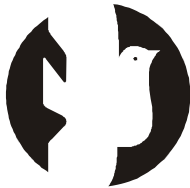


**MONALISA MUNIZ NASCIMENTO**



**UNIVERSIDADE  
SÃO FRANCISCO**

**CONSTRUÇÃO DE UM TESTE DINÂMICO  
INFORMATIZADO DE RACIOCÍNIO INDUTIVO PARA  
CRIANÇAS**

**Itatiba**

**2008**

MONALISA MUNIZ NASCIMENTO

CONSTRUÇÃO DE UM TESTE DINÂMICO INFORMATIZADO  
DE RACIOCÍNIO INDUTIVO PARA CRIANÇAS

Tese de Doutorado apresentada ao Programa  
de Pós-Graduação Stricto Sensu em Psicologia  
da Universidade São Francisco para obtenção  
do título de Doutor em Psicologia.

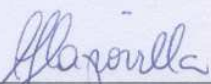
ORIENTADOR: PROF<sup>a</sup>. DR<sup>a</sup>. ALESSANDRA GOTUZO SEABRA CAPOVILLA

CO-ORIENTADOR: PROF. DR. RICARDO PRIMI

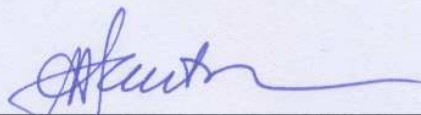
ITATIBA

2008

NASCIMENTO, Monalisa Muniz. **“Construção de um Teste Dinâmico e Informatizado do Raciocínio Indutivo para Crianças”**. Tese defendida e aprovada no Programa de Pós-Graduação *Stricto Sensu* em Psicologia da Universidade São Francisco em doze de dezembro de 2008 pela Banca examinadora constituída pelos professores:



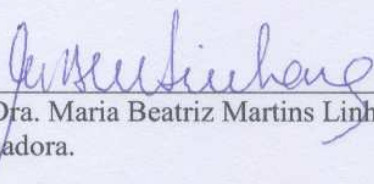
\_\_\_\_\_  
Profa. Dra. Alessandra Gotuzzo Seabra Capovilla.  
Orientadora e Presidente.



\_\_\_\_\_  
Profa. Dra. Acácia Aparecida Angeli dos Santos.  
Examinadora.



\_\_\_\_\_  
Prof. Dr. Carlos Henrique Sancineto da Silva Nunes.  
Examinador.



\_\_\_\_\_  
Profa. Dra. Maria Beatriz Martins Linhares.  
Examinadora.



\_\_\_\_\_  
Prof. Dr. Elizeu Coutinho de Macedo.  
Examinador.

## **Dedicatória**

Essa Tese é dedicada com muito carinho à Alessandra Gotuzo Seabra Capovilla, Ricardo Primi, minha mãe Ana, meu pai Edinaldo, meu irmão Charles e ao meu sobrinho que virá ao mundo em 2009, mas já é muito amado.

## **Agradecimentos**

Primeiramente agradeço à Deus pela vida e por ter me proporcionado uma família maravilhosa e colocado no meu caminho pessoas que somaram para o meu crescimento pessoal e profissional. Com relação ao doutorado, é justo que eu faça um agradecimento as pessoas que contribuíram nessa caminhada que foi trabalhosa, mas muito mais prazerosa. Com isso, agradeço:

Meus pais, Ana e Edinaldo, por serem exemplos positivos de pai e mãe e ao meu irmão Charles que sempre faz eu lembrar que o importante é ser feliz. Vocês com certeza são minhas fortalezas e não há nada no mundo, que seja melhor do que ter vocês como família. Meu eterno e incondicional agradecimento e amor por vocês.

Ao Ricardo Primi, mestre, co-orientador na tese e companheiro em todos os momentos. Fico sem palavras para expressar o quanto você é importante na minha vida, o quanto você é responsável pelo meu crescimento profissional e pessoal. A admiração que tenho por você é imensa, assim como a gratidão. Já disse isso na dissertação e repito, você é um ser humano insubstituível e agora acrescento, invejável, não somente pelo profissional que é, mas principalmente por ser um humano evoluído.

À minha queridíssima e estimada orientadora Alessandra Gotuzo Seabra Capovilla. Além de toda sua inteligência que é admirável, o que mais me encanta é o seu jeito de ser e de agir com as pessoas. Conviver e poder trabalhar junto com você é um privilégio.

Obrigada por ter aceito me orientar e por ter apoiado o projeto que hoje virou a tese. Agradeço ainda, toda sua paciência, compreensão e amizade. Valeu Prô!

Às minhas amigas Natália Martins Dias e a Bruna Tonietti Trevisan, vocês foram dois anjinhos nessa minha caminhada. Saibam que a ajuda de vocês durante todo o doutorado foi preciosa e que a amizade que construímos é uma peça valiosa na minha vida. Bom, nesse momento não poderia nunca esquecer do amigo Rodolfo Hipólito e a sua ajuda, paciência e bom humor na digitação do Geração. Com certeza seu lugar no céu já foi conquistado. Ah, sem esquecer do Gabriel Tortella, que quando precisei, me ajudou com toda boa vontade. Obrigaduooooooooooooooooo....

Nessa dupla ajuda e amizade, agradeço imensamente duas amigas que eu adoro. Tatiana Freitas da Cunha e Maiana Farias de Oliveira Nunes. Em relação à Mai, por termos a oportunidade de convivermos mais de perto, quero agradecer todo o seu apoio e companheirismo. Estendo esse agradecimento ao Tchê e já registro que sentirei muito a falta de vocês, mas torço para que sempre sejam felizes..

Agradeço também a amizade de Fabián Javier Marín Rueda, José Maria Montiel, Daniel Bartholomeu e Fabiano Koich Miguel. Independente do caminho que cada um de nós seguirmos, a amizade e as lembranças da convivência com vocês, desde o mestrado, ficarão para sempre no meu coração.

Aos amigos dos Laboratórios, LabAPE, LaPE e LaPSaM, que ainda não foram citados, em especial José Maurício Haas Bueno, Lucas Carvalho, Marjorie Silva, Arthur Berberian, Gleiber Couto, Maria Beatriz Zanarella Cruz, Rodolfo Augusto Ambiel, Priscila Rodrigues, Renata Franco, Maria de Fátima Xavier, Adriana Suehiro, Heitor Cozza, Marcelo Martinelli. Galera...dentro dos melhores momentos de mestrado e doutorado, com certeza estão os que passamos juntos, seja trabalhando, brincando, aproveitando as festinhas, etc...Porém o melhor de tudo é saber que vocês são amigos para todas as horas, sejam elas boas ou más. Também deixo registrada minha gratidão à Yeda Cirera Oswald, uma pessoa muito querida, que apostou na minha competência quanto professora.

Expresso também meus sinceros agradecimentos à minha banca de qualificação composta pelos doutores Carlos Henrique Sancineto da Silva Nunes, Elizeu Coutinho de Macedo, Fermino Fernandes Sisto e Maria Beatriz Martins Linhares. A escolha por cada um de vocês foi pensando nas competências que apresentam e na possibilidade de fazerem o trabalho enriquecer. Saibam que a contribuição que vocês propiciaram foi valiosíssima. Muito obrigada! Nesse ponto se faz necessário que eu agradeça com imensa gratidão à Professora Acácia Aparecida Angeli dos Santos pela sua generosidade em participar da banca de defesa, bem como pelas suas contribuições inigualáveis.

Também tenho um imenso prazer em agradecer à Prof<sup>ª</sup>. Dr<sup>ª</sup>. Adelinda de Araújo Candeias da Universidade de Évora em Portugal. O estágio desenvolvido sobre o tema dessa tese sob sua supervisão é de valor inestimável. O período que trabalhamos juntas

contribuiu enriquecedoramente tanto para a tese quanto para a minha vida profissional e acadêmica. Muito obrigada!

Essa caminhada percorrida foi inundada de conhecimentos, muitos vindos do corpo docente que compõe o Programa de Pós-Graduação em Psicologia da Universidade São Francisco. Meu carinhoso agradecimento aos Professores Doutores Acácia Aparecida Angeli dos Santos, Alessandra Gotuzzo Seabra Capovilla, Ana Paula Porto Noronha, Anna Elisa de Villemor Amaral, Carlos Henrique Sancineto da Silva Nunes, Claudette Maria Medeiros Vendramini, Cláudio Garcia Capitão, Fermino Fernandes Sisto, Makilim Nunes Baptista, Maria Cristina Rodrigues Azevedo Joly e Ricardo Primi. Meu agradecimento também fica registrado à Secretaria da Pós Graduação, em especial à Marcela Gressoni pela sua disposição e bom humor.

Esse trabalho em hipótese nenhuma teria se desenvolvido e finalizado sem a colaboração das diretoras e coordenadoras das escolas participantes. A permissão para a coleta dos dados nas escolas foi imprescindível. No entanto, os grandes responsáveis por esse trabalho ter se concretizado foram as crianças voluntárias. Não foi fácil durante 4 meses, todos os dias, manhã e tarde, fazer a coleta dos dados. Mas ao final de cada dia, me sentia rejuvenescida, pois o contato com as crianças, a espontaneidade delas e a empolgação para realizarem os testes, gratificaram o esforço.

Por fim, agradeço à Fundação de Amparo à Pesquisa do Estado de São Paulo-FAPESP pelo apoio financeiro a esse trabalho.



Deixo para vocês essa letra de música composta por uma pessoa muito querida,  
Rafael Gasparin Corrêa, conhecido como Rafinha.

### **FIM DE ANO**

Estou vendo filme  
Cada estrada que eu passei  
Cada história que eu vivi  
Os futuros que eu sonhei  
Cada dia, cada mês  
Cada estação que eu dividi com vocês  
Abri meu coração  
Se um dia alguém falou  
Que o que foi não volta mais  
A vida continua a escolha é sua  
Sou feliz de olhar pra trás  
Como não lembrar de vocês  
Como não chorar outra vez  
Só de lembrar  
Poder te abraçar  
Se a estrada se abrir  
E eu tiver que seguir  
Vou levar comigo cada amor de amigo  
Como não querer de novo  
Posso até ser bobo com vocês  
Viveria tudo outra vez  
Portas se abriram  
Tive dias de frio  
Histórias surgindo  
Corações se unindo  
Cada choro  
Cada riso  
Cada emoção que eu dividi com vocês  
Abri meu coração  
Se um dia alguém falou

Que o que foi não volta mais  
A vida continua a escolha é sua  
Sou feliz de olhar pra trás  
Como não lembrar de vocês  
Como não chorar outra vez  
Só de lembrar  
Poder te abraçar  
E os laços que eu plantei  
Que formaram raízes no meu chão  
Vou colher por toda vida, amigos do meu coração  
E se o mundo te esquecer  
Quando o dia amanhecer  
Quero estar aqui também  
Vamos juntos pro ano que vem

Esse trabalho consistiu na construção de um teste para verificar o potencial de aprendizagem de uma criança, porém a mensagem que eu quero deixar é a seguinte:

*Avaliar o potencial de aprendizagem ou mensurar uma habilidade já existente, é extremamente importante quando bem avaliado. No entanto será inútil, caso se encerre no apenas avaliar. Então, espero, que todos nós possamos ter a sensibilidade e a competência de sabermos avaliar, mas que além disso, sejamos indivíduos facilitadores no processo do desenvolvimento do potencial de qualquer pessoa. Mas para isso é preciso acreditarmos que todos nós, seja pouco ou muito, temos potencial para nos desenvolvermos. E esteja certo, que agindo dessa forma, você fará alguém feliz e será feliz.*

**Monalisa Muniz 06/11/2008**

## Resumo

Nascimento, M. M. (2008). *Construção de um teste dinâmico informatizado de raciocínio indutivo para crianças*. Tese de Doutorado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba, pp. 335.

O raciocínio indutivo refere-se ao resultado geral de observações individuais de regularidades que se percebe em uma determinada situação. Esse raciocínio pode ser avaliado por meio de testes dinâmicos, que procuram verificar o potencial do indivíduo para aprender. Um formato possível desse tipo de teste envolve três etapas, pré-teste, intervenção e pós-teste. Há poucos instrumentos dinâmicos para mensurar o potencial de aprendizagem do raciocínio indutivo, tanto no Brasil quanto no exterior. Considera-se importante, a utilização da testagem dinâmica do raciocínio indutivo, especialmente em crianças, visto que tal habilidade tem sido relacionada ao desempenho escolar e a outros aspectos cognitivos importantes, tais como as funções executivas. Nesse contexto, o objetivo do presente trabalho foi a construção e análise de algumas propriedades psicométricas, como precisão, validade, ajuste dos itens, entre outras, de um teste dinâmico informatizado do raciocínio indutivo para crianças de 7 a 11 anos de idade. O estudo consistiu de duas partes. A primeira envolveu a construção dos itens e da intervenção que compuseram o teste dinâmico e que foram fundamentados na Teoria Prescritiva do Raciocínio Indutivo proposta por Karl Josef Klauer, bem como o desenvolvimento do *software* do teste. Nessa primeira etapa construiu-se um total de 68 itens e, após análises teóricas e psicométricas sobre dificuldade dos itens, índices de ajustes (*infit* e *outfit*) e correlação *item-theta*, investigadas numa amostra de 417 crianças de 6 a 11 anos, selecionou-se 16 itens para compor o pré-teste e 16 para formar o pós-teste. Ainda nessa primeira etapa, por meio de análises de juízes, definiu-se a intervenção e construiu-se o *software* do teste dinâmico agrupando pré-teste, intervenção e pós-teste. Com o teste dinâmico informatizado prosseguiu-se à segunda parte do estudo, com a finalidade de investigar a precisão, a capacidade do teste em mensurar potencial de aprendizagem e verificar evidências de validade baseada na estrutura interna, por meio da análise fatorial; na relação com outras variáveis, utilizando a Prova de Raciocínio Abstrato-RA da BPR-5i que avalia raciocínio indutivo, o Teste de Geração Semântica que avalia controle inibitório, o Teste Stroop que mensura atenção seletiva e a nota escolar para verificar desempenho acadêmico. Por fim, averiguou-se a validade incremental do teste dinâmico para prever desempenho acadêmico. Para tanto participaram 320 crianças, sendo todas avaliadas, de forma individual, no teste dinâmico, e dentre elas, 303 foram avaliadas, de forma coletiva, no RA, e 172, de forma individual, no Geração Semântica e no Stroop. O desempenho acadêmico foi verificado para as 320 crianças. Os principais resultados apontaram baixa precisão do pré e pós-teste provavelmente pela baixa quantidade de itens, unidimensionalidade para o pré-teste e dois fatores ligados ao raciocínio indutivo para o pós-teste, bem como capacidade do teste dinâmico em mensurar potencial de aprendizagem para resolução de tarefas com raciocínio indutivo. Houve pouca evidência de validade

teste-critério com os testes RA e Stroop, mas grande relação com desempenho acadêmico, validade essa que foi reforçada pela validade incremental do teste dinâmico em prever desempenho acadêmico. Os resultados encontrados apontam situação favorável e promissora para o teste dinâmico informatizado do raciocínio indutivo, porém há a necessidade de mais pesquisas que possam esclarecer os dados encontrados.

**Palavras-chave:** avaliação dinâmica, teste psicológico informatizado, inteligência, inteligência fluida e avaliação de crianças

## Abstract

Nascimento, M. M. (2008). *Development of a Computerized Dynamic Inductive Reasoning Test for Children*. Doctoral Dissertation, Graduate Program of Psychology, University of São Francisco, Itatiba, pp. 335.

Inductive reasoning refers to a general result of individual observations of regularities perceived in a particular situation. This type of reasoning can be assessed throughout dynamic tests that intend to assess individual's learning potential. A possible format of such test entail three phases: pre-test, intervention and post-test. In Brazil, as well as in other countries, there are few dynamic instruments to measure learning potential. Specially with children, the use of dynamic testing is considered important since this ability is related to school achievement and other cognitive variables such as executive functions. The aim of the present study was to construction and to analyze the psychometric properties, such as reliability, validity, item fit and others, of a computerized dynamic inductive reasoning test for children's aged from 7 to 11 years. The study was composed of two parts. The first involved the items and intervention development based on Karl Josef Klauer's Prescriptive Theory of Inductive Reasoning as well as the development of a software for the test. In this first phase it was created 68 items which were answered by 417 children aged from 6 to 11 years. Based on psychometric item analysis (infit and outfit, difficulty index and item-theta correlation) 16 items were selected for each form of the pre and post test for the second phase. Also in this first phase the intervention was analyzed by specialists. Based on the result of this phase pre test, intervention and post test was assembled in a computer software that manages its administration. The second phase aimed to study the test reliability, it's capacity to measure learning potential and to verify it's validity by the analysis of the internal structure based on factor analysis, correlations with external variables such as Abstract Reasoning from BPR-5i (RA) that measures inductive reasoning, Semantic Generation Test that measures inhibitory control, Stroop Test that measures selective attention and school grades as a indicator of scholastic achievement. It was also analyzed test incremental validity in predicting school grades. In his phase 320 children answered individually to the test. From this group 303 also answered the RA and 172 answered Stroop and Semantic Reasoning Test. The main results show a low reliability of pre and post test probably due to low number of items, unidimensionality for the pre test and two-factor solution for the post test as well as its capacity to measure learning potential to solve problems of inductive reasoning. Low coefficients of correlation were found with external variables (RA and Stroop) but stronger correlations as well as incremental validity were found in predicting school grades. In general results show a positive and promising outlook for the test but also point to the need of more research to further clarify some patterns observed in the data.

**Key-words:** dynamic assessment, computerized testing, fluid intelligence, children psychological assessment

## SUMÁRIO

<b>APRESENTAÇÃO .....</b>	<b>1</b>
<b>CAPÍTULO 1 – INTRODUÇÃO .....</b>	<b>5</b>
1.1 – Avaliação dinâmica: conceitos, aplicabilidade, estudos e procedimentos.....	5
1.2– Inteligência e raciocínio indutivo.....	29
1.3 - Teoria Prescritiva e Treinamento do Raciocínio Indutivo.....	59
1.4- Raciocínio indutivo e variáveis critérios: desempenho acadêmico, funções executivas, especificamente atenção seletiva e controle inibitório. ....	89
1.5 - Construção de instrumento psicológico e seus requisitos psicométricos .....	99
<b>CAPÍTULO 2- METODOLOGIA, RESULTADOS E DISCUSSÃO DOS ESTUDOS 1 E 2. ....</b>	<b>112</b>
2.1 - Estudo 1. Construção de itens indutivos e elaboração da intervenção para a resolução desses itens .....	114
2.1.1 – Estudo 1 – Parte 1.....	114
2.1.1.1 – Estudo 1 – Parte 1 – Estudo Piloto.....	114
2.1.1.2 – Participantes .....	115
2.1.1.3- Materiais.....	115
2.1.1.3.1- Teste de Raciocínio Indutivo para Crianças .....	115
2.1.1.4 - Procedimento.....	122
2.1.1.5. Resultados e discussão do Estudo-piloto.....	123
2.1.2 – Estudo 1 – Parte 1 – Aplicação principal. Seleção da amostra de itens para a composição do pré e pós-teste do instrumento dinâmico .....	133
2.1.2.1 – Participantes .....	133
2.1.2.2 - Materiais.....	134
2.1.2.2.1- Teste de Raciocínio Indutivo para Crianças .....	134
2.1.2.4 - Procedimento.....	135
2.1.2.5 - Resultados e discussão do Estudo 1, Parte 1- Aplicação principal .....	136
2.1.2.5.1. Estatísticas descritivas e inferenciais no Teste de Raciocínio Indutivo em função de série.....	137

2.1.2.5.2. Análise fatorial do Teste de Raciocínio Indutivo.....	153
2.1.2.5.3. Análise da precisão do Teste de Raciocínio Indutivo.....	161
2.1.2.5.4. Análise do Teste de Raciocínio Indutivo pela Teoria de Resposta ao Item.....	162
2.1.3 – Estudo 1 – Parte 2. Elaboração e avaliação por juízes da intervenção para resolução de itens indutivos .....	183
2.1.3.1 – Participantes .....	183
2.1.3.2 – Materiais .....	183
2.1.3.2.1 – Desenvolvimento da intervenção para raciocínio indutivo.....	183
2.1.3.3 – Procedimento .....	188
2.1.3.4 - Resultados e discussão do Estudo 1, Parte 2.....	189
2.2 - Estudo 2- Análises psicométricas da estrutura interna e externa do instrumento dinâmico. ....	194
2.2.1 - Participantes .....	194
2.2.2 - Materiais.....	195
2.2.2.1 - Teste Dinâmico Informatizado de Raciocínio Indutivo .....	195
2.2.2.2- Raciocínio Abstrato da Bateria de Provas de Raciocínio-BPR5i (Primi, Almeida & Cruz, em construção).....	197
2.2.2.3 – Teste de Geração Semântica (Capovilla & Macedo, no prelo).....	199
2.2.2.4 – Teste de Stroop Computadorizado (Capovilla, Montiel, Macedo & Charin, no prelo).....	201
2.2.2.5 – Desempenho Acadêmico.....	204
2.2.3 – Procedimento .....	205
2.2.4 Resultados e Discussão do Estudo 2 .....	206
2.2.4.1 Análise do Teste Dinâmico de Raciocínio Indutivo pela Teoria de Resposta ao Item.....	208
2.2.4.2 - Análise das estatísticas descritivas e inferenciais no Teste Dinâmico de Raciocínio Indutivo em função de tipo de paradigma, momento da testagem e série.....	211



2.2.4.3 - Análise das respostas aos itens no pré e pós-teste dinâmico e verificação da eficácia da intervenção em função do tipo de paradigma explorando os erros cometidos. .....	233
2.2.4.4 - Análise da evidência de validade baseada na estrutura interna .....	253
2.2.4.5 - Análise das evidências de validade baseada nas relações com outras variáveis	258
<b>CAPÍTULO 3- CONSIDERAÇÕES FINAIS .....</b>	<b>304</b>
<b>4. REFERÊNCIAS .....</b>	<b>315</b>
<b>5- ANEXOS .....</b>	<b>331</b>

