáticas, habilidades de fluência (geração de idéias, imaginação), habilidade de visualização e abstração. Dificilmente pode-se argumentar em favor da unidimensionalidade de um teste dessa natureza. Novamente, a solução é conceber a unidimensionalidade como uma questão de grau e adequação.

Um segundo aspecto relativo à situação de teste psicológico e educacional diz respeito à natureza da escala resultante do escore total. Nos testes com questões objetivas, o escore total é expresso no conjunto dos inteiros. As questões abertas dos testes em geral são corrigidas ou identificando-se sub-tarefas de cada item que possam ser avaliadas dicotomicamente (MAIA, 2010) ou então classificadas em uma dentre várias categorias segundo um esquema de rubrica (AMANTES, 2009, COELHO, 2010). Quando se pretende usar determinado teste como medida educacional e o teste contém questões abertas, a fidedignidade do teste quase sempre depende da adoção de um protocolo de correção bem estabelecido. Desta forma, o escore total de testes pode ser mapeado no conjunto dos inteiros quer contenham questões objetivas ou abertas. Como dizem Wright & Linacre (1989), todas as observações, e o escore total é uma observação, são sempre ordinais. No entanto, as habilidades são medidas, e como eles dizem, medidas são sempre intervalares.

A teoria clássica de testes (CTT) supõe que subjacente ao escore total ordinal há uma escala intervalar. Mais ainda, se um teste contém um número suficientemente grande de itens pode-se admitir que a escala de escore total, que no conjunto dos inteiros varia de zero até o escore máximo, na verdade é uma escala intervalar. Uma alternativa à teoria clássica dos testes foi proposta pelo matemático dinamarquês

George Rasch, no final da década de 50 e consiste em partir-se das observações ordinais para construir, via modelamento matemático das observações, uma escala intervalar em que se mensura a habilidade mobilizada pelo teste. A família de modelos matemáticos adequados para esse fim é conhecida como modelos Rasch. Na década de 60, modelos de dados matematicamente similares aos modelos Rasch, porém mais gerais do que eles, foram introduzidos e tornaram-se conhecidos como teorias de resposta ao item (TRI), mas que não serão aqui discutidas (PRIMI, & PASQUALI, 2003).

A idéia base do modelo Rasch é que o processo de medida psicológica ou educacional não é de natureza distinta do processo de medidas das chamadas ciências duras, como a física, química e geologia. Em Física, mesmo no ensino médio, ensina-se que toda medida é uma soma de duas componentes: o valor verdadeiro da medida mais as variações aleatórias que ocorrem devido a existência de múltiplas fontes não controláveis de variação. Uma medida em física é sempre uma variável aleatória, composta da soma de uma variável real e uma variável aleatória, por exemplo, o comprimento de uma mesa é uma variável aleatória, no seguinte sentido, cada vez que ele é medido nas mesmas condições obtém-se um resultado que obedece a uma lei de distribuição estatística em torno do valor verdadeiro do comprimento. Por esta razão é que o comprimento da mesa, normalmente, é indicado por uma estimativa do valor verdadeiro e do desvio padrão associado à distribuição de erros aleatórios das medidas. É importante destacar que os erros aleatórios são inerentes ao processo de medida e por princípio, não são controláveis. Fontes de erros controláveis introduzem erros sistemáticos, vieses na medida em que podem e devem ser controlados pelo pesquisador.

Uma segunda característica no processo de mensuração em física é que ele é de natureza objetiva, de tal forma que pode se comparar duas grandezas de mesma natureza, independente do instrumento utilizado para medir cada uma delas. Por exemplo, pode se comparar dois comprimentos distintos independente de qual instrumento foi utilizado para mensurar cada um deles.

Rasch, baseando-se nestas duas características, propôs que todo processo de medida educacional ou psicológica fosse de natureza probabilística, para contemplar

a existência de múltiplas fontes não controláveis do desempenho nos testes, e obedecesse a um princípio de objetividade específica. Este último princípio estabelece as condições para se comparar a habilidade de duas pessoas distintas, independentes dos itens de testes que elas respondam, e de comparar a dificuldade de dois itens de teste independentemente de quais pessoas responderam a eles. Rasch descobriu que os modelos matemáticos adequados para isto são modelos probabilísticos de dois parâmetros, um que descreve a habilidade dos sujeitos e outro que descreve a dificuldade dos itens. Descobriu, ainda, que estes modelos permitem separar e estimar, separadamente, os parâmetros das pessoas e os parâmetros dos itens. O modelo Rasch mais simples é aquele usado para modelar um teste de questões dicotômicas (certo ou errado, falso ou verdadeiro, aplica-se ou não se aplica, etc.), que pode ser formulado como se segue.

Chamando de β_n a habilidade da pessoa n e θ_i a dificuldade do item i , o logaritmo natural da razão da probabilidade de acerto do item i pela pessoa n pela probabilidade de erro do item i pela pessoa n é igual a diferença entre a habilidade da pessoa n e o dificuldade do item i . Como a soma da probabilidade de acerto e com a probabilidade de erro totaliza um, pode-se expressar matematicamente o modelo pela expressão

$$\left[\ln \left(\frac{p}{1-p} \right) = \beta_n - \theta_i \right] \quad (\text{Eq. 1})$$

Onde p é a probabilidade de acerto do item i pela pessoa n . O termo do lado esquerdo é conhecido como logit. Esta expressão mostra que, ambos os parâmetros, a habilidade da pessoa e a dificuldade do item, estão em uma mesma escala adimensional (sem unidade de medida). Esta escala construída é conhecida logit. Uma forma alternativa de expressar o modelo é expressando a probabilidade de acerto do item i pela pessoa n como função da diferença entre o parâmetro da pessoa e o parâmetro do item:

$$\left[P_{ni} = \frac{e^{\beta_n - \theta_i}}{1 + e^{\beta_n - \theta_i}} \right] \quad (\text{Eq. 2})$$

Se a habilidade da pessoa é igual à dificuldade do item ($B_n = \theta_i$), a probabilidade de resposta correta é 0,5. Se a habilidade da pessoa é superior a dificuldade do item ($B_n > \theta_i$) a probabilidade vai para um (1) e no caso oposto ($B_n < \theta_i$) a probabilidade torna-se zero (0).

Os parâmetros dificuldade do item e habilidade da pessoa podem ser estimados a partir dos escores brutos dos testes e dos itens. Como os parâmetros das pessoas e dos itens são separáveis cada um deles pode ser estimado separadamente: a dificuldade dos itens é estimada a partir do total de respostas corretas para cada item e a habilidade da pessoa pode ser estimada a partir do número de respostas corretas dadas por cada pessoa.

Existem diversos algoritmos para se estimar a dificuldade dos itens e a habilidade das pessoas. Em geral são algoritmos complexos e que demandam certo esforço de computação. Felizmente, existem diversos pacotes de aplicativos que implementam estes algoritmos e permitem modelar um conjunto de dados empíricos usando o modelo Rasch. Nesta pesquisa utilizou-se o Software Winsteps (LINACRE & WRIGTH, 2000), que implementa dois algoritmos de estimativa dos parâmetros, o JMLE (“Joint Maximum Likelihood Estimation”) e o PROX (“Normal Approximation Algorithm”) proposto por Cohen (1979). Estes pacotes fornecem diversas estatísticas para se testar o ajuste do modelo aos dados empíricos.

Quando o modelo Rasch se ajusta adequadamente aos dados, há uma relação linear bem estabelecida entre os escores brutos da pessoa e as correspondentes medidas de habilidades. Os escores brutos são indicadores ordinais, porém as medidas da habilidade estão em uma escala intervalar. A vantagem da escala intervalar é que as medidas são diretamente comparáveis entre si. Por exemplo, uma pessoa com uma habilidade de 3 logit tem três vezes mais habilidade que uma pessoa com uma habilidade de 1 logit. Isso é muito diferente das escala ordinais, como por exemplo, quando o escore expresso como porcentagem do escore máximo possível (POMP). Uma pessoa com escore POMP igual a 0,3 não tem três vezes mais habilidade do que uma pessoa com o escore POMP igual a 0,1. Porcentagens podem refletir a posição correta da pessoa ou item, mas não o intervalo correto entre a habilidade da pessoa ou a dificuldade dos itens. Como a

escala que mensura a habilidade das pessoas é a mesma que mensura a dificuldade dos itens, esta também é uma escala intervalar, o que permite uma comparação direta entre as dificuldades dos itens. Por exemplo, um item com uma dificuldade de 3 logit é três vezes mais difícil que um item com uma dificuldade de 1 logit.

Em todos algoritmos para estimativas de parâmetros das pessoas e dos itens há, em geral, mais equações que parâmetros a determinar, o que faz com que a solução não seja única. Para fixar uma solução, os programas, em geral, fazem alguma restrição nos valores dos parâmetros. A restrição padrão consiste em tornar nula a média da dificuldade dos itens. Alternativamente pode se tornar nula a média da habilidade das pessoas. Esses algoritmos, também, permitem a determinação dos erros padrão das estimativas dos parâmetros. Isto significa que estima-se a habilidade da pessoa e o erro padrão associado a essa habilidade. Da mesma forma estima-se a dificuldade de um item e o erro padrão associado com esta dificuldade.

As estimativas são mais precisas se o número de pessoas e de itens for grande. E, além disso, a distribuição de habilidade das pessoas e distribuição de dificuldade dos itens devem-se superpor de forma marcante. Se a dificuldade dos itens se distribui por uma pequena faixa de valores na escala logit, as pessoas com habilidade muito maior ou muito menor que os valores limites da faixa de distribuição de itens serão mal mensuradas.

Dentre as estatísticas para se testar o ajuste do modelo aos dados empíricos duas são as mais importantes: a estatística Infit e a estatística Outfit . Para entendê-la é necessário recordar o fato de que o modelo Rasch é um modelo probabilístico. Se a habilidade da pessoa é maior do que a dificuldade do item a probabilidade da pessoa acertar o item é maior do que 0,5 (ou 50%), de forma que a maioria das pessoas acertará o item, mas um certo número de pessoas errará o item. Da mesma forma, se habilidade da pessoa é menor que a dificuldade do item, a probabilidade de que ela acerte o item será menor que 0,5 (ou 50%). Significa que a maioria das pessoas errará o item e algumas o acertarão. As estatísticas Infit e Outfit avaliam estes comportamentos não esperados, isto é , errar um item quando deveria acertá-lo ou acertar um item quando deveria errá-lo. A diferença das duas é que a

estatística infit dá mais peso estatístico aos comportamentos não esperados nas situações em que a habilidade da pessoa é próxima da dificuldade do item. Já a estatística Outfit é confere mais peso estatístico, aos comportamentos não esperados nas situações em que a habilidade da pessoa é muito maior ou muito menor do que a dificuldade do item.

2.7 – Modelo de regressão múltipla multinível para estudos longitudinais

Nas últimas décadas, modelos longitudinais têm sido cada vez mais utilizados no campo da pesquisa educacional (SINGER & WILLET, 2003; GOLDSTEIN, 1995).

Um modelo longitudinal é usado para modelar a evolução temporal de uma variável de resposta de interesse. No caso educacional, a variável de resposta de interesse, em geral, recai sobre alguma medida de desempenho escolar, de alguma habilidade, de alguma atitude em relação a algum aspecto da vida escolar, algum estado de motivação, etc.. Para se realizar um estudo longitudinal é necessário mensurar a variável de resposta para cada um das pessoas da amostra do estudo em diversas ocasiões ou ondas de medidas.

Há duas questões, segundo Singer & Willet (2003), que normalmente são respondidas nos estudos longitudinais na área da educação. A primeira pode ser formulada da seguinte maneira: “*como a variável de resposta muda ao longo do tempo?*”. Esta questão indaga sobre o padrão de mudança que ocorre em cada pessoa. Por exemplo, a mudança é linear ou não linear? A mudança ao longo do tempo é sistemática ou flutua em torno de uma evolução sistemática?

A segunda questão pode ser formulada da seguinte forma: “*nós podemos prever as diferenças existentes nessas mudanças?*”. Esta questão discute se há diferenças nos padrões de mudanças entre as diversas pessoas e se estes padrões podem ser previstos por variáveis que caracterizam os diversos grupos de pessoas. Diferentes tipos de pessoas experimentam diferentes padrões de mudanças? Quais preditores estão associados com quais padrões de mudanças?

A primeira questão, segundo Singer & Willet (2003), é de natureza descritiva já que almeja essencialmente descrever matematicamente as trajetórias da mudança que ocorrem em cada pessoa. A segunda questão, segundo os mesmos autores, é de natureza relacional, visto que pretende descobrir relações entre as características das pessoas e seus padrões de mudanças. Convém lembrar que mudança não significa apenas crescimento, ela pode significar um decréscimo, por exemplo, uma pessoa pode ficar mais ou menos interessada em uma disciplina escolar à medida que o ano letivo passa.

Um estudo longitudinal, segundo os mesmos autores, que formule questões desta natureza, deve utilizar uma metodologia hierárquica de análise de dados. No nível hierárquico mais baixo, deve examinar as mudanças intra-individuais, isto é, os padrões de mudança que ocorrem em cada sujeito. No segundo nível de análise, deve examinar as diferenças inter-individuais nos padrões de mudança. O que se pretende é detectar nas pessoas a heterogeneidade nas mudanças e determinar a relação entre variáveis preditoras que caracterizam grupos de pessoas e a forma da trajetória de mudança de cada pessoa.

Há duas estratégias para se realizar a análise nos dois níveis: a primeira consiste em fazer estudos de regressão múltipla em cascata, isto é, modela-se a trajetória de mudança de cada pessoa utilizando-se uma regressão múltipla. Neste estágio modelam-se diretamente os dados da variável de resposta de cada pessoa em função do tempo. Cada pessoa passa a ser caracterizada por um conjunto de parâmetros obtidos na regressão múltipla intrapessoal.

Em um segundo estágio, toma-se como dados os parâmetros estimados em cada regressão múltipla realizada no estágio um. Neste estágio usa-se uma regressão múltipla para modelar a relação de cada tipo de parâmetro com as variáveis que exprimem as diferenças inter-individuais, variáveis que caracterizam grupos de pessoas. Essas variáveis podem ser de natureza orgânica (gênero, idade etc..) ou contextuais, contemplando tanto o contexto social (status sócio econômico, nível de escolaridade do pai e da mãe etc.) quanto o contexto escolar (turma, professor, série etc..). Essas variáveis preditoras podem ser: contínuas; de natureza intervalar,

como por exemplo, idade medida em dias; categóricas, como por exemplo, status sócio econômico; ou dicotômicas como, por exemplo, gênero.

A técnica de regressão múltipla é uma técnica de análise de dados muito geral e muito flexível. Ela é utilizada quando se pretende investigar as relações entre uma variável de resposta intervalar ou quantitativa, e um conjunto de variáveis independentes, que podem ser tanto quantitativas (intervalar), quanto categóricas (ordinal) ou mesmo dicotômicas (nominal). A forma de relação não se restringe à linear, pelo contrário, não há nenhuma restrição matemática na forma funcional da relação entre a variável de resposta e as variáveis independentes. Essa generalidade e a simplicidade dos algoritmos de estimação dos parâmetros tornaram a regressão múltipla como uma técnica poderosa e prevalente nas pesquisas praticamente em todas as áreas da ciência.

Em particular, a regressão múltipla pode ser usada para se analisar a evolução de certa grandeza ao longo do tempo. Por isto é natural quando se pensa em estudar mudança educacional considerar a possibilidade de se empregar a regressão múltipla. Por exemplo, pode-se empregar a regressão múltipla para se analisar a evolução da aprendizagem em física ocorrida entre os estudantes de uma turma. Este exemplo é interessante porque ele coloca a questão de qual é a unidade de análise que deve ser utilizada na investigação, ou seja, qual é a variável de resposta que será modelada pela regressão múltipla: a aprendizagem média dos estudantes? Ou a aprendizagem de cada estudante?

A segunda estratégia consiste em fazer uma única regressão utilizando-se o modelo multinível de regressão multivariada. Na década de 70, começou a ser desenvolvido o modelo multinível de regressão múltipla, uma nova forma de abordar o problema da mudança educacional. Essa nova forma partiu do reconhecimento de que as variáveis educacionais normalmente estão organizadas em um sistema hierárquico. Por exemplo, desempenho de cada estudante muda ao longo do ano letivo, mas os estudantes estão agrupados em turmas, as turmas em séries e escolas, as escolas em unidades regionais de administração educacional, etc. O modelo multinível de regressão múltipla foi desenvolvido para lidar com estes dados hierarquizados. O seu emprego na análise de dados educacionais tem mostrado que ele é muito mais

potente para explicitar as relações existentes entre as variáveis do que o uso do modelo de regressão múltipla em cascata. Como no caso da regressão múltipla, o modelo é bem geral e não impõe restrição matemática a forma funcional da relação entre a variável de resposta, que deve ser intervalar, e as variáveis independentes, que podem ser de natureza intervalar (ou contínua), categórica (ou ordinal) ou dicotômica (ou nominal).

Nesta dissertação será utilizado o modelo multinível de regressão múltipla para se analisar a evolução temporal da competência em física escolar, medida em uma escala intervalar, ao longo do ano letivo, de um conjunto de estudantes de uma mesma série e de uma escola, porém, que se distribuem por diversas turmas. A apresentação que será feita do modelo multinível de regressão múltipla considerará a existência de dois níveis de análise dos dados, o nível intra-individual e o nível inter-individual. Devido ao pequeno número de ocasiões de medida da competência em física escolar, será considerado apenas um modelo de evolução linear.

O modelo multinível de regressão múltipla para modelar a evolução da variável de resposta Y em função do tempo, limitada a uma evolução linear, pode ser formulado da seguinte maneira:

Modelo intra-individual (Modelo de 1º Nível)

$$[Y_{ij} = \pi_{0i} + \pi_{1i}Tempo_j + \varepsilon_{ij}]. \quad (\text{Eq. 1})$$

Modelo inter-individual (Modelo de 2º Nível)

$$[\pi_{0i} = \gamma_{00} + \gamma_{01}V_{1i} + \gamma_{02}V_{2i} + \gamma_{03}V_{3i} + \zeta_{0i}], \quad (\text{Eq. 2})$$

$$[\pi_{1i} = \gamma_{10} + \gamma_{11}V_{1i} + \gamma_{12}V_{2i} + \gamma_{13}V_{3i} + \zeta_{1i}]. \quad (\text{Eq. 3})$$

Nestas equações, Y_{ij} é o valor da variável Y medida no sujeito i na ocasião j ; $Tempo_j$ corresponde ao tempo da j -ésima ocasião de medida; V_{1i} , V_{2i} e V_{3i} correspondem a três variáveis preditoras medidas no sujeito i .

O modelo intra-individual prevê que a variável de resposta de cada pessoa muda de forma sistemática e linearmente com o tempo, e muda de forma aleatória entre as ocasiões. A mudança sistemática é caracterizada por dois parâmetros que podem diferir de pessoa para pessoa: o intercepto, π_{0i} , que corresponde ao valor da variável de resposta no tempo zero, e a inclinação, π_{1i} , que corresponde à taxa de variação da variável Y com o tempo. A mudança aleatória é descrita pela variável ε_{ij} , que descreve um processo aleatório intrapessoal e entre as ocasiões. Isto deve ser interpretado da seguinte forma, o valor mensurado da variável Y no sujeito i na ocasião j não corresponde ao valor verdadeiro desta variável nesta ocasião. O valor verdadeiro da variável Y no sujeito i na ocasião j é aquele obtido substituindo-se o valor do tempo correspondente à ocasião na equação da reta de mudança do sujeito i . O valor medido Y_{ij} corresponde a esse valor verdadeiro mais uma variação aleatória, devida a múltiplas e incontroláveis fontes de variação intrapessoal, por exemplo, se Y representa uma medida do desempenho escolar, avaliado por um teste, o escore do sujeito no teste não é o seu escore verdadeiro, mas a soma deste com a variação aleatória ε_{ij} . Neste exemplo, fontes incontroláveis de variação do desempenho podem, por exemplo, ser o nível de atenção do sujeito, variação no nível de ruído da sala, etc.. A variável aleatória pode ser descritas pelos parâmetros da distribuição de valores a que obedece. No modelo multinível de regressão múltipla a hipótese distribucional padrão é que a variação aleatória de todas as pessoas e em todas as ocasiões é distribuída normalmente com média zero, e um mesmo valor de variância:

$$[\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon}^2)]. \quad (\text{Eq.4})$$

O modelo de inter-individual descreve como os parâmetros que caracterizam a evolução sistemática individual, o intercepto e a inclinação, dependem das variáveis orgânicas e contextuais, que são as variáveis preditoras. O modelo descrito prevê uma dependência linear sistemática de cada um desses dois parâmetros em relação às variáveis preditoras e uma variação aleatória devido às inúmeras e incontroláveis fontes de inter-individual, neste caso, descrita pelas variáveis aleatórias ζ_{0i} e ζ_{1i} . Como no modelo do primeiro nível, essas variáveis aleatórias do segundo nível obedecem a uma distribuição normal bivariada com média zero, e

uma matriz de variância-covariância que especifica a variância de cada um dos dois processos aleatórios e a covariância entre eles:

$$\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{21} \\ \sigma_{21} & \sigma_1^2 \end{bmatrix} \right). \quad (\text{Eq.5})$$

Uma forma alternativa de representar este modelo é o chamado modelo composto obtido pela substituição das equações (Eq.2) e (Eq.3) na equação (Eq.1):

$$Y_{ij} = [\gamma_{00} + \gamma_{01}V_{1i} + \gamma_{02}V_{2i} + \gamma_{03}V_{3i} + (\gamma_{10} + \gamma_{11}V_{1i} + \gamma_{12}V_{2i} + \dots + \gamma_{13}V_{3i})Tempo_j]_{\text{estrutural}} + [\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}Tempo_j]_{\text{aleatório}}. \quad (\text{Eq. 6})$$

Este modelo mostra claramente que a evolução da variável de resposta Y é composta de uma evolução componente estrutural e um componente aleatório. A componente estrutural descreve uma evolução linear com o tempo para cada pessoa. Diferentes pessoas possuem diferentes valores das variáveis orgânicas e contextuais e, portanto, diferentes interceptos e inclinações. Cada grupo de pessoa que compartilham os mesmos valores das variáveis orgânicas e contextuais possuem uma mesma evolução sistemática. A componente aleatória mostra que há três processos aleatórios compondo um processo aleatório que modifica a evolução da variável de resposta. A variância desse processo aleatório varia com o tempo:

$$[Var = \sigma_\varepsilon^2 + \sigma_0^2 + 2\sigma_{12}Tempo_j + \sigma_1^2Tempo_j^2]. \quad (\text{Eq.7})$$

Desta forma, ainda que aparentemente o modelo multinível de regressão múltipla se assemelhe a uma regressão múltipla ordinária, estes modelos diferem um do outro radicalmente. Na regressão múltipla ordinária a variância do processo aleatório é a mesma para todas as pessoas e em todas as ocasiões (processo homocedástico). Entretanto, no modelo multinível de regressão múltipla o processo aleatório é heterocedástico, a variância muda com o tempo.

Os algoritmos para se estimar os parâmetros do modelo multinível de regressão múltipla são mais complicados do que são os algoritmos utilizados para estimar os

parâmetros do modelo da regressão ordinária. Felizmente há diversos programas que implementam os algoritmos que tornam fácil o uso desse modelo na pesquisa educacional. Nesta dissertação utiliza-se um desses programas, o MLwin (RASBASH et al., 2005), mas aqui, não se discutirá detalhes da implementação do algoritmo utilizado¹². Este programa permite trabalhar com dados faltantes. Em estudos longitudinais, é muito comum que em uma determinada onda de medida não se consiga mensurar todas as pessoas que participaram das ondas anteriores. Por exemplo, se cada onda de mensuração significa a aplicação de um teste sobre o conteúdo de uma disciplina, devido a ausência eventual de alguns estudantes, ocorrem dados faltantes. O programa MLwin fornece a estimativa dos parâmetros do modelo e o erro padrão associado a cada um desses parâmetros. Fornece, também, para cada modelo uma estimativa da deviência.

A deviência é uma estatística muito utilizada para testar o ajuste global de um modelo linear aos dados empíricos. O seu principal uso é para comparar dois modelos aninhados. Um modelo é dito aninhado se pode ser obtido de um modelo mais amplo por restrição de parâmetros. Por exemplo, considerem-se duas versões do modelo da equação (Eq.6), denominado Modelo A e Modelo B, caracterizados pela seguinte restrição de parâmetros:

$$\begin{array}{l} \text{Modelo A: } \gamma_{02} = \gamma_{03} = \gamma_{12} = \gamma_{13} = \mathbf{0} \\ \text{Modelo B: } \gamma_{03} = \gamma_{13} = \mathbf{0} \end{array} \quad (\text{Eq.8})$$

O Modelo A incorpora apenas o efeito da variável preditora V_1 , enquanto o Modelo B incorpora os efeitos das variáveis preditoras V_1 e V_2 . O Modelo B é mais geral que o modelo A, que pode ser obtido a partir do Modelo B fazendo-se as restrições

$$\gamma_{02} = \gamma_{12} = \mathbf{0} . \quad (\text{Eq.9})$$

¹² O leitor interessado pode consultar o texto de Goldstein (2003) e os manuais do MLwin (Rasbash, et al., 2009a, 2009b) e Browne et al.(2009).

Assim o modelo A está aninhado dentro do Modelo B. Pode-se testar se o modelo A se ajusta melhor aos dados do que o Modelo B, utilizando-se a diferença da estatística de deviance dos dois modelos. Esta diferença obedece a uma distribuição de Chi-quadrado com o número de graus de liberdade igual à diferença de parâmetros estimados nos dois modelos, no exemplo acima esta diferença corresponde a dois parâmetros. A hipótese nula é de que os dois modelos se ajustam igualmente bem aos dados empíricos. O teste consiste, então, em se determinar a probabilidade, admitindo-se como verdadeira a hipótese nula, de se obter um valor tão alto quanto à diferença de deviance constatada. Se esta probabilidade for menor do que o nível de confiança escolhido, usualmente, $p = 0,05$, há evidência suficiente para rejeitar a hipótese nula e admitir-se que o modelo menos parcimonioso (com maior número de parâmetros) se ajusta melhor aos dados do que o modelo mais parcimonioso (com menor número de parâmetros). Caso não haja evidência para se rejeitar a hipótese nula, os dois modelos se ajustam igualmente bem aos dados empíricos e, neste caso, dá-se preferência ao modelo mais parcimonioso, aquele que estima menor número de parâmetros.

Neste capítulo apresentou-se os referenciais teóricos que permitiram a introdução do conceito de competência em física escolar, como uma representação daquilo que os estudantes aprendem e desenvolvem ao longo do ano letivo nesta disciplina do currículo escolar. Apresentou-se também, um referencial teórico que justifica a construção de uma escala intervalar para mensurar a competência em física escolar e o modelo estatístico utilizado para a construção desta escala. Apresentou-se ainda, o modelo da estrutura das habilidades intelectuais humanas de Carroll, destacando-se as habilidades cognitivas amplas do seu estrato II e que são mensuradas pela bateria BaFaCaLo. Finalmente, apresentou-se o modelo multinível de regressão múltipla que foi utilizado para modelar a evolução temporal da competência em física escolar e investigar a sua relação com as habilidades cognitivas amplas do estrato II de Carroll. No próximo capítulo, serão apresentados os principais aspectos metodológicos da pesquisa relatada nesta dissertação.

3. ASPECTOS METODOLÓGICOS DA PESQUISA

A pesquisa relatada nesta dissertação insere-se no campo temático da relação entre a aprendizagem escolar e as habilidades cognitivas amplas. A pesquisa focaliza na relação entre a aprendizagem em física escolar e as habilidades cognitivas amplas. A aprendizagem em física escolar é representada pelo desenvolvimento da competência em física escolar. Nesta pesquisa são investigadas duas questões. A primeira questão é de natureza instrumental, indaga sobre quais variáveis organismicas e contextuais afetam a evolução temporal da competência em física escolar e sobre como elas afetam esta evolução, e a segunda, investiga se, após controlar para os efeitos das variáveis organísmicas e contextuais, há uma relação entre a evolução temporal da competência em física escolar e algumas das habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll (1993).

A pesquisa foi organizada em torno de dados empíricos existentes decorrentes tanto do registro burocrático ordinário da vida escolar dos estudantes quanto da aplicação da bateria BaFaCaLo a um grupo de estudantes de uma escola pública federal no escopo da pesquisa feita por Gomes (2010).

Neste capítulo serão descritos os principais aspectos metodológicos da pesquisa aqui relatada. Inicialmente na seção 3.1 apresenta-se o contexto escolar em que foi realizada esta pesquisa, e na seção 3.2, o contexto específico da disciplina física que é analisada nesta investigação. Na seção 3.3 expõem-se as fontes dos dados utilizados, e na seção seguinte, as variáveis consideradas na pesquisa. Ao descrever as variáveis intervenientes na pesquisa apresenta-se com mais detalhes o longo e complexo processo de construção da escala intervalar para mensurar a variável competência em física escolar. Finalmente, na seção 3.5 expõe-se o processo de análise dos dados.

3.1 - Contexto da pesquisa

A pesquisa foi conduzida em oito turmas da 1ª série do ensino médio do ano de 2007, na disciplina Física de uma escola pública federal, mantida por uma

universidade federal, e que oferece aos estudantes o ensino médio e técnico. Esta escola organiza – se em setores acadêmicos por disciplina. O corpo docente da escola é constituído, em sua maioria, de professores das carreiras existentes na universidade em regime de 40 horas semanais com dedicação exclusiva e de professores substitutos. No setor de Física, na época da pesquisa, existiam quatro professores efetivos com dedicação exclusiva. Completavam o quadro docente seis professores substitutos com contrato provisório de dois anos com dedicação de 20 horas semanais.

A escola oferecia 198 vagas para o primeiro ano do ensino médio ou médio conjugado com o técnico, sendo que, 90 destas vagas eram destinadas aos estudantes do ensino fundamental II de uma escola mantida pela mesma universidade. O restante das vagas era colocado em concurso público, porém era distribuída em três categorias sócio-econômicas, de acordo com o tipo de escola (pública ou privada) de origem e o tempo de permanência na escola de origem. Esta forma de seleção visava garantir a heterogeneidade sócio-cultural e econômica no corpo discente da escola.

A primeira série era constituída de 8 turmas, com estudantes vindos de escolas públicas, de escolas particulares, repetentes e da escola de ensino fundamental da universidade. O número de estudantes por turma variava entre 26 a 29 estudantes. Não havia um critério explícito de enturmação dos estudantes considerando gênero, faixa etária e condição sócio-econômica. A secretaria da escola era responsável pela enturmação e este procedimento pretendia tornar as turmas o mais heterogêneas em termos da escola de origem (pública, privada ou da própria instituição) e gênero (feminino e masculino) ¹³.

¹³ Informações cedidas pela secretaria da escola e pela psicóloga Fabíola O. Lima, responsável pelo acompanhamento dos estudantes do ensino médio e técnico, e complementadas por Vila (2009).

A distribuição dos estudantes por gênero para as turmas da 1ª série é mostrada no QUADRO 3. Percebe-se segundo o QUADRO 3, que o número de estudantes do sexo masculino (58%) era maior que os do sexo feminino (42%). E que a quantidade de estudantes do sexo masculino por turma era maior que a quantidade do sexo feminino.

QUADRO 3
Relação de estudantes por turma e gênero na 1ª série

Série	Turmas	Estudantes / turmas	Feminino (0)	Masculino (1)
1ª Série	T1	26	10	16
	T2	29	11	18
	T3	27	10	17
	T4	28	13	15
	T5	28	13	15
	T6	27	11	16
	T7	27	12	15
	T8	28	13	15
-	Total	220	93	127

Os cursos técnicos de nível médio oferecidos em 2007 pela escola pesquisada foram: Eletrônica, Instrumentação e Controle, Patologia Clínica e Química. Estes cursos tinham vagas limitadas que não atendiam o total de estudantes da escola. Por esta razão, além dos cursos técnicos, ofertava-se o ensino médio regular.

As vagas oferecidas nos diversos cursos técnicos eram preenchidas somente pelos estudantes que ingressaram por concurso e a prioridade de escolha era exercida segundo o desempenho dos estudantes nas disciplinas importantes para o respectivo curso. Os estudantes, que vieram do ensino fundamental II da própria instituição, eram automaticamente matriculados nas vagas para o ensino médio regular.

A avaliação anual de desempenho dos estudantes era expressa em cada disciplina, em uma nota inteira entre zero e cem, que era o valor acumulado das notas parciais obtidas em três trimestres. No 1º trimestre os estudantes eram avaliados em 30 pontos e no 2º e 3º trimestres, em 35 pontos. Para a aprovação o estudante precisaria atingir a nota anual mínima de 60 pontos e uma frequência global de 75%, considerando todas as disciplinas. O estudante, reprovado por duas vezes consecutivas na mesma série, era jubilado da escola, exceto para o terceiro ano que, quando reprovado, era desligado imediatamente da escola.

3.2 - A disciplina física na 1ª série do ensino médio em 2007

Na série pesquisada, a disciplina de Física foi lecionada por quatro professores, dois efetivos e dois substitutos sendo cada um responsável por duas turmas. O professor P₁ era responsável pelas turmas 1 e 4, o professor P₂ era responsável pelas turmas 2 e 6, o professor P₃ era responsável pelas turmas 3 e 5 e o professor P₄ era responsável pelas turmas 7 e 8.

A carga horária da disciplina era de três aulas semanais de 50 minutos cada e duas aulas geminadas de laboratório a cada quinze dias. As aulas teóricas e práticas eram ministradas pelos mesmos professores e aconteciam no período diurno, manhã e tarde. Tanto as aulas teóricas como as práticas eram centradas nos estudantes. Os estudantes passavam a maior parte do tempo das aulas, realizando atividades individuais ou em grupos. Eventualmente, os professores faziam exposições dos tópicos curriculares para a turma toda, mas estas exposições, quando aconteciam, não ocupavam toda duração de uma aula.

As atividades dos estudantes, realizadas em sala de aula ou em casa, eram quase sempre baseadas em folhas de atividades teóricas ou práticas. Estas folhas podiam conter roteiros de atividades práticas, orientações de estudos no livro texto adotados pela escola, listas complementares de exercícios e textos explicativos sobre alguns tópicos não abordados no livro texto. As folhas de atividades eram elaboradas pelo coordenador da disciplina e pela equipe de professores da série. A disciplina física adotava o livro texto Amaldi (1995). A equipe de professores exigia que os estudantes mantivessem em dia dois cadernos com os registros das atividades

realizadas: em um caderno eram registrados os relatórios das atividades práticas e, no outro, todas as demais atividades realizadas.

Em cada trimestre, o desempenho dos estudantes era avaliado através de quatro testes e do exame dos dois cadernos de registro das atividades realizadas pelos estudantes. Dos testes, um era a prova trimestral comum a todas as turmas e elaborada pela equipe de professores. Os testes restantes eram específicos por turma e elaborados pelo próprio professor da turma. Estes testes específicos continham somente questões abertas. A prova trimestral que era o teste comum a todas as turmas continham apenas questões fechadas do tipo Verdadeiro (V) ou Falso (F) e questões de múltipla escolha.

O programa da disciplina Física nas três séries da escola pesquisada pretendia ser organizado na forma de um currículo em espiral e recursivo. O objetivo esperado era que os estudantes tivessem contato com os mesmos tópicos em todas as séries, mas com abordagens e níveis de profundidade diferenciados por série. O programa de física utilizado na primeira série da escola pesquisada no ano de 2007 é apresentado no ANEXO A. Nele constam alguns tópicos da mecânica, termodinâmica, eletricidade, eletromagnetismo, relatividade restrita e modelo de Bohr.

3.3 - Fontes de dados

O interesse dessa pesquisa é estudar a possível relação entre o desenvolvimento da competência em física escolar e as seis habilidades cognitivas amplas mensuradas por Gomes (2010) e identificadas como equivalentes a aquelas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll (1993). A escola na qual essa pesquisa foi realizada foi escolhida por oportunidade, visto que nela havia sido aplicada uma bateria de testes cognitivos (GOMES, 2010) e os dados estavam disponíveis para análise. A competência em física escolar será uma variável intervalar construída a partir das notas trimestrais dos estudantes, conforme já foi discutido no capítulo 2. Na escola escolhida era possível o acesso às notas trimestrais dos estudantes, que foram submetidos aos testes cognitivos e a alguns dados referentes ao contexto sócio-econômico e cultural dos mesmos.

As variáveis contextuais resultaram de um cruzamento de três fontes de dados. A primeira fonte de dados foi a secretaria da escola que informou sobre a história acadêmica dos 220 estudantes da 1ª série do ensino médio do ano de 2007. Para cada um dos estudantes foram obtidos dados relativos às notas trimestrais, ao gênero, à idade, à frequência, à turma e ao professor. O QUADRO 4, nos informa sobre o número de estudantes por turma e gênero.

A segunda fonte de dados foi o banco de dados disponibilizado pelo Prof. Gomes e relativo à sua pesquisa sobre a bateria de testes cognitivos denominada BaFaCaLo (GOMES, 2010). No banco de dados havia informações referentes aos resultados de 18 testes de indicadores cognitivos, além de informações sobre o gênero, a idade, a renda familiar, a escolaridade do pai e a escolaridade da mãe de 292 estudantes, em que 231 eram do ensino médio e dos cursos técnicos e 61 eram do Programa de Educação de Jovens e Adultos (EJA) da escola pesquisada. Foram utilizados os dados de 121 estudantes da primeira série na presente pesquisa, sendo que 110 foram desconsiderados, pois eram estudantes da 2ª e da 3ª série e do EJA. O QUADRO 4, apresenta a relação dos estudantes da primeira série do ensino médio por turma e gênero que responderam aos testes BaFaCaLo.

QUADRO 4
Relação do número de estudantes do ensino médio e técnico, segundo o gênero, turma e série, fornecida pelo professor Cristiano

Série	Turma	Estudantes por Turmas	Masculino (1)	Feminino (0)
1ª série	T1	20	11	9
	T2	15	10	5
	T3	13	9	4
	T4	19	10	9
	T5	21	10	11
	T6	1	1	0
	T7	14	8	6
	T8	18	10	8
-	Total	121	69	52

A terceira fonte de dados foi obtida junto à comissão de vestibular da universidade mantenedora da escola pesquisada, que forneceu dados relativos ao questionário sócio-econômico respondido pelos candidatos aos vestibulares de 2008, 2009 e 2010¹⁴. Destes questionários foram selecionados dados de 494 candidatos aparentemente oriundos da escola pesquisada, sendo que destes, 111 eram candidatos ao vestibular de 2008, 180 eram candidatos ao vestibular de 2009 e 203 eram candidatos ao vestibular de 2010. Das variáveis disponíveis nos questionários sócio-econômicos foram selecionadas aquelas referentes à data de nascimento, à idade (a redundância existia nos dados), ao ano de conclusão do ensino médio, ao tipo de escola que estudou no ensino médio, à condição de 'treineiro', ao gênero, ao curso escolhido, à escolaridade do pai, à escolaridade da mãe e à renda familiar. Foram aproveitados os dados 328 candidatos, sendo 102 relativos à 2008, 120 relativos à 2009 e 106 relativos à 2010. Esta fonte foi utilizada na esperança de conseguir dados relativos às variáveis demográficas e contextuais dos estudantes

¹⁴ Os dados foram gentilmente cedidos pelo Prof. Arnaldo Vaz.

da primeira série da escola pesquisada que não responderam aos questionários da BaFaCaLo e complementar os dados faltosos daqueles que responderam aos questionários.

3.3.1 - Preparação dos dados

Esta fase iniciou – se pelo alinhamento dos dados fornecidos pela primeira e segunda fonte de dados. O alinhamento consistiu na construção de um banco de dados com um único registro, por estudante, contendo todas as informações relativas a cada estudante¹⁵. A terceira fonte de dados foi utilizada para complementar os dados faltosos da segunda fonte e obter dados relativos aos estudantes que não estavam arrolados na segunda fonte. Além disto, ela foi utilizada para checar os dados da segunda fonte. Os questionários da segunda fonte foram digitados por diversas pessoas e os procedimentos de checagem não foram muito rígidos. Desta forma, nos casos em que os dados da segunda e da terceira fontes não coincidiam, optou-se por reter no banco de dados os da terceira fonte, visto que, nesta fonte, os dados existentes foram informados em questionário on line, pelo próprio candidato ao vestibular. Ao final deste processo de alinhamento o nome dos estudantes foi mascarado e a partir deste momento cada estudante foi identificado por um código numérico de quatro dígitos.

Em seguida, os dados foram preparados para o estudo multinível. Em um estudo longitudinal os dados são coletados em diversas ocasiões. Eventualmente em uma das ocasiões não se consegue os dados de alguns sujeitos, o que gera o problema dos dados faltosos. Uma das vantagens dos algoritmos de ajuste dos modelos multiníveis é a facilidade em lidar com dados faltantes. De fato, o MLwin, o software utilizado nesta pesquisa para ajuste dos modelos multiníveis, consegue estimar os parâmetros dos mesmos na presença de dados faltantes. No nosso caso, faltavam dados de notas trimestrais de 18 estudantes, sendo que não haviam dados de nenhuma das três notas trimestrais de oito estudantes. Além disso, faltavam dados

¹⁵ O Banco de dados foi implementado em uma planilha Excel com uma linha por estudante e uma coluna por variável.

de pelo menos uma nota trimestral de outros oito estudantes e de duas notas trimestrais de um estudante. Apesar do MLwin ser capaz de estimar os modelos na presença destes dados, considerando-se o tamanho da amostra (220 estudantes), optou – se, nesta pesquisa, por retirar estes 18 estudantes da amostra analisada. Após esta eliminação, ficamos com uma amostra contendo dados de 202 estudantes, sendo que destes, 115 eram meninos e 87 eram meninas. Comparando-se o QUADRO 5 com o QUADRO 3 percebe-se que o número de estudantes do sexo masculino por turma continua maior que a do sexo feminino, exceto para a turma T5.

QUADRO 5
Relação de Estudantes por turma e gênero na primeira série do ensino médio

Turmas	Sujeitos / turma	Masculino (1)	Feminino (0)
T1	24	15	9
T2	25	15	10
T3	25	16	9
T4	26	14	12
T5	25	12	13
T6	26	15	11
T7	24	13	11
T8	27	15	12
Total	202	115	87

Finalmente foi preparada uma tabela para a análise multinível, com um registro por estudante e por ocasião de medida. Assim, nesta tabela, cada estudante ocupou três linhas, uma por ocasião de medida. As únicas variáveis com valores distintos nas três ocasiões eram a competência em física escolar e o tempo. As demais variáveis eram repetidas em cada ocasião. Este procedimento foi necessário, pois esta é a estrutura de dados utilizadas para a análise no MLwin.

3.4 - As variáveis da pesquisa

Esta pesquisa considera variáveis organísmicas, específicas de cada sujeito, e variáveis contextuais, que caracterizam o ambiente social, econômico, escolar e cultural do estudante. A principal variável organísmica, que é a variável de resposta cuja evolução temporal é investigada, é a competência em física escolar. Esta variável é uma variável latente, isto é, não é diretamente observável, porém admite-se que ela se refere a um aspecto real da estrutura mental dos estudantes. Outras variáveis organísmicas importantes são as habilidades cognitivas do estrato II do modelo de Carroll, que também se concebem como variáveis latentes não diretamente observáveis. No entanto, nesta pesquisa, o nível destas habilidades é indicado pelo seu escore normalizado, obtido a partir da soma dos escores brutos nos testes marcadores das habilidades cognitivas especializadas do estrato I. Esta opção metodológica possibilita um viés teórico de tomar essas habilidades cognitivas amplas como variáveis manifestas, diretamente observáveis. As habilidades cognitivas amplas, consideradas na pesquisa são a inteligência fluida (Gf), a inteligência cristalizada (Gc), a habilidade visuo-espacial (Gv), a fluência (Gr), a memória de curto prazo (Gy) e a rapidez cognitiva (Gs). O gênero e a idade completam as variáveis organísmicas.

As variáveis contextuais caracterizam o ambiente escolar, o ambiente sócio-econômico e o ambiente cultural do estudante. O ambiente escolar é caracterizado pelas seguintes variáveis: a frequência do estudante, a sua turma e o seu professor. O ambiente sócio-econômico é caracterizado por uma variável relacionada à renda da família do estudante, e o ambiente cultural, por variáveis que descrevem a escolaridade de seu pai e a escolaridade de sua mãe.

Nas subseções seguintes descrevem-se todas essas variáveis de forma mais detalhada. Nesta descrição e no capítulo seguinte, agrupamos as variáveis organísmicas e contextuais em variáveis demográficas que englobam o gênero, a idade e renda familiar e variáveis contextuais, que englobam as variáveis relativas à frequência do estudante, a sua turma, o seu professor, a escolaridade de seu pai e a escolaridade de sua mãe. Esta separação é instrumental, pois o foco principal dessa pesquisa é a investigação da possível relação entre a evolução temporal da

competência em física escolar e as habilidades cognitivas amplas, após controlar para os efeitos demográficos e contextuais.

O QUADRO 6 apresenta o tipo e a natureza de todas variáveis preditoras utilizadas na pesquisa.

QUADRO 6
Tipo e natureza das variáveis preditoras da pesquisa

-	Tipo da variável			Natureza da Variável		
	Organísmica	Demográfica	Contextual	Categórica	Dicotômica	Contínua
Variáveis Preditoras da pesquisa						
Competência em física escolar (CFIS)	X	-	-	-	-	X
Habilidades cognitivas amplas do Estrato II de Carroll (Gf, Gc, Gv, Gr, Gy, Gs)	X	-	-	-	-	X
Gênero (Ge)	X	X	-	-	X	-
Idade (Cid)	X	X	-	X	-	-
Renda familiar (Re)	X	X	-	X	-	-
Frequência (categoria de assiduidade) (CA)	-	-	X	-	X	-
Turma (T _n)	-	-	X	X	-	-
Professor (P _n)	-	-	X	X	-	-
Escolaridade do Pai (Escpai)	-	-	X	X	-	-
Escolaridade da Mãe (Escmae)	-	-	X	X	-	-

3.4.1 - Competência em física escolar

A variável competência em física escolar será construída a partir das notas dos estudantes em cada um dos trimestres. Na disciplina Física lecionada para a 1ª série na escola pesquisada, a nota trimestral refletia o resultado do desempenho dos estudantes em provas e testes, nas atividades de classe e extraclasse, nas atividades práticas e na valorização de comportamento e atitude dos estudantes. Em

cada trimestre, a nota é um número inteiro e pode ser considerada uma variável categórica, no entanto, se assim for, a variável terá muitas categorias (30 a 35 dependendo do trimestre), e muitas destas categorias são pouco populadas (menos de cinco pessoas por categoria) ou mesmo estão vazias. E isto faz com que o conjunto de dados se torne extremamente esparso.

Além disto, a aplicação dos critérios de atribuição de notas pelos professores variava entre as turmas, mesmo entre as turmas de um mesmo professor, mas também variavam ao longo dos três trimestres dentro de uma mesma turma. Essa hipótese de diversidade de aplicação de critérios é baseada em conhecimento de senso comum entre os professores. Mesmo a literatura educacional reconhece essas fontes de variação no comportamento dos professores e defende a utilização de procedimentos de equalização das avaliações para torná-las comparáveis entre si (LIVINGSTON & KIM, 2009). Visando minorar os efeitos descritos anteriormente, nesta pesquisa, as notas serão colapsadas em um número menor de categorias, denominadas de conceitos.

Na universidade mantenedora da escola pesquisada, o desempenho dos estudantes da graduação e pós-graduação tem uma dupla representação como nota e como conceito. A nota é um número inteiro de zero à cem, enquanto o conceito é expresso no sistema de seis categorias (A, B, C, D, E e F). O conceito é atribuído a partir da nota alcançada pelo estudante. O QUADRO 7 mostra as faixas de notas e os conceitos correspondentes usados na graduação e pós graduação na universidade.

QUADRO 7
Conceitos utilizados na universidade segundo a faixa de notas

Conceito	Valor relativo	Faixa de nota
F	0	Nota < 40
E	1	$40 \leq \text{nota} < 60$
D	2	$60 \leq \text{nota} < 70$
C	3	$70 \leq \text{nota} < 80$
B	4	$80 \leq \text{nota} < 90$
A	5	$90 \leq \text{nota} \leq 100$

Nesta pesquisa, tanto quanto possível, será usado o mesmo critério da universidade. Eventualmente, se alguma categoria de conceito estiver pouco populada ou vazia, ela poderá ser colapsada, agregando-a a categoria anterior ou à posterior. Desta forma, espera-se que nenhuma das categorias de conceito esteja pouco populada, o que atenuará o problema dos dados esparsos. No entanto, a conversão das notas em conceito pode resultar em variáveis categóricas que têm significados diferentes nas diversas turmas. A categoria de conceito tem o mesmo significado para dois grupos populacionais quaisquer, se a dificuldades de obtenção deste conceito for similar nos dois grupos. O indicador usual de dificuldade das diversas categorias é a proporção de pessoas em cada uma delas. Para minimizar esta dificuldade, utilizaremos um procedimento de equalização adequado para pequenas amostras.

Dentre os diversos procedimentos para equalização de pequenas amostras e para medidas repetidas (várias ondas de dados coletados na mesma população), a equalização via ajuste de um modelo Rasch aos dados é considerada a mais adequada (LINACRE, 2009; CAVANAGH & WAUGH, 2011) Este procedimento é particularmente adequado para estudos longitudinais com medidas repetidas, devido ao fato de que o modelo obedece ao princípio da objetividade específica (RASCH, 2008). Isto significa que em uma mesma ocasião de medida, é possível comparar as habilidades das pessoas independentemente de quais sejam os estímulos a que elas respondem e é possível comparar a dificuldade dos estímulos independente de quais sejam as pessoas que respondam a eles. Entre duas ocasiões diferentes é possível comparar as dificuldades dos estímulos, se um conjunto de pessoas

responde a eles, nas duas ocasiões, e é possível comparar as habilidades da mesma pessoa em duas ocasiões, se ela responde aos mesmos estímulos, em cada uma das ocasiões.