## LISTA DE FIGURAS

Figura 1- Exemplo de como os problemas são apresentados na avaliação inicial. ....	23
Figura 2- Exemplo de como o item é apresentado na fase de avaliação final com <i>feedback</i> . .....	24
Figura 3- Teoria CHC das Habilidades Cognitivas (McGrew, 2007) .....	37
Figura 4- Integração das teorias fatoriais no modelo CHC (McGrew, 2007) .....	44
Figura 5- Exemplo de item com o paradigma de generalização de atributos (Klauer, Willmes & Phye, 2001). .....	64
Figura 6- Exemplo de item com o paradigma Diferenciação de atributos (Klauer & colaboradores, 2001). .....	65
Figura 7- Exemplo de item com o paradigma Classificação cruzada (Klauer & colaboradores, 2001). .....	67
Figura 8- Exemplo de item com o paradigma Reconhecimento de relações (Klauer & colaboradores, 2001). .....	69
Figura 9- Exemplo de item com o paradigma Discriminação de relações (Klauer & colaboradores, 2001). .....	70
Figura 10- Exemplo de item com o paradigma Sistema de construção (Klauer & colaboradores, 2001). .....	71
Figura 11- Facetas do modelo do raciocínio indutivo (adaptado de Klauer & colaboradores, 2001). .....	73
Figura 12- Genealogia das tarefas em raciocínio indutivo (adaptado de Klauer & colaboradores, 2001). .....	77
Figura 13- Tarefa de generalização e discriminação de atributos. ....	118
Figura 14- Tarefa de Classificação Cruzada.....	119
Figura 15- Tarefa de Reconhecimento e Diferenciação de Relações.....	120
Figura 16- Tarefa de Construção de Sistemas.....	121
Figura 17- Exemplo de um item da Prova RA da BPR5i. ....	199
Figura 18- Layout da tela para a figura de “cadeira” do Teste de Geração Semântica. ....	201
Figura 19 - Layout da tela para a parte 1 do Teste de Stroop Computadorizado. ....	202

Figura 20- Layout de uma tela da parte 2 do Teste de Stroop Computadorizado. ....	203
Figura 21- Layout da tela para a parte 3 do Teste de Stroop Computadorizado. ....	204
Figura 22 – Diferença dos ganhos considerando tipo (paradigma) e série.....	231
Figura 23 – Indivíduos que obtiveram ganhos e perdas de habilidade.....	260
Figura 24- Diferentes possibilidades de quantidade de ganho dependente da linha de base (pré-teste).....	261
Figura 25- Normalização das variações de ganho entre sujeitos com linha de base baixa e alta. ....	262

## LISTA DE TABELAS

Tabela 1- Processos do pensamento indutivo com os formatos de problemas (adaptado de Klauer & colaboradores, 2001). .....	75
Tabela 2- Propriedades psicométricas do teste de raciocínio indutivo por meio da análise de Rasch. ....	125
Tabela 3- Propriedades psicométricas dos itens do teste de raciocínio indutivo por meio da análise de Rasch.....	127
Tabela 4- Propriedades psicométricas dos itens do teste de raciocínio indutivo, considerando os recodificados e os excluídos, por meio da análise de Rasch. ....	130
Tabela 5 – Média e desvio padrão das idades por série e tipo de escola.....	134
Tabela 6 – Estatísticas descritivas dos escores das Formas A e B. ....	138
Tabela 7- Estatísticas descritivas dos escores das Formas A e B separando por série. ....	139
Tabela 8- Resultados da ANOVA da diferença entre séries nos escores das Formas A e B do teste de raciocínio indutivo.....	141
Tabela 9- Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore total da Formas A. ....	142
Tabela 10- Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore total da Formas B. ....	143
Tabela 11- Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Forma A considerando o paradigma SD.....	144
Tabela 12- Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Formas B considerando o paradigma SD. ....	145
Tabela 13- Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Forma A considerando o paradigma CC. ....	146
Tabela 14- Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore total da Forma B considerando o paradigma CC. ....	147
Tabela 15- Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Forma A considerando o paradigma RE.....	148
Tabela 16- Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Forma B considerando o paradigma RE.....	149

Tabela 17– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Formas A considerando o paradigma SC. ....	150
Tabela 18– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Forma B considerando o paradigma SC. ....	151
Tabela 19– Análise de correlação entre os escores das Formas A e B com a variável série. ....	153
Tabela 20- Fatores extraídos da análise fatorial por meio do programa Microfact pedindo dois fatores para as Formas A e B. ....	155
Tabela 21- Fator extraído da análise fatorial por meio do programa Microfact pedindo 1 Fator para as Formas A e B. ....	159
Tabela 22– Índices de precisão para as Forma A e B em cada série avaliada.....	162
Tabela 23– Propriedades psicométrica por meio da análise Rasch para dos itens escolhidos para compor o instrumento dinâmico. ....	165
Tabela 24– Descrição dos itens do pré e pós-teste referentes ao paradigma SD.....	169
Tabela 25– Descrição dos itens do pré e pós-teste referentes ao paradigma CC. ....	171
Tabela 26– Descrição dos itens do pré e pós-teste referentes ao paradigma RE.....	173
Tabela 27– Descrição dos itens do pré e pós-teste referentes ao paradigma SC.....	175
Tabela 28- Alternativa correta de cada item.....	178
Tabela 29– Informações sobre as sessões das aplicações dos testes utilizados no Estudo 2 .....	206
Tabela 30– Índices psicométricos do pré e pós-teste por meio Teoria de Resposta ao Item com o modelo de Rasch.....	209
Tabela 31– Índices de dificuldade do teste e habilidade dos sujeitos. ....	212
Tabela 32- Índices de habilidade dos sujeitos separando-os por série. ....	214
Tabela 33– Diferenças de média entre as séries para as habilidades dos sujeitos geral e em cada paradigma .....	216
Tabela 34- Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o escore geral do Pré-teste.....	217
Tabela 35- Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o escore geral do Pós-teste. ....	218

Tabela 36- Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma SD no pré-teste.....	219
Tabela 37- Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma SD no pós-teste.....	220
Tabela 38- Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma CC no pré-teste.....	221
Tabela 39- Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma CC no pós-teste. ....	222
Tabela 40- Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma RE no pré-teste.....	223
Tabela 41- Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma RE no pós-teste.....	224
Tabela 42- Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma SC no pré-teste. ....	225
Tabela 43- Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma SC no pós-teste.....	226
Tabela 44- Resultados da diferença entre grupos emergidos da Análise de Medidas Repetidas. ....	228
Tabela 45- Resultados da Análise de Medidas Repetidas investigando o ganho de habilidade por meio da diferença significativa entre pré e pós teste considerando os paradigmas.....	230
Tabela 46- Alternativa correta e porcentagem de escolha as alternativas de cada item....	234
Tabela 47- Porcentagem de erros e índices de dificuldade dos itens do pré e pós teste....	238
Tabela 48- Porcentagem de erros dos itens do pré e pós teste – 1ª série.....	239
Tabela 49- Porcentagem de erros dos itens do pré e pós teste – 2ª série.....	240
Tabela 50- Porcentagem de erros dos itens do pré e pós teste – 3ª série.....	241
Tabela 51- Porcentagem de erros dos itens do pré e pós teste – 4ª série.....	242
Tabela 52- Porcentagem de erros dos itens do pré e pós teste – 5ª série.....	243
Tabela 53 – Fator extraído para o pré-teste por meio da análise tetracórica. ....	254
Tabela 54– Fatores extraídos para o pós-teste por meio da análise tetracórica.....	257

Tabela 55– Estatística descritiva da amostra para na Prova de Raciocínio Abstrato. ....	263
Tabela 56– Correlações obtidas entre os escores do teste dinâmico com as notas no RA.	264
Tabela 57– Estatística descritiva do desempenho da amostra geral e por série no teste Stroop. ....	267
Tabela 58– Diferenças de média entre as séries no teste Stroop verificadas pela ANOVA. ....	270
Tabela 59- Correlação entre os escores do Teste Dinâmico e o Teste de Stroop sem separação por série.....	271
Tabela 60– Correlação entre os escores dos paradigmas com o teste Stroop.....	273
Tabela 61– Correlações por série entre os escores do teste dinâmico, incluindo os paradigmas, com o teste Stroop.....	274
Tabela 62- Estatística descritiva do tempo de reação no Teste de Geração Semântica da amostra geral e separada por série.....	277
Tabela 63– Diferença de média no teste Geração Semântica verificada pela ANOVA....	279
Tabela 64- Análise correlacional entre o Teste de Geração Semântica e os escores no teste dinâmico. ....	280
Tabela 65– Análise correlacional entre o teste Geração Semântica e os escores dos paradigmas.....	281
Tabela 66– Estatística descritiva da amostra total e por série em relação as matérias.....	284
Tabela 67– Correlações obtidas entre o desempenho acadêmico e escores do teste dinâmico. ....	286
Tabela 68- Correlações obtidas entre o desempenho acadêmico e os paradigmas. ....	288
Tabela 69- Correlações entre desempenho acadêmico e os escores totais e ganho do teste dinâmico separando por série. ....	289
Tabela 70- Correlações obtidas entre o desempenho escolar e os paradigmas do raciocínio indutivo.....	292
Tabela 71- Estatísticas descritivas das amostras com maior e menor facilidade para aprender. ....	295
Tabela 72- Diferença de média entre os grupos com maior e menor facilidade de aprender. ....	296

Tabela 73– Correlações obtidas entre o teste dinâmico e a nota da professora por meio da percepção em sala de aula .....	297
Tabela 74- Correlações obtidas entre o teste dinâmico e a nota da professora por meio da percepção em sala de aula separada por série.....	298
Tabela 75– Correlações obtidas entre o desempenho acadêmico e escores do teste dinâmico para a 5ª série. ....	299
Tabela 76- Correlações obtidas entre o desempenho acadêmico e escores dos paradigmas do teste dinâmico para a 5ª série.....	300
Tabela 77– Análise de regressão hierárquica com o teste estático e o dinâmico para a predição do desempenho acadêmico. ....	302



## APRESENTAÇÃO

O presente trabalho visou atender à necessidade de criação de novos instrumentos psicológicos que não apenas objetivem avaliar o conhecimento já adquirido do indivíduo, mas que, principalmente, possam compreender o potencial de aprendizagem, a capacidade e as mudanças nas habilidades cognitivas que podem vir a ocorrer durante uma intervenção. Os testes convencionais, também denominados de estáticos e que são representados principalmente pelos testes psicométricos de inteligência, são muito úteis em diversas situações quando é preciso avaliar o conhecimento que o indivíduo possui sobre um determinado conteúdo ou quantificar a habilidade já existente. No entanto, eles são restritos quando o objetivo não é somente verificar o que um indivíduo já sabe, mas principalmente o que ele pode vir a saber caso seja submetido a intervenções, ou seja, verificar qual a sua capacidade para aprender.

Testes dinâmicos permitem fazer essa avaliação da capacidade do indivíduo, pois neste tipo de teste são fornecidos ao examinando *feedback* sobre o seu desempenho e instruções sobre qual o caminho para se chegar à resposta correta. Assim, responder ou não a uma determinada tarefa depende da capacidade do indivíduo de aprender com as orientações que lhe são oferecidas para a resolução da tarefa.

Esse tipo de teste tem sido considerado uma forma mais justa de avaliar a capacidade de aprendizagem do indivíduo e esse fator é ainda mais relevante em crianças, visto que elas estão em processo pleno de aquisição de conhecimentos e habilidades. Logo,

parece ser mais adequado e útil avaliar a sua capacidade de aprendizagem e, com isso, promover intervenções orientadas de acordo com a capacidade da criança de aprender (Grigorenko & Sternberg, 1998).

Neste contexto, o presente projeto teve a finalidade de construir um teste dinâmico informatizado para crianças, voltado para a avaliação do raciocínio indutivo, e buscou evidências de validade pela relação com outras variáveis. Para este tipo de raciocínio, a pessoa induz regras, por exemplo, entre dois termos e aplica esta regra a uma segunda situação. Segundo Munhoz (2004), o raciocínio indutivo é um mecanismo para a aprendizagem, mas falta definir claramente a sua relação com aprendizagem geral e com outras funções investigadas pela Psicologia, como por exemplo atenção seletiva e controle inibitório o que corrobora a relevância do presente estudo.

O raciocínio indutivo é um fator específico da inteligência fluida, que, por sua vez, em artigos da neuropsicologia, tem sido associada às funções executivas (Blair, 2006; Fernandez-Duque, Baird & Posner, 2000). As funções executivas referem-se, de forma geral, à capacidade do sujeito de engajar-se em comportamentos orientados a objetivos, ou seja, à realização de ações voluntárias, independentes, autônomas, auto-organizadas e orientadas para metas específicas (Ardila & Ostrosky-Solís, 1996; Gazzaniga, Ivry & Mangun, 2002). Diante disso o presente estudo também pretendeu buscar evidência de validade do teste dinâmico de raciocínio indutivo pela correlação com testes que avaliam

funções executivas, mais especificamente que envolvem o controle inibitório e a atenção seletiva.

Dessa forma, o presente projeto se propôs a aproximar essas duas áreas de pesquisa, psicométrica e neuropsicológica, desenvolvendo um instrumento para avaliação dinâmica do raciocínio indutivo, componente da inteligência fluida. Isto será feito focalizando o fator indutivo e buscando evidências de validade, por meio da análise da correlação entre os desempenhos nesse instrumento e em testes de funções executivas e desempenho escolar.

Considerando-se seu caráter dinâmico, o teste foi formado por um pré-teste, no qual se avalia o desempenho de base sem ajuda; uma fase de intervenção<sup>1</sup>, em que o indivíduo recebe informações sobre os passos que devem ser seguidos para a resolução das questões-problema; e por um pós-teste, no qual se verifica se o indivíduo atingiu ou não um nível mais alto de desempenho. Dessa forma pode ser avaliado o seu potencial de aprendizagem nessa situação.

É importante destacar, ainda, que o instrumento foi informatizado e, portanto, além de possibilitar o registro de mais informações, tal como o tempo de reação durante execução da tarefa, permitiu que se controlasse automaticamente todas as fases da avaliação. Nesse sentido, alguns autores como Primi (2003) relatam que uma das grandes vantagens da aplicação computadorizada é a ampliação dos modos possíveis de

---

<sup>1</sup> No instrumento construído na presente tese optou-se por denominar essa fase, entre o pré e pós testes, de intervenção, porém na literatura encontram-se outras designações, por exemplo mediação, treinamento e instrução e que foram mantidas

administração e coleta de dados, que geralmente são dificilmente introduzidos no formato tradicional de aplicação.

No presente estudo a introdução encontra-se organizada em cinco seções: (a) a avaliação dinâmica, destacando conceitos, aplicabilidade, estudos e procedimentos, (b) o raciocínio indutivo, (c) a teoria prescritiva do raciocínio indutivo de Josef Klauer, que fundamenta os itens e a intervenção do teste dinâmico, (d) relação entre variáveis critério e potencial de aprendizagem do raciocínio indutivo e e) propriedades psicométricas na construção de um teste. Em seguida constam dois estudos, um de construção dos itens e intervenção para o teste dinâmico e outro abarcando estudos psicométricos para esse teste. Cada estudo contemplou separadamente seus objetivos, metodologia, resultados e discussão. Na última parte são feitas as considerações finais e por fim, são apresentadas as referências utilizadas na tese.

## CAPÍTULO 1 – INTRODUÇÃO

### 1.1 – Avaliação dinâmica: conceitos, aplicabilidade, estudos e procedimentos

A avaliação psicológica é um processo que tem por objetivo coletar, por meio de diversas técnicas, aspectos relevantes de um indivíduo para que se possa compreender, de maneira mais global, os comportamentos dessa pessoa (Urbina, 2006; Tavares, 2004). Uma das técnicas utilizadas na avaliação psicológica para coletar informações sobre um indivíduo é o teste psicológico. O teste é uma amostra da conduta de uma pessoa, pelo qual se podem fazer inferências sobre a conduta e funcionamento cognitivo e emocional do indivíduo avaliado (Muniz, 2004).

De acordo com a resolução nº 002/2003 do Conselho Federal de Psicologia (CFP), os testes psicológicos são instrumentos de avaliação ou mensuração de características psicológicas, constituindo-se um método ou uma técnica de uso privativo do psicólogo.

“Os testes são procedimentos sistemáticos de observação e registro de amostras de comportamento e respostas de indivíduos com o objetivo de descrever e/ou mensurar características e processos psicológicos, compreendidos tradicionalmente nas áreas emoção\afeto, cognição\inteligência, motivação, personalidade, psicomotricidade, atenção, memória, percepção, dentre outras, nas suas mais diversas formas de expressão, segundo padrões definidos pela construção dos instrumentos” (CFP, 2003, p. 2).

Pode-se subdividir os testes psicológicos em quatro tipos, quais sejam, psicométrico, projetivo, neuropsicológico e dinâmico. Ressalta-se que tradicionalmente os testes psicológicos são classificados apenas como psicométrico ou projetivo e os outros dois tipos podem ser considerados ramificações dos psicométricos. Porém os testes dinâmicos e neuropsicológicos, apesar de poderem apresentar características semelhantes ao psicométrico, propiciam outros dados além de quantificar uma determinada habilidade. Os testes psicométricos, também conhecidos como objetivos pautam-se em padronizações e maior objetividade para mensurar os construtos psicológicos. Já os projetivos ou impressionistas têm a finalidade de caracterizar as pessoas, buscando uma descrição mais individualizada e, apesar de alguns testes projetivos apresentarem certo grau de objetividade, esse tipo de teste é marcado pela subjetividade dos avaliadores (Cronbach, 1996).

Os testes neuropsicológicos, especialmente da neuropsicologia cognitiva, são fundamentados na teoria cognitiva. Eles podem complementar substancialmente as avaliações feitas com testes psicométricos, pois nessa concepção não basta apenas saber o quanto um indivíduo está deficitário ao compará-lo a um grupo normativo composto por pessoas da sua idade, é preciso saber em que etapa e como os processos mentais ocorrem independentemente de ter ou não déficit. Os testes da neuropsicologia cognitiva têm tal objetivo. Por fim, testes dinâmicos, tal como o proposto no presente estudo, têm o objetivo de avaliar o potencial de aprendizagem do indivíduo e não apenas mensurar o quanto esta

pessoa já sabe. A fundamentação desse tipo de teste também é a teoria cognitiva (Linhares, Escolano & Enumo, 2006).

A abordagem cognitiva surgiu em 1967 por meio do livro *Cognitive Psychology* de Ulrich Neisser ([www.wikipedia.org](http://www.wikipedia.org)) retomando uma metodologia experimental para o estudo da inteligência, buscando explicações em situações de laboratório, centrando-se nas tarefas e nos processos internos usados, e criando situações para testagem de hipóteses prévias. Com essa metodologia pôde-se definir um conjunto de componentes cognitivos mais ou menos gerais a toda a realização intelectual (Almeida, 1988; Sternberg, 1992b).

De acordo com Sternberg (1992a), a Psicologia Cognitiva procura uma melhor compreensão sobre os componentes, estratégias e representações envolvidos nos processos mentais. Além disso, investiga as interações desses fatores, componentes, estratégias e representações, interações essas que propiciam as diferenças individuais mensuráveis das capacidades intelectuais humanas.

Almeida (1988) classificou os estudos da abordagem cognitiva em três grandes grupos: estudos centrados nos mecanismos básicos da cognição, estudos centrados nos componentes cognitivos da realização e estudos de aplicação da situação real. O primeiro grupo trabalha com técnicas relacionadas aos estudos do tempo de reação para averiguar o funcionamento dos aspectos básicos do processamento da informação, relacionando os resultados aos diferentes níveis de aptidão dos sujeitos. Aqui vale destacar que o tempo de reação, para ser considerado uma medida de inteligência, necessita ser verificado baseado

no acerto, pois resolver um problema de maneira rápida, porém errada não agrega positivamente para o desempenho. O segundo grupo verifica a realização de tarefas específicas, buscando elaborar um modelo de processamento da informação, decompondo tarefas em sub-tarefas ou componentes. Já o terceiro grupo utiliza os conhecimentos da Psicologia cognitiva em circunstâncias concretas do cotidiano do indivíduo. Em geral fazem comparações entre peritos e iniciantes em determinada tarefa, objetivando identificar as diferentes estratégias empregadas por esses indivíduos, com a finalidade de desenvolver treinos cognitivos.

Os defensores da teoria cognitiva, na concepção de Sternberg (1992b), alegam que a inteligência pode ser compreendida em termos de componentes que constituem o processamento de informação. Voltam-se, também, para o processo pelo qual ocorre o comportamento inteligente, enfatizando a velocidade do processamento e buscando sustentação empírica a partir de tarefas de laboratórios.

Como a teoria cognitiva prima em conhecer o processo de realização de uma determinada tarefa e não apenas o seu produto final, contribuiu para a fundamentação das avaliações dinâmicas, das quais se originaram os testes dinâmicos. No entanto, o instrumento dinâmico em questão é mais restrito à investigação do processamento que o indivíduo pode utilizar para resolver uma determinada tarefa, pois a intervenção desenvolvida para o teste foi padronizada, ou seja, todo sujeito foi ensinado a raciocinar da mesma maneira indutivamente. Essa questão será melhor abordada no capítulo 3.



Retomando, a avaliação dinâmica, assistida ou interativa inclui o conceito de aprendizagem mediada, que consiste em uma modalidade interativa, incluindo suporte instrucional (fornecimento de pistas para organizar a tarefa; fornecimento de pistas específicas; modelos; demonstrações; instrução passo a passo e outras) no processo de avaliação de uma habilidade geral ou específica. O objetivo dessa avaliação é verificar o quanto modificável pode ser o desempenho inicial de base a fim de atingir níveis mais altos de funcionamento após a fase de intervenção e, conseqüentemente, visualizar o potencial para a aprendizagem (Ferriolli, Linhares, Loureiro & Marurano, 2001; Linhares, 1995; Linhares, Maria, Escolano & Gera, 1998). Ressalta-se que no presente projeto optou-se por seguir a terminologia avaliação dinâmica adotada, entre outros autores por Grigorenko e Sternberg (2002), no entanto há autores que utilizam a designação de avaliação assistida, como exemplo Linhares, Escolano e Enumo (2006).

O principal objetivo da avaliação dinâmica é avaliar as mudanças constantes no processo de pensamento, utilizando-se de um método baseado em compreender como fazer para aprender coisas novas e o que fazer para remover obstáculos dessa aprendizagem. Refere-se então à avaliação do pensamento, percepção, aprendizagem e resolução de problemas por meio de um processo ativo de ensino com a finalidade de modificar o funcionamento cognitivo (Haywood & Tzuriel, 2002). Mas é importante esclarecer que a mudança cognitiva tende a ocorrer mediante intervenções mais direcionadas ao problema e com maior durabilidade, ou seja, a avaliação dinâmica sem esse enfoque tenderá somente a produzir mudança temporária no funcionamento cognitivo do examinando.

Reforçando sua base na teoria cognitiva, a avaliação dinâmica vem sendo estudada desde a década de 1970 por meio dos trabalhos de Feuerstein e seus colegas em Israel, embasados nas idéias desse autor (Linhares & colaboradores, 2006). Para Feuerstein o organismo humano é capaz de se modificar para se adaptar as demandas do ambiente. Essa teoria de Feuerstein versa sobre a modificabilidade cognitiva, originada das concepções teóricas da plasticidade cognitiva e da necessidade de compreender melhor o mau desempenho de crianças nos testes estáticos (Feuerstein e cols, 1980). A modificabilidade cognitiva é concebida como uma habilidade do indivíduo em auto-modificar o funcionamento cognitivo e adaptá-lo perante mudanças, sendo que essa mudança não é específica a uma determinada habilidade, e sim estrutural, que é como o organismo interage com as informações. Tal modificabilidade ocorreria a partir de intervenções, sendo que o propósito de avaliação seria descobrir o potencial do indivíduo para a mudança cognitiva e identificar suas dificuldades para então elaborar um tratamento com as intervenções com a finalidade de desenvolver a capacidade de aprender (Feuerstein, Rand, Hoffman & Miller, 1980).

Segundo Simões (1995) a avaliação dinâmica e o teste dinâmico, que é um instrumento para a avaliação dinâmica, propiciam investigar a probabilidade da modificação cognitiva do indivíduo. No entanto, é importante ressaltar que, mesmo detectando as dificuldades de um indivíduo e propiciando intervenções para desenvolver a capacidade de aprender, todo ser humano tem suas limitações e é equivocado achar que, por exemplo, todo indivíduo pode atingir nível alto de inteligência por meio de intervenções.

Além da teoria cognitiva, pode-se também considerar a modalidade sócio-cognitiva, dos seguidores das idéias de Vygotsky, como base teórica para a avaliação dinâmica. Tal modalidade estuda os processos cognitivos envolvidos no pensamento, a eficiência das resoluções de problemas, a manutenção e transferência de aprendizagens, a sensibilidade à mediação fornecida pelo examinador e os indicadores dos potenciais para aprendizagem (Haywood & Tzuriel, 2002; Linhares, 1995; Tzuriel, 2001). Um conceito bastante importante dessa teoria é a zona de desenvolvimento proximal (ZDP), que se refere à distância entre o nível de desenvolvimento real e o potencial, sendo que este último ocorre por meio de orientações advindas de uma pessoa mais capacitada do que o indivíduo que está aprendendo (Vygotsky, 1970/1991). Um dos seguidores de Vygotsky, Leontiev, resumiu o significado da ZPD, apontando que esse termo destaca não o que o indivíduo sabe, mas sim o que ele pode vir a saber (Bronnfenbrenner, 1977, citado por Grigorenko & Sternberg, 1998). Essa maior compreensão da capacidade do indivíduo para aprender é um ponto central nas avaliações dinâmicas.

De acordo com Linhares e colaboradores (2006), a avaliação dinâmica é um conceito amplo e se utiliza de diferentes técnicas e procedimentos para ser desenvolvida. Por exemplo, nos Estados Unidos e na Europa, principalmente em pesquisas, é comum um método mais estruturado, padronizado, tanto na aplicação quanto na pontuação dos resultados. Tende-se a utilizar um procedimento “teste-assistência-reteste”, o qual permite avaliar os ganhos do sujeito após a ajuda, bem como verificar a manutenção da aprendizagem quando a tarefa é rerepresentada sem ajuda.

Esse tipo de avaliação dinâmica tem sido bastante utilizada com crianças que tendem a apresentar uma dificuldade cognitiva, possibilitando não apenas diagnosticar tais crianças, mas ampliar a compreensão de seus potenciais e suas dificuldades, contribuindo para o planejamento de intervenções mais adequadas. Pela sua eficiência é uma importante ferramenta para a pesquisa e para o trabalho clínico nas áreas de Psicopatologia, Neuropsicologia, Educação, Educação Especial, estudos de diferenças culturais e desenvolvimento humano (Haywood & Tzuriel, 2002; Linhares & colaboradores, 2006; Moore-Brown, Huerta, Uranga-Hernandez & Peña, 2006).

Apesar de ter um caráter mais clínico quando usado em crianças que apresentam algum tipo de deficiência cognitiva, os contextos de aplicação são diversos, como por exemplo, clínica, escolar e hospitalar (Linhares & colaboradores, 2006). A avaliação dinâmica pode ser usada com pré-escolares, universitários, pacientes clínicos, indivíduos em penitenciária, e também para a avaliação de programas cognitivos, como o *Bright Start* (Haywood, Brooks & Burns, 1992). Este último é um programa de treinamento cognitivo aplicado em diversos países (Linhares & colaboradores, 2006; Haywood & Tzuriel, 2002). Esse tipo de treinamento será mais bem desenvolvido na seção 3, que trata sobre a Teoria do Treinamento do Raciocínio Indutivo de Klauer, que fundamentou o teste desenvolvido no presente estudo.

Haywood e Tzuriel (2002) observam que, mesmo diante de poucas pesquisas empíricas e do pequeno corpo teórico sobre avaliação dinâmica, podem ser apontar

algumas conclusões em relação a essa prática. Uma delas se refere à melhora do desempenho no teste com a mediação, mas tal melhora dependerá do que será ensinado e como será ensinado, considerando as limitações da pessoa e a eficiência da intervenção. O mesmo pode-se dizer em relação à generalização da transferência de aprendizagem para diversos contextos, que dependerá dos caminhos traçados na intervenção, pois a mudança cognitiva não ocorre ao acaso, é preciso proporcionar recursos para isso. Outras evidências apontam que o potencial de aprendizagem pode ser verificado em aprendizagens subseqüentes e em situações de desempenho na escola e no cotidiano, e que as observações em procedimentos dinâmicos apresentam mais informações sobre a capacidade do indivíduo, sendo que o reconhecimento do potencial tem sido útil em estudos de diagnósticos e reabilitação cognitiva em pessoas que sofreram lesão cerebral, pois auxilia o planejamento de intervenções. Por fim, tais autores relatam que os testes estáticos fornecem para as informações sobre um indivíduo, porém é incorreto interpretar que um escore obtido em um teste seria suficiente para saber o potencial do indivíduo.

No Brasil, a pesquisa com avaliação dinâmica vem sendo desenvolvida de forma mais intensa a partir do início da década de 1990. A esse respeito, podem-se citar os grupos de trabalho das universidades, USP/Ribeirão Preto, coordenado por Maria Beatriz Martins Linhares, Centro de Estudos e Pesquisa em Reabilitação da Universidade Estadual de Campinas (Cepre/ UNICAMP), sob a coordenação de Cecília Guarnieri Baptista, e o grupo sob a coordenação de Sônia Regina Fiorim Enumo na Universidade Estadual do Espírito Santo (UFES), dentre outros.

Os trabalhos realizados por esses grupos de pesquisas no Brasil têm sido desenvolvidos com alunos da primeira série do Ensino Fundamental, crianças com dificuldades de aprendizagem, crianças prematuras, crianças com deficiência visual, e crianças com problemas de comunicação. Também têm investigado problemas na habilidade lingüística de narrativa em crianças com câncer, avaliação de programas de intervenção com crianças com necessidades educativas especiais, programas de criatividade, programa com sistemas de comunicação alternativa e ampliada, intervenção no ensino regular e avaliação de fatores não-intelectuais em provas assistidas para crianças com necessidades educativas especiais (Enumo, Ferrão, Motta, Moraes & Linhares, 2006).

A avaliação dinâmica implica algumas especificidades no processo de avaliação psicológica. A primeira refere-se à sua função diagnóstica, ao complementar informações obtidas em avaliações estáticas. A segunda diz respeito a sua utilidade para captar diferenças individuais ligadas ao potencial de aprendizagem dos indivíduos. Já a terceira é sobre a condição de oferecer suporte instrucional dessa avaliação, o que, além de auxiliar o indivíduo na resolução da tarefa, faz com que ele tenha uma maior consciência dos processos cognitivos que devem ser utilizados no momento de tal resolução, propiciando contingência para melhor desempenho nas atividades (metacognição) (Goetz, Hall & Fetsco, 1990, citado por Linhares, 1995).

Como pode ser visto até o momento, em relação à avaliação dinâmica, a idéia básica subjacente é a verificação do potencial de aprendizagem do indivíduo. Essa investigação

pode ser facilitada por meio de procedimentos como os testes dinâmicos. Segundo Grigorenko e Sternberg (1998), o teste dinâmico é uma parte da avaliação dinâmica, sendo mais restrito à interação examinador/examinado, ou seja, quem é submetido ao teste e quem o auxilia nessa tarefa. Nas avaliações dinâmicas são utilizadas outras técnicas como a entrevista. Conforme descrito por esses mesmos autores, a finalidade da avaliação dinâmica é intervir e mudar e do teste dinâmico procura identificar quanto a pessoa poderia mudar caso tivesse instruções com sugestões de resolução de uma determinada tarefa, a partir de uma interação examinador/examinando. Em alguns métodos essa interação é limitada por apenas um simples *feedback*, em outros momentos pode romper essa barreira e ser uma forma de intervenção (Grigorenko & Sternberg, 1998).

Assim, os testes convencionais, estáticos, atentam para quantificar habilidades já desenvolvidas, mensurando uma capacidade latente e somente como ela é realizada no desempenho, o qual pode ser afetado por diversas variáveis, como por exemplo o histórico educacional que a pessoa teve a possibilidade de adquirir, o suporte familiar, entre outros fatores. A extensão do quanto uma habilidade poderia se desenvolver depende dessa capacidade latente, que seria uma sensibilidade às instruções, e do tipo de instrução fornecida para ajudar no maior desenvolvimento dessa capacidade latente (Sternberg & Grigorenko, 2002). Como já relatado, os testes dinâmicos apresentam essa característica de fornecer instruções.

Existem dois formatos mais comuns para os testes dinâmicos, um quando a instrução ocorre entre o pré-teste e o pós-teste, e outro em que a instrução ocorre em resposta à solução do examinado para cada item do teste. No primeiro formato o examinando faz o pré-teste, o qual é semelhante aos testes estáticos, e somente após completar todo o pré-teste são dadas instruções sobre a habilidade que o teste quer mensurar. Essas informações podem ser transmitidas em grupo ou individualmente e, quando são de maneira individualizada, refletem particularmente forças e fraquezas do indivíduo. Quando as instruções são em grupo, estas são as mesmas para todos os indivíduos. Após as instruções, os examinandos são testados novamente no pós-teste, o qual pode ser equivalente ao pré-teste ou igual a ele (o que é menos usual) (Sternberg & Grigorenko, 2002).

No segundo formato, o teste é sempre feito individualmente e os examinadores oferecem instruções item por item. Quando um item é corretamente solucionado, apresenta-se o próximo. Caso o indivíduo não solucione o item, o examinador oferece uma série de sugestões, sem resolver a questão, e também decide quantos e como será o tipo de sugestão, de acordo com o que ele julga ser necessário que o sujeito necessite para compreender a tarefa (Sternberg & Grigorenko, 2002).

Linhares e colaboradores (2006) acrescentam que os procedimentos com testes dinâmicos para avaliação assistida/dinâmica podem ser utilizados em pelo menos dois grupos diferentes de tarefas. O primeiro com testes psicométricos tradicionais, como



Matrizes Progressivas Coloridas de Raven (Raven, 1987) e a Escala de Inteligência Wechsler para Crianças III (1991), que são inicialmente aplicados na forma padronizada e depois com assistência ou mediação do examinador. O segundo grupo são tarefas construídas especialmente para serem aplicadas em processos de avaliação dinâmica, como o teste interativo ou dinâmico *Children's Analogical Thinking Modifiability Test – CATM* (Tzurriel & Klein, 1987, 1990) e tarefas de problemas por meio de perguntas de busca e restrição de alternativas (*constraint – seeking questions*) (Linhares, Santa Maria, Escolano & Gera, 1998).

O teste Raven implica diversas operações mentais, dentre elas o raciocínio analógico, classificação, discriminação, análise-síntese e atenção seletiva (Angeli, Alves, Custódio, Duarte & Duarte, 1999). Quando utilizado de forma dinâmica, permite verificar possíveis deficiências nas crianças em relação a alguma operação mental ou ao seu funcionamento cognitivo. Portanto uma das formas a se proceder com o Raven “dinâmico” é aplicá-lo inicialmente da forma convencional, e depois prosseguir com uma avaliação mais qualitativa em relação ao itens respondidos incorretamente, de modo a verificar como a criança pensou e, assim, averiguar possíveis deficiências. É possível, ainda, aplicar o instrumento oferecendo assistência para novas tentativas de respostas. Esse mesmo procedimento pode ser elaborado para o teste WISC, no entanto essa utilização que não segue a padronização dos testes inviabiliza o uso das tabelas normativas.

Já o teste dinâmico de Modificabilidade Cognitiva (CATM) foi elaborado para crianças pré-escolares ou com deficiências cognitivas e envolve o raciocínio analógico na resolução de problemas. É composto por 18 blocos coloridos de fórmica, diferentes na cor, forma e tamanho, e três conjuntos com 14 problemas cada, destinados às fases de pré-treino, treino/ensino e pós-treino. Esse teste está organizado em quatro etapas. A primeira é a fase preliminar para familiarizar a criança com a aplicação e os materiais utilizados com os conceitos básicos de cor, forma e tamanho e ensinar regras básicas de resolução de problemas. A segunda é a fase pré-treino, com a finalidade de verificar a linha de base para a avaliação da modificabilidade cognitiva do raciocínio analógico. A terceira fase é a do treino, com o objetivo de ensinar a criança a procurar o que é relevante para a solução analógica, compreendendo as regras de princípios de transformação e analogias, sistematizando os blocos corretos e melhorando a eficiência do treinamento. Por fim a quarta fase é a do pós-treino, em que se suspende a ajuda e avalia-se o nível de desempenho da criança, comparando posteriormente com seu desempenho no pré-treino. O registro e a análise dos dados são quantitativas e qualitativas, levando em conta o desempenho da criança nas diferentes fases a fim de identificar os ganhos com a assistência. Esses dados servem como uma indicação da modificabilidade cognitiva, em vez de uma evidência da realização duradoura e estável (Linhares e cols, 1998).

Ao que diz respeito às tarefas de perguntas de busca, esse procedimento permite verificar a eficiência ou não de estratégias relevantes de formular questões para obter informações, analisar alternativas, eliminar possibilidades irrelevantes e solucionar o

problema corretamente (Linhares & colaboradores, 1998). Esse tipo de tarefa é a base do Jogo de Perguntas de Busca com Figuras Geométricas-PBFG (Linhares, 1996) e o Jogo de Perguntas de Busca com Figuras Diversas (Gera & Linhares, 1998), que objetivam investigar as estratégias utilizadas pela criança na elaboração de busca de informações para tentar solucionar o problema. Para isso a criança formula uma pergunta e exclui alternativas para saber qual das oito figuras foi escolhida pela examinadora em cada arranjo para ser a figura-alvo. O PBFG foi adaptado em 2000 pelas pesquisadoras Enumo e Batista, a fim de avaliar crianças com deficiência visual. As modificações ocorridas dizem respeito a apenas quatro arranjos e não oito, e também o atributo cor foi substituído pela textura.

Normalmente os procedimentos de avaliação/teste dinâmico são realizados de maneira individual, o que facilita a interação entre examinando/examinador, bem como o registro e a avaliação qualitativa do teste. Porém, já há pesquisas sobre aplicações em grupos (Tzuriel & Feuerstein, 1992) e alguns testes computadorizados (Tzuriel & Shamir, 2002). Esses últimos autores desenvolveram um estudo avaliando o método dinâmico assistido pelo computador em relação ao desempenho cognitivo, comparando com o método somente com o examinador, para verificarem a eficiência do computador como uma ferramenta assistencial na avaliação dinâmica.

O método de avaliação dinâmica de Tzuriel e Shamir (2002) foi baseado na teoria da experiência de aprendizagem mediada de Feuerstein e no método para crianças de Tzuriel. O instrumento, especialmente desenvolvido para esse estudo, foi embasado no

*Children's Seriation Thinking Modifiability-CSTM* para ser aplicado em crianças de 3 a 5 anos de idade e é composto por pré-teste, treino e pós-teste, sendo baseado principalmente em processos de seriação e funções de comparação e quantificação. Esse teste foi elaborado no formato com e sem computador. Participaram do estudo 30 crianças avaliadas com o método computadorizado e examinador concomitantemente e 30 crianças apenas com o examinador, sendo grupos compostos por sujeitos com idade média de seis anos, sendo que o nível de inteligência delas foi controlado por meio da distribuição de frequência no teste Matrizes Progressivas Coloridas de Raven. No grupo experimental aplicou-se o programa de computador e no grupo controle aplicou-se o CSTM somente com o examinador. É importante ressaltar que o programa computadorizado foi desenvolvido de maneira a ser sempre necessária a ajuda de um examinador, mas os principais procedimentos são computadorizados, sendo que o aplicador pode responder as questões de maneira flexível em relação a aprendizagem durante todo o processo, inclusive durante o pré e pós treino.

Os resultados apontaram que o procedimento computadorizado mediado pelo examinador gerou mais ganhos na performance das mudanças cognitivas do que a mediação somente com o examinador, e que, em itens mais complexos, a forma computadorizada também foi mais eficiente para a aprendizagem. Também indicaram que, para uma melhor efetividade, é preciso associação entre o objetivo do ensino, as características da aprendizagem, o programa do computador e as decisões do educador, sendo que a intervenção do adulto com o computador é a mais eficaz. Os autores por

observações nas testagens, concluíram ainda, que a aplicação com o computador forneceu um ambiente mais estimulante e motivador, o que ajudou na efetividade do teste dinâmico.

Na última década, a utilização do computador como ferramenta de aprendizagem e ensino tem crescido, devido a algumas vantagens relacionadas a ele, como por exemplo, a de mensurar precisamente o tempo de reação das respostas e a de poder ser lúdico. Pesquisas como as de Chnag e Osguthorp (1990) e Goldmacher e Lawrence (1992) sustentam essa posição. Nesses trabalhos a utilização do computador, como uma ferramenta para aprendizagens cognitivas, permitiu maiores ganhos aos participantes dos grupos utilizando o computador. As habilidades trabalhadas foram memória e percepção visual no primeiro estudo, e identificação de palavras e compreensão de leitura no segundo estudo.

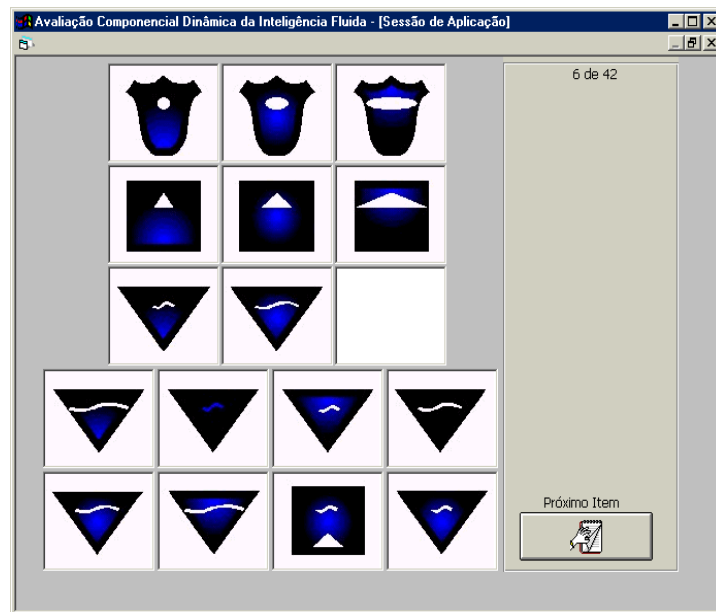
Um outro teste dinâmico informatizado foi desenvolvido por Primi, Sternberg e Grigorenko (1997). Esse teste tem como característica minimizar as intervenções de um tutor, ou seja, a aplicação inteira é informatizada, sendo a interação indivíduo-computador. Esse teste dinâmico é denominado de Avaliação Dinâmica do Raciocínio Indutivo da Inteligência Fluida-Gfri, apesar da terminologia avaliação, utilizada pelos autores, essa ferramenta é um teste dinâmico. Tal teste foi elaborado para a população universitária, podendo também ser aplicado a partir dos 15 anos de idade, tal como descrito a seguir

O programa foi feito para que a aplicação fosse coletiva e não requeresse do aplicador. A seção de aplicação é dividida em seis fases, quais sejam, identificação, instruções, avaliação inicial (Forma A, 12 itens), treinamento, avaliação final (Forma B, 12

itens com *feedback*) e resumo dos resultados. O início da aplicação é composto por três telas contendo as instruções para resposta ao teste. Após essas instruções os sujeitos iniciam o teste propriamente dito. Os itens são apresentados como mostra a Figura 1. Para escolher uma alternativa os sujeitos clicam com o *mouse* a alternativa escolhida. Isso faz com que a figura se mova para o espaço em branco da matriz. O sujeito pode desfazer sua escolha clicando sobre a alternativa escolhida no espaço da última célula. Isso faz com que ela volte ao lugar de origem.

Os sujeitos têm a possibilidade de eliminar as alternativas que julguem erradas. Esse recurso auxilia na solução do problema. Para eliminar uma alternativa, os sujeitos clicam com o botão direito do *mouse* sob a alternativa escolhida, o que faz com que ela desapareça do campo de visão. Eles podem anular essa ação clicando novamente sobre a alternativa eliminada.

Ao escolher a resposta os sujeitos devem clicar sobre o botão “Próximo Item” para passar ao próximo problema. A aplicação informatizada também permite o armazenamento de várias informações relativas à aplicação, além de gravar todas as respostas, coleta o tempo de reação da resposta.