O princípio da objetividade específica permite formular vários esquemas de equalização de múltiplas ondas de dados. Considere que sujeitos respondam aos mesmos n itens em duas ocasiões distintas. Neste caso, os dados podem ser organizados em uma tabela com n colunas (organização sequencial dos dados) e $2m$ linhas ou em uma tabela com m linhas e $2n$ colunas (organização paralela dos dados), conforme exemplificado no QUADRO 8. No primeiro caso, assume-se que a dificuldade dos estímulos é a mesma nas duas ocasiões de medida, porém a habilidade dos sujeitos pode ser diferente nestas duas ocasiões. No segundo caso, assume-se que a habilidade do sujeito permanece inalterada nas duas ocasiões, porém pode haver mudança na dificuldade dos estímulos entre as duas ocasiões de medidas (LINACRE, 2009). Estes dois esquemas gerais podem ser misturados se parte dos itens são comuns entre as ocasiões e parte são específicos por ocasião. Linacre (2009) sugere um esquema de organização dos dados, a equalização virtual mostrada no QUADRO 9, em que nenhum dos estímulos é comum às diversas ocasiões de medidas. No QUADRO 9, as colunas sombreadas são comuns às duas ocasiões de medidas e as não sombreadas são específicas por ocasião.

QUADRO 8
Duas formas de organização dos dados de duas ondas de medidas

		Sequencial					Paralela							
		Itens					1ª ocasião				2ª ocasião			
		sujeito	1	2	...	n								
1ª ocasião	1	X	X	...	X	sujeito	1	2	...	n	1	2	...	n
	2	X	X	...	X	1	X	X	...	1	X	X	...	X
	2	X	X	...	0	X	X	...	X
	m	X	X	...	X
	1	X	X	...	X	m	X	X	...	X	X	X	...	X
2ª ocasião	2	X	X	...	X									
									
	m	X	X	...	X									
	m	X	X	...	X									

Nesta pesquisa utilizaremos testes de aderência para verificar se as categorias de conceitos, em uma mesma turma, retêm o seu significado nos três trimestres. Se o grau de dificuldade das categorias for o mesmo nos três trimestres em cada turma, podemos usar a organização sequencial dos dados. A organização paralela dos dados pressupõe que os sujeitos não evoluem ao longo das ocasiões de medidas. Assim, este esquema de equalização é inadequado para a presente pesquisa, tendo em vista que seu objetivo básico envolve estudar a evolução temporal da competência em física escolar. Desta forma, se o grau de dificuldade das categorias

não for o mesmo nos três trimestres e em cada turma, podemos utilizar a equalização virtual, sugerida por Linacre (2009).

QUADRO 9
Organização dos dados para equalização virtual

		Itens									
		sujeito	1	2	3	...	n-1	n	n+1	n+2	n+3
1ª Ocasião	1	X	X	X	...	X	X				
	2	X	X	X	...	X	X				
				
	m	X	X	X	...	X	X				
2ª Ocasião	1	X	X	X	X	X	X	
	2	X	X	X	X	X	X	
	
	m	X			...	X	X	X	X	X	

A última dificuldade, no processo de construção da medida da competência em física escolar, é que em cada trimestre temos apenas uma variável categórica mensurando os sujeitos. Nestas circunstâncias, o modelo Rasch não consegue estimar ao mesmo tempo a dificuldade das categorias e os parâmetros das pessoas (competência em física escolar de cada sujeito). Para superar esta dificuldade, transformaremos a variável categórica em uma escala do tipo Guttman, (LINACRE, 2009). Numa escala desse tipo, as categorias são ordenadas segundo seu grau de dificuldade. Se um sujeito satisfaz o critério para pertencer à categoria X ele também satisfaz o critério para cada uma das categorias anteriores a X. O vetor de resposta de cada sujeito é então um vetor ordenado. Linacre chama a atenção para o fato de que o modelo Rasch não é capaz de estimar simultaneamente a dificuldade dos itens e a habilidade das pessoas se os vetores de respostas dos diversos sujeitos são todos eles ordenados. Mas, neste caso, Linacre sugere como um truque a introdução de um sujeito virtual ao qual atribuímos um vetor de respostas não ordenado. A incorporação deste sujeito virtual aos dados faz com que o modelo consiga estimar ao mesmo tempo a dificuldade dos itens e a habilidade das pessoas.

3.4.2 Habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll

Ao aplicar a bateria BaFaCaLo para os estudantes da primeira série da escola pesquisada e outros conjuntos de sujeitos, Gomes (2010) conseguiu identificar que os testes, que compõem a bateria, mensuravam seis habilidades do estrato II do modelo de Carroll (1993): inteligência fluida (Gf), inteligência cristalizada (Gc) , habilidade visuo-espacial (Gv), memória de curto prazo (Gy), rapidez cognitiva (Gs) e fluência(Gr). Detalhes específicos sobre a aplicação da bateria BaFaCaLo, podem ser obtidos em Gomes (2010). O QUADRO 10 mostra quais os testes da bateria BaFaCaLo que mensuram cada uma das seis habilidades amplas do estrato II de Carroll (1993). Observa-se que um mesmo teste pode ser indicador de diferentes habilidades. Por exemplo, o teste de memória visual (MV) é utilizado para medir inteligência visuo espacial (Gv) e a memória de curto prazo (Gy).

QUADRO 10
Seis habilidades amplas do estrato II de Carroll

Habilidade	Teste
inteligência fluida (Gf)	I; RL; RG e N
inteligência cristalizada (Gc)	V1; V2 e V3
habilidade visuo-espacial (Gv)	Vz; CF; I; MV; P2; P3; N e RL
fluência (Gr)	FI1; FI2 e FF
memória de curto prazo (Gy)	MV; MA1 e MA2
rapidez cognitiva (Gs)	P1; P2; P3 e N

Para cada habilidade obteremos um escore bruto somando-se os escores obtidos pelo estudante nos testes indicadores desta habilidade. Nesta pesquisa a medida de cada uma das seis habilidades cognitivas do estudante será o seu escore normalizado em cada habilidade.

Para cada habilidade o escore normalizado de cada estudante é a razão entre o desvio do seu escore bruto na habilidade e o desvio padrão do escore bruto da habilidade:

$$NG = \frac{G - G_{\text{m\u00e9dio}}}{\sigma_G}$$

onde, NG \u00e9 escore normalizado da habilidade, G \u00e9 o escore bruto da habilidade, $G_{\text{m\u00e9dio}}$ \u00e9 escore bruto m\u00e9dio da habilidade e σ_G \u00e9 o desvio padr\u00e3o da habilidade. Nesta pesquisa a m\u00e9dia e o desvio padr\u00e3o das habilidades ser\u00e3o calculados para o subconjunto dos estudantes da primeira s\u00e9rie que participaram da pesquisa de Gomes, no ano de 2007. O uso do escore normalizado faz com que todas as habilidades sejam mensuradas em uma mesma faixa de valores, usualmente, de -3,0 a 3,0.

3.4.3 - Vari\u00e1veis demogr\u00e1ficas

Nesta pesquisa consideraremos as seguintes vari\u00e1veis demogr\u00e1ficas: g\u00eanero, idade e renda familiar do estudante.

G\u00eanero (Ge)

O g\u00eanero de cada estudante foi identificado a partir do banco de dados disponibilizado pelo professor Gomes complementado pela identifica\u00e7\u00e3o do g\u00eanero do estudante a partir do seu nome. Foram identificados o g\u00eanero de cada um dos 202 estudantes que comp\u00f5em a amostra. A vari\u00e1vel g\u00eanero foi tratada como uma vari\u00e1vel dicot\u00f4mica de valor zero (0) para o menina e um (1) para o menino.

Idade do estudante (Cid)

A partir das tr\u00eas fontes de dados mencionados anteriormente, a idade da maioria dos estudantes foi obtida como um n\u00famero inteiro entre 14 e 19 anos. Desta forma tratamos a vari\u00e1vel como categ\u00f3rica com tr\u00eas categorias. A primeira categoria 14 anos, a segunda categoria 15 anos e a terceira categoria 16 anos ou mais. Utilizando-se das tr\u00eas fontes de dados conseguimos identificar a idade de 154 dos 202 estudantes que comp\u00f5e a amostra.

Renda Familiar (Re)

A variável Renda Familiar foi identificada a partir dos dados disponibilizados pela segunda e terceira fontes mencionadas na seção 3.3. As duas fontes utilizavam sistemas categóricos distintos. Para esta pesquisa utilizamos três classes de renda familiar: abaixo de cinco salários mínimos, entre cinco e dez salários mínimos e acima de dez salários mínimos. O cruzamento da segunda e terceira fontes de dados permitiu a identificação da classe de renda familiar de 116 estudantes.

3.4.4 - Variáveis de Contextos

Nesta pesquisa consideraremos variáveis de contexto a frequência do estudante (assiduidade), a turma à qual pertence, o seu professor, a escolaridade do pai e a escolaridade da mãe do estudante.

Categoria de Assiduidade (CA)

A Variável assiduidade do estudante foi acessada a partir dos dados disponibilizados da primeira fonte mencionada na seção 3.3. Partindo do princípio que alguns professores não têm o hábito de fazer chamada sistematicamente e alguns até mesmo a fazem esporadicamente, não podemos, em princípio, comparar a assiduidade de estudantes de turmas diferentes, no entanto podemos comparar a assiduidade de estudantes da mesma turma. Por exemplo, é fácil distinguir os estudantes mais assíduos dos menos assíduos. Com isto, nesta pesquisa, optou-se por tratar a assiduidade como uma variável dicotômica indicando se o estudante é tão assíduo ou menos assíduo que a média da turma ou se é mais assíduo que a média da turma. A assiduidade foi avaliada em cada ocasião de medida, ou seja, em cada trimestre.

Turma (T_n)

A variável turma foi identificada a partir dos dados disponibilizados pela primeira fonte mencionada na seção 3.3. Na escola havia oito turmas na 1ª série, identificadas nesta pesquisa pelos rótulos T_1 à T_8 . Assim, a variável turma é uma variável categórica com oito categorias.

Professor (P_n)

No ano de 2007, quatro professores, identificados nesta pesquisa pelos rótulos P_1 à P_4 , lecionaram física para as oito turmas da primeira série da escola pesquisada. Desta forma, a variável professor é uma variável categórica com quatro categorias.

Escolaridade do Pai (Escpai)

A variável escolaridade do pai foi identificada a partir dos dados disponibilizados da segunda e da terceira fontes mencionadas na seção 3.3. As duas fontes utilizavam sistemas de categorização distintos. Para esta pesquisa utilizamos quatro classes de escolaridade do pai para os estudantes: pai que não concluiu o ensino fundamental, pai que concluiu o ensino fundamental, pai que concluiu o ensino médio, pai que concluiu o ensino superior. Desta forma, a variável escolaridade do pai é uma variável categórica com quatro categorias. O cruzamento da segunda e terceira fontes de dados permitiu a identificação da classe de escolaridade do pai de 160 estudantes.

Escolaridade da Mãe (Escmae)

A variável escolaridade da mãe foi identificada a partir dos dados disponibilizados da segunda e da terceira fontes mencionadas na seção 3.3. As duas fontes utilizavam sistemas de categorização distintos. Para esta pesquisa utilizamos quatro classes de escolaridade da mãe para os estudantes: mãe que não concluiu o ensino fundamental, mãe que concluiu o ensino fundamental, mãe que concluiu o ensino médio, mãe que concluiu o ensino superior. Desta forma, a variável escolaridade da mãe é uma variável categórica com quatro categorias. O cruzamento da segunda e terceira fontes de dados permitiu a identificação da classe de escolaridade da mãe de 162 estudantes.

Variáveis que identificam as ondas de dados

A essência de um estudo longitudinal está em analisar como a variável de resposta, no nosso caso, a competência em física escolar, evolui ao longo das ondas de dados, ao longo das ocasiões de medidas. A ocasião de medida é uma variável categórica e, nesta pesquisa, possui três categorias que identificam cada trimestre.

O sucesso de um estudo longitudinal, segundo Singer & Willet (2003), depende de uma escolha sensível da métrica temporal. Ainda que, em cada trimestre, a nota do estudante seja apurada de forma distribuída ao longo do trimestre, consideramos que a apuração da nota só se completa na última semana do trimestre, a 14ª semana no primeiro trimestre, a 30ª no segundo e a 44ª no terceiro trimestre. Nesta pesquisa, decidimos tratar o tempo como uma fração do ano letivo. Considerou-se como instante inicial (tempo = 0) a semana imediatamente anterior à primeira semana de aula. E como instante final (tempo = 1), a 44ª semana de aula. Desta forma considerou-se a primeira ocasião no tempo = 0,3182; a segunda no tempo = 0,6818 e a terceira ocasião no tempo = 1.

3.5 - Análise dos dados

A análise dos dados nesta pesquisa será feita em duas etapas, uma análise exploratória descritiva e a construção de um modelo longitudinal multinível de regressão múltipla. A primeira etapa, análise exploratória, visa apenas ganhar um senso dos dados e de suas possíveis relações. A segunda etapa, a construção do modelo longitudinal multinível de regressão múltipla, visa investigar a relação entre a competência em física escolar e seis das habilidades cognitivas do estrato II do modelo de estruturas das habilidades intelectuais humanas, proposto por Carroll, controlando para os efeitos das variáveis demográficas e contextuais, tanto no nível intrapessoal como no nível interpessoal do modelo longitudinal.

O modelamento utilizou o aplicativo MLwin (RABASH, BROWNE & GOLDSTEIN, 2009). O ajuste do modelo aos dados foi feito em uma sequência de modelos partindo do mais simples, o modelo incondicional de médias, seguido do modelo incondicional de mudança e, a partir daí, acrescentando-se uma a uma as variáveis demográficas e contextuais. O ajuste do modelo mais complexo será analisado com base na variação da estatística deviança em relação a deviança do modelo mais simples anterior, e estatística Wald para os coeficientes de regressão (SINGLER & WILLET, 2003). Cada variável preditora será introduzida no modelo, simultaneamente, no intercepto e na inclinação. Caso o modelo não se ajuste bem aos dados, a variável preditora será testada separadamente, primeiro, no intercepto e em seguida, na inclinação.

Neste capítulo, foram apresentados os aspectos metodológicos importantes da pesquisa aqui relatada, bem como as opções metodológicas encaminhadas. Destaca-se, em particular, a descrição do processo de construção da variável de resposta analisada no modelo longitudinal, que é a competência em física escolar, uma medida intervalar que representa a aprendizagem em física escolar. Descreveu-se também a construção das variáveis intervalares relativas às habilidades cognitivas mensuradas pela bateria BaFaCaLo. Cada variável demográfica e contextual também foi descrita. No próximo capítulo, apresentam-se os resultados, sua análise e interpretação.

4. RESULTADOS

Este capítulo apresenta os resultados de uma pesquisa que focaliza a relação entre a aprendizagem em física escolar, representada pela competência em física escolar, e algumas habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll (1993), depois de controlado para os efeitos das variáveis demográficas e contextuais. A pesquisa foi organizada utilizando os dados cognitivos dos estudantes (GOMES, 2010) de uma instituição pública federal e os dados das avaliações escolares ordinárias existentes na secretaria da escola. Inicialmente, apresenta os resultados sobre a construção da escala intervalar para mensurar a competência em física escolar. Em seguida, apresenta a construção das escalas das habilidades cognitivas amplas (Gf, Gc, Gv, Gr, Gy e Gs), mensuradas pela bateria BaFaCaLo. Prossegue descrevendo os indicadores relativos às variáveis demográficas e contextuais. Finalmente, descrevem-se os resultados advindos da construção do modelo longitudinal multinível de regressão múltipla para a evolução temporal da competência em física escolar.

4.1 - A competência em física escolar

4.1.1 - Construção da medida da competência em física escolar

Nesta seção serão apresentados os procedimentos metodológicos que foram utilizados no processo de construção da variável competência em física escolar (CFIS). O primeiro passo desse processo consistiu em obter da secretaria da escola pesquisada os dados relativos às avaliações trimestrais de alunos de 8 turmas da primeira série do ensino médio do ano de 2007. As notas trimestrais são representadas por números inteiros, variando de 0 a 30, no primeiro trimestre e de 0 a 35, nos dois trimestres seguintes. No capítulo 2 já foram apresentados os argumentos que sustentam o uso das notas escolares usuais como um indicador bruto de desempenho que permite a construção da variável competência em física escolar.

O segundo passo do processo de construção da variável competência em física escolar consistiu em adotar uma escala de conceitos para categorizar as notas. É sabido que a nota atribuída pelo professor é usualmente considerada uma medida ordinal. Como a nota é um número inteiro entre zero e trinta ou entre zero e trinta e cinco, de acordo com o trimestre, isto significa que há um número muito alto de categorias de notas em cada trimestre (trinta e uma categorias no primeiro trimestre e trinta e seis categorias no segundo e terceiro trimestres). Como o número de estudantes por turma é pequeno (entre 24 e 27 estudantes), há poucos estudantes em cada categoria de notas em cada turma e muitas das categorias de notas estão vazias. Por esta, razão optou-se pela adoção de um número menor de categorias de desempenho dos estudantes, os conceitos. A escola pesquisada pertence a uma universidade que adota um duplo sistema de registro do desempenho escolar: o registro da nota e do conceito. As notas são atribuídas como um inteiro de 0 a 100. O conceito é atribuído segundo uma faixa de valores da nota obtida pelo estudante em cada disciplina como mostra no QUADRO 11.

QUADRO 11
Conceitos utilizados na universidade segundo a faixa de notas

Conceito	Valor relativo	Faixa de nota
F	0	Nota < 40
E	1	$40 \leq \text{nota} < 60$
D	2	$60 \leq \text{nota} < 70$
C	3	$70 \leq \text{nota} < 80$
B	4	$80 \leq \text{nota} < 90$
A	5	$90 \leq \text{nota} \leq 100$

O esquema de atribuição de conceitos adotado nesta pesquisa baseou-se em uma variação do esquema exposto anteriormente. Como o número de alunos que obteve conceito A e B nas avaliações trimestrais de física foi baixo, decidiu-se colapsar estas categorias em uma só. Assim, atribuiu-se a cada aluno e em cada trimestre um conceito segundo as faixas de notas mostradas no QUADRO 12.

QUADRO 12
Conceitos atribuídos aos estudantes segundo a faixa de notas

conceito	Valor relativo na variável categórica	Faixa de nota no 1º trimestre	Faixa de nota nos 2º e 3º trimestres
E	0	Nota < 12	Nota < 14
D	1	$12 \leq \text{nota} < 18$	$14 \leq \text{nota} < 21$
C	2	$18 \leq \text{nota} < 21$	$21 \leq \text{nota} < 24,5$
B	3	$21 \leq \text{nota} < 24$	$24,5 \leq \text{nota} < 28$
A	4	$24 \leq \text{nota} < 30$	$28 \leq \text{nota} < 35$

O QUADRO 13 mostra em cada turma a distribuição de estudantes pelos conceitos segundo a ocasião. Uma mera inspeção nesta tabela sugere que, em cada turma, as distribuições de estudantes pelos conceitos nas três ocasiões são diferentes entre si; e que em uma mesma ocasião, as distribuições de estudantes pelos conceitos nas diversas turmas, também são diferentes entre si. Em vista disto, decidiu-se testar estatisticamente estas duas hipóteses. Usaram-se dois procedimentos: Em cada turma testou-se a existência de similaridade entre as distribuições de estudantes pelos conceitos nas três ocasiões, utilizando-se um teste de aderência. O segundo procedimento consistiu em testar se em cada ocasião há similaridade da distribuição de estudantes pelos conceitos nas diversas turmas.

O primeiro procedimento visa verificar se os conceitos retêm o significado nas três ocasiões, neste caso reter o significado corresponde a ter a mesma dificuldade relativa, aferida pela proporção de estudantes em cada categoria. Optou-se pelo teste de aderência porque, segundo Triola (2005, p.435), “utiliza-se um teste de aderência para testar a hipótese de que uma distribuição de frequências observadas se ajusta (ou adere) a determinada distribuição teórica”. Não temos uma teoria que preveja uma distribuição esperada de estudantes pelos conceitos, portanto, optou-se por considerar a distribuição de estudantes pelos conceitos em uma dada ocasião como a distribuição esperada e testar se as distribuições nas outras duas ocasiões aderiam a esta distribuição esperada. Por exemplo, na turma 5, considerando-se a distribuição de estudantes pelos conceitos na terceira ocasião como a distribuição esperada, testou-se se a distribuição na primeira e segunda ocasião aderiam a esta

distribuição esperada usando-se o procedimento descrito por Triola (2005, p.435-40). Desta forma, em cada turma fizeram-se seis comparações, utilizando-se a correção de Bonferroni para o nível de significância (ABDI, 2007). Considerando o nível de significância de 5% ($\alpha = 0,05$) para um único teste, para seis comparações, adota-se $\alpha = 0,0083$ como novo nível de significância para a rejeição da hipótese nula. Em cada comparação a hipótese nula foi a de que a distribuição observada aderiria a distribuição esperada. Das oito turmas testadas, apenas duas turmas, T3 e T7, não apresentaram evidências para se rejeitar a hipótese nula.

QUADRO 13
Distribuição de estudantes por turma e conceito, segundo a ocasião

-	-	T1				-	T2			
-	Ocasões	1	2	3	Total	-	1	2	3	Total
Conceitos	0	6	4	4	14	-	1	4	3	8
	1	6	8	9	23	-	7	12	8	27
	2	2	6	6	14	-	6	6	8	20
	3	10	4	5	19	-	6	4	9	19
	4	2	4	2	8	-	8	2	0	10
	Total	26	26	26	78	-	28	28	28	84
-	-	-	-	-	-	-	-	-	-	-
-	-	T3				-	T4			
-	Ocasões	1	2	3	Total	-	1	2	3	Total
Conceitos	0	1	3	3	7	-	2	2	0	4
	1	4	4	5	13	-	9	11	8	28
	2	6	2	5	13	-	4	6	6	16
	3	10	9	6	25	-	8	3	9	20
	4	6	9	8	23	-	3	4	3	10
	Total	27	27	27	81	-	26	26	26	78
-	-	-	-	-	-	-	-	-	-	-
-	-	T5				-	T6			

	Ocasões	1	2	3	Total	-	1	2	3	Total
Conceitos	0	3	3	6	12	-	0	0	0	0
	1	9	7	6	22	-	4	6	6	16
	2	7	6	5	18	-	5	6	11	22
	3	1	3	6	10	-	10	10	7	27
	4	8	9	5	22	-	7	4	2	13
	Total	28	28	28	84	-	26	26	26	78
-	-	-	-	-	-	-	-	-	-	-
-		T7					T8			
-	Ocasões	1	2	3	Total	-	1	2	3	Total
Conceitos	0	3	1	2	6	-	2	2	2	6
	1	6	5	9	20	-	9	7	3	19
	2	8	8	5	21	-	7	5	3	15
	3	5	7	7	19	-	3	9	12	24
	4	2	3	1	6	-	6	4	7	17
	Total	24	24	24	72	-	27	27	27	81

O segundo procedimento testou, em cada ocasião, a similaridade das distribuições de conceitos entre as diversas turmas. Pelas mesmas razões anteriores, optou-se pelo teste de aderência. Em cada ocasião tomou-se a distribuição de estudantes pelos conceitos em cada turma, uma por vez, como a distribuição esperada e testou se as distribuições de estudantes pelos conceitos nas demais turmas aderiam a esta distribuição esperada. Em cada ocasião fez-se 56 comparações (8 turmas x 7 comparações), com um total de 168 comparações nas três ocasiões. Devido ao elevado número de comparações, a correção de Boferroni para o nível de significância não foi utilizada. Alternativamente, optou-se por realizar cada uma das 168 comparações com nível padrão de significância ($\alpha=0,05$) e contar o número de comparações em que há evidência para rejeitar a hipótese nula. Das 168 comparações, 92 apresentaram evidência para a rejeição da hipótese nula. Isso corresponde a 54,8% do total de testes realizados rejeitando a hipótese nula. Desta

forma, admitimos que em cada ocasião, a distribuição de estudantes pelos conceitos, nas diversas turmas, não são similares entre si e, portanto, os conceitos não retêm o significado ao longo do tempo e entre as turmas.

O próximo passo do processo de construção da variável competência em física escolar consistiu em extrair uma medida intervalar ajustando um modelo da família de modelos Rasch às variáveis categóricas disponíveis. A adoção de um procedimento de equalização permitiu construir uma medida de competência em física escolar que tem a mesma métrica nas três ocasiões.