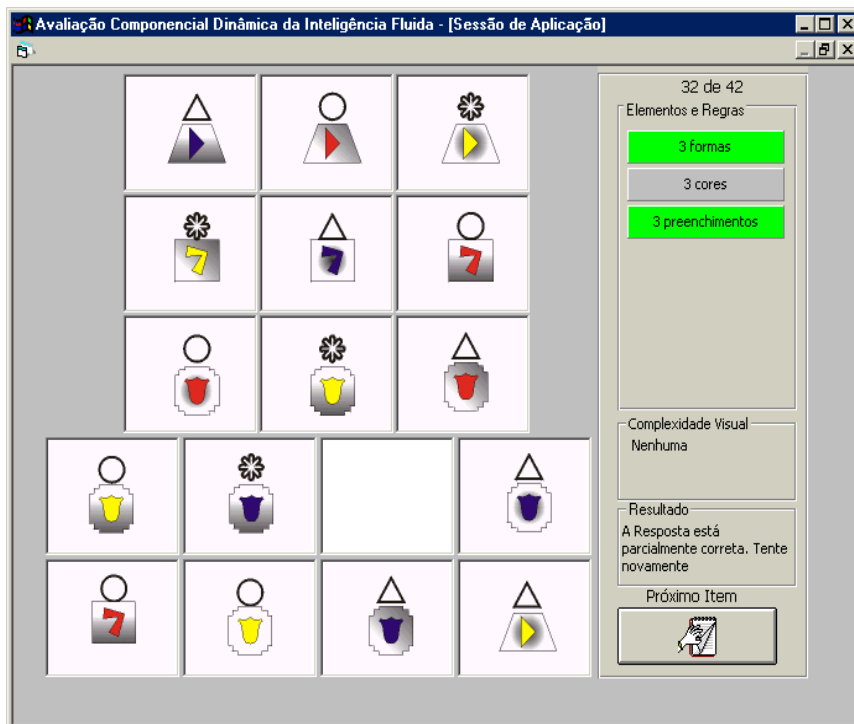


**Figura 1-** Exemplo de como os problemas são apresentados na avaliação inicial.

Após responder os doze itens da fase de avaliação inicial, começa o treinamento que explica a estrutura dos problemas. Nesta fase são fornecidas informações sobre como são compostos os problemas (número de elementos, regras e tipo de regras e complexidade perceptual) e sobre os passos do raciocínio analógico de acordo com o modelo de Sternberg (1977), isto é, codificação, inferência, mapeamento, aplicação e resposta.

Posterior esse treinamento, inicia-se a fase de avaliação final na qual mais doze problemas são apresentados, semelhantes aos primeiros, porém nessa final ocorre *feedback* aos sujeitos. Nessa etapa informa-se aos sujeitos o número de elementos, o número de regras e a complexidade perceptual. Além disso, permite-se que o sujeito saiba se a resposta está correta ou errada e, caso esteja errada, apresentam-se quais regras os sujeito conseguiu

alcançar. Porém para acertar a resposta, precisará descobrir as outras regras restantes.  
(Figuras 2).



**Figura 2-** Exemplo de como o item é apresentado na fase de avaliação final com *feedback*.

Os sujeitos podem clicar sobre a parte da tela com o nome da regra para ver novamente o exemplo discutido na seção de treinamento, caso não recordem o tipo de regra presente no problema. Nesta fase também é permitido que o sujeito mude a opção de resposta duas vezes, isto é, ele tem três tentativas. Entretanto a pontuação é proporcional à quantidade de tentativas usadas, ou seja, se o sujeito responder com duas tentativas, recebe metade dos pontos e, utilizando as três chances, recebe um quarto do total de ponto de sua



última escolha. Com o término da avaliação final é apresentada uma tela com os resultados preliminares dos sujeitos comparando com outras pessoas que já responderam.

Estudos com esse teste constataram melhora no desempenho entre o pré e o pós-treino, sustentando a hipótese de que é possível mensuração do potencial de aprendizagem em relação à inteligência fluida, mais especificamente no que se refere ao fator de raciocínio indutivo. Esse último estudo foi mais bem detalhado, pois a presente proposta da construção do teste dinâmico computadorizado também ocorrerá sem o auxílio de um examinador e avaliará o raciocínio indutivo, mas neste caso será desenvolvido para crianças de anos sete a 11 de idade. Porém não será uma adaptação desse teste proposto por Primi, Sternberg e Grigorenko (1997).

Concorda-se que a aplicação com o examinador e o recurso informatizado juntos parecem constituir uma situação mais adequada. No entanto, nem sempre isso é possível, principalmente em situação de aplicação de teste na forma coletiva, o que permite ter uma amostra maior em menor tempo. A eliminação da intervenção de um examinador contribui para que o teste dinâmico possa ser aplicado em grupo e de uma maneira mais padronizada a todos os indivíduos. Dessa forma, facilita a aplicação e a interpretação dos resultados, resguardando o principal objetivo que é mensurar o potencial de aprendizagem.

Voltando para a teoria de Vygotsky (1970/1980), o conceito de zona proximal de desenvolvimento é uma das bases para a teoria da avaliação dinâmica. Também se pode citar o conceito de aprendizagem mediada, que é uma interação entre dois indivíduos,

sendo que um possui habilidades mais bem desenvolvidas propiciando que o outro aprenda por meio dessa relação. Aliás a questão da mediação é uma idéia central na teoria de Vygotsky. A mediação caracteriza a relação do ser humano com o mundo, sendo que por meio desse processo as funções psicológicas superiores se desenvolvem. Para que ocorra a mediação há dois elementos fundamentais, o *instrumento* e o *signo*, os dois regulam ações, primeiro sobre os objetos e o segundo sobre o psiquismo das pessoas. No caso do teste dinâmico construído, pode-se referir a um tipo de mediação computacional, visto que ocorrerá por meio de instruções lingüísticas padronizadas oferecidas pelo software. Isso difere das avaliações dinâmicas mais tradicionais em que a interação é mais social com outro ser humano.

Ressalta-se que esse teste dinâmico não tem o caráter de avaliar uma mudança permanente em relação a alguma função cognitiva, mas sim funcionar como um *screening* do potencial de aprendizagem, devendo ser utilizado dentro de uma avaliação mais completa, como é o caso da avaliação dinâmica e não do teste dinâmico. Também é importante esclarecer, que essa forma computadorizada é uma alternativa dentre os formatos de testes dinâmicos. Dessa forma os testes dinâmicos que implicam tutores continuam tendo seus espaços preservados, pois compreende-se a grande importância desse formato interativo examinando-examinador.

A efetividade em avaliar o potencial de aprendizagem do presente teste será verificada no estudo ao comparar o desempenho entre pré e pós-teste, havendo entre eles

um procedimento de intervenção relacionado à habilidade a ser ensinada, no caso o raciocínio indutivo. O desempenho acadêmico também será uma variável investigada em relação ao raciocínio indutivo e ao potencial de aprendizagem da criança. O desempenho foi investigado tanto com notas escolares finais quanto com uma escala de percepção dos professores para classificarem os cinco alunos com maior e menor facilidade de aprender, sem considerar comportamento emocional. Essa escala foi aplicada para que houvesse um critério de desempenho menos contaminado por outras variáveis, por exemplo o comportamento ou a progressão continuada.

Para Depresbiteris (1997) a nota final da avaliação tem um caráter burocrático e que pode não refletir a aprendizagem ocorrida durante o período letivo, bem como a atribuição de nota pelo professor muitas vezes é feita sem considerar os objetivos da disciplina ou do projeto pedagógico. Essa situação parece se agravar com a implantação do sistema de progressão continuada, também referido como progressão automática, em todo o Brasil na educação básica (Lei de Diretrizes e Bases da Educação, 1996).

Por fim, outra variável critério a ser avaliada com o potencial de aprendizagem e com o raciocínio indutivo serão as funções executivas atenção seletiva e controle inibitório. Tais variáveis são estudadas na literatura e correlacionadas com a inteligência fluida, que tem como um dos principais fatores o raciocínio indutivo. Já em relação ao potencial de aprendizagem do raciocínio indutivo não foram identificados estudos com a finalidade de comparação com essas funções executivas. Ressalta-se que a discussão sobre o conceito de

funções executivas e habilidades relacionadas, bem como o seu papel em relação aos processos cognitivos é um tema recente de pesquisa, com várias divergências ainda existentes como, por exemplo, quais as habilidades que estão relacionadas à essas funções.

Outro propósito deste estudo que merece ser mencionado é que todo o teste dinâmico, incluindo os itens e o procedimento de intervenção, serão fundamentados na Teoria Prescritiva do Treinamento do Raciocínio Indutivo de Josef Klauer (Klauer,1990, Klauer & Phye, 1995, Klauer, Willmes & Phye, 2002). Esse autor desenvolveu um treinamento do raciocínio indutivo para crianças e será mais bem detalhado na Seção 3. Como a habilidade cognitiva do raciocínio indutivo é o objeto a ser mensurado pelo teste dinâmico, a próxima seção tratará mais especificamente sobre a inteligência, destacando a capacidade do raciocínio indutivo.

## 1.2– Inteligência e raciocínio indutivo

A inteligência humana é um construto central na história da Psicologia, sendo um dos mais estudados. Entre os anos de 1900 e 1990, um dos focos da ciência psicológica foi a busca pela conceituação da inteligência, quais os elementos que a compunham e como esses elementos se organizavam. Essa busca continua, visto que ainda não há uma única definição e consenso sobre a mesma, sendo que Primi (2002b) também constatou isso em 2002.

As opiniões sobre a inteligência sempre divergiram, como revela o estudo realizado pelos editores do *Journal of Educational Psychology* em 1921, que trouxe à tona muitas opiniões diferentes sobre a conceituação do tema (Sternberg, 1992). De acordo com Richardson (1991), durante o século XX diversos conceitos sobre inteligência foram formulados, porém as definições sempre foram genéricas, permitindo muitas discussões e opiniões divergentes acerca da conceituação do construto. Reforçando essa posição, Sternberg (1992) sugere que o conceito de inteligência geral é vago e evasivo. No entanto esse mesmo autor junto ao teórico Derthman em 1986, procurando minimizar essas discussões sobre o conceito de inteligência e buscando uma maior convergência entre os teóricos desse construto, fizeram uma análise de 24 definições da inteligência propostas por especialistas da área. A partir disso elaboraram um conceito sobre inteligência;

“a capacidade para aprender a partir da experiência, usando processos metacognitivos para melhorar a aprendizagem, e a capacidade para adaptar-se ao ambiente circundante, que pode

exigir diferentes adaptações dentro de diferentes contextos sociais e culturais” (Sternberg, 2000, p.400).

A inteligência pode ser explicada por três grandes correntes, quais sejam, a desenvolvimentista, a abordagem do processamento humano de informação ou Psicologia cognitiva, e a psicométrica ou fatorial (Almeida, 1988, 1994; Sternberg, 1977, 1979, 1980, 1981, 1983, 1986). O modelo desenvolvimentista é baseado nas teorias de Piaget e Vygotsky, e define as estruturas e a dinâmica da inteligência durante o desenvolvimento do ser humano. Esse modelo concebe a inteligência como um fenômeno que se desenvolve por etapas de maneira crescente e invariante na estruturação, porém de forma qualitativamente diferente entre os indivíduos. O término do desenvolvimento estrutural ocorre na adolescência. Destaca-se que o método de investigação nesse modelo é a observação e a descrição dos fenômenos como por exemplo a fala, a interação com o ambiente e a aprendizagem.

No modelo da Psicologia Cognitiva o termo inteligência abrange e unifica todas as dimensões cognitivas e tem como objetivo sistematizar os processos mentais envolvidos na cognição, tais como a sensação, a percepção e a memória. Essa teoria, hoje em dia, estuda os processos cognitivos envolvidos na resolução das tarefas que compõem os testes tradicionais de inteligência usados pela psicomетria, empregando métodos experimentais de análise, como por exemplo o trabalho desenvolvido por Primi (1995).

Por fim o modelo psicométrico tradicional investiga a estrutura e a organização subjacente à inteligência, utilizando-se frequentemente do método de análise fatorial. Por

esse objetivo mais estrutural, tal modelo tem sido alvo de críticas, sugerindo que ele não investiga o processo cognitivo, mas somente o produto advindo de diferenças individuais nos testes de inteligência.

Mais recentemente outra abordagem está se consolidando no estudo da inteligência, a neurociência cognitiva. A finalidade da abordagem é a identificação dos processos cognitivos da inteligência, relacionando-os com os aspectos estruturais e funcionais do cérebro (Hunt, 1999; Posner & DiGirolamo, 2000).

Independente da abordagem que investiga a inteligência, o raciocínio é uma variável freqüentemente estudada. Até mesmo no senso comum, utiliza-se a expressão “raciocinar bem” para caracterizar uma pessoa inteligente. Raciocínio é uma atividade do pensamento, crucial, durante toda a nossa vida, sendo um ponto central em toda estrutura do pensamento (Wilhelm, 2005). Para Vinod Goel (2005), raciocínio é uma atividade cognitiva que deriva inferências a partir de informações fornecidas. Todo raciocínio envolve eduções com base em uma ou mais premissas, as quais fornecem alguma informação para aceitar outra proposição que é a conclusão.

A indução é um raciocínio fundamental na ciência quando tratam-se de descobertas, mas para testar teorias é necessário o raciocínio dedutivo (Wilhelm, 2005). Para Holyoak e Morrison (2005), as diferenças entre os dois tipos de raciocínios são sutis e enfatizam conclusões a partir de algumas premissas iniciais. Na lógica, uma inferência é dedutiva se a verdade da premissa garante a verdade da conclusão por virtude da forma do argumento. Já

se a premissa verdadeira não garante conclusões confiáveis, a inferência é chamada de indutiva.

Especificamente na pesquisa aqui desenvolvida o raciocínio, particularmente o indutivo, é a habilidade proposta para ser ensinada. Seu arrazoado teórico será pautado tanto na perspectiva psicométrica quanto na cognitiva, pois essas teorias juntas contribuem tanto para uma melhor compreensão do raciocínio, quanto para a construção de instrumentos para a avaliação dessa habilidade. A abordagem cognitiva permite aprofundar como o indivíduo raciocina e, pela abordagem psicométrica, verifica-se a estrutura do raciocínio, sua validade e predição para a vida do ser humano.

Atualmente está ocorrendo uma interface do modelo psicométrico com a Psicologia Cognitiva, o que vem permitindo avanços na compreensão da inteligência, bem como a possibilidade de realização de novos estudos empíricos. Esses dois modelos se complementam e preenchem uma lacuna, respondendo à crítica de que o modelo psicométrico preocupa-se apenas com o resultado final e não com os processos cognitivos subjacentes às respostas. Os principais elementos dessa aproximação podem ser encontrados nos trabalhos de Embretson (1983, 1985, 1994, 1999). Portanto, a seguir primeiramente será explorado o raciocínio indutivo pela abordagem psicométrica, depois pela cognitivista e, por fim, será argumentada a relevância de utilizar essas duas abordagens na construção de instrumento, no caso uma medida do raciocínio. No decorrer da exposição, o leitor perceberá que apesar do foco ser o raciocínio indutivo, não há como



deixar de abordar o dedutivo, mesmo porque na abordagem psicométrica essa separação entre esses dois processos é difícil (Wilhelm, 2005).

Os estudos e teorias fatoriais da inteligência, que são a grande marca da psicometria, podem ser divididos entre a tradição britânica e a tradição americana. A britânica tem como destaques as teorias de Spearman, Burt e Vernon, já a americana é representada por Thurstone, Cattell e Horn. Hoje em dia a teoria psicométrica mais moderna e difundida é a do modelo Cattell-Horn-Carroll, que integra essas duas tradições (Carroll, 1997). Para uma melhor compreensão desse histórico da inteligência serão abordadas essas passagens e evoluções entre as teorias fatoriais.

Das teorias fatoriais da inteligência, a proposta por Spearman (1927) é uma das mais difundidas nessa área. O autor concebeu a inteligência como um fator geral ou simplesmente *fator g*, definindo tal construto como um “estoque de energia mental”. A partir do *fator g* formulou a Teoria dos Dois Fatores da Inteligência, conhecida como Bi-Fatorial. Tal teoria pressupõe que, para a realização de qualquer atividade intelectual, é necessário o fator geral e um fator específico condizente com a tarefa proposta. No entanto, o fator específico não é totalmente independente, pois possui correlação com o geral.

Em termos cognitivos o fator *g* pode ser descrito por três processos básicos. O primeiro corresponde à apreensão da experiência e refere-se à capacidade relacionada à percepção, rapidez e acuidade para discriminar os estímulos e apresentar consciência desse processo. O segundo, educação das relações, é a capacidade de estabelecer relações entre duas

ou mais idéias percebidas no ambiente ou evocadas da memória (esse processo é o correspondente ao raciocínio indutivo). Já o terceiro processo denominado de educação de correlatos, diz respeito à capacidade de derivar novas idéias ou relações a partir de uma idéia ou uma relação previamente estabelecida (correspondente ao raciocínio dedutivo).

Thurstone (1934), a partir de uma análise fatorial verificando componentes da inteligência, formulou a Teoria das Aptidões Primárias. Nela, ele descarta a existência de um fator geral, como proposto por Spearman, e propõe nove fatores de capacidades mentais independentes entre si, ou seja, sem relação entre eles. Tais fatores são compreensão verbal, fluência verbal, raciocínio indutivo, raciocínio numérico, raciocínio aritmético, raciocínio dedutivo, visualização espacial, memória e rapidez perceptiva.

A divergência entre as teorias de Spearman e Thurstone levaram outros pesquisadores a estudar o construto inteligência e a sua estrutura por meio da análise fatorial. Então, na década de 1950, surgiram alguns modelos tentando conciliar as duas teorias. Os modelos hierárquicos e o da inteligência fluida e cristalizada foram os mais importantes para o desenvolvimento da teoria psicométrica da inteligência (Almeida, 1988). Segundo Sisto (2006) ainda há pesquisadores que apóiam a Teoria de Thurstone sobre as Aptidões Primárias, porém o próprio Thurstone (Thurstone & Thurstone, 1941 citado por Sisto, 2006), por meio de outros trabalhos, verificou a existência de um fator geral da inteligência, concordando com Spearman.

Com relação ao modelo hierárquico, Cyril Burt foi o primeiro a propor uma estrutura baseando-se na teoria de Spearman, dividindo as capacidades em quatro níveis. O primeiro, sensações e reações; o segundo, percepções e coordenação motora; o terceiro, memória e formação de hábitos; e o quarto, compreensão e raciocínio. No entanto, outra teoria hierárquica emergida da análise fatorial, proposta por Vernon, foi mais disseminada e também classificou a inteligência em quatro níveis (Almeida, 1988). No primeiro nível encontra-se o *fator g* de Spearman. O segundo é composto por dois fatores, o verbal-educativo e o perceptivo-mecânico. E no terceiro nível agrupam-se fatores semelhantes aos da teoria de Thurstone. Por fim o quarto nível é constituído por fatores específicos como os relacionados as tarefas de leitura e ortografia (Vernon, 1961).

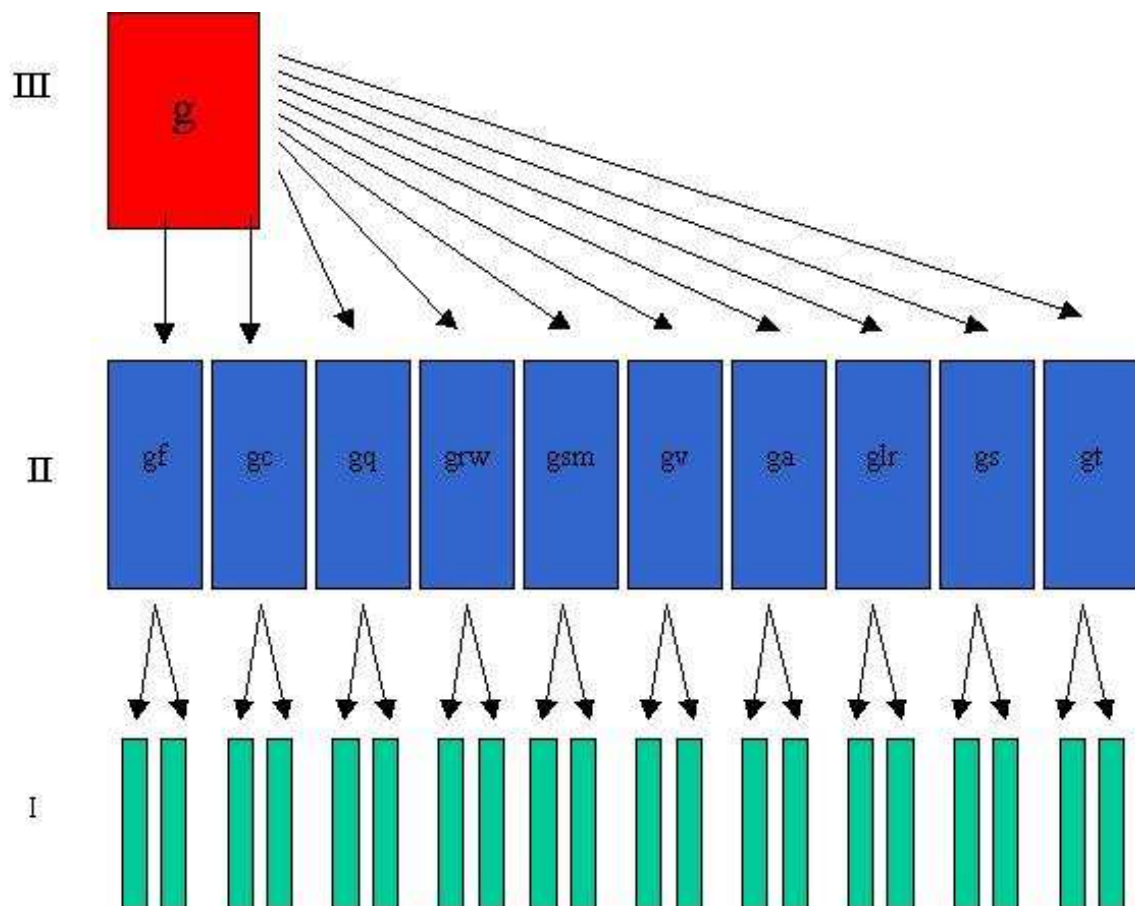
No que diz respeito à Teoria da Inteligência Fluída e Cristalizada, esta foi descrita por Cattell nos anos de 1941 e 1957. Estas duas capacidades foram deduzidas dos fatores primários de Thurstone, também fazendo uso da análise fatorial. Este modelo dicotomizou a inteligência em Fluida e Cristalizada. A primeira refere-se à capacidade de raciocinar em situações novas e a segunda, ao conhecimento acumulativo (Flanagan, McGrew & Ortiz, 1999). Cattell e Horn (1966) acrescentaram a essa Teoria Fluida-Cristalizada as capacidades de processamento visual-Gv, memória a curto-prazo-Gsm, memória de longo prazo-Glr e rapidez de processamento cognitivo-Gs.