A equalização das notas entre as ocasiões dentro da mesma turma e entre as diversas turmas é um procedimento controverso (DIVGI, 1987). Como o número de estudantes em cada turma é pequeno, o procedimento deve ser adequado para amostras pequenas. Conforme relatado no capítulo 3, o procedimento de equalização consistiu em categorizar as notas segundo um esquema de atribuição de conceitos, expressarem os conceitos em uma escala do tipo Guttman e ajustar um modelo Rasch a esta escala usando o procedimento de equalização virtual sugerido por Linacre (manual do Winsteps, 2011, p. 327). Este procedimento é adequado quando não há itens de avaliação comuns entre as diversas formas a serem equalizadas.

De fato, como mostrado anteriormente, a distribuição de estudantes pelos conceitos, em cada ocasião, difere entre as turmas; e em cada turma ela difere entre as ocasiões. Portanto, as variáveis categóricas que descrevem os conceitos alcançados pelos estudantes, em cada ocasião e em cada turma, não são comuns entre si, o que justificou a adoção do esquema de equalização virtual sugerido por Linacre (manual do Winsteps, 2011, p.327). O programa Winsteps pode ser usado para ajustar o modelo Rasch a uma única variável categórica, mas apenas consegue estimar a dificuldade de cada categoria, mas não, a habilidade do sujeito (LINACRE, 2009, p. 521). Uma escala do tipo Guttman, mesmo com muitos itens, não é estritamente analisável pelo mesmo programa, porém a introdução de respondente virtual, como um vetor de resposta não Guttman (vetor de resposta não ordenado) é suficiente para permitir a construção da medida e estimar a dificuldade dos itens (LINACRE, 2009, p. 501).

Nesta pesquisa os conceitos foram usados para a construção de uma escala do tipo Guttman, composta de quatro indicadores dicotômicos por ocasião, designados pela letra I seguida de um número ordinal, desde I1 até I12. Os indicadores I1, I5 e I9 indicam se na ocasião de avaliação o estudante ultrapassou o ponto de transição do conceito E para o conceito D. Os indicadores I2, I6 e I10 indicam se ele ultrapassou o ponto de transição do conceito D para C e assim sucessivamente. O QUADRO 14 mostra o esquema de conversão de conceito nos quatro indicadores na primeira ocasião. Nas demais ocasiões a conversão de conceitos em indicadores segue o mesmo esquema, apenas trocando os indicadores para I5 a I8 na segunda ocasião e para I9 a I12 na terceira ocasião.

QUADRO 14
Esquema de transformação dos conceitos atribuídos aos estudantes na primeira ocasião em indicadores de uma escala do tipo Guttman

Conceito	I1	I2	I3	I4
E	0	0	0	0
D	1	0	0	0
C	1	1	0	0
B	1	1	1	0
A	1	1	1	1

Assim, os conceitos alcançados nos três trimestres transformaram-se em 12 indicadores dicotômicos por estudante, que formam, em cada ocasião, uma escala do tipo Guttman com quatro itens cada. Conforme discutido por Linacre no manual do Winsteps (2011), não é possível ajustar um modelo Rasch a uma escala do tipo Guttman sem introduzir um estudante fictício com um vetor de resposta não ordenado. Segundo Linacre, a interferência deste estudante virtual é pequena e é minimizada a medida que a amostra cresce. Como não se pode considerar que os conceitos em cada turma retinham seu significado ao longo das três ocasiões e que os conceitos relativos a cada ocasião diferiam em distribuição e significado, de uma turma para outra, organizou-se um quadro contendo 96 itens, 12 por turma, e 205 sujeitos, conforme o esquema mostrado no QUADRO 15. No QUADRO 15, são mostrados apenas 24 itens, correspondentes às três ocasiões de medida em duas turmas. Cada coluna contém os dados de uma ocasião de medida e dos estudantes de uma turma. Desta forma, o esquema de organização de dados corresponde à equalização de 24 formas distintas de um teste de 4 itens, sem itens comuns às diversas formas. Em cada turma, se o estudante não obteve conceito em determinada ocasião os itens correspondentes foram deixados vazios. O Winsteps automaticamente controla e retira os dados faltantes.

QUADRO 15
Relação dos conceitos atribuídos aos estudantes por ocasião segunda uma escala Guttman

Turma	Sujeito	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13	I14	I15	I16	I17	I18	I19	I20	I21	I22	I23	I24
T1	Suj-1	x	x	x	x																				
T1																				
T1	Suj_27	x	x	x	x																				
T1	Suj-1					x	x	x	x																
T1																
T1	Suj_27					x	x	x	x																
T1	Suj-1									x	x	x	x												
T1												
T1	Suj_27									x	x	x	x												
T2	Suj-1													x	x	x	x								
T2								
T2	Suj_25													x	x	x	x								
T2	Suj-1																	x	x	x	x				
T2				
T2	Suj_25																	x	x	x	x				
T2	Suj-1																					x	x	x	x
T2
T2	Suj_25																					x	x	x	x

Ao final do QUADRO 15, foi acrescentado um estudante inexistente, com escore 0111 (vetor de resposta não ordenado segundo a escala Guttman) em cada ocasião e em cada turma. Este acréscimo foi necessário para que o Winsteps estimasse os parâmetros do modelo para os itens que formam cada uma das escalas Guttman e para que os parâmetros dos 96 itens fossem estimados na mesma escala.

Após isto, usou-se o Winsteps para ajustar um modelo Rasch aos dados. O programa Winsteps fornece um conjunto de estatísticas para se avaliar o ajuste do modelo aos dados. Duas delas são importantes, o Infit e o Outfit. O infit é uma

estatística de ajuste ponderada pela informação e é mais sensível aos comportamentos inesperados que afetam as respostas a itens cuja dificuldade é próxima do nível de habilidade do respondente. No ajuste feito, o Infit médio para os itens foi de 0,87 com desvio padrão de 0,39. O máximo Infit dos itens foi de 1,48 e o mínimo igual a 0,39. Linacre sugere, por um lado, que itens com infit entre 0,5 e 1,5 são produtivos para a mensuração (LINACRE, 2011, p. 272), enquanto que itens com Infit acima de 1,5 degradam o processo de construção da medida. Por outro lado, itens com Infit abaixo de 0,50 são menos produtivos para o processo de mensuração, porém não o degradam. Mas podem levar a fidedignidade ou a índices de separação a normalmente bons. No resultado obtido, o infit máximo está na faixa produtiva, e o mínimo abaixo do valor desejável. Há 31 itens ou 32,3% dos itens com infit entre 0,49 e 0,39. A correlação escore bruto dos itens para medida foi de -0,94, com valor desejável de -1,0. O Modelo explica 79% da variância dos dados. O modelo ajustou-se de forma satisfatória aos dados categóricos.

Deste procedimento resulta uma medida intervalar que representa em uma mesma escala a competência em física escolar de cada estudante nas três ocasiões de medida. É possível comparar, em cada turma, a competência em física escolar de um estudante com as de seus colegas. No entanto, não é possível comparar, em uma mesma ocasião, a competência de um estudante da turma 1 com a de qualquer estudante de outra turma. Isto se deve ao fato de que as dificuldades médias dos itens, que formam a escala Guttman em cada ocasião e em cada turma, não são iguais entre si. O QUADRO 16 mostra a localização de cada item na escala de logit, comum a todos os itens, por turma e por ocasião. O Winsteps força a média da medida dos itens a ser nula. No entanto, a média dos itens, que compõem a escala de conceitos, em cada ocasião e turma, mostrada nas três últimas linhas do quadro, é diferente de zero.

QUADRO 16
 Categorização conceitual por turma em relação às dificuldades dos itens

Ocasião	Cat	Itens	M11	m12	M13	M14	M15	M16	M17	M18
1	E	1	- 4,04480	- 4,82410	- 4,02610	- 5,02520	- 4,93700	- 4,02100	- 4,41040	- 5,12890
	D	2	- 0,23090	- 1,11250	- 1,35420	- 0,65270	- 0,80370	- 1,28330	- 1,10360	- 0,98740
	C	3	- 0,37150	- 0,98720	- 0,76990	- 0,79910	- 1,91420	- 0,55170	- 1,53160	- 1,46160
	B	4	- 5,26210	- 4,38630	- 5,16270	- 4,94430	- 2,83190	- 5,00270	- 4,52550	- 3,61790
2	E	5	- 4,75460	- 5,40040	- 3,62040	- 5,32550	- 4,68500	- 4,65680	- 4,40050	- 4,75350
	D	6	- 0,80120	- 0,61710	- 0,57480	- 0,72810	- 0,98320	- 1,23210	- 1,42040	- 0,97050
	C	7	- 1,34480	- 1,51700	- 0,04330	- 1,44940	- 1,22690	- 0,84710	- 1,18820	- 0,81230
	B	8	- 4,01900	- 4,20130	- 4,65520	- 3,64790	- 3,41640	- 5,26160	- 4,79390	- 5,08780
3	E	9	- 4,99790	- 5,11940	- 4,21670	- 5,12050	- 4,37100	- 4,88970	- 4,96810	- 3,32350
	D	10	- 0,79890	- 1,23130	- 1,08600	- 1,12740	- 0,90780	- 1,65550	- 0,69590	- 0,92620
	C	11	- 1,30780	- 1,30830	- 0,75690	- 0,93740	- 0,91260	- 1,49370	- 1,09160	- 0,19510
	B	12	- 4,38280	- 5,43240	- 4,20590	- 5,15180	- 4,37600	- 5,01630	- 4,88570	- 5,21920
1	Média	- 0,33948	- 0,14078	- 0,13808	- 0,01638	- 0,24865	- 0,06253	- 0,13578	- 0,25920	
2	Média	- 0,04800	- 0,07480	- 0,12583	- 0,23908	- 0,25623	- 0,05495	- 0,04030	- 0,04403	
3	Média	- 0,02655	- 0,09750	- 0,08498	- 0,03967	- 0,00245	- 0,00880	- 0,07833	- 0,29115	

Uma redução das medidas ao mesmo referencial, tornando comparáveis as competências de estudantes diferentes em uma mesma ocasião, é possível se formos que a média dos quatro itens que compõem a escala de conceito em cada ocasião e em cada turma seja nula. Ao fazer esta correção, a medida do sujeito em cada ocasião é deslocada por um valor igual à média dos itens respectivos. O QUADRO 17 mostra o exemplo de três estudantes, com suas medidas não

corrigidas e a medida corrigida. Após a correção, a construção da medida de competência em física escolar em três ocasiões de medida está terminada.

QUADRO 17
Três situações de sujeitos para as três ocasiões com as medidas corrigidas e não corrigidas

Sujeito	Ocasião	Medida não corrigida	Correção	Medida corrigida
1101	1	3,0566	0,3395	2,7171
1101	2	0,2314	-0,0480	0,2794
1101	3	4,9602	-0,0265	4,9867
1201	1	2,1808	-0,1408	2,3216
1201	2	1,5735	-0,0748	1,6483
1201	3	1,5706	0,0975	1,4731
1301	1	2,2723	0,1381	2,1342
1301	2	1,8791	0,1258	1,7533
1301	3	1,6413	-0,0850	1,7263

4.1.2 – Uma análise exploratória da competência em física escolar

Para uma melhor percepção da variabilidade da variável competência em física escolar ao longo das três ocasiões e, conseqüentemente, tomar uma primeira noção de como essa variável modifica-se ao longo das ocasiões, fez-se necessário o procedimento de uma análise exploratória, a partir da obtenção das médias e dos desvios padrão da variável competência em física em cada ocasião e da construção de gráficos com a representação do perfil de $CFIS_{\text{médio}}$ ao longo das três ocasiões.

A partir da TAB. 1, verifica-se que a variável competência em física escolar média ($CFIS_{\text{médio}}$) decai de 0,13 da 1ª para a 2ª ocasião e em seguida, na 3ª ocasião, sofre um acréscimo de 0,11, indicando uma suspeita de que, em média, a competência física escolar dos estudantes na 1ª ocasião foi ligeiramente maior se comparada com a 2ª e 3ª ocasiões; e que, em média, os estudantes na 2ª ocasião, tiveram os menores valores de competência em física escolar. A dispersão da competência em física escolar média ($CFIS$) em cada ocasião é próxima, demonstrando que a variabilidade de $CFIS_{\text{médio}}$ dos estudantes em torno da média é pequena. A partir do

GRÁF. 1, observa-se que as barras de erro estão superpostas entre as três ocasiões sugerindo que não há diferença entre as médias entre as três ocasiões. As conclusões tiradas a partir das observações feitas, tanto na TAB 1 como no GRÁF.1, não são suficientes, para se desistir da exploração da evolução da competência em física escolar, pois a variação para cada estudante é diferente da variação média da competência em física escolar.

TABELA 1
Média e desvio padrão da competência em física escolar dos estudantes da 1ª série do ensino médio por ocasião

Ocasião	1ª	2ª	3ª
Média	0,68	0,55	0,66
Desvio padrão	3, 390	3, 356	3, 236
Número de Estudantes	202	202	202

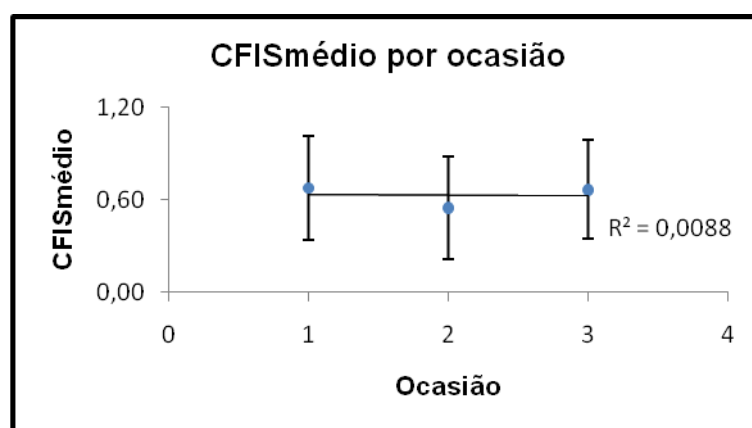


GRÁFICO 1 – Média da competência em física escolar dos estudantes da 1ª série do ensino médio

4.2 - Habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll

Os 18 testes da Bateria de Fatores Cognitivos de Alta Ordem (Bateria BaFaCaLo), desenvolvida por Gomes (2010) foram aplicados, no primeiro trimestre de 2007, a 116 dos 202 estudantes investigados nesta pesquisa. Esta Bateria avaliou seis das oito habilidades cognitivas amplas do estrato II do modelo de estrutura das

habilidades intelectuais humanas, proposto por Carroll (1993) (inteligência fluida, inteligência cristalizada, habilidade visuo-espacial, memória de curto prazo, rapidez cognitiva e fluência) além de um fator geral (g) no estrato III do modelo de Carroll (1993). Os resultados dos testes destas seis habilidades cognitivas foram utilizados nesta pesquisa para investigar uma possível relação entre as habilidades cognitivas amplas do estrato II do modelo de Carroll e a competência em física escolar. Uma descrição mais detalhada dos testes referentes às seis habilidades cognitivas apresentadas a seguir pode ser encontrada em Gomes (2010).

Na bateria BaFaCaLo:

- a inteligência fluida é identificada por quatro testes: rapidez numérica (N), raciocínio lógico (RL), raciocínio geral (RG) e indução (I). O escore bruto para cada estudante foi formado somando-se os escores brutos obtidos nos quatro testes citados acima. Considerando a amostra aqui estudada e os escores brutos respectivos, a inteligência fluida tem média 62,41, desvio padrão 13,531, escore bruto máximo igual a 93,0 e mínimo igual a 34,0;
- a inteligência cristalizada é identificada por três testes: compreensão verbal 1 (V1), compreensão verbal 2 (V2) e compreensão verbal 3 (V3). O escore bruto para cada estudante foi formado somando-se os escores brutos obtidos nos três testes V1, V2 e V3. Considerando a amostra aqui estudada e os escores brutos respectivos, a inteligência cristalizada tem média 21,52, desvio padrão 8,212, escore bruto máximo igual a 54,0 e mínimo igual a 7,0;
- a inteligência visuo-espacial é identificada por oito testes: visualização (VZ), flexibilidade de fechamento (CF), indução (I), memória visual (MV), velocidade perceptiva 2 (P2), velocidade perceptiva 3 (P3) rapidez numérica (N) e raciocínio lógico (RL). O escore bruto para cada estudante foi formado somando-se os escores brutos obtidos nos oito testes mencionados acima. Considerando a amostra aqui estudada e os escores brutos respectivos, a inteligência visuo-espacial tem média 167,27 e desvio padrão 25,403, escore bruto máximo igual a 221,0 e mínimo igual a 91,0;

- a fluência é identificada por três testes: fluência ideacional 1 (FI1), fluência ideacional 2 (FI2) e fluência figural (FF). O escore bruto para cada estudante foi formado somando-se os escores brutos obtidos nos três testes já mencionados. Considerando a amostra aqui estudada e os escores brutos respectivos, a fluência tem média 31,57, desvio padrão 11,588, escore bruto máximo igual a 72,0 e mínimo igual a 12,0;
- a memória de curto prazo é identificada por três testes: memória visual (MV), memória associativa 1 (MA1) e memória associativa 2 (MA2). O escore bruto para cada estudante foi formado somando-se os escores brutos obtidos nos três testes citados acima. Considerando a amostra aqui estudada e os escores brutos respectivos, a memória de curto prazo tem média 24,83, desvio padrão 7,872, escore bruto máximo igual a 42,0 e mínimo igual a 7,0;
- a rapidez cognitiva é identificada por quatro testes: velocidade perceptiva 1 (P1), velocidade perceptiva 2 (P2), velocidade perceptiva 3 (P3) e rapidez numérica (N) . O escore bruto para cada estudante foi formado somando-se os escores brutos obtidos nos quatro testes citados acima. Considerando a amostra aqui estudada e os escores brutos respectivos, a rapidez cognitiva tem média 108,95, desvio padrão 19,467, escore bruto máximo igual a 155,0 e mínimo igual a 60,0.

A partir dos resultados dos escores (valores brutos) de cada habilidade cognitiva descrita nos parágrafos anteriores foram construídas as medidas das habilidades cognitivas seguindo o procedimento descrito na seção 3.4.2. As habilidades cognitivas, seus escores mínimos e máximos, a medida da habilidade, a medida máxima e a mínima estão representadas no QUADRO 18.

QUADRO 18
Habilidades e seu escore mínimo e máximo, a medida da habilidade normalizada, seus valores mínimo e máximo

Habilidade	Escore mínimo	Escore máximo	Medida	Medida Mínima	Medida Máxima
Inteligência Fluida (Gf)	34,0	93,0	NGf	-2,116	2,274
Inteligência Cristalizada (Gc)	7,0	54,0	NGc	-1,785	3,926
Inteligência Visuo Espacial (Gv)	91,0	221,0	NGv	-3,035	2,117
Fluência (Gr)	12,0	72,0	NGr	-1,718	3,495
Memória de Curto Prazo (Gy)	7,0	42,0	NGy	-2,236	2,138
Rapidez Cognitiva (Gs)	60,0	155,0	NGs	-2,529	2,393

O QUADRO 18 exhibe uma das vantagens de se usar os escores padronizados ao invés do escore bruto como medida da habilidade: amplitude da faixa de valores das medidas padronizadas é menor do que a faixa dos escores brutos. Deve-se cuidar para que as faixas de variação das variáveis sejam similares porque, desta forma, é possível a comparação entre elas. Conforme sugerido por FOX & LEVIN (2004), o escore padronizado possibilita: (1) ter uma medida do quanto o valor bruto se desvia da média de uma distribuição em uma unidade de desvio padrão (σ); (2) fazer comparação entre o escore bruto de cada estudante em relação à média do grupo; e (3) ter uma noção da variabilidade do escore bruto de cada estudante em relação à média do grupo.

4.3 - Análise exploratória do efeito das variáveis demográficas e contextuais

Nesta seção, é apresentada uma descrição das variáveis demográficas (gênero, idade, renda familiar) e contextuais (turma, professor, escolaridade do pai e escolaridade da mãe). Será também apresentada uma análise exploratória da possível relação entre essas variáveis e a competência em Física escolar.

4.3.1 - Gênero

O gênero do estudante, identificado a partir dos dados da fonte 2, mencionada na seção 3.3, foi representada como uma variável dicotômica (0 ↔ meninas;

1↔meninos). Na amostra, há 87 meninas e 115 meninos. O QUADRO 19 representa a distribuição de gênero dos estudantes por turma. Percebe-se que em cada turma há mais meninos que meninas. A única turma que apresenta um maior número de meninas em relação aos meninos é na turma T5. No entanto, a proporção de meninos e meninas em cada turma é praticamente a mesma. Um teste de associação (Chi-quadrado) foi usado para verificar se há alguma relação entre turma e gênero. O resultado ($\chi^2 = 2,26$; $df = 7$; $N = 202$; $p = 0,94$) indica que o gênero não está associado à turma. Assim se houver algum efeito de turma na evolução da competência em física escolar ele não será confundido com um possível efeito de gênero.

QUADRO 19
Relação do número de meninas e menino por turma

Turmas	Estudantes	Meninos	Meninas
T1	24	15	9
T2	25	15	10
T3	25	16	9
T4	26	14	12
T5	25	12	13
T6	26	15	11
T7	24	13	11
T8	27	15	12
Total	202	115	87

A TAB. 2 apresenta a média e o desvio padrão da competência em física escolar para meninas e meninos segundo a ocasião. Por um lado, a média da competência em física escolar dos meninos é maior que a das meninas nas três ocasiões. Por outro lado, nas mesmas ocasiões, a competência em física escolar dos meninos está mais dispersa do que a das meninas.

TABELA 2
Competência em física escolar segundo gênero e ocasião de medida

Gênero	Ocasião 1		Ocasião 2		Ocasião 3	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
Feminino	0,543	3,391	0,486	3,314	0,653	3,190
Masculino	0,779	3,400	0,591	3,401	0,672	3,240
Total	0,677	3,390	0,546	3,356	0,664	3,236

O uso do teste ANOVA mostra que não há diferença entre a média da competência em física escolar de meninos e meninas em nenhuma das três ocasiões (1ª ocasião: $F = 0,241$, $df1 = 1$, $df2 = 200$; $p = 0,624$), (2ª ocasião: $F = 0,049$, $df1 = 1$, $df2 = 200$, $p = 0,826$) e (3ª ocasião: $F = 0,002$, $df1 = 1$, $df2 = 200$, $p = 0,968$). Apesar de esse resultado indicar que não seria esperado um efeito de gênero sobre a evolução da competência em física escolar, esta variável será testada no modelo multinível de regressão múltipla, dada a importância que o gênero tem nos estudos educacionais.

4.3.2 - Idade

A variável idade foi obtida pelo cruzamento dos dados disponibilizados pelas três fontes descritas na seção 3.3. Foram identificadas as idades de 154 de 202 estudantes da amostra. As idades identificadas variaram de 14 a 19 anos, sendo que 28,6% dos estudantes declararam possuir 14 anos, 52,6% declararam possuir 15 anos e 18,8% declararam possuir idade de 16 ou mais anos. A TAB. 3 apresenta valores da média e do desvio padrão da competência em física escolar dos estudantes da primeira série no ano de 2007 em relação à idade e a ocasião. Para cada idade, as médias parecem diferir segundo a ocasião.

TABELA 3
Média, desvio padrão da competência em física escolar e número de estudantes da 1ª série do ensino médio de 2007 por categoria de idade em anos.

	Idade	14	15	16 ou mais anos	Dados Faltosos
-	Número de estudantes	44	81	29	48
1ª Ocasião	Média	2,03	1,10	0,45	-1,01
	Desvio padrão	3,045	3,525	2,959	2,982
2ª Ocasião	Média	1,80	0,82	0,24	-0,89
	Desvio padrão	2,911	3,478	2,30	3,153
3ª Ocasião	Média	1,62	0,88	0,27	-0,35
	Desvio padrão	3,250	3,258	2,943	3,130

Usando o teste ANOVA, comparou-se a diferença entre as médias de CFIS dos estudantes agrupados segundo a idade e a ocasião. Observou-se que não há diferenças significativas entre as médias de CFIS nas três ocasiões (1ª ocasião: $F = 2,159$, $df1 = 2$, $df2 = 151$ e $p = 0,119$), (2ª ocasião: $F = 2,200$, $df1 = 2$, $df2 = 151$ e $p = 0,114$) e (3ª ocasião: $F = 1,648$, $df1 = 2$, $df2 = 151$ e $p = 0,196$). Este resultado sugere ser desnecessário incluir a idade como um possível preditor em um modelo multinível de regressão múltipla para a evolução temporal da competência em física escolar. Além disso, não se conseguiu informações confiáveis sobre a data de nascimento da maioria dos estudantes, o que torna a qualidade deste dado muito duvidosa e quaisquer conclusões nele baseadas.

4.3.3 - Renda familiar

Ainda que na literatura educacional, o status sócio econômico seja considerado um dos mais importantes preditores de desempenho acadêmico (FULLER & CLARKE 1994; LEES 1994; WHITE, 1982), o fato de se ter conseguido dados relativos a essa

variável de apenas 116 em 202 estudantes, levou a tomar a decisão de desconsiderar a variável nesta pesquisa.