Em 1968, Horn adicionou a capacidade de processamento auditivo-Ga e, em 1985, a capacidade de conhecimento quantitativo-Gq (Flanagan, McGrew & Ortiz, 1999). Horn

também propôs uma estrutura multidimensional hierárquica em dois níveis com fatores gerais e específicos, porém não concebe a existência de um fator geral. Para ele o fator *g* é apenas um artefato das correlações, sem realmente ter um fator único subjacente a outros fatores, ou mesmo nas tarefas de testes. Os fatores gerais são representados pela Inteligência Fluída-Gf, Inteligência Cristalizada-Gc, processamento visual-Gv, processamento auditivo-Ga, memória de curto prazo-Gsm, conhecimento quantitativo-Gq, velocidade de processamento e velocidade de decisão-Gs, já os específicos são atrelados às tarefas apresentadas aos sujeitos (McGrew, 1997).

Carroll (1993), a partir dos resultados de uma meta análise, propôs a Teoria dos Três Estratos da Inteligência, procurando integrar os vários fatores da inteligência. Nesse estudo o autor fez um levantamento dos últimos sessenta anos de pesquisas realizadas pela psicometria sobre a inteligência e selecionou 1500 artigos, obtendo 460 conjuntos de dados para serem analisados. Os resultados encontrados sugeriram uma estrutura fatorial hierárquica da inteligência, composta por três estratos e que estão correlacionados. O estrato mais básico é o terceiro e assemelha-se ao fator geral de Spearman, sendo descrito como um complexo processo cognitivo superior que é comum às capacidades e atividades intelectuais. O segundo estrato é composto pelas capacidades que compõem a inteligência, incluindo a inteligência fluída, cristalizada, processos de aprendizagem e memória, percepção visual, percepção auditiva, habilidade de recuperação, velocidade cognitiva e velocidade de processamento e decisão. Por fim o terceiro estrato é um total de 75

habilidades específicas ligadas às capacidades do segundo estrato. Por meio da Figura 3 pode-se visualizar a estrutura hierárquica do modelo CHC.



**Figura 3-** Teoria CHC das Habilidades Cognitivas (McGrew, 2007)

Mais recentemente, a teoria dos três estratos foi integrada ao modelo da inteligência fluída e cristalizada (Gf-Gc), passando a ser conhecida como teoria de Cattell-Horn-Carroll (CHC) das habilidades cognitivas (Carroll, 1997). Apesar de Cattell e Horn não conceberem a existência de um fator geral, o modelo CHC mantém a idéia de um fator

geral proposto por Spearman e corroborado por Carrol. A teoria CHC consiste em uma visão multidimensional composta por dez fatores ligados a áreas amplas do funcionamento cognitivo: inteligência/raciocínio fluido (Gf), inteligência/raciocínio cristalizado (Gc), raciocínio/conhecimento quantitativo (Gq), leitura-escrita (Grw), memória a curto prazo (Gsm), inteligência/processamento visual (Gv), inteligência/processamento auditivo (Ga), armazenamento e recuperação associativa a longo prazo (Glr), rapidez de processamento cognitivo (Gs), tempo/velocidade de decisão/reação (Gt) (McGrew & Flanagan, 1998). Tais fatores são detalhados a seguir e fazem parte do segundo estrato dessa teoria.

1. A inteligência fluida (Gf) refere-se às operações mentais de raciocínio que o indivíduo realiza diante de situações novas que não podem ser realizadas automaticamente e que dependem minimamente de conhecimentos adquiridos. Essas operações envolvem relacionar idéias, induzir conceitos abstratos e solucionar problemas, empregando principalmente raciocínio indutivo e dedutivo.
2. A inteligência cristalizada (Gc) refere-se à extensão e profundidade dos conhecimentos adquiridos de determinada cultura, bem como a aplicação dos conhecimentos aprendidos previamente. Esse fator representa a habilidade de raciocínio adquirida pelo investimento da capacidade geral em experiências de aprendizagem. É um fator ligado à linguagem verbal.

3. O conhecimento quantitativo (Gq) refere-se ao estoque de conhecimentos quantitativos declarativos<sup>2</sup> e procedural<sup>3</sup> armazenados por um indivíduo, destacando a habilidade em utilizar informação quantitativa e manejo de símbolos numéricos. Está mais ligado ao conhecimento matemático do que ao raciocínio matemático, que se localiza como um fator específico de Gf.
4. A leitura e escrita (Grw) referem-se ao conhecimento adquirido em habilidades básicas exigidas no entendimento de textos e vocábulo escrito, incluindo a capacidade elementar (decodificação em leitura, ortografia) e complexa (compreensão de texto, composição de histórias).
5. A memória de curto prazo imediata (Gsm) é a capacidade de manutenção de informações na consciência por um espaço de tempo curto para poder readquiri-las logo em seguida.
6. O processamento visual (Gv) é capacidade de gerar, perceber, analisar, armazenar, relembrar, manipular, transformar e raciocinar com representações ou padrões visuais. Está vinculada aos diferentes aspectos do processamento de imagens (geração, transformação, armazenamento e recuperação).

---

<sup>2</sup> “Conhecimento declarativo – reconhecimento e entendimento da informação real sobre os objetos, as idéias e os eventos no ambiente (“saber que”, não “saber como”)” (Sternberg, 2000 p.430).

7. O processamento auditivo (Ga) é a capacidade ligada à percepção, análise e síntese de padrões sonoros, não tem relação direta com a compreensão, mas afeta o desenvolvimento. Está ligado à discriminação de padrões sonoros (incluindo a linguagem oral), particularmente quando emitidos em situações mais complexas envolvendo distorções, assim como à percepção de nuances em estruturas musicais complexas.
8. A capacidade de armazenamento e recuperação da memória de longo prazo (Glr) está ligada à extensão e fluência de informação ou conceitos readquiridos da memória de longo prazo por associação. Está ligada também ao processo de armazenamento e recuperação posterior por associação.
9. A velocidade cognitiva geral (Gs) refere-se à habilidade de realizar uma tarefa cognitiva de forma rápida e automática, mantendo a atenção focalizada e a concentração. Comumente está ligada à circunstância em que há um intervalo fixo e a pessoa necessita realizar o maior número possível de tarefas simples.
10. A velocidade de processamento/rapidez de decisão (Gt) refere-se à velocidade de reação ligada à capacidade de responder às tarefas cognitivas mais complexas, envolvendo problemas de compreensão, raciocínio e solução

---

<sup>3</sup> “Conhecimento procedural (procedimento) – informação quanto ao modo de executar uma seqüência de operações; compreensão e consciência de como realizar tarefas, habilidades ou procedimentos específicos (“saber como” não “saber que”) (Sternberg, 2000 p.430).



de problemas associados à rapidez em reagir ou tomar decisões. Enquanto Gs denota a capacidade de trabalhar rapidamente por um tempo mais longo (sustentabilidade), Gt diz respeito à reação rápida a um problema envolvendo processamento e decisão (imediatez).

Mesmo com a integração dos dois modelos e sendo a teoria fatorial da inteligência, na atualidade, mais completa, o modelo não esgota todas as capacidades de inteligência existentes ou que ainda poderão ser descobertas. Foi o caso, por exemplo, da inteligência emocional, descrita em 1990 pelos pesquisadores Paul Salovey e John Mayer, e que até o momento não integra o modelo CHC, visto que ainda são necessárias mais pesquisas para consolidar esse construto. No entanto, após o surgimento do modelo CHC, têm sido propostas algumas habilidades a serem integradas no modelo, quais sejam Conhecimento geral/domínio específico-Gkn, Velocidade psicomotora-Gps, Habilidade psicomotora-Gp, Habilidade Olfativa-Go, Habilidade Tátil-Gh e Habilidade Cinestésica-Gk. As três primeiras são classificadas como estruturas internas, pois já são domínios mais bem estabelecidos na natureza CHC, já as três últimas são consideradas estruturas externas, pois sugerem domínios novos ou que apresentam apenas investigações parciais (McGrew, 2007). Estudos com essas novas habilidades sugerem a ampliação do modelo CHC (Bowman, Markham & Roberts, 2002; Stankov, 2000). Essas seis novas habilidades são melhores descritas a seguir.

1. Conhecimento Geral (Domínio Específico) (Gkn) refere-se à amplitude e profundidade de conhecimentos adquiridos sobre um domínio específico que não representam experiências universais gerais típicas de um indivíduo com a cultura. Por exemplo, o conhecimento do inglês como uma segunda língua ou conhecimento em geografia.

2. Velocidade Psicomotora (Gps) é a habilidade de executar movimentos motores com o corpo de maneira rápida e fluente, como, por exemplo, escrever uma palavra repetidamente o mais rápido possível. Ressalta-se que a precisão não é importante.

3. Habilidade Psicomotora (Gp) diz respeito a movimentos motores do corpo com coordenação ou força. A habilidade para executar exige destreza, precisão e fluência entre vários movimentos. Desempenhar-se bem na atividade de dança, por exemplo, demonstra essa habilidade.

4. Habilidade Olfativa (Go) depende dos receptores sensoriais do sistema olfatório principal. Envolve habilidades como memória de odores e sensibilidade para diferenciar odores.

5. Habilidade Tátil (Gh) é semelhante à olfativa, porém depende dos receptores sensoriais do sistema tátil e de *inputs* sobre o aparato tátil. É uma habilidade que permite, por exemplo, a realização de discriminações finas de pressão sobre a superfície da pele.

6. Habilidade Cinestésica (Gk), por sua vez, depende dos receptores sensoriais que detectam a posição corporal, peso ou movimento dos músculos, tendões e juntas. Inclui a habilidade para detectar resposta ou movimento do corpo ou partes do corpo.

A Figura 4 sintetiza visualmente a integração de teorias fatoriais que originaram o modelo CHC. A Figura 1a é o modelo fatorial geral de Spearman que contribuiu para a teoria hierárquica dos Três Estratos de Carrol. A Figura 1b é o modelo fatorial Habilidades mentais Primárias de Thurstone que foi referência tanto para a teoria hierárquica dos Três Estratos quanto para a teoria hierárquica do modelo Inteligência Fluida e Inteligência Cristalizada, Gf-Gc, de Cattell e Horn e que é ilustrada pela Figura 1c. Por fim a Figura 1e representa o modelo CHC que uniu os modelos hierárquicos Gf-Gc e Três Estratos que como já apontado o primeiro se embasou no modelo de Thurstone e o segundo no modelo de Thurstone e Spearman. Por isso que se enfatiza que o modelo CHC é o mais completo e abrange as principais teorias fatoriais.

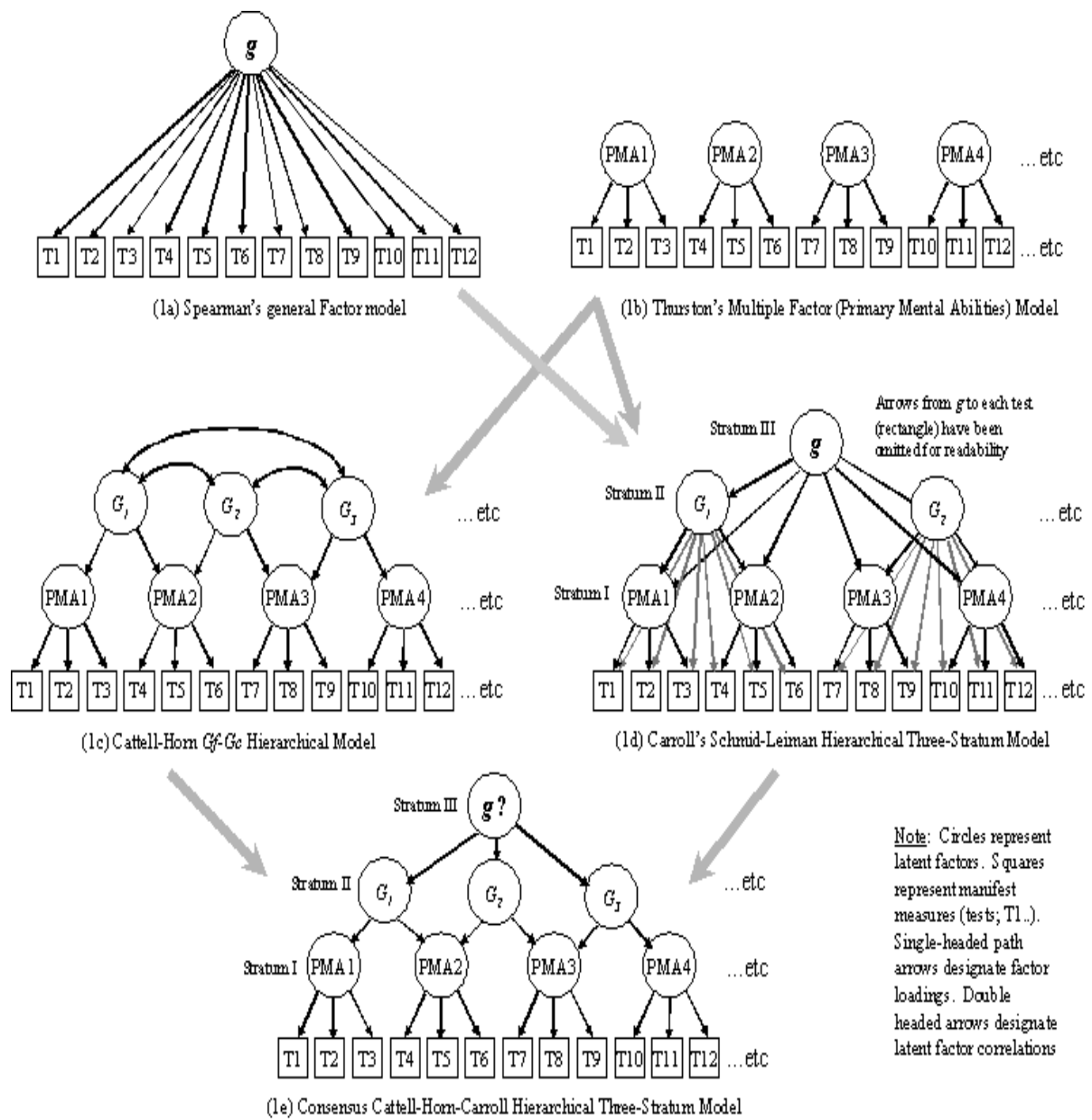


Figure 1: Major stages in the evolution of psychometric theories from Spearman's  $g$  to Cattell-Horn-Carroll (CHC) theory

Figura 4- Integração das teorias fatoriais no modelo CHC (McGrew, 2007)

Ressalta-se que os estudos de Carrol (1993) revisaram dados entre o período de 1927 e 1980, portanto não incluíram artigos da década de 1990. Diante disso foi criado o *Carrol Human Cognitives Abilities Project*, com o objetivo de continuar os estudos que geraram o CHC refazendo análises com métodos mais avançados, como a análise fatorial confirmatória, e analisar matrizes de estudos mais recentes, tendo o marco inicial a década de 1980. Já foram levantadas 118 referências potenciais para serem incluídas nesse novo tratamento dos dados sobre a inteligência (McGrew, 2007).

Mais especificamente em relação à estrutura do raciocínio indutivo e dedutivo, Wilhelm (2005) elencou contribuições de algumas teorias psicométricas. De acordo com esse autor, Spearman criou a análise *tetrad* para avaliar as correlações entre as capacidades de raciocínio e iniciou, assim, a análise fatorial, definindo o fator g como educação de relações (indução) e correlações (dedução), que são melhores refletidas em medidas de raciocínio. Na teoria de Thurstone a distância entre dedutivo e indutivo foi abandonada e as variâncias associadas foram alocadas em Raciocínio, Competência Verbal, Numérico e Espacial, sendo que o fator Raciocínio é marcado por tarefas de raciocínio indutivo e vários outros fatores apresentam cargas substanciais de raciocínio.

Na teoria de Carrol (1993), o fator inteligência fluida é largamente identificado por habilidades de raciocínio distintas no primeiro estrato da teoria, quais sejam, raciocínio seqüencial, indução e raciocínio quantitativo. Raciocínio seqüencial é mensurado por tarefas que requerem dos participantes a análise de premissas, regras ou condições para

conclusões que propriamente sigam as questões. A indução é avaliada por tarefas que fornecem aos indivíduos materiais que são governados por alguma regra ou princípio, similaridades ou diferenciações, requerendo inferência ou detecção dessas características dos estímulos e aplicar a regra inferida. Raciocínio quantitativo é medido por tarefas que pede ao participante relacionar conceitos envolvendo relações numéricas ou matemáticas. Horn e Noll (1994 e 1997, citados por Wilhelm, 2005) interpretam inteligência fluida como raciocínio indutivo e dedutivo. Esses raciocínios podem ser mensurados tanto por conteúdo verbal quanto espacial.

Com relação à diferença do raciocínio entre a teoria de Carrol e as de Horn e Cattell, pode-se observar que a primeira localiza o raciocínio quantitativo dentro da Gf. Acrescenta-se que no modelo mais atual, o CHC, que juntou as teorias de Carrol e de Horn e Cattell, os fatores de raciocínio dedutivo, indutivo e quantitativo, juntamente com velocidade de raciocínio e raciocínio piagetino, compõem o Fator Inteligência Fluida (McGrew, 2007)

Wilhelm (2005) ainda argumenta que a distinção entre raciocínio indutivo e dedutivo é controversa, ou seja, a idéia de que raciocínio indutivo parte de premissas específicas para conclusões gerais e o dedutivo ao contrário. Existem argumentos de que tanto o indutivo quanto o dedutivo dependendo do item, podem apresentar tanto premissas do geral para o específico quanto vice-versa. Por exemplo, o argumento “Quase todos os suecos são loiros. Jan é sueca. Portanto Jan é loira” é um argumento indutivo que aparenta pouco da diferenciação do dedutivo, pois parte de uma premissa geral. Porém caracteriza-se

como indutivo visto que a conclusão é inferida por regras que não necessariamente são verdadeiras, característica do raciocínio indutivo, pois no dedutivo a conclusão é necessariamente verdadeira. Em um outro exemplo, “Jan é sueca. Jan é loira. Portanto alguns suecos são loiros” parte-se de premissas específicas, o que corresponderia ao raciocínio indutivo, mas chega-se a uma conclusão específica e correta, característica do raciocínio dedutivo.

De acordo com Wilhelm (2005), Carrol também discute a estrutura das habilidades de raciocínio, verificando objeções e dificuldades, argumentando que os testes de raciocínio são freqüentemente complexos e requerem os dois tipos de raciocínio no processo de pensamento. Ressalta ainda que esses dois tipos são apreendidos e desenvolvidos juntos e muitas medidas desconhecem a quantidade de habilidades espaciais, lingüísticas e quantitativas embutidas nos itens. Apesar desse autor ter proposto três fatores, quais sejam, dedutivo, indutivo e quantitativo conforme anteriormente mencionado, em várias reanálises verificou-se um fator geral de raciocínio, pois não havia medidas suficientes dos três fatores. Outro dado que também dificulta essas análises de diferenciação é que vários estudos analisados não tinham a intenção de verificar a estrutura do raciocínio, bem como há variações de procedimento, tempo, instruções e as medidas são classificadas depois e não *a priori* de sua criação.

Nos estudos de Carrol (1993), as tarefas dedutivas são freqüentemente verbais, as indutivas, figural-espaciais e as de raciocínio quantitativo, numéricas. A distinção entre os

três fatores provavelmente decorre desse conteúdo. Conseqüentemente, o fator dedutivo pode não ser distinguido do verbal, possuindo alta carga no fator cristalizado, enquanto o fator indutivo com conteúdo figural-espacial tem altas cargas em inteligência fluida, bem como tem cargas altas em inteligência geral. A média da carga de tarefas indutivas em *g* é 0,57 e a de tarefas dedutivas é 0,41. Inteligência fluida é melhor definida por indução nas reanálises de Carrol, as cargas fatoriais de tarefas indutivas no fator inteligência fluida, nessas análises, ocorrem 19 vezes com carga média de 0,64, enquanto que as tarefas dedutivas aparecem somente 6 vezes e com carga média 0,55 (citado por Wilhelm, 2005).

Dadas essas questões, as propostas de raciocínio indutivo, dedutivo e quantitativo parecem estar fortemente relacionadas com os tipos de raciocínio verbal, figural-espacial e quantitativo. Existem basicamente cinco teorias para explicar a estrutura da habilidade de raciocínio. A primeira propõe um fator de raciocínio geral, contabilizando comunalidades de tarefas de raciocínio variando o conteúdo (verbal, figural-espacial e quantitativo) e a operação (indutivo e dedutivo). A segunda especifica dois fatores que se correlacionam, indutivo e dedutivo, sem especificar conteúdo. A terceira, coloca três fatores correlacionados, sendo raciocínio verbal, quantitativo e figural-espacial, sem distinguir dedutivo e indutivo. A quarta apresenta um fator geral que engloba raciocínio verbal e quantitativo. Finalmente, a quinta pressupõe dois fatores correlacionados de raciocínio dedutivo e indutivo com fatores de conteúdos verbais e quantitativos. Avaliando esses modelos percebe-se uma confusão entre processo e conteúdo de tarefas, bem como



dificuldades para encontrar medidas apropriadas, pois as tarefas tendem a misturar demandas (Wilhelm, 2005).

Para avaliar as estruturas de raciocínio, Wilhelm (2000 citado por Wilhelm, 2005) selecionou medidas desse construto fundamentadas em demandas cognitivas e conteúdos envolvidos. Doze medidas foram incluídas no estudo e as nomenclaturas D e I representam dedutivo e indutivo, e F, N e V representam, respectivamente, conteúdo figural, numérico e verbal. A combinação entre o raciocínio e conteúdo dos itens foi especificada em tarefas, como DF1-circuitos elétricos, DF2-relações espaciais, DN1-solução de equações, DN2-raciocínio aritmético, DV1-proposições, DF2-silogismos, IF1-classificação figural, IF2-matrizes, IN1-números em séries, IN2-descobertas de números que não se ajustam em uma série, IV1-analogias verbais e IV2-significado de palavras.