4.3.4 - Turma

Em cada turma foram calculados a média e seu erro padrão. Para cada turma, as médias foram plotadas contra o tempo e uma reta de regressão simples foi ajustada aos dados. O GRÁF. 2, 3, 4, 5, 6, 7, 8 e 9 apresentam a evolução temporal da competência média em Física escolar para cada turma. Nos gráficos, a média é representada pela sua estimativa e a barra de erro corresponde ao seu intervalo de confiança de 95 %. Em cada gráfico, é apresentada uma estatística, o quadrado do coeficiente de correlação, que permite avaliar o ajuste da linha de tendência aos dados. Há retas muito bem ajustadas às médias, sugerindo uma clara tendência linear de evolução média, como são os casos das turmas 6 e 8, mas há casos, como da turma 1, em que não há nenhuma tendência linear. Estes resultados sugerem que deve haver um efeito de turma sobre a evolução temporal da competência em física escolar.

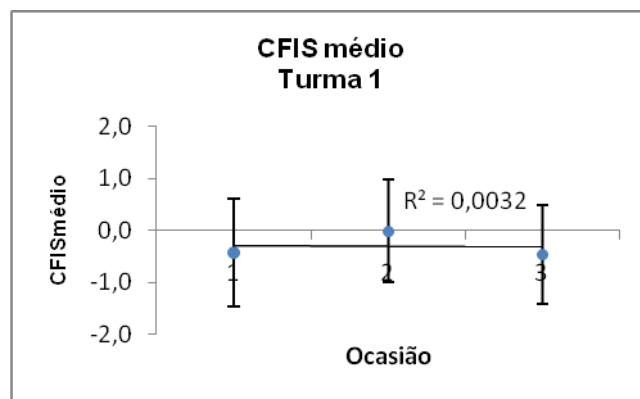


GRÁFICO 2 – Competência em física escolar média da turma 1

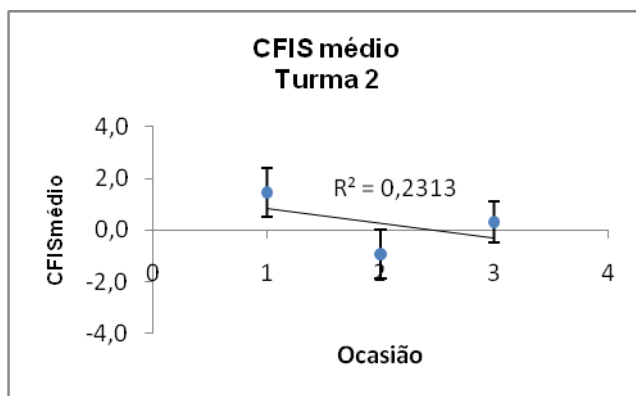


GRÁFICO 3 – Competência em física escolar média da turma 2

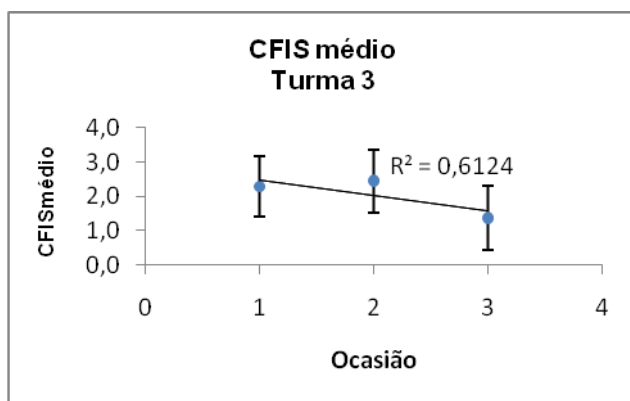


GRÁFICO 4 – Competência em física escolar média da turma 3

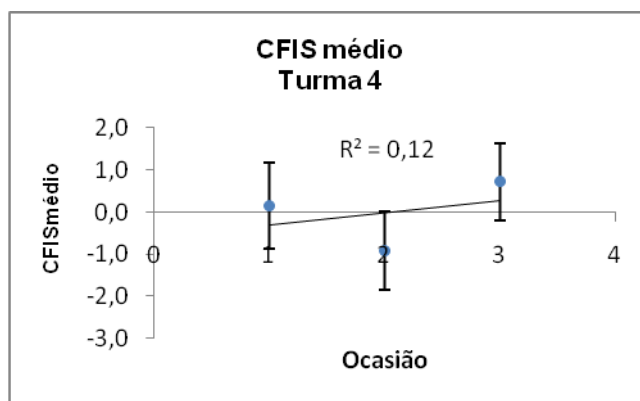


GRÁFICO 5 – Competência em física escolar média da turma 4

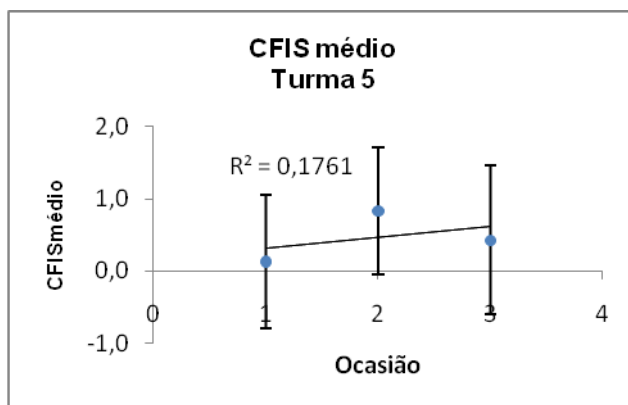


GRÁFICO 6 – Competência em física escolar média da turma 5

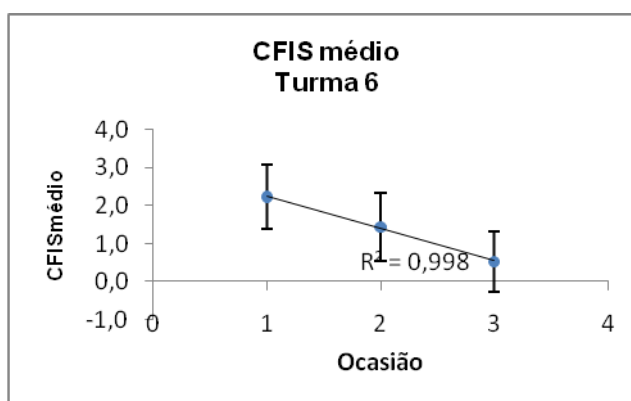


GRÁFICO 7 - Competência em física escolar média da turma 6

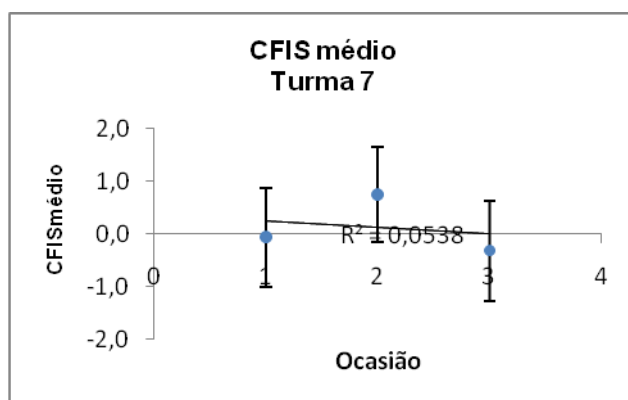


GRÁFICO 8 - Competência em física escolar média da turma 7

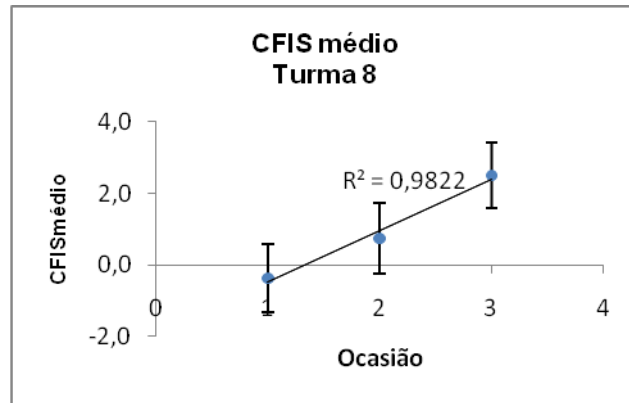


GRÁFICO 9 - Competência em física escolar média da turma 8

Usando o teste ANOVA, comparou-se a diferença entre as médias de CFIS para turmas. Observou-se que há diferenças entre as médias de competência em física escolar. Observou-se que há diferenças significativas entre as médias da competência em física escolar nas três ocasiões (1ª ocasião: $F = 2,662$, $df_1 = 7$, $df_2 = 204$ e $p = 0,012$), (2ª ocasião: $F = 3,089$, $df_1 = 7$, $df_2 = 204$ e $p = 0,004$) e (3ª ocasião: $F = 2,415$, $df_1 = 7$, $df_2 = 204$ e $p = 0,021$). Estes resultados confirmam que se deve esperar um efeito de turmas sobre a evolução temporal da competência em física escolar e, por esta razão, a variável turma deve ser usada como um possível preditor no modelo multinível de regressão múltipla.

4.3.5 - Professor

Quatro professores distintos lecionaram para as oito turmas da série investigada. Considerando o tempo de docência e a titulação, os quatro professores podem ser separados em duplas distintas. Seria razoável, então, esperar que tais diferenças se refletissem na evolução temporal da competência em física escolar de de seus alunos. A TAB. 4 mostra, para cada ocasião, as estatísticas da competência em física escolar dos estudantes agrupados segundo o professor.

TABELA 4
Competência em física escolar segundo o professor e a ocasião

Professor	Número de estudantes	Média			Desvio padrão		
		1 ^a	2 ^a	3 ^a	1 ^a	2 ^a	3 ^a
-	-						
1	50	-0,130	-0,483	0,156	3,524	3,312	3,232
2	51	1,852	0,281	0,420	3,123	3,365	2,751
3	50	1,203	1,641	0,895	3,253	3,177	3,410
4	51	-0,220	0,745	1,179	3,277	3,308	3,500
Total	202	0,677	0,546	0,663	3,390	3,356	3,236

Observa-se na TAB. 4 que, em cada ocasião, as médias da competência em física escolar parecem ser mais variadas do que os desvios padrão. Comparou-se a diferença entre tais médias usando o teste ANOVA. Verificou-se que há diferenças significativas entre as médias de competência em física escolar na 1^a ocasião ($F = 4,842$, $df_1 = 3$, $df_2 = 198$ e $p = 0,03$) e na 2^a ocasião ($F = 3,647$, $df_1 = 3$, $df_2 = 198$ e $p = 0,014$), porém não há diferença entre as médias da competência em física escolar (CFIS) na 3^a ocasião ($F = 1,023$, $df_1 = 3$, $df_2 = 198$ e $p = 0,384$). O teste de homogeneidade da variância de CFIS dos grupos formados pelos estudantes de cada professor indica que na 1^a e 2^a ocasiões a variância dos grupos é homogênea (1^a ocasião: $p = 0,517$; 2^a ocasião; $p = 0,998$). Testes Post-hoc (Bonferroni) mostraram que na primeira ocasião as médias de CFIS entre os alunos dos professores 1 e 2 e 2 e 4, foram significativamente diferentes entre si. Na segunda ocasião de medida, houve diferença significativa na média da competência em física escolar (CFIS) apenas entre os alunos dos professores 1 e 3.

A existência dessas diferenças reforça a expectativa de que as diferenças entre os professores se reflitam no desenvolvimento da competência em física escolar de seus alunos e sugere que a variável categórica Professor seja incluída como um possível preditor no modelamento da evolução temporal da competência em física escolar (CFIS) com o tempo.

4.3.6 - Escolaridade do pai

A partir dos dados disponibilizados pelas fontes 2 e 3 mencionadas na seção 3.3, identificou-se a escolaridade do pai de 160 dos 202 estudantes da amostra. A tabela 6 mostra a média e o desvio padrão da competência em física escolar, segundo a ocasião e a categoria de escolaridade do pai.

A partir da TAB. 5, percebe-se que na 1ª ocasião, a competência média dos estudantes cujo pai não tem ensino fundamental completo (categoria 1) é maior do que a média dos demais estudantes. Já nas ocasiões 2 e 3 os filhos de pai com nível superior (categoria 4) possuem média de competência em física escolar maior do que a dos demais estudantes.

TABELA 5
Competência em física escolar dos estudantes segundo a escolaridade do pai e a ocasião

-	Categoria de Escolaridade do Pai (Escpai)	1	2	3	4	Dados faltosos
	Número de Estudantes	15	35	64	46	42
1ª ocasião	Média	4,44	0,11	0,90	1,75	-0,71
	Desvio Padrão	3,320	18,353	3,424	2,888	42,528
2ª ocasião	Média	0,91	0,65	0,078	1,33	0,18
	Desvio Padrão	3,874	2,968	3,345	3,153	3,666
3ª ocasião	Média	0,83	0,44	0,22	1,37	0,70
	Desvio Padrão	2,961	3,185	3,365	3,196	3,270

Comparando a diferença entre as médias da competência em física escolar (CFIS) para estudantes agrupados segundo o nível de escolaridade do pai usando o teste ANOVA, observou-se que não há diferenças entre as médias da competência em física escolar (CFIS) na 1ª ocasião ($F = 1,837$, $df_1 = 3$, $df_2 = 159$ e $p = 0,143$), porém há diferenças significativas entre médias na 2ª ocasião ($F = 5,237$, $df_1 = 3$, $df_2 = 159$ e $p = 0,002$) e na 3ª ocasião ($F = 3,021$, $df_1 = 3$, $df_2 = 159$, $p = 0,031$).

O teste de homogeneidade da variância indicou que não há diferença na variância da competência em física escolar entre os grupos de estudantes definidos pelo nível de escolaridade do pai. Verificaram-se diferenças entre a competência em física

escolar dos estudantes agrupados segundo o nível de escolaridade do pai. O test Pos-Hoc para a 2ª ocasião indica que os filhos de pai, com ensino fundamental incompleto, têm médias significativamente superiores às médias dos filhos de pai com ensino fundamental completo ($\Delta\mu=2,636$). Os filhos de pai com ensino fundamental completo têm média inferior à dos filhos de pai com ensino superior completo ($\Delta\mu=-2,642$). Já na 3ª ocasião, os filhos de pai com ensino fundamental completo têm média inferior à dos filhos de pai com ensino superior completo ($\Delta\mu= -1,877$).

Estes resultados indicam que o nível de escolaridade do pai deve ser usado como um possível preditor da evolução temporal da competência em física escolar no modelo multinível de regressão múltipla.

4.3.7 - Escolaridade da mãe

A partir dos dados disponibilizados pelas fontes 2 e 3, descritas na seção 3.3, identificou-se a escolaridade da mãe de 162 dos 202 estudantes da amostra. A TAB. 7 mostra a média e o desvio padrão da competência em Física escolar, segundo a ocasião e a categoria de escolaridade da mãe.

A partir da TAB.6, constata-se que na 1ª ocasião, a competência média dos estudantes cuja mãe não tem ensino fundamental completo (categoria 1) é a maior do que a média dos demais estudantes. Já na 2ª ocasião a categoria de estudantes que tem a maior a média da competência em física escolar é a de filhos de mãe com ensino fundamental completo (categoria 3), e na 3ª ocasião é a de filhos de mãe com ensino médio completo (categoria 3).

TABELA 6
Competência em física escolar segundo a escolaridade da mãe em cada ocasião

-	Categoria de Escolaridade da Mãe (Escmae)	1	2	3	4	Dados Faltosos
		Número de Estudantes	11	23	74	54
1ª ocasião	Média	2,20	-0,72	0,72	1,16	-0,57
	Desvio Padrão	6,328	21,926	16,894	16,697	40,309
2ª ocasião	Média	-0,23	1,06	0,33	0,55	0,85
	Desvio Padrão	30,126	3,406	24,600	25,124	3,463
3ª ocasião	Média	-1,45	0,23	0,87	0,32	1,56
	Desvio Padrão	3,307	3,063	3,364	25,107	2,936

Comparando a diferença entre as médias da competência em física escolar (CFIS) para estudantes agrupados segundo o nível de escola da mãe e usando o procedimento ANOVA, observou-se que não há diferenças significativas entre as médias de CFIS nas três ocasiões (1ª ocasião: $F = 2,507$, $df_1 = 3$, $df_2 = 161$ e $p = 0,061$; 2ª ocasião: $F = 1,875$, $df_1 = 3$, $df_2 = 161$ e $p = 0,136$; 3ª ocasião: $F = 1,883$, $df_1 = 3$, $df_2 = 161$ e $p = 0,135$). Este resultado sugere ser desnecessário incluir a escolaridade da mãe como um possível preditor em um modelo multinível de regressão múltipla para a evolução da competência em física escolar.

4.4 - Modelo multinível de regressão múltipla de mudanças

Para investigar a evolução temporal da competência em física escolar e a sua relação com as habilidades cognitivas amplas do estrato II de estrutura das habilidades intelectuais humanas do modelo proposto por Carroll (1993) levando em conta os efeitos demográficos e contextuais, utilizou-se um modelo multinível de regressão múltipla em dois níveis. O primeiro nível descreve a evolução temporal da competência em física escolar (CFIS)_{ji} do estudante (i) em termos de uma soma de efeitos fixos e efeitos aleatórios. Os efeitos fixos descrevem a influência do tempo sobre a competência em física escolar do estudante e refletem as mudanças intra-individuais. Eles descrevem a evolução temporal da competência em física escolar do estudante i, nomeada, a partir daqui como competência em física escolar

verdadeira. Os efeitos aleatórios descrevem como, em cada ocasião de medida, a competência em física escolar do estudante i observada difere da evolução prevista para o estudante i pelos efeitos fixos. Singer e Willet (2003, p.54) argumentam que a parte aleatória intra-individual não reflete apenas erros de mensuração, e sim, todos os efeitos não observados.

A disponibilidade de apenas três ondas de dados implicou na adoção de um modelo linear para representar os efeitos fixos. Desta forma, no primeiro nível a evolução temporal média da competência em física escolar é caracterizada por dois parâmetros: um intercepto (π_{0i}) e uma inclinação (π_{1i}). O intercepto (π_{0i}) representa o valor verdadeiro da competência em física escolar do estudante i no início do ano letivo. A inclinação (π_{1i}) representa a taxa verdadeira de crescimento ou de decréscimo estimada para o estudante i , isto é, representa a rapidez de mudança da competência em física escolar ao longo do ano letivo, sendo que, a mudança pode significar tanto o aumento quanto a diminuição da competência em física escolar. Assim o modelo de primeiro nível pode ser escrito como

$$CFIS_{ji} = \pi_{0i} + \pi_{1i}t_j + \varepsilon_{ji}. \quad (\text{Eq. 1})$$

A variável ε_{ji} descreve os efeitos aleatórios, o desvio da competência em física escolar verdadeira do estudante i na ocasião j em relação a evolução temporal verdadeira de sua competência em física escolar.

O segundo nível que analisa as variações inter-individuais modela como o intercepto e a inclinação do sujeito i se divide em efeitos fixos e aleatórios. Também neste nível podemos investigar como as variáveis cognitivas demográficas e contextuais afetam a evolução temporal da competência em física escolar verdadeira de cada estudante. Cada uma dessas variáveis pode interferir tanto na inclinação quanto no intercepto relativo a cada estudante, e em cada uma delas afeta tanto a evolução verdadeira quanto os efeitos aleatórios. No primeiro caso, significa que elas podem interferir tanto na competência em física escolar inicial verdadeira quanto na taxa de mudança verdadeira do sujeito i . No segundo caso, quando a introdução de uma variável preditora no segundo nível melhora o ajuste do modelo a variância não explicada diminui, significando que as diferenças inter-individuais na competência

em física escolar verdadeira diminuem. Nesta pesquisa, as equações que descrevem o segundo nível têm a forma geral:

$$\begin{aligned}\pi_{0i} &= \gamma_{00} + \gamma_{01}V_1 + \gamma_{02}V_2 + \zeta_{0i} \\ \pi_{1i} &= \gamma_{10} + \gamma_{11}V_1 + \gamma_{12}V_2 + \zeta_{1i}\end{aligned}\quad (\text{Eq. 2})$$

Nesta equação γ_{00} representa o intercepto médio esperado para todos os estudantes, γ_{01} e γ_{02} representam respectivamente o efeito da variável preditora V_1 e V_2 sobre o intercepto, γ_{10} representa a inclinação média esperada para todos os estudantes, γ_{11} e γ_{12} representam respectivamente o efeito da variável preditora V_1 e V_2 sobre a inclinação. A variável aleatória ζ_{0i} representa o desvio do intercepto verdadeiro de cada estudante em relação à média prevista para todos os estudantes que tenham as características V_1 e V_2 , enquanto que a variável aleatória ζ_{1i} representa o desvio da inclinação verdadeira de cada estudante em relação à média prevista para todos os estudantes que também tenham as características V_1 e V_2 . Finalmente, as variáveis aleatórias ζ_{0i} e ζ_{1i} representam os efeitos inter-individuais de variáveis não observadas, das variáveis preditoras observadas, mas ainda não introduzidas no modelo e de erros de mensuração da competência em física escolar verdadeira.

Inicialmente, ajustou-se o modelo incondicional médio (SINGER & WILLET, 2003, p.92) que estimou e particionou a variância da competência em física escolar verdadeira entre as pessoas, mas não se levando em conta o tempo. Esse ajuste, assim como os outros que se seguem, foi feito usando-se o aplicativo MLwin (RASBASH et al., 2009). Em seguida, ajustamos o modelo incondicional de crescimento (SINGER & WILLET, 2003 p.92), que difere do anterior por levar em conta explicitamente o efeito do tempo, ou seja, esse é o primeiro modelo que descreve a evolução da competência em física escolar verdadeira. Verificou-se que o segundo modelo ajustava-se melhor aos dados do que o primeiro. De fato, a deviência diminuiu de 2952,607 para 2943,406 quando estimou-se 3 parâmetros adicionais. Como a variação da deviência em modelos aninhados segue uma distribuição de Chi-quadrado com o número de graus de liberdade igual ao número de parâmetros estimados, a variação da deviência acima corresponde a uma

probabilidade $p = 0,027$, que é uma evidência em favor da rejeição da hipótese nula de que os dois modelos se ajustam igualmente bem aos dados empíricos.

Os trabalhos prosseguiram do tipo bottom-up em que cada variável preditora é acrescentada uma por vez. Inicialmente, ao introduzir cada variável, era produzido um efeito no intercepto e também na inclinação. Para testar o ajuste do novo modelo, que é o modelo anterior acrescido da nova variável introduzida, verificava-se se a deviância do novo modelo era menor do que a anterior. Se este primeiro critério era atendido, prosseguia-se na análise do ajuste do novo modelo aos dados. Neste caso, fazia-se o teste do ajuste deste novo modelo em relação ao anterior, calculando-se a probabilidade de se obter um Chi-quadrado tão alto quanto a diferença de deviância dos dois modelos com um número de graus de liberdade igual à diferença de parâmetros estimados nos dois modelos. Se essa probabilidade fosse menor que 0,05, considerava-se que o modelo novo se ajustava sensivelmente melhor aos dados do que o modelo anterior. Em seguida, e se esse era o caso, verificava-se se os efeitos, ou seja, os coeficientes estimados para a variável introduzida eram significativamente diferentes de zero, usando-se o teste da estatística z (escore normalizado). O novo modelo só era considerado melhor que o anterior se, o ajuste global era melhor e os coeficientes eram significativamente distintos de zero. Neste caso retinha-se este novo modelo como o modelo atual, caso contrário, testava-se a variável separadamente no intercepto e na inclinação verificando se o ajuste do novo modelo era melhor que o modelo anterior usando-se os mesmos procedimentos descritos anteriormente.

A introdução de uma variável categórica, no modelo o MLwin, implica na introdução de uma variável dicotômica para cada categoria, exceto uma delas, que era tomada como a categoria de referência. Por exemplo, ao se introduzir a variável turma, uma variável categórica com oito categorias, tomando-se a categoria 1 (correspondente a turma 1) como referência, o MLwin introduzia automaticamente sete variáveis dicotômicas na equação do intercepto, e também na equação da inclinação. Neste caso, o coeficiente de cada uma dessas variáveis no intercepto é interpretado como o quanto o intercepto aumenta ou diminui, segundo o sinal do coeficiente, se o estudante pertencer a esta turma e não à turma de referência. Podia-se escolher qualquer categoria da variável como a categoria de referência. Ao mudar a categoria

de referência, o valor da deviência do novo modelo não se alterava, porém os coeficientes podiam manter seu estado de significância ou se modificar, apresentando ou não valores significativamente distintos de zero.

Ao se introduzir a variável turma no modelo e alterar a turma de referência, percebeu-se que apenas a turma 8 era significativamente distinta das demais, isto levou a se definir uma nova variável dicotômica descrevendo o pertencimento ou não do estudante à turma 8. Retirando-se a variável turma do modelo e introduzindo em seu lugar esta nova variável, obteve-se o mesmo valor de deviência com muito menos parâmetros estimados. O mesmo aconteceu quando da introdução das variáveis professor e escolaridade do pai. O resultado de ambas sugeria que apenas uma das categorias era responsável pelo melhor ajuste do modelo. Dessa forma foram criadas duas novas variáveis dicotômicas, a primeira descrevia se o estudante era ou não aluno do professor três. A segunda variável dicotômica descrevia se o estudante era filho ou não de pai com escolaridade de nível superior.

A TAB. 7 mostra as variáveis que introduzimos no modelo multinível, os coeficientes dessas variáveis em cada um dos modelos, os erros padrão desses coeficientes, os componentes da variância e a deviência de cada um dos modelos retidos, na ordem em que foram obtidos. Nessa tabela, se uma variável está multiplicada pelo tempo, significa que ela está modificando a inclinação, caso não esteja multiplicada pelo tempo, ela está modificando o valor do intercepto. Abaixo estão apresentadas todas as etapas para se chegar no melhor modelo hierárquico para os dados . Os valores dos parâmetros estão especificados na TAB. 7.

TABELA 7
Sequência de modelos para a evolução temporal da competência em física escolar

	Modelo A	Modelo B	Modelo C	Modelo D	Modelo E	Modelo F	Modelo G
Resposta	CFIS	CFIS	CFIS	CFIS	CFIS	CFIS	CFIS
Parte Fixa							
Y_{00}	0.629 (0,201)	0.647 (0, 312)	0.782 (0, 308)	0.514 (0, 326)	0.681 (0, 383)	-0.356 (0, 355)	-0.326 (0,35)
Y_{10}		-0.027 (0, 329)	-0.529 (0, 324)	-0.558 (0, 324)	-0.736 (0, 368)	0 (0,403)	0.006 (0, 403)
NGf						2.452 (0, 313)	2.245 (0, 324)
NGy							0.476 (0, 218)
Tempo.PT8			2.86 (0,55)	3.024 (0, 551)	2.984 (0,62)	2.603 (0, 647)	2.574 (0, 645)
AP3				1.116 (0, 464)	0.912 (0, 508)	1.209 (0,45)	1.142 (0, 442)
EPNS					1.097 (0, 478)	1.325 (0, 437)	1.275 (0, 429)
Tempo.NGf						-0.95 (0,38)	-0.948 (0,38)
Parte aleatória							
1º Nível Intraindividual							
σ_{ϵ}^2	4.27 (0,3)	3.465 (0, 345)	3.45 (0, 343)	3.45 (0, 343)	3.622 (0, 405)	3.202 (0, 424)	3.201 (0, 424)
2º Nível Interindividual							
σ_0^2	6.753 (0,82)	11.83 (2, 103)	11.308 (2, 051)	11.111 (2, 032)	10.535 (2, 281)	4.024 (1, 769)	3.749 (1, 738)
σ_{01}		-5.927 (2, 091)	-4.588 (1, 962)	-4.575 (1, 955)	-4.571 (2, 255)	-1.425 (2, 028)	-1.353 (2, 015)
σ_1^2		6.961 (2,642)	4.413 (2, 428)	4.363 (2, 424)	4.162 (2, 823)	2.655 (2,85)	2.667 (2, 852)
Nº de parâmetros estimados	3	6	7	8	9	11	12
Deviância	2952.607	2943.406	2918.436	2912.719	2306.356	1549.084	1544.429
Estatística pseudo R^2	11,023	22,256	19,171	18,924	18.319	9,881	9,617

Modelo A

O modelo A é o modelo incondicional de médias. Neste modelo foram estimados três parâmetros, a média, a variância de primeiro e de segundo níveis. A média, nomeada como competência em física escolar verdadeira é a mesma para todos os estudantes em todas as ocasiões. A média obtida foi $\mu = 0,629 (0,201)$, sendo significativamente diferente de zero ($z = 0,629/0,201 = 3,13$; $p < 0,001$). A variância de segundo nível, que descreve a variação inter-individual, obtida é $\sigma_0^2 = 6,753 (0,82)$, sendo significativamente distinta de zero ($z = 6,753/0,82 = 8,24$;

$p < 0,001$) e a variância de primeiro nível, que descreve a variação intra-individual $\sigma_{\varepsilon}^2 = 4,27(0,3)$, sendo, também, significativamente distinta de zero ($z = 4,27/0,3 = 14,23$; $p < 0,001$). A variância total de 11,023, sendo que 61,3% dela está entre os estudantes e apenas 38,7% está no nível intra-individual. Isto significa que os estudantes diferem mais uns dos outros do que diferem de si mesmo nas três ocasiões de medida. A deviança para este modelo foi de 2.952,61.

Modelo B:

O modelo B é o modelo incondicional de crescimento. Este modelo difere do anterior por introduzir uma relação linear entre a competência em física escolar e o tempo. Com a introdução da variável tempo, foram estimados seis parâmetros: o intercepto e a inclinação, que são efeitos fixos, e as quatro componentes da variância, que são efeitos aleatórios. A deviança caiu para 2.943,41. Assim houve uma variação de deviança de - 9,2 com três parâmetros adicionais. Usando-se a distribuição de Chi-quadrado com $\chi^2 = 9,2$; $df = 3$ obtém-se uma probabilidade $p = 0,03$. Desta forma, o modelo incondicional de crescimento ajusta-se sensivelmente melhor aos dados do que o modelo incondicional de médias.

O intercepto, que representa a competência em física escolar verdadeira no início do ano letivo, foi estimado em $\gamma_{00} = 0,647 (0,312)$, e é significativamente diferente de zero. A inclinação, que representa a taxa de mudança verdadeira, foi estimada em $\gamma_{10} = -0,027 (0,329)$ e não é significativamente diferente de zero. Apesar de aparentemente o modelo indicar uma taxa de mudança na competência em física escolar verdadeira que, em média, é nula, isto não significa que a competência em física escolar verdadeira de cada estudante não mude com o tempo. De fato, a componente da variância ligada à inclinação foi estimada como $\sigma_1^2 = 6,961(2,642)$, sendo significativamente diferente de zero ($p=0,031$). Isto significa que os estudantes têm inclinação que diferem muito entre si, sendo que alguns têm inclinação positiva e outros, negativa. O desvio padrão associado à inclinação é 2,64, e o intervalo de confiança de 95% para a inclinação é $[-5,20;5,14]$. A componente da variância associada ao intercepto também é significativamente diferente de zero e foi estimada $\sigma_0^2 = 11,83 (2,103)$. A competência em física escolar verdadeira no início do ano letivo é, em média, relativamente alta e ainda varia

bastante de um estudante para outro. De fato, o intervalo de confiança de 95% para o intercepto é $[-6,09;7,39]$. Apesar de o modelo estar mais bem ajustado aos dados e ser mais substantivo do ponto de vista educacional houve um preço a ser pago, a estatística pseudo R^2 (22,256) aumentou.

Os modelos nomeados a partir do modelo C foram aqueles obtidos pela introdução de variáveis preditoras. Começamos com o preditor gênero e escolaridade da mãe. Foi observado que para cada um dos preditores introduzidos a deviência aumentava e os coeficientes não ficavam significativos. Para as variáveis preditoras turma e escolaridade do pai verificou-se que a introdução de cada uma delas no modelo diminuía a deviência indicando um melhor ajuste global do modelo aos dados. No entanto, conforme explicado, nem todos os coeficientes relativos às categorias destas variáveis eram significativos mesmo mudando a categoria de referência. Procedeu-se como explicado anteriormente e criou-se as variáveis estudante pertencente à turma 8 (PT8), aluno do professor 3 (AP3) e filho de pai com escolaridade de nível superior (EPNS). Em seguida será discutido cada um dos modelos.

Modelo C:

Este modelo difere do modelo B por ter sido introduzido a variável estudante pertencente à turma 8 (PT8). Esta variável produz efeito significativo apenas na inclinação. De fato a variação de deviência foi de $-24,97$ com um parâmetro adicional estimado, o que indica que o modelo C ajusta-se sensivelmente melhor aos dados do que o modelo B ($p < 0,000$). Apesar de a variável estar afetando a inclinação, a estimativa da competência em física escolar verdadeira no início do ano letivo passou de $\gamma_{00} = 0,647$ (0,312) para $\gamma_{00} = 0,782$ (0,308), e este novo valor é significativamente diferente de zero ($p = 0,01$).

A taxa de variação da competência em física escolar verdadeira dos estudantes que não pertencem a turma 8 (a maioria dos estudantes) passou a ser $\gamma_{10} = -0,529$ (0,324), mas não é significativamente diferente de zero ($p = 0,051$). No entanto, a taxa de variação da competência em física escolar verdadeira para os estudantes que pertencem a turma 8 é a soma do componente γ_{10} mais o componente γ_{11} , que é

$\gamma_{11} = 2,86$ (0,55), sendo significativamente diferente de zero. Desta forma a inclinação dos estudantes da turma 8 é $\pi_{11T8} = 2,331$. A variância no nível intra-individual praticamente não mudou, no entanto a variância total no nível inter-individual diminuiu 16,3%, para 15,72.

Modelo D

Este modelo difere do modelo C por ter sido introduzido a variável aluno do professor 3 (AP3). Esta variável produz efeito significativo apenas no intercepto. De fato a variação de deviência foi de -5,717 com um parâmetro adicional estimado, o que indica que o modelo D ajusta-se sensivelmente melhor aos dados do que o modelo C ($p = 0,017$). O efeito dessa variável ocorre apenas sobre a estimativa da competência em física escolar verdadeira no início do ano letivo que para os estudantes que não são alunos do professor três é $\gamma_{00} = 0,514$ (0,326), que não é significativamente diferente de zero. No entanto, os estudantes que são alunos do professor 3 tem um acréscimo na competência em física escolar verdadeira no início do ano letivo de $\gamma_{01} = 1,116$ (0,464), que é significativamente diferente de zero ($p = 0,008$). Desta forma, a competência em física escolar verdadeira no início do ano letivo para estes estudantes é $\pi_{01P3} = 1,63$. A variância no nível intra-individual não se alterou, a variância total no nível inter-individual caiu apenas 1,6%, para 15,47 e a estatística pseudo R^2 diminuiu para 18,92.

Modelo E

Este modelo difere do modelo D por ter sido introduzido a variável 'filho de pai com escolaridade de nível superior' (EPNS). Esta variável produz efeito significativo apenas no intercepto. De fato a variação de deviência foi de - 606,363 com um parâmetro adicional estimado, o que indica que o modelo E ajusta-se sensivelmente melhor aos dados do que o modelo D ($p < 0,000$). O efeito dessa variável ocorre apenas sobre a estimativa da competência em física escolar verdadeira no início do ano letivo que, para os estudantes que não são filhos de pais de nível superior e não são alunos do professor 3 é $\gamma_{00} = 0,681$ (0,383), que não é significativamente diferente de zero.

No entanto, os estudantes que são filhos de pais de nível superior, independente de serem ou não alunos do professor 3, tem um acréscimo na competência em física escolar verdadeira no início do ano letivo de $\gamma_{01} = 1,097 (0,478)$, que é significativamente diferente de zero ($p = 0,011$). Desta forma, a competência em física escolar verdadeira no início do ano letivo para estes estudantes é $\pi_{01EPNS} = 1,778$. A variância no nível intra-individual aumentou 5% chegando a 3,622, este aumento pode ser entendido como uma redistribuição de variância do segundo nível para o primeiro nível, já que a variância total no nível inter-individual caiu 5%, para 14,697 e a estatística pseudo R^2 diminuiu para 18,319.

Modelo F

O modelo anterior foi o último modelo em que se introduziram as variáveis demográficas e contextuais. Os dois últimos modelos contêm os efeitos das variáveis cognitivas do estrato II do modelo de Carroll. Este modelo difere do modelo E por ter sido introduzida a variável habilidade inteligência fluida (NGf) tanto no intercepto quanto na inclinação, e seu efeito é significativo nestes dois parâmetros do modelo. De fato, a variação de deviência foi de - 757,272 com dois parâmetros adicionais estimado, o que indica que o modelo F ajusta-se sensivelmente melhor aos dados do que o modelo E ($p < 0,000$).

O efeito dessa variável ocorre sobre a estimativa da competência em física escolar verdadeira no início do ano letivo para todos os estudantes, que não são alunos do professor 3 e não são filhos de pais com escolaridade de nível superior, é $\gamma_{00} = -0,356 (0,355)$, que não é significativamente diferente de zero. No entanto, para todos os estudantes que possuem $NGf \geq 0,15$, possuem a competência em física escolar verdadeira positiva no início do ano letivo. De modo similar, todos os estudantes que não pertencem a turma 8, tem uma estimativa nula de taxa de mudança de sua competência em física escolar verdadeira, mas esta estimativa é alterada pela valor de NGf do estudante.

Os estudantes que possuem NGf negativo, tem uma inclinação positiva, ou seja, sua competência em física escolar verdadeira aumenta com o tempo. E todos os que possuem NGf positivo têm uma diminuição da competência em física escolar

verdadeira. O coeficiente $\gamma_{03} = 2,452 (0,313)$ e o coeficiente $\gamma_{13} = - 0,95 (0,38)$ são ambos significativamente diferentes de zero. Além deste modelo globalmente se ajustar melhor aos dados, há uma melhora na explicação da variância. A variância intra-individual decresce 11,6%. A variância total inter-individual diminui 54,6%, totalizando 6,679 e a estatística pseudo R^2 diminuiu para 9,881.

Modelo G

Este modelo difere do modelo F por ter sido introduzida a variável habilidade memória de curto prazo (NGy) no intercepto e seu efeito é significativo nesta variável. De fato a variação de deviência foi de -4,655 com um parâmetro adicional estimado, o que indica que o modelo G ajusta-se sensivelmente melhor aos dados do que o modelo F ($p = 0,031$). O efeito dessa variável apenas ocorre sobre a estimativa da competência em física escolar verdadeira no início do ano letivo que, para todos os estudantes que não são alunos do professor 3 e não são filhos de pais com escolaridade de nível superior e tem $NGf = 0$ (escore de inteligência fluida igual à média populacional), é $\gamma_{00} = -0,326 (0,35)$, que não é significativamente diferente de zero. No entanto, para todos os estudantes que possuem $NGy \geq 0,685$, a competência em física escolar verdadeira é positiva no início do ano letivo. De uma forma geral, todos os estudantes que possuem NGy positivo sofrem um acréscimo na sua competência em física escolar verdadeira inicial, enquanto que, os que têm NGy negativo sofrem um decréscimo. A variância intra-individual não muda e a variância total inter-individual diminui 3,9%, para 6,416 e a estatística pseudo R^2 diminui para 9,617.

As demais variáveis cognitivas foram testadas, mas nenhuma delas melhorou o ajuste deste modelo aos dados. Desta forma adotou-se modelo G como modelo final que descreve a relação entre a competência em física escolar e as habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll (1993), levando em conta o efeito de algumas variáveis demográficas e contextuais. Combinando-se as equações de primeiro nível (Eq 1) e as de segundo nível (Eq 2) obtém-se o chamado Modelo Composto formado pelas componentes estrutural e estocástica. O componente estrutural

descreve os efeitos fixos das variáveis preditoras da competência em física escolar. No nosso caso a componente estrutural é

$$CFIS_{ij} = [-0,326(0,35) + 2,245(0,324) * NGF + 0,476(0,218) * NGy + 1,275(0,429) * EPNS + 1,142(0,442) * AP3] + [0,006(0,403) + 2,574(0,645) * PT8 - 0,948(0,38) * NGF] \times Tempo_j. \quad (Eq.3)$$

Esta componente estrutural descreve a competência em física escolar verdadeira de todo estudante i em cada instante de tempo j . Nem todos os estudantes possuem a mesma evolução temporal da competência em física escolar verdadeira, visto que eles diferem entre si devido as variáveis contextuais e aos distintos níveis de habilidades de inteligência fluida e memória de curto prazo. No entanto, todos os estudantes que possuem as mesmas características, isto é, compartilham o mesmo contexto e têm níveis iguais de habilidades, possuem a mesma evolução temporal da competência em física escolar verdadeira.

A componente estocástica do modelo (r_{ij}) descreve como a competência em física escolar de cada estudante difere da trajetória de sua competência em física escolar verdadeira ao longo do tempo. A componente estocástica do modelo pode ser escrita como

$$r_{ij} = \varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}Tempo_j. \quad (Eq.4)$$

Nesta equação r_{ij} representa o resíduo do sujeito i na ocasião j . Ela mostra claramente que os resíduos, em princípio, são independentes e identicamente distribuídos, respeitando o princípio da homocedasticidade através das ocasiões e dos indivíduos. Isto implica em reconhecer que cada erro é desenhado independentemente para uma distribuição com média zero e com uma componente residual para a variância. No nosso modelo as três variáveis estocásticas ε_{ij} , ζ_{0i} e ζ_{1i} possuem as seguintes características distribucionais:

$$\varepsilon_{ij} \sim N(0, 3,201(0,424)) \quad (Eq.5)$$

$$\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3,749(1,738) & -1,353(2,015) \\ -1,353(2,015) & 2,667(2,852) \end{bmatrix} \right). \quad (Eq.6)$$

Observando os resultados acima, nota-se que a componente da variância relativa ao nível intra-individual (ϵ_{ij}) em larga medida não foi explicada pelos preditores que foram introduzidos. De fato, nenhum dos preditores utilizados assume valores distintos nas três ocasiões para cada sujeito i (covariado). Isto significa que provavelmente um aspecto importante do ensino permanece não acessado por esta pesquisa que, se por um lado é uma limitação da mesma, por outro, deixa aberta uma investigação futura.

Nas componentes da variância interpessoal, a situação foi diferente. A componente da variância ligada a variável aleatória ζ_{1i} , que modula a inclinação foi completamente explicada pelos preditores introduzidos, o mesmo acontecendo com a covariância entre o intercepto e a inclinação. A componente de variância ligada ao ζ_{0i} , que modula o intercepto, foi praticamente explicada pelos preditores. No entanto, ainda resta alguma diferença individual em relação à competência em física escolar no início do ano letivo que não foi explicada. Aqui também há uma janela de oportunidade para investigações futuras.

Do ponto de vista dos resultados, o mais importante é compreender a parte estrutural do modelo ajustado aos dados. A próxima seção analisa mais detalhadamente a relação entre a competência em física escolar verdadeira e as habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll (1993).

4.5 - Análise do Resultado

A partir da equação que descreve a parte estrutural, pode-se obter uma equação mais simples que descreve a relação entre a competência em física escolar verdadeira e as habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll (1993). Para tanto, fixamos as variáveis contextuais no seu nível de referência. Desta forma, a equação 7 apresenta esta relação controlada para todos os efeitos demográficos e contextuais relevantes.

$$CFIS = \{ [-0,326(0,35) + 2,245(0,324) * NGf + 0,476(0,218) * NGy] + [0,006(0,403) - 0,948(0,38) * NGf] * TEMPO \} \quad \text{Eq.7)}$$

A equação é simples, pois é de primeiro grau em cada uma das variáveis. Para cada valor de NG_f e de NG_y , há uma relação linear entre a competência em física escolar e o tempo. Alterando-se o valor de NG_f , altera-se, ao mesmo tempo, a competência em física escolar verdadeira inicial (intercepto) e sua taxa de mudança (inclinação). Mudando-se o valor de NG_y altera-se apenas o valor da competência em física escolar verdadeira inicial (intercepto). No entanto, a equação é de difícil apreciação. Desta forma, para se examinar melhor o efeito dos níveis de NG_f e NG_y sobre a evolução temporal da competência em física escolar verdadeira, optou-se por criar categorias de níveis de habilidades cognitivas.

Para cada habilidade, foram criadas quatro categorias de níveis baseadas no ranqueamento ordinal dos estudantes, na respectiva habilidade. Procurou-se criar categorias com mesmo número de estudantes, tomando-se o cuidado de evitar que estudantes com o mesmo nível de habilidade fossem alocados para categorias distintas. Para fazer esta classificação, utilizou-se a amostra BaFaCaLo composta pelos 116 estudantes que responderam aos testes da bateria BaFaCaLo. A amostra da Eq. 7 é uma subamostra da BaFaCaLo e composta por estudantes para os quais o modelo representado pela equação 7 se aplica. A TAB. 8 mostra as estatísticas descritivas para as quatro categorias da Inteligência fluida para as amostras BaFaCaLo e Eq. 7.

TABELA 8
Estatísticas descritivas segundo as categorias de inteligência fluida para as amostras BaFaCaLo e Eq. 7

Amostra	Categoria	N	Média	Desvio padrão	Erro padrão da média	Intervalo de Confiança de 95% para a média		Mínimo	Máximo
						Limite inferior	Limite superior		
Eq7	Muito Baixa	16	-1,316	0,373	0,093	-1,515	-1,118	-1,97	-0,7
	Baixa	11	-0,316	0,213	0,064	-0,460	-0,173	-0,63	0,04
	Alta	11	0,341	0,156	0,047	0,236	0,446	0,12	0,64
	Muito Alta	13	1,289	0,424	0,118	1,033	1,546	0,79	2,13
	Total	51	-0,079	1,061	0,149	-0,378	0,220	-1,97	2,13
BaFaCaLo	Muito Baixa	30	-1,317	0,383	0,070	-1,460	-1,174	-2,12	-0,7
	Baixa	26	-0,315	0,191	0,037	-0,392	-0,238	-0,63	0,04
	Alta	29	0,336	0,170	0,032	0,271	0,400	0,12	0,64
	Muito Alta	31	1,218	0,474	0,085	1,044	1,391	0,71	2,27
	Total	116	-0,002	1,007	0,093	-0,187	0,183	-2,12	2,27

O método de construção das categorias garante que elas não se superpõem nas duas amostras, como se pode observar nas colunas mínimo e máximo. Para cada uma das amostras, as colunas relativas ao intervalo de confiança de 95% para a média mostram que as categorias possuem médias significativamente distintas entre si. Examinando os valores das médias, percebe-se que cada categoria da amostra Eq. 7 não difere significativamente da categoria correspondente na amostra BaFaCaLo.

A TAB 9 mostra as estatísticas descritivas para as quatro categorias da memória de curto prazo para as amostras BaFaCaLo e Eq. 7.

TABELA 9
Estatísticas descritivas segundo as categorias de memória de curto prazo para as amostras BaFaCaLo e Eq. 7

Amostra	Categoria	N	Média	Desvio padrão	Erro padrão da média	Intervalo de Confiança de 95% para a média		Mínimo	Máximo
						Limite inferior	Limite superior		
Eq 7	Muito Baixa	14	-1,256	0,383	0,102	-1,477	-1,035	-1,86	-0,74
	Baixa	9	-0,471	0,132	0,044	-0,572	-0,370	-0,61	-0,36
	Alta	20	0,226	0,276	0,062	0,097	0,355	-0,24	0,64
	Muito Alta	8	1,386	0,477	0,169	0,987	1,785	0,76	2,01
	Total	51	-0,122	0,947	0,133	-0,388	0,144	-1,86	2,01
BaFaCaLo	Muito Baixa	29	-1,199	0,415	0,077	-1,356	-1,041	-2,24	-0,74
	Baixa	19	-0,506	0,104	0,024	-0,556	-0,456	-0,61	-0,36
	Alta	39	0,134	0,279	0,045	0,044	0,225	-0,24	0,64
	Muito Alta	29	1,313	0,466	0,086	1,136	1,490	0,76	2,14
	Total	116	-0,009	0,984	0,091	-0,190	0,172	-2,24	2,14

Examinando a TAB.9 percebe-se, em cada amostra, que as categorias não se superpõem como era de se esperar, e como se pode observar pelas colunas de mínimo e máximo. Também se pode observar pelas colunas relativas ao intervalo de confiança de 95% para a média que, em cada uma das amostras, as categorias possuem médias significativamente distintas entre si. Examinando os valores das médias, percebe-se que cada categoria da amostra Eq.7 não difere significativamente da categoria correspondente na amostra BaFaCaLo.

Os GRÁF. 10, 11, 12 e 13 apresentam a evolução temporal da competência em física escolar verdadeira ao longo do ano letivo segundo a categoria de inteligência fluida para cada categoria de memória de curto prazo (NGy). Cada gráfico corresponde a uma única categoria de memória de curto prazo. Em cada gráfico, cada uma das linhas corresponde a evolução temporal da competência em física escolar verdadeira (CFIS) para uma única categoria de inteligência fluida.

Observa-se, a partir dos gráficos, que independente da categoria de memória de curto prazo (NGy), para a mesma categoria de habilidade de inteligência fluida (NGf), as taxas de evolução temporal da competência em física escolar verdadeira (CFIS), ao longo do ano letivo, são as mesmas. Quanto maior a categoria de memória de curto prazo (NGy), maior é a competência em física escolar verdadeira inicial em cada categoria de inteligência fluida (NGf). Para cada categoria de memória de curto prazo (NGy), quanto maior a categoria de habilidade de inteligência fluida (NGf), maior é a competência em física escolar verdadeira inicial. E para as duas maiores categorias de habilidade de inteligência fluida, a taxa de evolução da competência em física escolar verdadeira (CFIS) tem um crescimento negativo e para as duas menores categorias, a taxa de evolução temporal da competência em física escolar verdadeira (CFIS) tem um crescimento positivo. Se prolongássemos as retas de regressão da competência em física escolar verdadeira para além do ano letivo, observa-se que elas tendem a se aproximar, e a distância entre retas das competências em física escolar das diversas categorias diminuiria, significando que as diferenças entre as competências em Física escolar dos estudantes ficariam menores. Assim teríamos, estudantes mais homogêneos.