Estes testes e outras medidas de critério foram aplicados em 279 estudantes do terceiro colegial com idade média de 17,7 anos e desvio padrão de 1,2 anos. Os resultados demonstraram cinco modelos estatísticos de ajustes, por meio da análise fatorial confirmatória, para a explicação estrutural das habilidades de raciocínio. O modelo um representou um fator geral das habilidades de raciocínio. O modelo dois separa tarefas dedutivas e indutivas. O modelo três distingue os fatores de acordo com o conteúdo das tarefas, ou seja, verbal, quantitativo e figural-espacial, correlacionados entre si. O modelo quatro especificou um fator geral das habilidades de raciocínio e dois fatores subjacentes, um com tarefas verbais e outro com tarefas quantitativas. Por fim o quinto modelo se

referiu a um fator com tarefas de raciocínio indutivo e dedutivo e um segundo fator contendo o modelo quatro já descrito. Ressalta-se no modelo dois que os fatores encontrados de indutivo e dedutivo foram correlacionados, estimando exatamente 1, ou seja, correlação perfeita, indicando que os dois fatores mensuram o mesmo raciocínio.

A melhor solução encontrada foi a do modelo que postula os três conteúdos e em cada um deles submergem tarefas de raciocínio indutivo e dedutivo, que no caso foi o modelo três. Ressalta-se que as diferenças entre os modelos não foram substanciais, e que com outras medidas e outros participantes tais resultados podem se modificar. O resultado mais importante desse estudo refere-se à não possibilidade de distinção entre raciocínio indutivo e dedutivo. Essa unificação também foi verificada no estudo de Wilhelm (2000 citado por Wilhem, 2005) com análises de escalonamento multidimensional, análise fatorial exploratória e pelo procedimento da análise tetrádica de Spearman.

Com relação ao raciocínio pela perspectiva cognitiva, existem várias teorias que descrevem e explicam o processo de pensamento que nele ocorre. Nas pesquisas contemporâneas, as teorias são embasadas em domínio de processo duplo, ocorrendo dois processos de informação, sendo o primeiro um sistema associativo, heurístico, implícito, experimental e intuitivo, enquanto o segundo é baseado em regra, analítico, explícito e racional (Epstein, 1994; Evans, 1989; Hammond, 1996; Sloman, 1996; Stanovich, 1999, citados por Wilhelm, 2005).

A teoria cognitiva do modelo mental de Johnson-Laird (1983) procura explicar e descrever como as pessoas podem raciocinar frente as informações. Essa teoria tem sido aplicada tanto para raciocínio dedutivo quanto indutivo. De acordo com essa teoria, a dificuldade nas tarefas de raciocínio ocorre devido ao número de modelos mentais que podem ser compatíveis com as premissas de uma determinada situação.

Um exemplo que ilustra a dificuldade na tarefa é apresentado a seguir. Diante da premissa “A está à esquerda de B. B está à esquerda de C. C está à esquerda de D. D está à esquerda de E”, pode-se facilmente derivar o modelo mental “A B C D E”. Mas a premissa “A está à esquerda de B, B está à esquerda de C, C está à esquerda de E, D está à esquerda de E” pode derivar mais de um modelo, no caso “A B C D E” e “A B D C E”. No primeiro C está à esquerda de D e no segundo D está à esquerda de C. Portanto, essa segunda premissa, por derivar mais de um modelo, tende a ser mais difícil do que a primeira premissa, em que se deriva apenas um modelo.

Essa é uma teoria cognitiva do processo de pensamento, o que dificulta a construção de um teste psicométrico, visto que é difícil especificar a natureza e o número de modelos mentais. Então não é possível saber ao certo quantos e quais modelos cada indivíduo possui e, portanto, qual a dificuldade de uma tarefa.

De acordo com Johnson-Laird (2005), o indivíduo realmente pensa por modelos mentais. Por exemplo, percepções criam modelos do mundo que podem ou não ser reais, a compreensão de uma descoberta gera modelos do mundo que descrevem o que se fala, e

assim por diante. O pensamento é capaz de antecipar o mundo e escolhas de um curso de ações, confiando nas manipulações internas desses modelos mentais. Assim por exemplo, “Eu tinha um livro no restaurante do hotel e agora o perdi. Então, ou eu deixei no restaurante, ou ele caiu da minha bolsa no caminho de volta para minha sala, ou está em algum lugar aqui na sala. Ele não podia ter caído da minha bolsa, pois ela é profunda e eu caminhei devagar na volta para a sala. Então está aqui ou no restaurante” (p.185).

Esta seqüência é uma dedução lógica, em que as alternativas iniciais são A (estar no restaurante), B (cair da bolsa) ou C (estar na sala). Depois de eliminar a alternativa B, então a conclusão é A ou C. A conclusão é válida, pois as premissas são verdadeiras. Mas outros pensamentos podem ocorrer, mudando o modelo mental, como foi o caso do pensamento que o livro não podia ter caído da bolsa, pois ela é profunda, o que pode levar a erros nas conclusões. Nessa teoria, o raciocínio dedutivo e indutivo podem ser pensados pelos mesmos estágios de processamento de informação. Primeiro as premissas devem ser compreendidas, depois há uma descrição parcimoniosa dos modelos a serem construídos e, por fim, tais modelos são avaliados, mantidos, modificados ou rejeitados.

Estabelecer medidas de habilidades de raciocínio é difícil, pois a construção geralmente é dirigida pela aplicação de critérios psicométricos, ou seja, constróem-se itens, aplica-se numa amostra e verificam-se parâmetros, como por exemplo, de correlação, estrutura fatorial e a partir disso, conclui-se se o teste está bom ou não, sem muitas vezes analisar o sentido psicológico dos dados. Na verdade, a construção deveria ser

fundamentada por indicadores estritamente pautados em teorias, derivadas de um modelo cognitivo do processo de pensamento (Wilhelm, 2005). Teorias sobre processos de raciocínio em geral, mais especificamente a teoria do modelo mental tem sido largamente aplicada para explicar como o indivíduo pode raciocinar. No caso das tarefas de raciocínio psicométrico, estas têm sido pouco consideradas, pois apenas mensuram e não trazem informações que possibilitem verificar como ocorreu o raciocínio até chegar na resposta. Uma teoria geral do processo de raciocínio deveria fundamentar a construção do instrumento e ajudar a explicar as análises confirmatórias, ou seja, deve-se conciliar a teoria dos processos de raciocínio, provenientes da abordagem cognitiva, com as aplicações psicométricas, sendo que os itens de testes devem ser embasados em teorias do pensamento do processo de raciocínio (Wilhelm, 2005).

Killonen e Cristal (1990) resumizam que, desde Spearman (1923), o raciocínio tem sido definido como abstrato, um processo de alto nível e sem definição precisa. O desenvolvimento de bons testes de habilidade de raciocínio tem sido quase uma forma de arte, devido mais a processos de tentativa e erro empíricos do que ao delineamento sistemático de requisitos que o teste deva satisfazer. Apesar de evidências empíricas indicarem os melhores testes para se avaliar habilidade de raciocínio, ainda é muito limitado o conhecimento teórico para se saber qual teste é bom para qual raciocínio. Há, por exemplo, freqüentemente testes que pressupõem avaliar o mesmo tipo de habilidade de raciocínio, mas que são diferentes em termos de atributos e características.

Por meio de métodos baseados na Psicologia cognitiva pode-se compreender melhor as diferentes estratégias utilizadas pelos indivíduos para resolver um mesmo problema. No entanto, ainda não há muita informação sobre os processos envolvidos nas medidas de habilidades de raciocínio, tais como as usadas em pesquisas psicométricas, apesar do conhecimento sobre as estruturas e sobre a validade de predição na vida real. Os dois métodos, psicométrico e experimental-cognitivo, devem ser usados juntos, pois os benefícios serão mútuos. Diferenças nas correlações entre itens de raciocínio usados nas pesquisas cognitivas e as variáveis latentes do teste de habilidade podem revelar diferenças importantes entre as tarefas experimentais, bem como a variabilidade das diferenças dos itens nos testes psicométricos de raciocínio possivelmente podem ser explicados pela aplicação de várias teorias de processos de raciocínio como a teoria do modelo mental (Wilhelm, 2005).

Uma pesquisa que demonstra esse ganho mútuo quando se associam as abordagens cognitiva e psicométrica é a de Primi (2002a). O autor investigou quatro fatores de complexidade dos itens, baseado na teoria cognitiva do processamento de informação, para a resolução de problemas de inteligência fluida-Gf. Para isso criou uma estrutura de itens que representasse uma combinação ortogonal desses quatro fatores, de forma que seus impactos na complexidade do item e no tempo de reação fossem identificados. Além disso, verificou a relação das complexidades dos itens com a Gf, mais especificamente com memória de trabalho. Participaram do estudo 313 estudantes de graduação, ambos os sexos e com idade variando entre 17 e 22 anos. Foram construídos 64 itens, divididos em duas

formas, A e B, com 32 itens cada, garantindo que fossem estruturalmente idênticas, sendo aplicado em amostras independentes. Cada item consistiu numa matriz 3 x 3 com oito alternativas de respostas e foram baseados em figuras geométricas com a estrutura definida por número de elementos (quantidade de objetos no item), número de regras (quantas transformações ocorrem no item), tipo de regra (qual tipo de transformação) e organização perceptual (fácil ou difícil de visualizar). O teste possuía apenas uma alternativa correta e também apresentava três alternativas para treino.

Os resultados sobre as propriedades psicométricas das formas A e B do teste mostraram coeficientes de consistência interna alta, respectivamente 0,84 e 0,85. Com relação à dificuldade dos itens pela Teoria de Resposta ao Item por meio do modelo de Rasch, houve variação de 0,89 para os fáceis, a 0,05 para os difíceis, estando dentro dos parâmetros aceitáveis, segundo os autores. Tais itens formaram uma amostra representativa do espectro da complexidade das tarefas. As correlações entre item e o escore total apresentaram-se de moderadas para altas. Ressalta-se que a forma B foi sutilmente mais fácil que a forma A, no entanto as propriedades psicométricas das duas formas mostraram-se equivalentes. A fim de verificar a contribuição de cada fator de complexidade do item para prever a sua dificuldade, foi realizada uma análise de regressão múltipla por passos e constatou-se que somente o fator organização perceptual contribuiu para prever significativamente a dificuldade. Numa segunda análise foi usado o mesmo método de análise de regressão, no entanto excluindo itens espaciais e formando uma variável denominada quantidade de informação, na qual se somaram os fatores número de

elementos e número de regras. Foi observado um maior efeito da organização perceptual e um aumento significativo da predição de itens de complexidade pela variável quantidade de informação.

Investigando o tempo de reação, esta variável mostrou-se estatisticamente significativa, interagindo com todos os fatores de complexidade dos itens, sendo que o efeito mais importante ocorreu com a organização perceptual. Outro resultado interessante encontrado na pesquisa diz respeito à estratégia de eliminação de respostas feitas pelos indivíduos, informação que é possível de averiguar na aplicação computadorizada. Aqueles que usam freqüentemente a estratégia de eliminar respostas tendem a apresentar altos escores no teste, já os que não a utilizam demonstram maior variabilidade nos escores do teste. As tarefas com matriz geométrica de itens indutivos constituem ferramenta para avaliar a Gf e os dois testes, forma A e B, construídos neste estudo, por meio de suas propriedades psicométricas, sugeriram que são boas medidas para avaliar esse tipo de inteligência (Primi, 2002a).

A identificação das origens das dificuldades de itens que auxiliam compreensão da Gf foi um dos pontos mais importantes do trabalho, sendo que as duas variáveis que mais contribuíram para o aumento da complexidade dos itens foram a quantidade de informações e a organização perceptual, esta última apresentando um efeito maior. Como a organização perceptual está diretamente ligada ao processo de abstração, essa capacidade de abstrair é importante para a Gf. Itens mais complexos de organização perceptual requerem o



processamento de controle de atenção de atributos ou elementos seletivos, tendo em vista a associação deste processo de codificação seletiva e o processo de abstração, sendo que o processamento visual também é exigido. Todos esses processos fazem parte da memória de trabalho e, conseqüentemente, são importantes para Gf, já que esta forma de inteligência é bastante relacionada com este componente (Primi, 2002a).

Pesquisas para a compreensão do raciocínio devem se pautar em teorias dos processos a ele subjacentes e usar métodos confirmatórios, investigando o significado e a estrutura dos modelos. O mais importante é buscar a compreensão do processo de pensamento e das medidas de habilidades do raciocínio. Existem vários métodos para o desenvolvimento e treinamento de habilidade de raciocínio e com resultados encorajadores, como é o caso das pesquisas de Klauer (1990) e Klauer e colaboradores (2002). Embora seja difícil discriminar entre raciocínio indutivo e dedutivo psicometricamente, é possível que treinamentos apropriados fornecem ganhos diferenciais para ambos. O processo cognitivo das tarefas de raciocínio indutivo e dedutivo pode ser diferente, mas a mensuração pode ser observada em medidas que não são adequadas, ou seja, que não estejam medindo corretamente um ou outro. Isso não exclui a opção de que ambos os processos de pensamentos podem ser afetados por diferentes intervenções (Wilhelm, 2005)

Concordando com esse posicionamento de aproximar a abordagem psicométrica e a cognitiva para a construção mais adequada de instrumentos de avaliação do raciocínio, ou qualquer outro construto, a presente pesquisa tem como objetivo principal construir um

teste dinâmico de raciocínio indutivo para crianças de 7 a 11 anos. Para tanto, se fundamentará na Teoria Prescritiva do Raciocínio Indutivo de Klauer (1990), já mencionada. A abordagem psicométrica contribuirá para investigar possíveis estruturas desses itens construídos previamente de acordo com uma teoria cognitiva, bem como ajudará a investigar as propriedades psicométricas desse teste e a sua relação com outras variáveis, especificamente o desempenho acadêmico e as funções executivas.

Assim, a priori, tanto os itens do teste quanto da intervenção serão embasados na teoria de Klauer. Esses fundamentos teóricos permitirão a construção de um teste específico para raciocínio indutivo, minimizando a interferência de outro tipo de raciocínio. A fundamentação teórica referida será o assunto da próxima seção.

### **1.3 - Teoria Prescritiva e Treinamento do Raciocínio Indutivo**

Segundo Almeida e Morais (1997), a evolução da inteligência, na perspectiva cognitivista, proporcionou o surgimento de programas de treino cognitivo que abrangem as funções cognitivas, resolução de problemas ou ensino de alguma tarefa. Na Europa já existem vários treinamentos, inclusive o proposto pelos mesmos autores denominado de “Promoção Cognitiva”, desenvolvido para alunos do sétimo ao nono ano escolar em Portugal. De acordo com esses autores, os programas de treino cognitivo podem ser divididos em três principais tipos. O primeiro é da perspectiva piagetiana ou pós-piagetiana, que visa à modificação das estruturas ou esquemas operatórios por meio de confrontos cognitivos entre indivíduo e objeto, ou sócio-cognitivo entre indivíduo, objeto e ambiente. A segunda perspectiva envolve os treinamentos de processos cognitivos que são fundamentados na “Teoria do processamento de informação”. A terceira perspectiva é a heurística, que são treinos de estratégias mais globais de resolução de problemas.

O programa “Promoção Cognitiva” aproveita elementos das três perspectivas, no entanto é mais fundamentado na abordagem do processamento da informação. Almeida e Morais (1997) descrevem que é possível ocorrer a modificabilidade da inteligência, operacionalizando-a em diversos processos cognitivos, fazendo com que o indivíduo aprenda a pensar e aprenda a aprender. O objetivo global do programa é treinar processos cognitivos requeridos na situação de resolução de problemas, sendo que um dos objetivos

específicos é contextualizar os processos das tarefas para transferências ao cotidiano como, por exemplo, escolares, interpessoais e lúdicas.

Outro treinamento bastante reconhecido é o *Bright Start*-Currículo cognitivo para jovens crianças de Haywood, Brooks e Burns (1992). Este programa tem o objetivo de realçar o desenvolvimento social, emocional e cognitivo de crianças com idade entre três e seis anos, sendo elaborado para ensinar os processos básicos do pensamento e resolução de problemas. Quatro teorias fundamentam esse programa. Uma delas é a transacional, sobre a natureza e o desenvolvimento da inteligência, destacando que a aprendizagem também depende da capacidade inata, de origem genética e dos processos da cognição. A outra é a teoria de Jean Piaget, que os processos cognitivos se desenvolvem seqüencialmente em processos mais maduros. Há também a de Vygotsky, referindo-se ao contexto social e à zona de desenvolvimento proximal, a qual ressalta a importância do meio social para o desenvolvimento cognitivo e o modelo de um adulto como intermediador da aprendizagem. Por fim há a teoria da Modificabilidade cognitiva estrutural de Feuerstein, que propõe uma mudança cognitiva e também a aprendizagem mediatizada.

O presente estudo não pretende desenvolver um programa de treinamento de habilidades cognitivas, mas sim um teste dinâmico de raciocínio indutivo que abarcará de forma mais restrita um treinamento, e que por isso foi denominado de intervenção. Portanto optou-se por fundamentar tanto os itens do teste quanto a intervenção na Teoria Prescritiva do Treinamento de Raciocínio Indutivo para crianças de Josef Klauer (1990), que também é

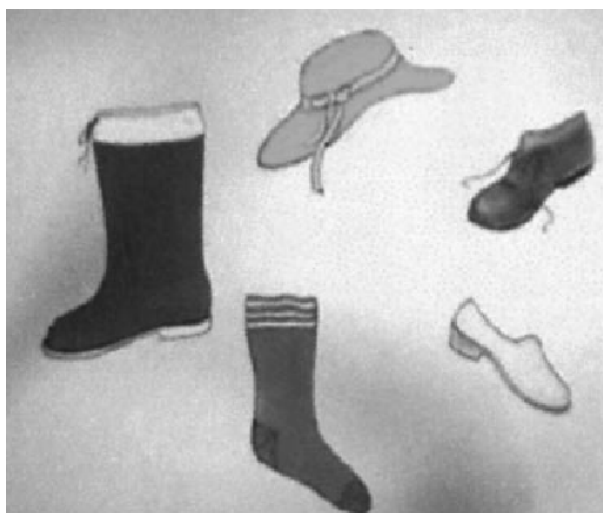
um programa de treinamento cognitivo, como os dois citados anteriormente. Este autor traz uma concepção mais clara sobre os processos cognitivos envolvidos no raciocínio indutivo e, a partir disso, propõe um treinamento voltado para a aprendizagem de como raciocinar indutivamente. O programa de treinamento é o ponto central dessa teoria, o qual é orientado para o processo de pensamento mais do que para o produto do pensamento. A transferência desse procedimento de raciocinar para outros contextos é fundamental, sendo que a base para isso são as estratégias procedurais, em que leva-se em consideração comparações sistemáticas e analíticas para a descoberta de regularidades e irregularidades.

A teoria proposta por Klauer (1990) é realmente prescritiva, pois não são descritas todas as possibilidades procedurais de como o indivíduo raciocina, apenas é exposto um procedimento de como pode ser resolvido um problema usando pensamento indutivo, mesmo porque os seres humanos podem utilizar estratégias diferenciadas para a resolução de um mesmo problema. Essa teoria é testada quando o indivíduo se depara com novos problemas, mas que podem ser solucionados utilizando o raciocínio indutivo, tendo a oportunidade de demonstrar habilidades de transferir estratégias treinadas previamente para novas tarefas de resolução de problemas requerendo raciocínio indutivo (Klauer & Phye, 1995). Para uma melhor compreensão dessa teoria, inicialmente será abordado o conceito do raciocínio indutivo segundo essa teoria e, depois, o treinamento desta capacidade, também pela mesma concepção.

Como anteriormente descrito, o raciocínio indutivo é um resultado geral de observações individuais, ou seja, regularidades que se percebe em uma determinada situação. Por exemplo, observando apenas cisnes brancos, indutivamente poderia se concluir que todos os cisnes são brancos, já que o fato de todos serem brancos é uma regularidade. Ou seja, no raciocínio indutivo partimos de idéias individuais que levam a conclusões gerais, porém, tais idéias podem ou não serem corretas. No caso dos cisnes a idéia é incorreta, pois apesar de observar somente cisnes brancos, também existem de outras cores. No entanto, em outras situações o processo do raciocínio indutivo pode levar a conclusões corretas (Klauer & Phye, 1995). “A regularidade representa um papel importante do pensamento, pois as regularidades e uniformidades fornecem a base dos conceitos e categorias que servem como conhecimento básico para pensamento abstrato e raciocínio” (Klauer & Phye, 1995, p. 37). O raciocínio é indutivo ao se detectar regularidades por similaridades e diferenças de atributos ou por relações com conteúdos acadêmicos, que são linguagem, figuras, pinturas, números, etc. Esta definição distingue os pensamentos do raciocínio indutivo de outro tipos de raciocínio, sendo uma definição fundamentada em uma teoria que especifica os processos cognitivos que constituem o raciocínio indutivo (Klauer & Phye, 1995).

O raciocínio indutivo pode ser explicado por seis paradigmas relacionados, quais sejam generalização, discriminação, classificação cruzada, reconhecimento de relações, diferenciação de relações e construção de sistemas (Klauer & Phye, 1995). Cada um desses paradigmas será explicado a seguir.

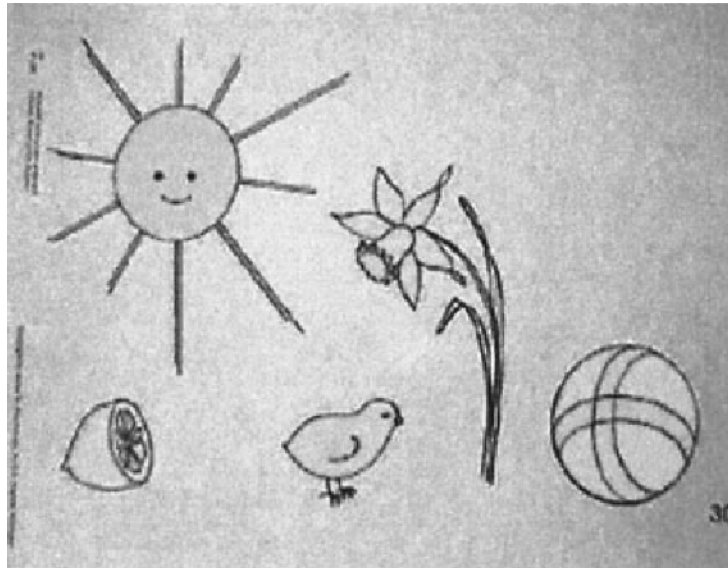
Por generalização (GE) entende-se a necessidade de estabelecer similaridades de atributos para diferentes objetos formando um grupo, pois apesar de haver diferenciações, partes desses objetos são comuns. Problemas envolvendo generalização podem ocorrer em três formas, a saber, construção de classe, complemento de classe e encontro de similaridades. Para a resolução desses problemas a pessoa procede analiticamente, comparando as propriedades dos objetos, procurando descobrir algo em comum entre eles. O sujeito pode fazer isso sistematicamente, ou seja, ir comparando um a um, da esquerda para a direita, por exemplo. Durante a análise o indivíduo utiliza estratégias de monitoramento, checando quais atributos dos itens excluídos não possuem, bem como aprendendo a distinguir esse monitoramento da solução do problema. A Figura 5 apresenta um exemplo de item de generalização de atributos, em que o indivíduo precisa descobrir quais são os três objetos que apresentam atributos iguais. Nesse caso ele precisa identificar que há três calçados, agrupando-os.