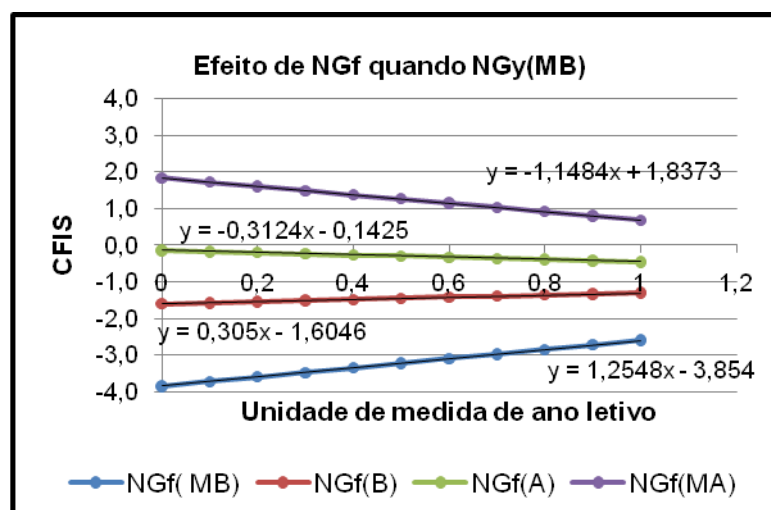


GRÁFICO 10 – Evolução temporal da competência em física escolar na categoria de NGy (MB) para categorias de NGf(MB), NGf(B), NGf(A) e NGf (MA)

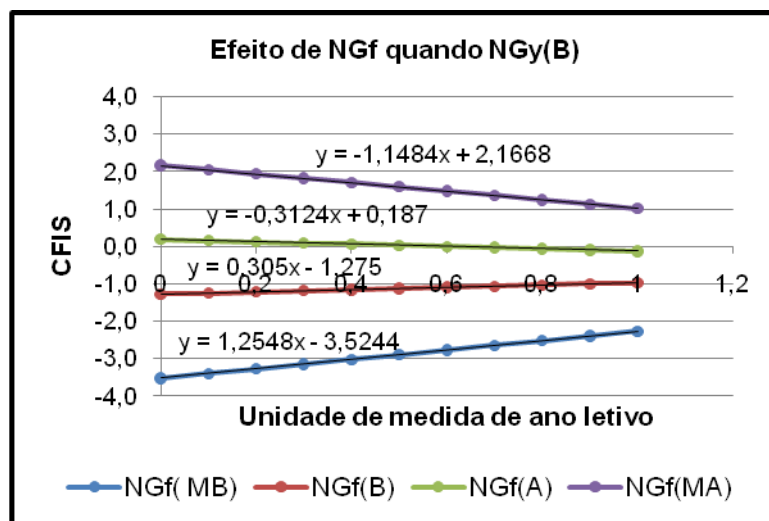


GRÁFICO 11 - Evolução temporal da competência em física escolar (CFIS) na categoria de NGy(B) para categorias de NGf(MB), NGf(B), NGf(A) e NGf(MA)

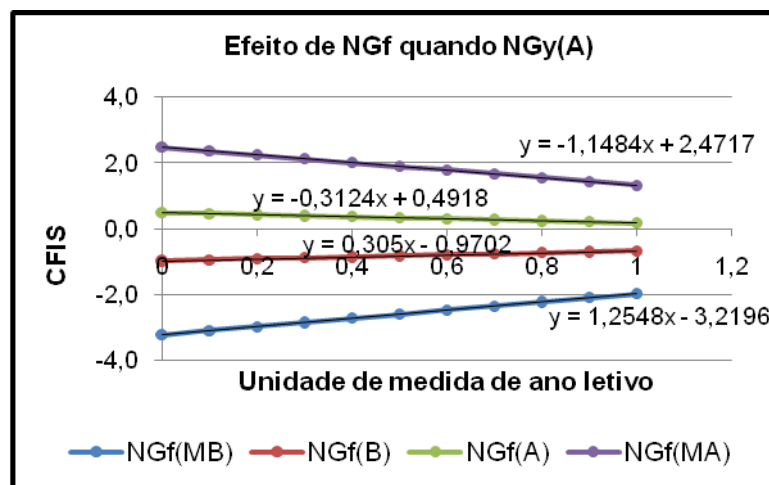


GRÁFICO 12 - Evolução temporal da competência em física escolar (CFIS) na categoria de NGy(A) para categorias de NGf(MB), NGf(B), NGf(A) e NGf(MA)

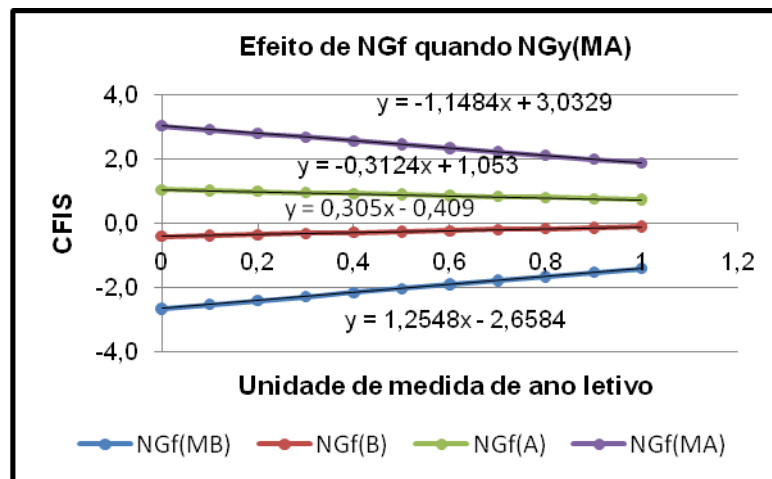


GRÁFICO 13 - Evolução temporal da competência em física escolar (CFIS) na categoria de NGy (MA) para categorias de NGf(MB), NGf(B), NGf(A) e NGf(MA)

QUADRO 20

Inclinação e intercepto segundo a categoria de habilidade de memória de curto prazo (NGy) e de inteligência fluida (NGf)

		NGy							
		Intercepto				Inclinação			
		MA	A	B	MB	MA	A	B	MB
NGf	MA	3,033	2,472	2,167	1,837	-1,148	-1,148	-1,148	-1,148
	A	1,053	0,492	0,187	-0,143	-0,312	-0,312	-0,312	-0,312
	B	-0,409	-0,970	-1,275	-1,605	0,305	0,305	0,305	0,305
	MB	-2,658	-3,220	-3,524	-3,854	1,255	1,255	1,255	1,255

A partir do QUADRO 20, podemos verificar que para uma mesma categoria de NGf, independente da categoria de NGy, o valor da inclinação é o mesmo e isto não acontece para o caso do intercepto. Para as categorias de muito alta (MA) e alta (A) Inteligência fluida, as inclinações são negativas. Para as categorias de Muito Baixa (MB) e Baixa (B) inteligência Fluida as inclinações são positivas e os interceptos negativos.

4.5 - Conclusão do capítulo

A partir da pesquisa podemos refletir sobre três aspectos apresentados e que são relevantes e importantes para a área educacional. O primeiro aspecto é que de todas as variáveis demográficas e contextuais utilizadas na pesquisa, as que apresentaram efeito sobre a evolução temporal da competência em física escolar foram relativas ao nível de escolarização do pai, ao professor e a turma do estudante. Este resultado reafirma o alerta de que ao se estudar aprendizagem escolar deve-se tomar o cuidado de se controlar os efeitos das variáveis que descrevem o contexto escolar e cultural do estudante.

O segundo aspecto é que, após controlar os efeitos demográficos e contextuais, foi possível encontrar uma relação significativa entre a evolução temporal da competência em física escolar e duas das habilidades cognitivas amplas, a inteligência fluida (Gf) e a memória de curto prazo (Gy). A inteligência fluida (Gf) afetou tanto a competência em física escolar inicial do estudante, quanto a sua taxa de evolução temporal.

Os estudantes, com habilidade de inteligência fluida nas categorias alta ou muito alta, possuem uma taxa de evolução temporal da competência em física escolar negativa, enquanto que os estudantes, com habilidade de inteligência fluida nas categorias de baixa ou muito baixa inteligência fluida, apresentam uma taxa de evolução temporal da competência em física escolar positiva. A habilidade memória de curto prazo (Gy) somente parece distinguir os estudantes quanto ao estado inicial de sua competência em física escolar, significando, somente, que esta habilidade diferencia os estudantes no início do ano letivo. Podemos especular que este resultado aparentemente contraditório, reflete como o ensino dessa disciplina está organizado na escola pesquisada: se por um lado, ele favorece os estudantes com habilidade de Inteligência fluida baixa ou muito baixa, por outro, o ensino de física, tal como estava organizado, parece ter propiciado aos estudantes com habilidade de inteligência fluida alta ou muito alta o desengajamento, ou descomprometimento com aprendizagem de física escolar, progressivo ao longo do ano letivo.

O terceiro aspecto, por fim, diz respeito à importância educacional do resultado encontrado. Pode-se perguntar se este resultado pode ser estendido para outras

turmas, para outras séries, e mesmo, para outras escolas. Como a evolução temporal da competência em física escolar foi controlada para os efeitos das variáveis demográficas e contextuais, deve-se esperar que o resultado seja válido para outros contextos escolares. Assim, em outras turmas e escolas, após retirar os efeitos demográficos e contextuais, deve-se, também, observar uma relação significativa entre a evolução temporal da competência em física escolar e as habilidades cognitivas, inteligência fluida e memória de curto prazo.

Este capítulo apresentou os resultados da investigação sobre a relação entre a aprendizagem em física escolar, representada pela competência em física escolar, e algumas habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll (1993), depois de controlado para os efeitos das variáveis demográficas e contextuais. No próximo capítulo, serão sintetizadas as principais conclusões e discutidas as limitações e as implicações educacionais deste trabalho.

5. CONCLUSÃO

Esta dissertação relata uma investigação sobre a existência da relação entre a evolução temporal da aprendizagem em física escolar, através da sua representação como competência em física escolar, e seis das habilidades cognitivas amplas do estrato II do modelo de estrutura das habilidades intelectuais humanas, proposto por Carroll (1993). Essa investigação se situa na área temática que investiga a relação entre a aprendizagem escolar e as habilidades cognitivas amplas. Duas questões foram investigadas: a dependência da evolução temporal da competência em física escolar em relação às variáveis demográficas e contextuais e, após controlar para os efeitos de tais dependências, estudou a relação entre a evolução temporal da competência em física escolar e algumas das habilidades cognitivas amplas do estrato II do modelo de Carroll (1993).

Para realizar esta investigação utilizaram-se os dados cognitivos dos estudantes já existentes devido a uma investigação anterior (GOMES, 2010) e dados das avaliações escolares ordinárias existentes nos registros acadêmicos da escola. A partir dos dados das avaliações escolares, de natureza ordinal, construiu-se uma escala intervalar para mensurar a competência em física escolar em três ocasiões distintas. Usando-se os resultados dos testes marcadores das habilidades cognitivas do estrato I do modelo de Carroll, foram construídas escalas intervalares (baseadas nos escores padronizados) para as seis habilidades amplas mensuradas pela bateria BaFaCaLo (GOMES, 2010). Foram construídos, também, indicadores dicotômicos ou categóricos para as variáveis demográficas e contextuais. Os dados foram modelados como um estudo longitudinal utilizando-se o modelamento multinível de regressão múltipla.

Neste capítulo, apresenta-se uma síntese dos resultados e das conclusões, discutem-se as limitações da pesquisa e as implicações educacionais que podemos vislumbrar a partir dos resultados.

5.1 - Resultados e conclusões

Nesta investigação introduzimos o conceito de competência em física escolar como uma representação daquilo que as avaliações escolares indicam. Competência em física escolar é concebida como uma amálgama de capacidades, habilidades, conhecimentos, valores, atitudes e crenças relativas ao domínio da física escolar de ensino médio, sendo inseparáveis os aspectos cognitivos, afetivos, motivacionais, volitivos, sociais, atitudinais e comportamentais. Competência em física escolar é, pois uma competência para a ação, especializada para o domínio de conteúdo da física escolar do ensino médio. Não é possível decompor a competência em física escolar nos seus diversos aspectos ou componentes para domínios mais especializados da física escolar. Metodologicamente, a competência em física escolar foi mensurada em uma escala intervalar, construída via modelamento Rasch das avaliações escolares ordinárias realizadas na disciplina física ao longo dos três trimestres do ano letivo.

A evolução temporal da competência em física escolar foi modelada utilizando-se o modelo multinível de regressão múltipla. O modelo intra-individual prevê que a competência em física escolar de cada estudante muda de forma sistemática e linearmente com o tempo, e muda de forma aleatória entre as ocasiões. A mudança sistemática é caracterizada por dois parâmetros que podem diferir de pessoa para pessoa: o intercepto, que corresponde ao valor da competência em física escolar no início do ano letivo, e a inclinação que corresponde à taxa de variação da competência em física escolar com o tempo. O modelo inter-individual modela como o intercepto e a inclinação, de cada estudante, dependem das variáveis demográficas, contextuais, e das habilidades cognitivas amplas.

As variáveis demográficas consideradas nesta pesquisa foram o gênero, a idade, e a renda familiar. A variável renda familiar não foi usada no modelamento devido à qualidade dos dados existentes. Nenhuma das outras duas variáveis afetou a evolução temporal da competência em física escolar, seja no seu valor no início do ano letivo, seja na taxa de mudança ao longo do ano letivo. As variáveis contextuais consideradas na investigação foram frequência do estudante, sua turma e seu professor de física, o nível de escolaridade do seu pai e de sua mãe. Pelo fato de

nem todos os professores das turmas investigadas fazerem um registro sistemático da frequência dos estudantes, a variável frequência (assiduidade) não foi incluída no modelo multinível.

A idade do estudante e o nível de escolarização de sua mãe não tiveram efeitos sobre a evolução temporal da competência em física escolar. Os alunos de um dos professores de física, o professor identificado como P₃, diferiram significativamente dos demais estudantes quanto ao seu nível inicial de competência em física escolar. Claramente, isso não é um efeito devido ao professor, mas um reflexo do processo de enturmação adotado pela escola e que afetou a turma deste professor. Da mesma forma, os estudantes cuja escolaridade do pai é de nível superior completo diferiram significativamente dos demais estudantes, quanto ao nível de sua competência em física escolar no início do ano letivo. Isto reflete diferenças individuais pré existentes que são, provavelmente, devido a história de aprendizagem desses estudantes.

Finalmente, os estudantes de uma das turmas, a turma T8, diferiram significativamente dos demais estudantes, quanto a taxa de evolução temporal da sua competência em física escolar. Este efeito, não reflete necessariamente um efeito do professor, já que os estudantes da outra turma do mesmo professor não diferiram significativamente dos demais estudantes. A origem desse efeito permanece inexplicada, porém podemos especular que, provavelmente, ele se deve ao ambiente da aprendizagem específico construído nesta turma pela interação entre professor e estudantes, entre os próprios estudantes, e ainda entre o estudante e os recursos didáticos. Apenas um estudo qualitativo, não realizado por falta de dados, poderia esclarecer a origem deste efeito.

Após identificar que variáveis demográficas e contextuais afetavam a competência em física escolar, foi introduzida, uma por vez, as seis variáveis seguintes: a Inteligência Fluida (Gf), a Inteligência Cristalizada (Gc), a Habilidade Visuo-Espacial (Gv), a Fluência (Gr), a Memória de Curto Prazo (Gy) e a Rapidez Cognitiva (Gs). Destas seis habilidades, apenas duas afetaram a evolução temporal da competência em física escolar.

A memória de curto prazo (Gy) afeta o nível de competência em física escolar no início do ano letivo sendo, portanto, um fator ligado a diferenças inter-individuais pré-existentes.

A inteligência fluida afeta tanto o nível inicial quanto a taxa de evolução temporal da competência em física escolar. Interpretamos que estes efeitos refletem diferenças inter-individuais pré-existentes ao início do ano letivo, e que provavelmente se deve à história de aprendizagem anterior, não apenas ao ambiente escolar e também não apenas em física.

De fato, como se adota nesta pesquisa uma perspectiva interacionista entre a aprendizagem escolar e habilidades cognitivas amplas, as habilidades cognitivas amplas pré-existentes influenciam o desenvolvimento da aprendizagem escolar e devem ser por estas influenciadas. Nem todas as habilidades testadas influenciaram a aprendizagem em física, mas isto deve ser interpretado como uma característica do currículo real implementado na instituição.

Por exemplo, a física ensinada na instituição aborda de forma esparsa temas que envolvem a visualização tridimensional ou situações, que exijam modelamento geométrico mais complexo. Um exemplo de tema que envolve visualização tridimensional é o estudo das forças magnéticas que, porém, não é abordado na primeira série da instituição. Desta forma, era de se esperar que a habilidade visuo-espacial não influenciasse na aprendizagem de física escolar.

Nas últimas três décadas, as teorias de ensino / aprendizagem de ciências e de física, em particular, têm construído um consenso de que o conhecimento prévio é o melhor preditor da aprendizagem futura. Como a inteligência cristalizada reflete as aprendizagens realizadas anteriormente, os conhecimentos adquiridos, relativos às linguagens específicas dos diversos domínios e conhecimentos declarativos conceituais, e as competências construídas, poder-se-ia esperar que ela influenciasse a aprendizagem em física escolar, o que não ocorreu.

No entanto, no ensino fundamental brasileiro muito pouco de linguagem específica de física e de conhecimento declarativo de física é ensinado pelos professores de ciências e aprendido pelos estudantes. Assim, a inteligência cristalizada reflete muito

pouco da aprendizagem anterior de física, e deve influenciar, pouco, a aprendizagem de física na primeira série do ensino médio.

A fluência, uma habilidade ligada à geração de idéias novas, influencia, pouco, a aprendizagem de física escolar, provavelmente porque a abordagem de ensino utilizava parcimoniosamente atividades práticas com características de investigação e modelamento, isto é, baseadas em métodos de descoberta guiada, tradicionalmente utilizadas no ensino de ciências em inglês, (ADEY, 2001). Assim, os estudantes podem ter tido poucas oportunidades para gerarem novas idéias em física, além de, ao serem avaliados, terem sido pouco exigidos quanto a esta capacidade.

A rapidez cognitiva também não influenciou aprendizagem em física escolar. Esta habilidade, segundo Flanagan & Harrison, (2005) “*é a habilidade de executar automática e fluentemente tarefas relativamente fáceis ou tarefas cognitivas mais complexas, porém já aprendidas, especialmente quando é demandada alta eficiência mental*”. Ela foi ligada, por Gomes (2010), à atenção e concentração focal. O ambiente de aprendizagem implementado para se ensinar física na primeira série do ensino médio da escola investigada destinava-se a ensinar física para todos, de forma mais conceitual e focando o pensar e o pensamento científico (BORGES, BORGES, VAZ, 2001, 2002). Assim, não exigia de forma sistemática e persistente nem atenção, nem concentração focal, e favorecia as habilidades de resolução de problemas, que são mais ligadas à inteligência fluída. Pode-se, então, entender porque a rapidez cognitiva não afetou a evolução temporal da competência em física escolar.

A influência da inteligência fluída sobre a evolução da aprendizagem em física não deixa de ser surpreendente. Ela é de tal forma que os estudantes com menor nível de inteligência fluída têm maior taxa de evolução da competência em física escolar, e os estudantes com maior nível de inteligência fluída, têm uma taxa de evolução temporal da competência em física escolar negativa. Isto provavelmente se deve ao fato de que os estudantes com maior nível de inteligência fluída, no início do ano letivo, se engajam de forma mais pronunciada nas atividades dessa disciplina pelo seu caráter de novidade e desafio. No entanto, à medida que o ano letivo se

desenvolvia, as atividades ficavam mais rotineiras, previsíveis e, conseqüentemente, menos desafiadoras. Levando, provavelmente, os estudantes com alta inteligência fluida, a se desengajarem. O currículo implementado parece funcionar muito bem para os estudantes com menor nível de inteligência fluida. E, neste sentido, ele é bastante inclusivo.

5.2 - Limitações da pesquisa

O trabalho apresenta algumas limitações. A primeira delas refere-se à própria natureza do estudo que pode ser mais bem caracterizada como um estudo observacional do que um estudo quase experimental. A pesquisa foi desenhada para se analisar dados previamente existentes. Isto introduz uma limitação séria sobre que variáveis orgânicas e contextuais poderiam ser envolvidas no estudo.

Os dados preexistentes permitiam um modelamento longitudinal da competência em física escolar, mas não permitiam, simultaneamente, o modelamento das habilidades cognitivas amplas. Se, de fato, como postula à perspectiva interacionista a aprendizagem escolar e as habilidades cognitivas amplas se codeterminam uma a outra, um estudo para detectar esta relação de codeterminação deveria tratar longitudinalmente tanto a competência em física escolar quanto as habilidades cognitivas amplas. Ainda que isto não tenha sido feito, o estudo é inédito na medida em que liga a evolução da aprendizagem em física escolar às habilidades cognitivas amplas. Contrastando com os estudos que ligam o desempenho final alcançado em física e as habilidades cognitivas amplas.

As variáveis demográficas e contextuais foram tratadas de forma dicotômica ou categóricas, formas que são compatíveis com o nível de informação existente nos dados disponíveis. Por esta razão, nesta pesquisa elas foram usadas apenas para eliminar os vieses da relação entre a evolução temporal da competência em física escolar e as habilidades cognitivas amplas.

Uma última limitação desta pesquisa se refere ao fato de que ela utiliza uma amostra de oportunidade composta por estudantes de uma mesma série e de uma mesma instituição de ensino. Na escola investigada, há dados similares aos aqui utilizados, relativos aos estudantes da segunda e da terceira série do ensino médio. No

entanto, devido ao tamanho dessas amostras e a limitação de tempo de realização desse estudo, optou-se por não estender a pesquisa às outras duas séries do ensino médio, da escola investigada.

Finalmente, sobre a possível transferência dos resultados dessa pesquisa para outras escolas e outras séries da mesma ou de outras escolas, na medida em que sejam controlados os efeitos das variáveis demográficas e contextuais, o resultado na sua forma mais qualitativa, isto é, quando afirma existir uma relação entre a evolução temporal da competência em física escolar e as habilidades cognitivas amplas, deve permanecer válido para todas as séries e todas as escolas.

Entretanto, há aspectos contextuais importantes não controlados nesta investigação, tais como, as formas de abordar o conteúdo, os tipos de atividades disponibilizadas aos estudantes, aspectos mais intangíveis do ambiente escolar, como a relação entre professores e alunos, a relação entre os próprios estudantes, e entre estes e os materiais de ensino. Também não foram controladas variáveis orgânicas importantes, tais como, a motivação, o interesse e o engajamento, dentre outros. Desta forma, desencoraja-se uma leitura mais desatenta e generalizante do resultado desta investigação.

5.3 - Implicações da pesquisa

O principal resultado desta investigação mostra que, pelo menos em um ambiente escolar específico, há uma relação sistemática entre a evolução da aprendizagem em física escolar e habilidades cognitivas amplas, como era de se esperar a partir da perspectiva interacionista aqui adotada. Com base nesta perspectiva, pode-se esperar que os estudantes desta instituição se beneficiaram, ainda que de forma não equitativa, em estudar física na primeira série, pois a aprendizagem de física deve ter potencializado suas habilidades cognitivas amplas.

É verdade que não há evidências empíricas que suportam esta última afirmação e que ela se fundamenta apenas em uma perspectiva teórica. Fica para investigações futuras, talvez melhor desenhadas, investigar se isso de fato acontece. No entanto, é isto que afirmam os resultados do programa CASE (ADEY, 1999) que mostram haver uma potencialização do desenvolvimento das habilidades cognitivas amplas

devido ao estudo de ciências, e que é transferido para o desempenho escolar posterior em ciências e em outras disciplinas.

No Brasil contemporâneo, estamos em um período em que a tendência curricular dominante parece ser a homogeneização das cargas horárias das diversas disciplinas refletindo o entendimento de que todas as disciplinas escolares são igualmente importantes. A política educacional que sustenta esta tendência curricular dominante não se fundamenta de forma sólida em evidências empíricas que demonstrem serem todas as disciplinas escolares igualmente valiosas para a formação dos jovens. O resultado dessa dissertação, por mais modesto e limitado que seja, serve para alertar que algumas disciplinas potencializam o desenvolvimento das habilidades cognitivas amplas. Neste sentido, ainda que de forma discreta e moderada, clama por uma política educacional que valorize no currículo escolar aquelas disciplinas que têm esse potencial. O resultado sugere a importância de mais estudos sobre a relação de codeterminação entre a aprendizagem escolar e as habilidades cognitivas amplas.