**Figura 5-** Exemplo de item com o paradigma de generalização de atributos (Klauer, Willmes & Phye, 2002).

A discriminação (DI) é o processo de averiguar diferenças entre objetos a respeito de atributos. Somente um tipo de problema exemplifica esse paradigma, ou seja, identificar o objeto que não pertence a um grupo. Nesse tipo de item também verifica-se o que há em comum entre os objetos, para conseqüentemente identificar a diferença. O contraste desse tipo de item com o de generalização, anteriormente apresentado, é que aqui apenas um objeto é incomum aos outros. Novamente para encontrar a resposta correta se comparam atributos, que é a base para testar as hipóteses. O processo de monitoramento também está implícito. A Figura 6 demonstra um item de diferenciação de atributos, em que a pessoa precisa descobrir qual figura não se ajusta com as outras. Nessa tarefa da Figura 6 o indivíduo precisa discriminar que somente a bola é um objeto inanimado.

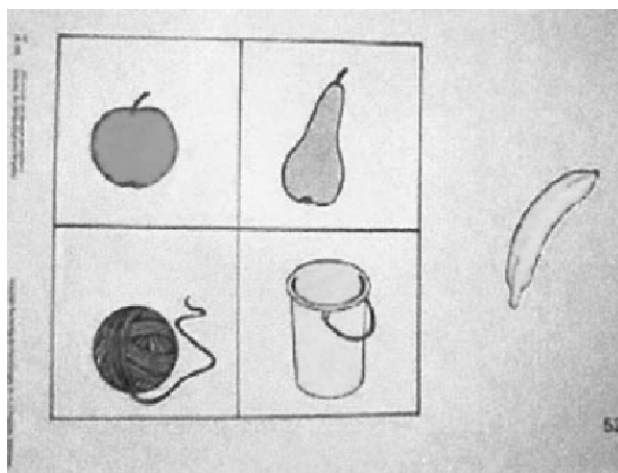




**Figura 6-** Exemplo de item com o paradigma Diferenciação de atributos (Klauer & colaboradores, 2002).

Com relação à classificação cruzada (CC), esse tipo de tarefa é organizada em esquemas de atributos em matrizes de quatro partes, ou quatro objetos, dois em cima e dois embaixo, onde pelo menos dois atributos são considerados ao mesmo tempo. Por exemplo, o objeto do primeiro quadrante precisa ter no mínimo um atributo igual ao objeto do segundo quadrante, ao seu lado, e outro atributo igual ao terceiro quadrante, que está abaixo dele. Isso deve ocorrer para cada objeto de um quadrante. A classificação cruzada requer a determinação de atributos diferentes e comuns, e a vantagem reside que todas as possibilidades podem ocorrer, ou seja, similaridade ou diferença em ambas as características, ou a mistura de similaridade e diferença.

Nesse tipo de item a pessoa também compara os objetos, mas cruza as informações entre todos os objetos para saber, por exemplo, onde uma determinada figura melhor se encaixa. As estratégias de monitoramento também são importantes. O tipo de item utilizado para esse paradigma é apresentado com quatro objetos em matriz de quatro, sendo um em cada quadrante, e todos apresentando atributos comuns com outros dois, como já especificado. Então mostra-se outro objeto e pergunta-se em qual quadrante ele se encaixaria melhor, levando o indivíduo a analisar essa figura exterior e seus atributos, comparando-a com cada figura do quadrante, até encontrar o quadrante que ela se encaixaria. Um exemplo de item de classificação cruzada pode ser observado na Figura 7. A pessoa precisa descobrir em qual item dos quadrantes a banana se encaixa. Aqui nessa ilustração a banana se enquadra no quadrante preenchido com a pêra, pois a banana é uma fruta como a laranja e apresenta formato comprido como o balde.

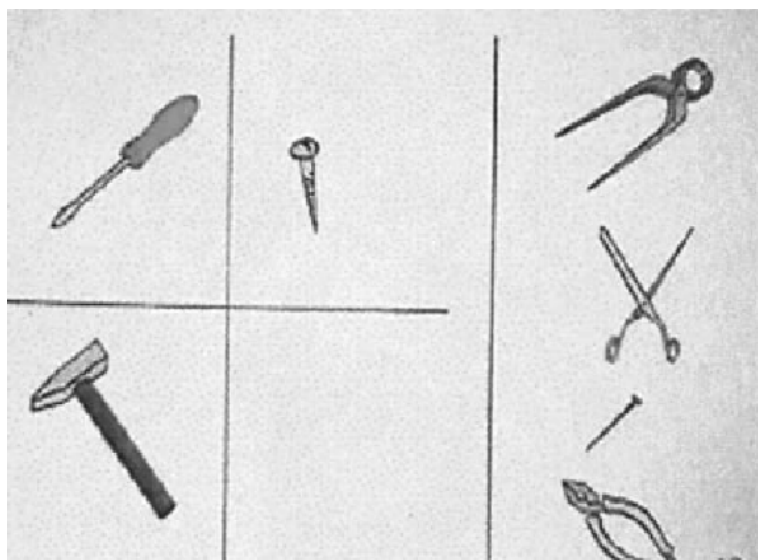


**Figura 7-** Exemplo de item com o paradigma Classificação cruzada (Klauer & colaboradores, 2002).

Já o reconhecimento de relações (RR) é possível quando pelo menos dois objetos estão presentes e a inferência de relação decorre de um procedimento de comparação entre pares mas, nesse caso, é preciso encontrar a similaridade entre relações e não mais atributos, o que torna a resolução do item mais complexa. Essas relações podem ser de acordo com a função de objetos, ou causa e efeito, ou algo único de um conjunto de problemas. Esses itens envolvem mais características abstratas, o que demanda maior capacidade de memória de trabalho, ou seja, maior capacidade de armazenar e trabalhar com as informações. Reconhecimento de relações ocorre em três formas de problemas, quais sejam, complemento de séries, arranjo de séries e analogias.

Por exemplo, em problemas com analogia é preciso determinar a relação específica entre um par de objetos que será o padrão de comparação. Várias relações podem ser

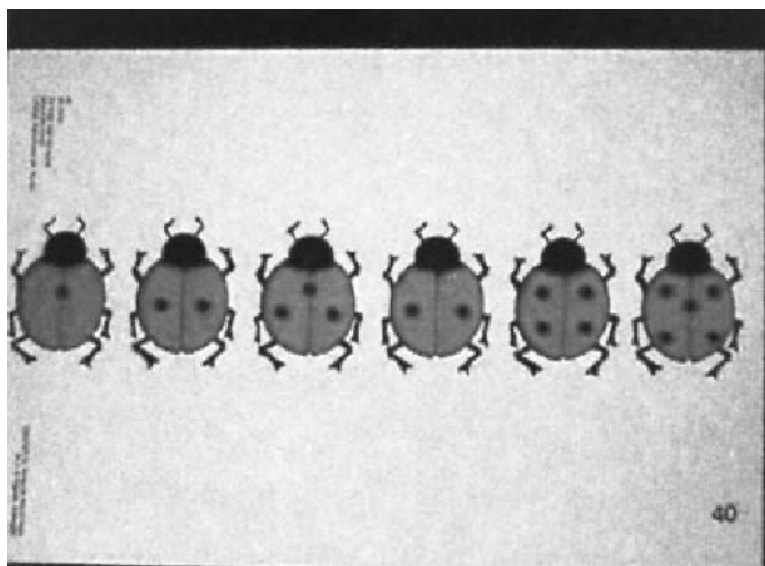
possíveis, então a estratégia de solução consiste em mapear a relação de um par de objetos e estabelecer, em um par incompleto, este mesmo tipo de relação. Nesse processo de mapear uma relação para aplicar em outra situação, o monitoramento metacognitivo para verificar se a escolha é a correta sempre deve estar presente. A Figura 8 mostra um exemplo de Reconhecimento de relações por analogia. A pessoa precisa descobrir qual dos objetos a direita completa o espaço vazio. Nesse caso, o objeto a ser escolhido pelo indivíduo deve ser o prego, pois assim como a chave de fenda enrosca o parafuso, o martelo prega o prego. Então chave de fenda está para parafuso, assim como martelo está para prego.



**Figura 8-** Exemplo de item com o paradigma Reconhecimento de relações (Klauer & colaboradores, 2002).

O paradigma diferenciação de relações (DR) requer o reconhecimento de diferenças em relações. Este tipo de item pode ser exemplificado somente por um tipo de problema, que são os distúrbios de séries, mas pode ocorrer em duas variações. No primeiro é necessário apenas reordenar os membros de um conjunto de problema definindo as séries corretas, já no segundo um objeto deve ser excluído. Nas duas variações de problema, a estratégia é encontrar a relação que ocorre entre os possíveis itens reconhecendo os objetos que deturpam a série. O monitoramento cognitivo para checar a solução correta é avaliar todas as similaridades de relações entre objetos que permaneceram ou entre os que foram reordenados. Pela Figura 9 pode-se visualizar um item de discriminação de relações e que

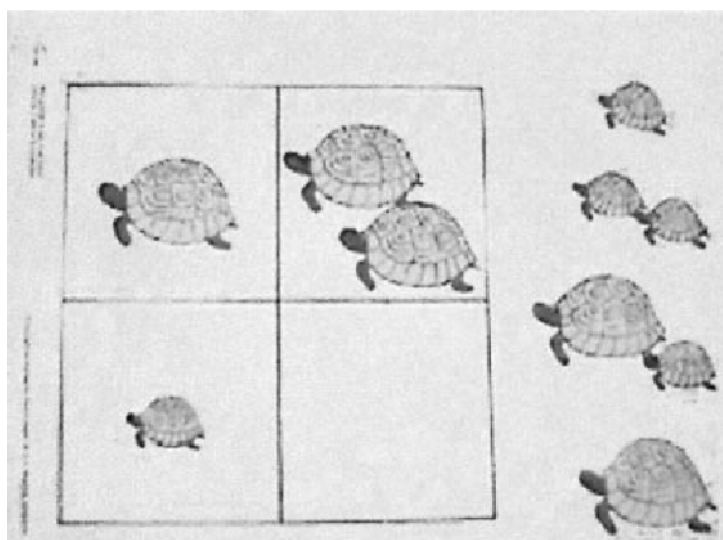
neste caso o indivíduo precisa excluir um objeto que está perturbando a série. No exemplo da Figura 9 é preciso identificar que a quarta joaninha está incorreta, não seguindo a seqüência da quantidade de bolinhas no corpo.



**Figura 9-** Exemplo de item com o paradigma Discriminação de relações (Klauer & colaboradores, 2002).

Por fim o último paradigma é a construção de sistemas (SC). Envolve um raciocínio semelhante ao usado para resolver problemas de classificação cruzada, mas no caso do SC existem pelo menos duas relações em que similaridade e diferenciação podem ser verificadas. Dois tipos de problemas representam esse paradigma, matrizes simples, já descrita anteriormente, e estendida. Esta última inclui mais do que quatro objetos em um conjunto de problemas e pelo menos um par de objetos tem relação comum com outro par de objetos, ao mesmo tempo em que tem diferenciação de relação com outro par de objetos.

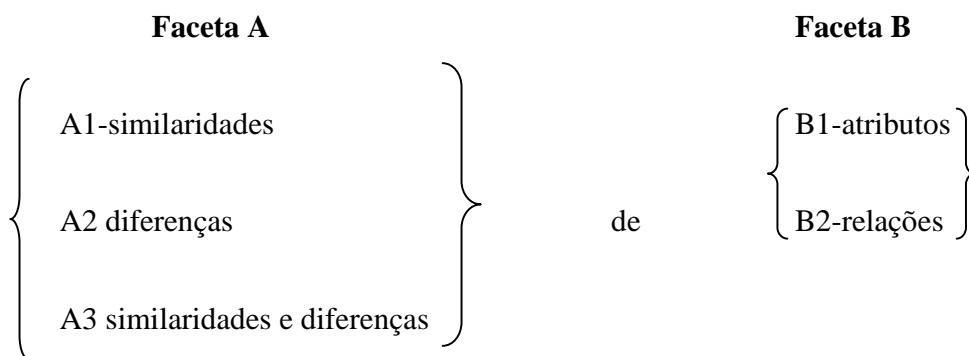
Para a solução desses tipos de item é preciso reconhecer onde cada relação é operacionalizada e onde existe similaridade e diferença. A estratégia metacognitiva nesses casos é uma checagem sistemática para similaridade de relações entre linhas e colunas, sempre iniciando na célula esquerda acima de uma matriz. A Figura 10 apresenta um item de sistema de construção simples para a pessoa descobrir qual a figura da direita que se encaixa melhor no quadrado vazio. Para resolver corretamente a tarefa da Figura 10 o indivíduo deve escolher as duas tartaruginhas pequenas das alternativas. Essa escolha é condizente com o que ocorre na linha de cima ao mesmo tempo em que difere em função do tamanho, ou seja, de uma tartaruga se passa a ser duas, porém menor no tamanho para ser semelhante a tartaruga da mesma linha.



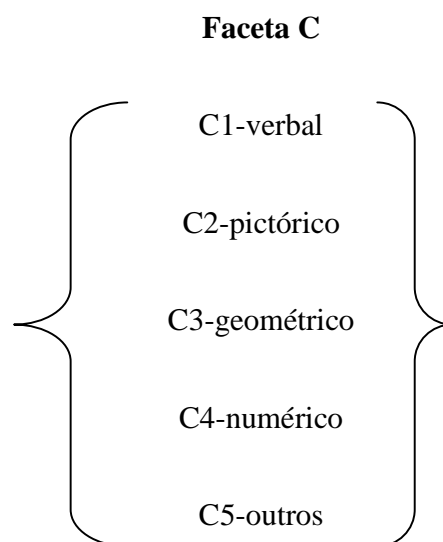
**Figura 10-** Exemplo de item com o paradigma Sistema de construção (Klauer & colaboradores, 2002).

Esses seis paradigmas que explicam o raciocínio indutivo encontram-se descritos em um modelo mais amplo, baseado na definição de que raciocínio indutivo consiste em detectar regularidades e irregularidades por meio da descoberta de três facetas denominadas de A, B e C, com 3, 2 e 5 elementos respectivamente. As facetas A e B são centrais e mostram apenas seis tipos de variações com comparações abstratas e analíticas. Na A são descobertas de similaridades e diferenças, e as regularidades são encontradas quando a atenção está voltada tanto para similaridades quanto para diferenças. A faceta B mostra que raciocínio indutivo não é uma comparação global e sim comparações de atributos para atributos ou relações para relações, e esse tipo de comparação se faz analiticamente, onde ocorre o processo de abstração considerando atributos individuais ou relações. Por fim, na C são especificados cinco tipos de conteúdos que podem ser construídos para avaliar o raciocínio indutivo (Klauer & Phye, 1995). A Figura 11 mostra as três facetas.





Relacionados com os objetos da Faceta C



**Figura 11-** Facetas do modelo do raciocínio indutivo (adaptado de Klauer & colaboradores, 2002).

Como o treinamento é para raciocínio indutivo, as tarefas precisam ser compatíveis para avaliar tal capacidade. No treinamento, apenas tarefas que realmente necessitam de raciocínio indutivo para resolvê-las é que são utilizadas. Ou seja, são tarefas em que o indivíduo precisa descobrir regularidades pelo processo de comparação de atributos ou relações de objetos ou eventos. A Tabela 1 mostra seis processos de raciocínio e as respectivas operações cognitivas envolvidas com raciocínio indutivo, bem como o formato do problema. Pode-se perceber que sempre está empregado um processo cognitivo (generalização ou discriminação) que são operações de pensamentos usados durante o raciocínio. *Generalização* é operacionalmente definida como um processo de reconhecimento de similaridade de atributos de objetos ou eventos e o mesmo para relações. *Discriminação* é definida como o processamento de diferenças entre objetos ou eventos baseado na habilidade de detectar diferenças entre atributos, o mesmo para relação (Klauer & Phye, 1995).

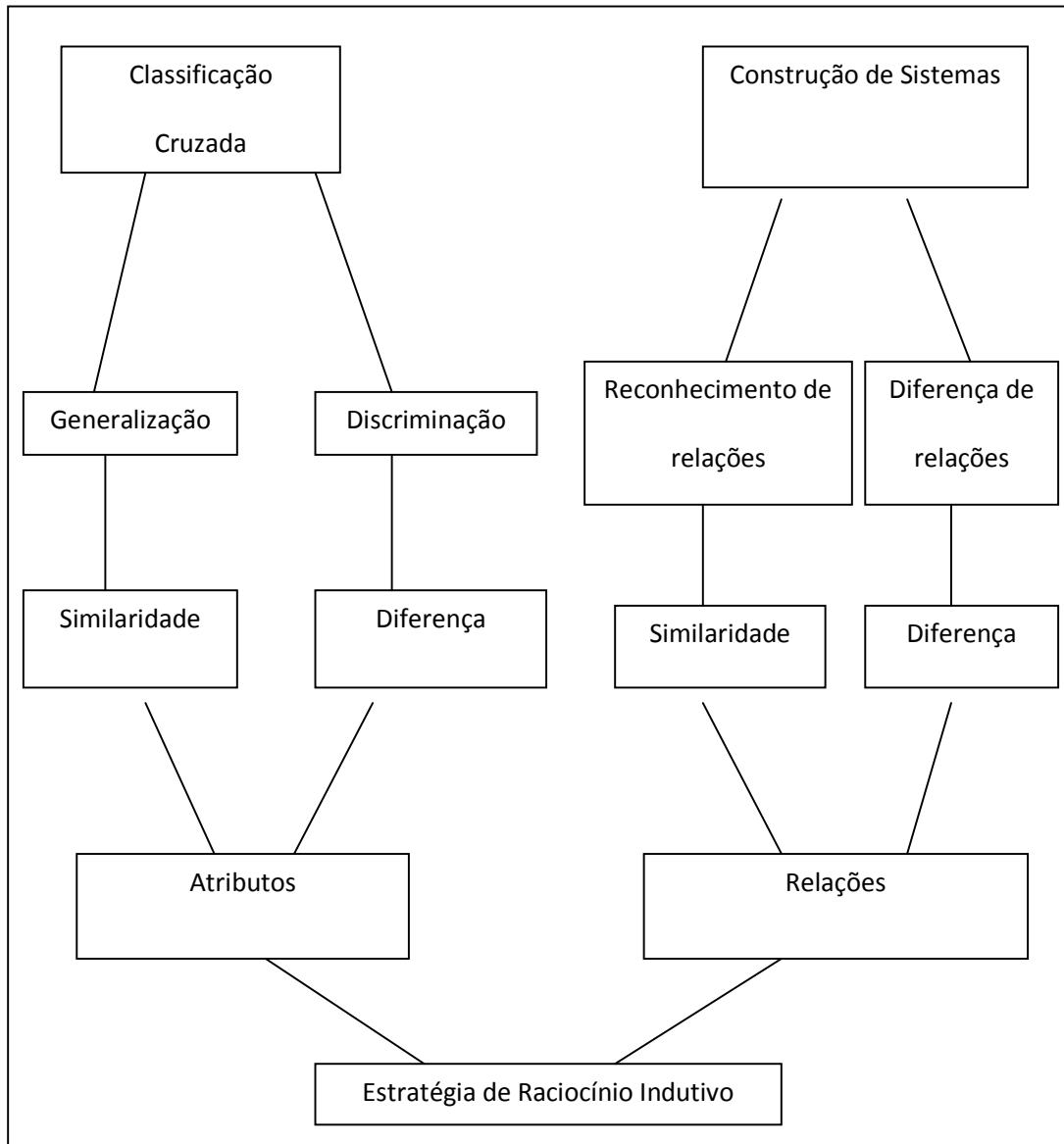
**Tabela 1-** Processos do pensamento indutivo com os formatos de problemas (adaptado de Klauer & colaboradores, 2002).

<b>Processo</b>	<b>Identificação da faceta</b>	<b>Formato dos problemas</b>	<b>Operação cognitiva requerida</b>
<b>Generalização</b>	a1b1	Formação de classe Expansão de classe Encontrar atributos comuns	Similaridade de atributos
<b>Discriminação</b>	a2b1	Identificar irregularidades	Discriminação de atributos (diferenciação de conceito)
<b>Classificação Cruzada</b>	a3b1	4 quadrículos 6 quadrículos 9 quadrículos	Similaridades e diferenças em atributos
<b>Reconhecimento de Relações</b>	a1b2	Complemento de séries Ordenação de séries analógicas	Similaridades de relações
<b>Diferenciação de relações</b>	a2b2	Distúrbio de séries	Diferenças em relações
<b>Construção de Sistemas</b>	a2b2	Matrizes	Similaridades e diferenças em relações

Esses mesmos autores descrevem que os dois processos cognitivos, generalização e discriminação, ocorrem de forma paralela, ou seja, ao ensinar ou ao aplicar um desses

processos, o outro tende a ocorrer, pois no momento que verificamos uma semelhança, também temos que se atentar ao que é diferente, para descobrir o que é comum entre dois ou mais objetos. Diante disso, para esse presente estudo, optou-se por juntar os paradigmas Generalização e Discriminação e também Reconhecimento de Relações e Diferenciação de Relações. Então, ao invés de construir itens referentes aos seis paradigmas, serão construídos quatro, quais sejam Generalização/Discriminação, Classificação Cruzada, Reconhecimento/Diferenciação de Relações e Construção de Sistemas. Ressalta-se que a complexidade dos itens ocorre entre os paradigmas, sendo gradual do primeiro até o sexto, Generalização/Diferenciação, Classificação Cruzada, Reconhecimento/Diferenciação de Relações e Construção de Sistemas. Os menos complexos são Generalização/Diferenciação e o mais difícil Construção de Sistemas.