Nos livros examinados no último Programa Nacional do Livro Didático (PNLD) na área de física, há vários livros com muita inovação temática, e abordando um grande número de temas interdisciplinares (Menezes et al., 2010). Isto é um reflexo curricular das mudanças iniciadas no ensino médio na segunda metade da década de 90. Estes livros, aparentemente, tendem para uma abordagem mais qualitativa e conceitual, em decorrência das próprias temáticas abordadas. Não há evidências empíricas ou argumentos baseados em teorias solidamente estabelecidas mostrando que estas abordagens do ensino de física favoreçam o desenvolvimento das habilidades cognitivas amplas, que aparentemente pretendem desenvolver em consonância com o que se espera do ensino médio brasileiro, como se estabeleceu no parecer do CNE, que aprovou as diretrizes curriculares nacionais para o ensino médio (BRASIL, 1998) e nas duas versões do ENEM.

O resultado dessa pesquisa sugere que tal abordagem pode ter um efeito de desengajamento entre os estudantes com maior nível de inteligência fluida. Assim, é necessário destinar esforços de pesquisa para verificar se estas novas formas de se ensinar física também contribuirão para essa antiga meta de se desenvolver as

habilidades intelectuais humanas, desenvolver todo o potencial humano de cada estudante.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ABDI, H. The Bonferroni and Sidák Corrections for Multiple Comparisons. In: SALKIND, N. ed., *Encyclopedia of Measurement and Statistics*, Thousand Oaks, CA: Sage, 2007.

ADEY, P., CSAPÓ, B., DEMETRIOU, A., HAUTAMAKI, J., SHAYER, M., Can we be intelligent about intelligence? Why education needs the concept of plastic general ability. *Educational Research Review*, p. 75–97, 2007.

ADEY, P. The Science of Thinking, and Science for Thinking: A Description of Cognitive Acceleration through Science Education (CASE). Geneva: International Bureau of Education, 1999.

ADEY, P. 160 years of science education: an uncertain link between theory and practice. *School Science Review*, 82(300), p. 41-48, mar. 2001.

ADEY, P., & SHAYER, M. Really raising standards – cognitive intervention and academic achievement, London: Routledge. 1994.

ADEY, P. It all depends on the context, doesn't it? Searching for general, educable dragons. *Studies in Science Education*, v. 29, p. 45-92, 1997.

ADEY, P., CSAPÓ, B., DEMETRIOU, A., HAUTAMAKI, J., SHAYER, M., Can we be intelligent about intelligence? Why education needs the concept of plastic general ability, *Educational Research Review*, p. 75–97, 2007.

ADEY, P., Issues Arising from the Long-Term Evaluation of Cognitive Acceleration Programs, *Research in Science Education*, 35, p. 3–22, 2005.

ADEY, P., ROBERTSON, A., & VENVILLE, G. Effects of a cognitive stimulation programme on Year 1 pupils. *British Journal of Educational Psychology*, 72, p. 1–25, 2002.

ADEY, P., SHAYER, M. & YATES, C. *Thinking science: The curriculum materials of the CASE project*. 3 ed. London: Nelson Thornes. 2001.

AIKENS, H. A., THORNDIKE, E. L. & HUBBELL, E. Correlations among perceptive and associative processes. *Psychological Review*, 9, p. 374-382. 1902.

AMALDI, U. *Imagens da Física; As idéias e as experiências do pêndulo aos Quarks*. São Paulo: Scpione, 1995.

AMANTES, Amanda. *Contextualização no ensino de física: efeitos sobre a evolução do entendimento dos estudantes*. 2009. 272f. Faculdade de Educação, Universidade Federal de Minas Gerais UFMG, Belo Horizonte, 2009.

ANASTASI, A. *Psychological Testing*. 3 ed. New York: Macmillan, 1968.

ANASTASI, A. & URBINA, S. *Testagem Psicológica*. Poto Alegre: Artmed. 2000.

ARBITMAN-SMITH, R., HAYWOOD, C., & BRANSFORD, J. Assessing cognitive change. In P. H. Brooks, R. Sperber, & C. McCauley (Eds.), *Learning and cognition in the mentally retarded*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1994.

ARRUDA, L.. Desvendando desigualdades de oportunidades em ciências e em Matemática relacionadas ao gênero do aluno - uma aplicação de modelagem multinível ao SAEB 99. *Revista brasileira de investigação em educação em ciências*, v.2, n.3, p. 84-96, 2002.

ATKINSON, C. R., & GEISER, S. Reflections on a century of college admissions tests. *Educational Researcher*, v. 38, n. 9, :p. 665-676, dez. 2009

BAARTMAN, L. K.J., BRUIJN, E.. Integrating knowledge, skills and attitudes: Conceptualising learning processes towards vocational competence, *Educational Research Review*, Março, 2011.

BERGER FILHO, R., *Currículo e Competências*. Disponível em http://academico.direitorio.fgv.br/ccmw/images/b/b3/Curriculo_e_competencias_Berger_Filho.pdf. Acessado em 03/06/2011.

BINET, A., & SIMON, T. *The development of intelligence in children*. Baltimore: Williams and Wilkins, 1916 (Reprinted 1973, New York: Arno Press).

BORGES, A. T.; BORGES, O. N.; VAZ, A. Planejamento da solução de um problema. ENCONTRO NACIONAL DE PESQUISA EM ENSINO DE CIÊNCIAS, III, 2001. In Moreira, M.A. (Org.). *Anais* Porto Alegre : ABRAPEC, 2001.

BORGES, O.N., GOMES, C. M. A. ;O currículo de ciências pode ajudar a desenvolver a inteligência do alunos? IX Encontro Nacional de Pesquisa em Ensino de Física, 2004.

BORGES, O.N., BORGES, A. T.; VAZ, A. Quatro planejamentos da solução de um problema. ENCONTRO DE PESQUISA EM ENSINO DE FÍSICA, VIII, ÁGUAS DE LINDÓIA, VIII, 2002. IN: VIANNA, D. M.; PEDUZZI, L. O. Q.; BORGES, O. N.; NARDI, R. (Orgs.). *Atas do ...*SÃO PAULO: SBF, 2002. p. 1-12. (CD-ROM).

BORGES, O. N., VAZ, A., BORGES, A.T., COELHO, G. R.. As percepções dos estudantes sobre a organização em espiral do currículo de física, Organização do Currículo - SNEF - 2007

BORGES, O. N., BORGES, A. T.. INOVAR – CURRÍCULOS: DESENVOLVENDO O PENSAR E O PENSAMENTO CIENTÍFICOS. PROJETO INTEGRADO DE PESQUISAS, APRESENTADO AO CNPQ, JULHO DE 2001. (MIMEO)

BOSWORTH, D. & DAVIES, R. & WILSON, R. *Managerial qualification and organisational performance: an analysis of employers'skill survey 1999*. UK: Department for Education and Skills, 2002.

BRASIL. CNE/CEB. *Diretrizes curriculares para o ensino médio, 1 de junho de 1998*. Parecer elaborado por Guiomar Namó de Mello.1998.

BROWNE, W.J. *MCMC Estimation in MLwiN, v 2.10*. Centre for Multilevel Modelling, University of Bristol, 2009.

CAMPOS, R. H. F. Helena Antipoff: razão e sensibilidade na psicologia e na educação. *Estudos Avançados*, São Paulo, v. 17, n. 49, dez. 2003. Disponível em http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40142003000300013&lng=en&nrm=iso. access on 13 July 2011. doi: 10.1590/S0103-40142003000300013.

CARROLL, J.B. (1993). *Human cognitive abilities: A survey of factor-analytical studies*. New York : Cambridge University Press. 1993.

CAVANAGH, F. & WAUGH, R. F.. *Applications of Rasch Measurement in Learning Environments Research. Advances in Learning Environments Research Boston, Sense Publisheers. 2011, v.2, p.43.*

CECI, J.A. How Much Does Schooling Influence General Intelligence and Its Cognitive Components? A Reassessment of the Evidence. *Developmental Psychology*, v. 27, n.5, p. 703-722, 1991.

COELHO, Geide Rosa. A evolução do entendimento dos estudantes em eletricidade: um estudo longitudinal, 2011, 174f. Faculdade de Educação, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.

DESECO, Key competencies, A developing concept in general compulsory education, Eurydice The information network on education in Europe, p. 146, outubro, 2002, E-mail: info@eurydice.org Internet: <http://www.eurydice.org>, Eurydice Brussels: Eurydice 2002 ISBN 2-87116-346-4 Descriptors: Skill, Learning, Compulsory Education, Primary Education, Secondary Education, Lower Secondary, Evaluation, Pupil, Certification, Remedial Teaching, Learning Difficulty, European Union.

DUBOIS, P.H. A test-dominated society: china, 1115 B.C.-1905 A.D. In C. W. Harris (Ed).*Proceedings of the 1964 Invitational Conference on Testing Problems* (pp.3-11) Princeton, NJ: Educational Testing Service. 1965.

FEUERSTEIN, R., RAND,Y., HOFFMAN, M. & MILLER,M. *Instrumental enrichment: an intervention program for cognitive modifiability*. Baltimore: University Park Press, 1980.

FISCHER, G. H. and MOLENAAR, I. W., eds.. *Rasch Models: Foundations, recent developments, and applications*. New York Springer, , 1995.

FLANAGAN, D.P., GENSHAFT, J.L. & HARRISON, P.L.. *Contemporary intellectual assessment: Theories, tests, and issues*. New York: The Guilford Press.1997

FLANAGRAN, D. P., HARRISON, P. L.. Contemporary intellectual assessment: theories, tests, and issues. , 2ª edição, Guilford Press, 2005 -p. 667, 2005.

GOLDSTEIN, H. Multilevel Statistical Models. 2. ed. London: Edward Arnold, 1995.

GOLDSTEIN, H. Multilevel Statistical Models. Kendall's Library of Statistics 3 ed. London: Arnold. 2003.

GOMES, M. C. A. 2002. Aprendizagem auto-regulada em leitura numa perspectiva de jogo de regras. Dissertação de Mestrado, Universidade Estadual de Campinas, Campinas, SP.2002.

GOMES, C. M. A.. Uma análise dos fatores cognitivos mensurados pelo Exame Nacional do Ensino Médio (ENEM). Tese de doutorado. Belo Horizonte, MG: Universidade Federal de Minas Gerais - Faculdade de Educação. 2005.

GOMES, C. M. A., BORGES, O. Qualidade Psicométricas de um conjunto de 45 testes cognitivos. *Revista de Psicologia*, v.20, p. 17-32, 2009.

GOMES, C. M. A.. Estrutura fatorial da Bateria de Fatores Cognitivos de Alta-Ordem (BaFaCaLo). *Avaliação psicológica*, v.9, n.3, p. 449-459. dez. 2010.

GOMES, C. M. A., BORGES, O. N. Propriedades psicométricas do conjunto de testes da habilidade visuo espacial, *Psicologia-USF*, v. 14, n. 1, p. 19-34, jan./ abr. 2009.

GROUWS, D. A., Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics. (ed) New York, NY, England: Macmillan Publishing, 1992.

HAUSER, R. M.. Causes and Consequences of Cognitive. Functioning Across the Life Course, *AERA Distinguished Lecture, Educational Researcher*, v. 39, n. 2, p. 95–109, Março, 2010, Downloaded from <http://er.aera.net> at CAPES on May 30, 2011.

HEMKER, B. T., SIJTSMA, K., MOLENAAR, I. W. and JUNKER, B. W. Stochastic ordering using the latent trait and the sum score in polytomous IRT models. *Psychometrika*, v.62, n. 3, p.331-347, set. 1997.

HOLLAND, P. W. & RUBIN, D. B. *Test Equating*, New York: Academic Press, 1982.

HOX, J. J., *Multilevel analysis. Techniques and applications*. New York: Routledge, 2010.

INEP. *Competências e habilidades*. s.d. Texto do portal do INEP/ENEM. Disponível em http://historico.enem.inep.gov.br/index.php?option=com_content&task=view&id=25&Itemid=55. Acessado em 15/07/2011.

JOHNSON, V.E., An Alternative to Traditional GPA for Evaluating Student Performance. *Statistical Science*, v 12, n. 4, p.251-269, nov. 1997. Disponível em <http://www.jstor.org/stable/2246210>. Acessado em 20/06/2011.

JUNKER, B. W. , ERIC, T. B. . An Alternative to Traditional GPA for Evaluating Student Performance]: Source: Statistical Science, Institute of Mathematical Statistics Stable URL , V. 12, N 4 P. 274-277, Nov., 1997. Published by:: <http://www.jstor.org/stable/2246214> .

HEMKER, B. T., SIJTSMA, K., MOLENAAR, I. W. and JUNKER, B. W. (1997). Stochastic ordering using the latent trait and the sum score in polytomous IRT models. *Psychometrika*. To appear. 1997.

KILPATRICK, J. et al. *Historia de la investigación en Educación Matemática. Educación Matemática y investigación*. Madrid: Editorial Sonteses. 1992

KIM, M., YOUN, S., SHIN, J., PARK, M., SONG, K., SHIN, T., CHI, J. , SEO, D., HONG, S. A Review of Human Competence in Educational Research: Levels of K-12, College, Adult, and Business Education. *Asia Pacific Education Review*, v. 8, n. 2, p. 343-363. 2007.

KLAUER, K. J. Teaching for analogical transfer as a mean of improving problem solving, thinking and learning. *Instructional Science*, 18(3), p.179–192, 1989.

KLAUER, K. J. A process theory of inductive reasoning tested by the teaching of domain-specific thinking strategies. *European Journal of Psychology of Education*, 5(2), p. 191–206, 1990.

KLAUER, K. J. Training inductive reasoning: a developmental programme of higher-order cognitive skills. In: Hamers J. H. M. & Overtoom M. Th. (Eds.), *Teaching thinking in Europe*. Inventory of European programmes. Utrecht: Sardes, p.77–81, 1997.

KLAUER, K. J., & PHYE, G. *Cognitive training for children*. A developmental program of inductive reasoning and problem solving. Seattle: Hogrefe and Huber, 1994.

LARKEY, P. D. An Alternative to Traditional GPA for Evaluating Student Performance Comment: Adjusting Grades at Duke University, *Statistical Science*, v. 12, n. 4, p. 269-271, nov. 1997. Disponível em <http://www.jstor.org/stable/2246211> . Acessado em 20/06/2011.

LEE, V. E. , BRYK, A. S. . A Multilevel Model of the Social Distribution of High School Achievement. *Sociology of Education*, v. 62, p.172-192, julho 1989

LEVIN, J., FOX, J.A. *Estatística Para Ciências Humanas*. 9. ed. São Paulo: Prentice Hall, 2004.

LIN, C., HU, W., ADEY, P., SHEN, J. The Influence of CASE on Scientific Creativity . *Research in Science Education*. 33, p.143–162, 2003.

LINACRE, J. M. , WINSTEPS® versão 3.69.0 Rasch measurement computer program. Beaverton, Oregon: Winsteps.com., 2009.

LINN, Comments on Atkinson and Geiser. Considerations for College Admissions Testing *Educational Researcher*, v 38, n. 9, pp. 677–679, dezembro, 2009, Downloaded from <http://er.aera.net> at CAPES on May 30, 2011

LIPMAN, M., SHARP, M., & OSCANYAN, F. *Philosophy in the classroom*. Philadelphia: Temple University Press, 1980.

LIVINGSTON, S. A., KIM, S.. The Circle-Arc Method for Equating in Small Samples, *Journal of Educational Measurement* Fall, Vol. 46, No. 3, pp. 330–343, 2009.

MACGREW, K. S. . CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research, *Intelligence*, v. 37, p. 1-10, August, 2008

MAIA, Elizabeth Vieira. 2010. Desenvolvimento da Aprendizagem sobre Estruturas de Concreto Armado. 164f. Faculdade de Educação da Universidade Federal de Minas Gerais. Belo Horizonte. 2010.

MANUAL DO MLWIN, 2009 Jon Rasbash, Fiona Steele, William J. Browne and Harvey Goldstein.

MARAÑÓN, R. C., & ANDRÉS-PUEYO, A. The study of human intelligence: a review at the turn of the millennium. *Psychology in Spain*, v 4 (1), p.167-182. 2000

MARINHO, Rafael Araújo Álvares. O Uso de Avaliações Escolares Ordinárias para Estudar a Evolução da Competência em Física. 2010. 92f. Faculdade de Educação da Universidade Federal de Minas Gerais. Belo Horizonte, 2010.

MEAD, R. A Rasch Primer: The measurement theory of Georg Rasch. Psychometrics services research memorandum. Maple Grove, Data Recognition Corporation, 2008.

MENEZES et al., Coleção quanta Física. 1ª edição. São Paulo, p. 96, 2010.

NELSON, G. Science Literacy for All in the 21st Century. *Educational Leadership*, v. 57 n.2. out. 1999. Disponível em <http://www.project2061.org/publications/articles/articles/ascd.htm>. Acessado em 14/06/2011.

NICHOLLS, J. G. Achievement Motivation: Conceptions of Ability, Subjective Experience, Task Choice, and Performance. *Psychological Review*, *American Psychological Association, inc.* v.91, n.3, p. 328- 346. 1984.

OECD, 2003. PISA 2003 Assessment Framework - Mathematics, Reading, Science and Problem Solving Knowledge and Skills. Disponível em

http://www.oecd.org/document/29/0,3746,en_32252351_32236173_33694301_1_1_1_1,00.html.

PASQUALI, L. Técnicas de Exame Psicológico – TEP. Fundamentos das Técnicas de Exames Psicológico. São Paulo: Casa do Psicólogo. 2001

PISA 2009 Assessment Framework – Key Competencies in Reading Mathematics and Science. Programme for International Student Assessment, Paris: OECD, p.3-292
<http://www.oecd.org/dataoecd/11/40/44455820.pdf>

PLANINIC, M.; IVANJEK, L.; SUSAC, A. Rasch model based analysis of the force concept inventory, Physical Review Special Topics–Physics Education Research 6, 2010.

PRIMI, R. et al . Competências e habilidades cognitivas: diferentes definições dos mesmos construtos. *Psic.: Teor. e Pesq.*, Brasília, v. 17, n 2, ago. 2001 Disponível em http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-37722001000200007&lng=en&nrm=iso. Acessado em 25 de junho de 2011.

PRIMI, R., PASQUALI, L., Fundamentos da Teoria da Resposta ao Item –TRI, Avaliação Psicológica (2),p. 99-110, 2003.

RASBASH, J., CHARLTON, C. & PILLINGER, R. *Manual Supplement to MLwiN v2.10*. Centre for Multilevel Modelling, University of Bristol, 2009.

RASBASH, J., CHARLTON, C., BROWNE, W.J., HEALY, M. & CAMERON, B. *MLwiN Version 2.02*. Centre for Multilevel Modelling, University of Bristol, 2005.

RASBASH, J., CHARLTON, C., BROWNE, W.J., HEALY, M. & CAMERON, B. *MLwiN Version 2.1*. Centre for Multilevel Modelling, University of Bristol, 2009.

RASBASH, J., STEELE, F., BROWNE, W.J. & GOLDSTEIN, H. *A User's Guide to MLwiN, v2.10*. Centre for Multilevel Modelling, University of Bristol, 2009.

RYCHEN, D. S. & SALGANIK, L. DeSeCo SYMPOSIUM - DISCUSSION PAPER, OECD, JAN. 2002.

RYCHEN, D. S. & SALGANIK, L. H. (Eds.). *Defining and Selecting Key Competencies*. Göttingen, Germany: Hogrefe & Huber. 2001.

ROGOSA, D.; BRANDT, D. & ZIMOWSKI, M., A Growth Curve Approach to the Measurement of Change. *Psychological Bulletin*, v.92, n.3, p.726–748, 1982.

SALGANIK, L. H.. Contributions to the Second DeSeCo Symposium, Definition and Selection of Key Competencies, *Hights from current Assessments*. Geneva, Switzerland, fev, 2002.

SCHELINI, P. W., Teoria das inteligências fluida e cristalizada: início e evolução. *Estudos de Psicologia*, Natal, v.11, n. 3, set./dez. P. 323-332, 2006.

SHAYER, M., & BEASLEY, F. Does Instrumental Enrichment work? *British Educational Research Journal*, n 13, p.101–119, 1987.

SINGER, J. D., WILLETT, J. B., *Applied Longitudinal Data Analysis Modeling Change and Event Occurrence*. Oxford: University Press, 2003.

SOARES, M. Alfabetização e letramento: caminhos e descaminhos, *Revista Pátio*, n. 29, p.18-21, 2004.

SOARES, T. M.. Modelo de três níveis hierárquicos para a proficiência dos alunos de 4a série avaliados no teste de língua portuguesa do SIMAVE/ PROEB-2002. *Revista Brasileira de Educação*, n. 29 Maio /Jun /Jul /Ago 2005.

SPEARMAN, C. E. "General Intelligence," objectively determined and measured. *American Journal of Psychiatry*, 15, 201–293. (1904).

SPEARRITT, D. Carroll's model of cognitive abilities: Educational Implications. *International Journal of Educational Research*, 25 (2), p.107-198, 1996.

STATISTICAL PACKAGE FOR SOCIAL SCIENCES (SPSS), versão 18.0, www.spss.com

STERNBERG, R. J. *Psicologia Cognitiva*. Porto Alegre: Artmed, 2000.

STERNBERG, R.J. & KOLLIGIAN, J.Jr.. Competence considered. New Haven: Yale University Press. 1990.

STOEL, R. D.; ROELEVELD, J.; PEETSMA, T.; WITTENBOER, G.; HOX, J., Issues in longitudinal research on motivation. *Learning and Individual Differences*, v. 16, p.159-174, 2006.

TERMAN, L. M. & MERRILL, M. A., *Measuring intelligence*. Boston: Houghton Mifflin. 1937.

TRIOLA, M. F. *Introdução à Estatística*. 9 ed. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora, 2005.

VAN DER LINDEN, W. J. & HAMBLETON, R. K., eds. Handbook of modern item response theory. New York: Springer, 1997.

VILAS, S. *Formas de sociabilidade entre alunos de uma escola de ensino médio/técnico*. Dissertação (Mestrado em Educação) - Faculdade de Educação da UFMG, Belo Horizonte, 2009.

WEINERT, E. F. Definition and Selection of Competencies: Concepts of Competence. Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo), Max Planck Institute for Psychological Research, Munich, Germany, Abril, 1999

WENTZEL, K. R..Adolescent Classroom Goals, standards for performance, and academic achievement: An interactionist perspective. *Journal of Educational Psychology*, v. 2, p. 131-142, 1989.

WHITE, R.H.. Motivation reconsidered: The concept of competence. *Psychological Review*, 66, 297-333. 1959

WHITE, K. R., The Relation Between Socioeconomic Status and Academic Achievement *Psychological Bulletin*, American Psychological Association v. 91, n. 3, p. 461-481, , Inc., 1982.

WRIGHT, B. D.; LINACRE, J. M. Observations are always ordinal; measurements, however, must be interval. *Archives of Physical Medicine and Rehabilitation*, 70(12), p. 857-60, nov,1989.

WITT, R., & LEHMAN, R.. Germany. In Country Contribution Process, briefing materials prepared for DeSeCo's 2^o International Symposium. Neuchâtel, Switzerland: Swiss Federal Statistical Office. JULHO, 2001.

YAN, Z. & FISCHER, K.. Dynamic Analysis of Micro development in learning a Computer Program, Tese de Doutorado, Graduate School of Education of Harvard University, 2000.

YOUNG, J. W., KOBRIN, J. L.. Differential Validity, Differential Prediction, and College Admission Testing: A Comprehensive Review and Analysis, College Entrance Examination Board, Research Report, v.6 New York, 2001.

ANEXO A - Planejamento do curso de física – 1º ano – 2007

1. Medidas e erros

Lab1 - Tempo de Reação

Lab 2 - Relação entre Medidas

Lab 3 - Atividade das Estrelas

2. Movimento Uniforme

Lab 4 - Análise do movimento de uma bolha

3. O movimento variado

Lab 5 - Estudo do Movimento Variado

Lab 6 – Cinemática: queda livre

4. Energia e Trabalho

Lab 7 – Transformações de energia

Lab 8 - Período do pêndulo Simples

5. Calor, Energia Interna e Temperatura

6. Calor: 1º Princípio da Termodinâmica

Lab 9 – Análise de um experimento de vazão (1a parte)

Lab 10 - Análise de um experimento de vazão (2a parte)

7. Transmissão da Energia

7.1 - Ondas: Energia Sonora

7.2 - Ondas: Energia Eletromagnética

8. O que é eletricidade? DDP como trabalho

9. Corrente elétrica, Geradores, Circuito

Lab 11 - Circuitos Simples

10. Leis de Ohm

Lab 12 - Medidas Elétricas

11. Circuitos em série e paralelo, Efeito Joule

Lab 13 - Corrente e Tensão

12. Ondas Eletromagnéticas -Espectro Eletromagnético

Lab 14 – Propriedades das ondas eletromagnéticas (Laser)

13. Relatividade Restrita

Lab 15 - Propriedades ondas eletromagnéticas (Micro-ondas)

14. Modelo do Átomo de Bohr