Já a Figura 12 mostra uma genealogia das tarefas no raciocínio indutivo. Observa-se que a base da figura é a estratégia de raciocínio indutivo, uma importante habilidade metacognitiva quando se treina esse raciocínio. A questão é que toda tarefa pode ser resolvida, primeiramente considerando similaridades e diferenças de atributos ou relações. Este é um ponto teórico importante, pois é uma ferramenta ensinada para a criança usar para solucionar problemas de raciocínio indutivo. O treinamento paradigmático primeiro ensina a base para aprender processos básicos de raciocínio indutivo e depois os sujeitos transferem essa aprendizagem aplicando em outros contextos. É um procedimento de comparação pareada e somente assume que o problema foi resolvido ao se reconhecer todos os atributos ou relações quando o conteúdo é apresentado para estudar.



**Figura 12-** Genealogia das tarefas em raciocínio indutivo (adaptado de Klauer & colaboradores, 2002).

A resolução de problemas raramente procede de acordo com essa estratégia mais analítica, pois na maioria das vezes a criança primeiro formula uma hipótese sobre a solução correta e a regra será feita com uma rápida inspeção global dos objetos para testar a hipótese (Kirby & Lawson, 1983, citados por Klauer & Phye, 2001). Caso a hipótese seja rejeitada, então são feitas novas comparações sistemáticas, em uma estratégia mais geral. No entanto sugere-se que a criança seja treinada em ambas, heurística geral e analítica, sendo que qualquer uma das estratégias envolve o controle das soluções de problemas por meio de similaridades e diferenças. Essas estratégias também podem ser consideradas procedurais, ou seja, o que o raciocínio requer para o indivíduo conseguir desempenhar bem a resolução de problemas indutivos.

Com relação ao programa de treinamento da Teoria de Klauer (1990), foi elaborado para promover o pensamento indutivo em crianças de 5 a 6 anos, no entanto, pode ser utilizado até o ensino médio, desde que os itens sejam mais complexos. O objetivo principal é o desenvolvimento da competência de raciocínio indutivo como uma ferramenta resolução de problemas. A fundamentação do programa é a Teoria de aptidão de Snow, que concebe a aptidão como potencial de aprendizagem que é desenvolvido pela prática, implicando mudança no conhecimento da criança. A teoria do processamento de informação da abordagem cognitiva também embasa o treinamento (Klauer & Phye, 1995).

A base do programa é que a aptidão de generalização, discriminação e monitoramento para verificar similaridades e diferenças são propedêuticas para o raciocínio

indutivo e resolução de problemas. O treinamento é paradigmático, pois utiliza-se de procedimentos gerais da estrutura do raciocínio que são aplicados em diferentes domínios de conteúdo. Ao se utilizar tais procedimentos para resolver diferentes classes ou domínios, pode dizer que houve a transferência paradigmática. E o treinamento é justamente para aperfeiçoar essa transferência (Klauer & Phye, 1995).

Como já mencionado, o treinamento promove uma mudança no conhecimento da criança em solução de problemas e o raciocínio indutivo é a ferramenta para isso. Tal modificação ocorre em três fases, primeiro o conhecimento declarativo, segundo o procedural e depois o estratégico (Klauer & Phye, 1995). No desenvolvimento do conhecimento declarativo a criança aprende sobre os tipos básicos de estruturas e como reconhecer problemas usando tal estrutura. É um conhecimento conceitual sobre objetos, atributos e ou relações, similaridades e diferenças que são elaborados e armazenados na memória de longo prazo. Como exemplo, saber que cada objeto apresenta atributos comuns com outros, ou que se pode comparar objetos por meio de uma relação estabelecida entre eles. O papel do treinador é sempre assistir a criança no desenvolvimento desse conhecimento e não avaliá-la.

No conhecimento procedural desenvolve-se procedimentos conscientes que são requeridos para usar o conhecimento declarativo na forma de raciocínio indutivo como ferramenta. O foco central dessa fase é o desenvolvimento de uma estrutura de procedimentos indutivos para a solução dos problemas, o que inclui estratégia apropriada e

controle, ou seja, monitoramento metacognitivo. Nessa fase a criança aprende o que e como fazer para resolver a tarefa, sempre utilizando análises de similaridades e diferenças sobre o problema indutivo apresentado. Mediante isso, deve-se tentar uma estratégia.

Por fim, a fase do conhecimento estratégico é a final, e envolve reconhecimento e uso espontâneo da estrutura de procedimento do raciocínio indutivo aprendido nas duas primeiras fases. Esse é o momento que a criança aplica o conhecimento desenvolvido em outras situações, como por exemplo, o reconhecimento das estruturas de processamento básico de problemas em diversos contextos, ou aplicar automaticamente uma estratégia de solução para resolver algum problema do dia-a-dia. Nesse estágio, se verificada a permanência da mudança cognitiva, porém para atingir essa fase existem alguns fatores que interferem para a realização dessa aprendizagem.

Primeiro é preciso a retenção das informações na memória de longo prazo, sendo necessário repetições das aplicações dos procedimentos para a resolução de um problema. Segundo, a transferência da aprendizagem também precisa ser praticada e planejada. Terceiro, a criança ter bem definido as características dos seis paradigmas do raciocínio indutivo, GE, DI, CC, RR, DR e SC. Com isso ela poderá aplicá-los em diferentes problemas de forma adequada. E quarto, acessar com facilidade os processos de raciocínio indutivo, de forma mais automática, o que demonstra a competência desenvolvida para utilizar o raciocínio indutivo para resolver problemas que o requeiram (Klauer & Phye, 1995). Ressalta-se que no teste construído no presente trabalho, a fase do conhecimento



estratégico é bem limitada, pois para essa etapa se pressupõem que já ocorreu uma mudança cognitiva, mas o teste dinâmico se limita a mensurar o potencial de aprendizagem, sem modificações na estrutura do pensamento.

Para o treinamento existem quatro principais métodos para interagir com a criança, auto-instrução verbal, descoberta guiada e verbalização e auto reflexão, mas em qualquer um deles, as crianças costumam preferir problemas dentro de contextos curtos. A auto-instrução verbal permite maior assistência do treinador, pois ele verbalmente permeia a solução do problema, fazendo comentários e ensinando os passos para se chegar à resolução. Caso a criança erre, o treinador demonstra como fazer corretamente, sendo um modelo para que o indivíduo também consiga ir pensando ou falando para si mesmo todos os processos envolvidos para chegar a resposta correta. Na descoberta guiada, o treinador oferece à criança oportunidades para que ela mesma descubra sozinha os diferentes tipos de problemas, características do problema, estratégias e monitoramento. O treinador intervém de forma mais amena, ou seja, emitindo sugestões ou pontuando alguma coisa. Os dois últimos métodos, verbalização e auto-reflexão são os que menos permitem ajuda. Na verbalização pede-se para a criança “pensar em voz alta” enquanto tenta resolver um problema, e na auto-reflexão pede-se para que o indivíduo faça comentários sobre os procedimentos que está utilizando. Certamente quando o treinamento é feito com crianças o primeiro método é o mais adequado (Klauer & Phye, 1995).

Algumas questões e procedimentos pedagógicos são empregados no desenvolvimento do programa de treinamento, quais sejam, paradigmas representam estruturas básicas do raciocínio, principalmente conceitos e resoluções de problemas; os paradigmas são aplicados em um grande número de conteúdos; cada paradigma será empregado com um número de problemas cuja solução pode ser aprendida na base de exemplos prévios; preferência por paradigmas que facilitem a transferência para várias áreas de conteúdo e paradigmas relacionados permitem um maior campo de aprendizagem (Klauer & Phye, 1995).

Essa teoria pode ser testada por meio de sujeitos que sejam treinados para aprender a estratégia de raciocinar indutivamente. Em 1990 Karl Josef Klauer publicou um artigo com a finalidade de mostrar a efetividade do seu programa de treinamento. Para isso foram analisados 30 experimentos, os quais geralmente são compostos por pré e pós teste entre o treinamento. Os resultados encontrados permitiram sustentar que o raciocínio indutivo pode ser aprendido, pois nos experimentos houve a transferência desse raciocínio, sendo que a medida no aumento do raciocínio indutivo é maior do que 1 desvio padrão quando referenciado aos grupos estudados.

Outro aspecto interessante é que se pode concluir teórica e empiricamente que, quando o indivíduo está sendo treinado para o processo de comparação do raciocínio indutivo, o qual requer a identificação de similaridades, esse processo de aprendizado se transfere para a solução de itens que requer o outro processo de comparação, a

diferenciação. Acrescenta-se ainda que esses dois processos são mutuamente exclusivos e um pode ser checado pelo outro. Nesse sentido, também foi investigado se a comparação entre atributos melhora a comparação entre relações e vice-versa. Constatou-se que o treinamento em função de uma ou de outra, atributo ou relações, pode gerar uma transferência pequena, positiva ou negativa, entre elas.

O conhecimento que é treinado também foi averiguado pelo mesmo autor, pois muitos treinamentos ensinam apenas o declarativo e a proposta dessa teoria é ensinar o procedural. Os dados apontam que após o treinamento há a transferência do conhecimento procedural, ou seja, o processo de pensamento para outros contextos, no caso outros testes. Ressalta-se que o treinamento é basicamente de raciocínio indutivo, ou seja, um domínio específico, portanto capacita os indivíduos a se desempenharem melhor em itens que requeiram tal capacidade. Sendo que a transferência ocorrida é diferente e maior do que em testes “*coaching*”, os quais ensinam a pessoa o que ela precisa responder em determinadas questões, isso é feito, por exemplo, nos famosos cursinhos para vestibulares.

A maioria dos trabalhos citados sobre a Teoria de Klauer (1990) demonstrou efeito positivo e significativo do treinamento para melhora do desempenho em testes de raciocínio indutivo a essa questão. Também descartou-se, empiricamente, efeitos como “*warming up*”, que é uma melhora que pode resultar após um período curto de treinamento, e efeito Hawthorne, que é uma condição nova introduzida, como um treinamento, e que pode levar a um maior ou menor desempenho geral, em algo que está sendo aprendido. Constatou-se

que esses efeitos não transferem ganhos em tarefas como as utilizadas no programa de treinamento.

Os ganhos no treinamento do raciocínio indutivo, também foram comprovados empiricamente, e uma das pesquisas que aborda isso é a de Klauer & colaboradores (2001), referentes a dois efeitos, um abrangente do domínio específico e outro, em menor intensidade, que abarca o conhecimento mais geral para resolver algum problema de teste de inteligência, como por exemplo, metacognição.

Em 2002 Klauer e colaboradores desenvolveram uma pesquisa aplicando seu programa de treinamento para estimular o raciocínio indutivo. Três objetivos principais delimitaram o estudo. O primeiro foi relatar o alcance do efeito do treinamento, verificando se esse induz a um domínio específico ou um efeito de transferência mais geral, bem como averiguar se o treinamento converge com medidas de raciocínio indutivo e diverge de medidas que não requerem este domínio. O segundo objetivo buscou testar se ocorria transferência por competência ou por desempenho, pois a competência demonstra o desempenho relevante e é algo duradouro, já somente a transferência de desempenho pode desaparecer depois de algum tempo, a não ser que a competência seja instalada. Por fim o terceiro objetivo procurou investigar se outras competências não observáveis, como a inteligência cristalizada influencia no treinamento da teoria prescritiva.

Os participantes da pesquisa foram 279 crianças, com a média de idade de sete anos. Essa amostra estava distribuída em 12 classes, sendo seis destas do grupo controle e seis do

grupo experimental, ou seja, que fez o treinamento. Todas as crianças foram pré-testadas aproximadamente uma semana antes do treinamento e pós-testadas duas semanas depois do treinamento. Seis meses depois outro pós-teste foi administrado e 219 crianças do grupo inicial foram avaliadas. Para avaliar a transferência do raciocínio indutivo, três testes foram administrados, quais sejam, *Cattell's Culture Fair Test-CFTI* e Matrizes Progressivas Coloridas de Raven para mensurar inteligência fluida. No Matrizes Coloridas foram adicionados os primeiros oito itens dos Matrizes Progressivas Standarizadas de Raven, para evitar que ocorresse o efeito teto. O terceiro teste foi para avaliar inteligência cristalizada e a tarefa era nomear figuras. O teste aplicado seis meses depois foi o CFT.

Com relação ao primeiro objetivo os resultados apontaram que o treinamento contribuiu significativamente mais para o desempenho no Raven e não para o teste de inteligência cristalizada. No entanto, em relação ao teste CFT o grupo experimental obteve melhor desempenho tanto nos subtestes indutivos quanto nos não indutivos. Os dados dos não indutivos não eram esperados, então é preciso investigar melhor o porque isso ocorreu.

O segundo objetivo foi verificado a partir da aplicação do teste CFT seis meses depois do treinamento. O grupo experimental continuou apresentando melhor desempenho que o grupo controle, então a competência foi instalada. No entanto com relação aos itens não indutivos, não houve diferença significativa entre os grupos, o que corrobora, com a questão do treinamento ser restrito a um domínio específico. Por fim para o terceiro objetivo, os resultados apontaram que o efeito do treinamento para a inteligência fluida é

bastante significativo e para a inteligência cristalizada insignificante. Os dados dessa pesquisa sustentam a Teoria Prescritiva do Raciocínio Indutivo de Klauer. No livro de Klauer e Phye (1995) são mencionados outros estudos que também comprovam a validade do treinamento do raciocínio indutivo proposto por Klauer em 1990.

Como pode-se perceber esse treinamento é bastante fundamentado nos processos cognitivos e apresenta evidências empíricas da sua validade. Assim como pontuado por Wilhelms (2005), é um treinamento interessante e que traz dados favoráveis de sua utilização. Além disso, trabalha um domínio específico que é o raciocínio indutivo, o mesmo elegido a ser trabalhado no presente projeto. No entanto, ressalta-se que a finalidade do trabalho atual não é desenvolver um treinamento tão minucioso quanto o proposto por Klauer, que implica várias semanas e seções de treinamento. A presente pesquisa tem a finalidade de construir um teste dinâmico informatizado do raciocínio indutivo, que envolve a construção de itens indutivos e um procedimento de intervenção curto, por isso o caráter dinâmico do teste, sobre como resolver tais tarefas de raciocínio indutivo. Diante disso a teoria de Klauer foi eleita como uma fundamentação adequada para a construção dos itens e da etapa de intervenção. Porém ressalta-se que aqui não será objetivo a mudança cognitiva, pois um teste dinâmico não permite isso, o que será verificado é o potencial de aprender, o que é uma estimativa do quanto uma pessoa poderia vir a mudar cognitivamente em relação a uma habilidade, caso fosse ensinada e instruída como em um treinamento.

Assim, como em qualquer teste estático que avalie alguma capacidade da inteligência, o teste dinâmico deste trabalho, relacionado ao construto da inteligência, também é afetado por variáveis que contribuirão para o bom ou mau desempenho do indivíduo. Por exemplo, fatores como ansiedade, falta de motivação, impulsividade também poderão afetar a capacidade que o sujeito já possui sobre o construto a ser mensurado por meio de testes dinâmicos. Assim, um indivíduo que já apresente uma maior capacidade de raciocínio viso-espacial, em tarefas desse tipo tenderá a apresentar melhor desempenho, necessitando de menos instruções para a resolução do que um indivíduo com menor capacidade nesse construto (Simões, 1995), com a devida ressalva ao peso das variáveis intervenientes mencionadas.

Outros pontos essenciais para o desempenho dos indivíduos nos testes dinâmicos são as instruções durante as tarefas e a habilidade do examinador para ensinar e perceber as dificuldades do indivíduo. As instruções bem formuladas e a habilidade de ensinar e identificar as dificuldades possibilitam auxiliar o sujeito da melhor forma possível frente suas deficiências (Tzuriel & Shamir, 2002, Sternberg & Grigorenko, 2002, Linhares, 1996, Simões, 1995, entre outros). No caso do presente teste, as instruções ou intervenções, como se decidiu denominar, é a parte mais importante, pois não haverá um examinador para conduzir o procedimento que é informatizado e padronizado. Então a exigência de desenvolver intervenções mais bem detalhadas foi muito maior, já que será por meio dela que poderá ser verificado o potencial de aprendizagem, bem como compreender melhor os resultados obtidos pela amostra investigada. Diante disso percebe-se novamente a

importância de um referencial teórico sólido para o desenvolvimento da intervenção. No caso concluiu-se que a Teoria de Klauer (1990) possibilitaria esse respaldo teórico, bem como prático.

Até esse momento da introdução foram desenvolvidas as principais variáveis desse estudo que compreende avaliação dinâmica, teste dinâmico, raciocínio indutivo e a teoria que fundamenta a construção dos itens e intervenção do raciocínio indutivo no teste dinâmico. A próxima seção tratará das variáveis critérios que serão utilizadas para os estudos de evidências de validade do teste dinâmico.



#### **1.4- Raciocínio indutivo e variáveis critérios: desempenho acadêmico, funções executivas, especificamente atenção seletiva e controle inibitório.**

Tradicionalmente o raciocínio indutivo não é especificamente trabalhado no ambiente escolar. A aprendizagem na escola é voltada para o desenvolvimento da inteligência cristalizada, pois é um momento em que o indivíduo adquire muitos conhecimentos relacionados a aspectos culturais. No entanto, algumas pesquisas apontam a associação entre raciocínio indutivo e desempenho acadêmico (Cruz, 2008; Santos & Primi, 2005; Primi, Santos & Vendramini, 2002; Primi & Almeida, 1998).

O trabalho de Primi, Santos e Vendramini (2002) teve como um dos objetivos verificar a predição do raciocínio indutivo em relação ao desempenho nos cursos de Medicina, Pedagogia, Odontologia, Administração, Psicologia, Letras, Engenharia Civil e Matemática. Os resultados obtidos mostraram que raciocínio indutivo é uma variável preditora do desempenho acadêmico na maior parte dos cursos, exceto Medicina e Pedagogia, sendo que o maior poder de predição foi para os cursos de Engenharia Civil e Matemática.

No contexto de escola, em vez de universidade, encontra-se o trabalho de Primi e Almeida (2000), que demonstrou correlação positiva entre os subtestes da Bateria de Provas de Raciocínio-5, BPR-5 (Primi & Almeida, 1998), que são tarefas de raciocínio indutivo, com o desempenho acadêmico de indivíduos, alunos desde a 5ª série do ensino

fundamental até a última série do ensino médio. Esse resultado pôde ser constatado com amostras brasileiras e portuguesas.

A relação entre raciocínio indutivo e desempenho acadêmico na escola também foi investigada na pesquisa de Santos e Primi (2005). A medida de desempenho foi obtida por meio de um questionário respondido pelas professoras que abarca dificuldades na aprendizagem como hiperatividade, leitura, escrita e cálculo. Os resultados apontaram correlação negativa entre raciocínio indutivo e dificuldade em leitura e cálculo, ou seja, quanto menor o desempenho nesse raciocínio, maior a dificuldades em tais áreas.

Um estudo mais recente, investigando essa relação entre raciocínio indutivo e desempenho acadêmico com crianças de 7 a 11 anos estudantes da primeira à quinta série do ensino fundamental foi o de Cruz (2008) e que utilizou a prova de Raciocínio Abstrato da BPR-5i, também utilizada na presente pesquisa. Participaram do estudo 289 crianças e os resultados mostraram correlações moderadas e positivas entre o desempenho acadêmico da segunda série nas matérias de Português, Matemática, História, Geografia, Ciências e Artes com raciocínio indutivo. Com as crianças da quarta série também foram encontradas correlações positivas, mas baixas, entre desempenho acadêmico nas matérias de Português, História, Geografia e Ciências com raciocínio indutivo. Na quinta série essa relação baixa também ocorreu, mas com as matérias de Português, Geografia, Matemática, Ciências e Artes.

No que se refere ao potencial de aprendizagem do raciocínio indutivo avaliado por um teste dinâmico, o estudo de Primi (2005) encontrou algumas correlações entre a magnitude de aprendizagem e o desempenho acadêmico de estudantes universitários, porém todas baixas. Já no ambiente escolar pode-se citar o trabalho de Escolano e Linhares (2000), com crianças cursando a primeira série do ensino fundamental. As autoras identificaram predição baixa de um teste dinâmico denominado de Jogo das Perguntas de Busca com Figuras Geométricas-PBFG (Linhares, 1991) em relação ao desempenho acadêmico, porém considerando somente as crianças da amostra com desempenho acadêmico baixo a predição foi mais robusta. Para Sternberg e Grigorenko (2002), o potencial de aprendizagem não tem se demonstrado um bom preditor de desempenho acadêmico, assim como pode ser verificado nas pesquisas de Primi (2005) e Escolano e Linhares (2000) citadas anteriormente.

Sternberg e Grigorenko (2002) afirmam que tais correlações podem não ocorrer porque há variáveis intervenientes, por exemplo, o escore do teste dinâmico é criado e administrado diferentemente da composição das notas escolares. O teste dinâmico não tem a finalidade de verificar o desempenho já adquirido, como o acadêmico. Porém os autores ressaltam que não se deve criticar os testes dinâmicos por eles não fazerem tal predição, e sim usar testes apropriados que possam predizer, como testes psicométricos estáticos.

Outro ponto que pode explicar a ausência de correlação entre testes dinâmicos e desempenho escolar é que, muitas vezes, o construto avaliado no teste dinâmico não é

semelhante ao que é ensinado na escola. Um ponto que deve ser levado em consideração é o tipo de tarefa que se pretende avaliar no teste dinâmico, pois algumas variáveis são mais fáceis de ensinar e podem ter maior semelhança com conteúdos ensinados na escola, enquanto outras variáveis parecem ser mais difíceis de ensinar, tanto em um teste dinâmico, quanto na escola (Sternberg & Grigorenko, 2002). O tipo de tarefa do presente estudo aborda o raciocínio indutivo e, como já relatado, esse raciocínio não é especialmente promovido no ambiente escolar.

Além do desempenho acadêmico, outras variáveis-critério investigadas serão controle inibitório e atenção seletiva, habilidades relacionadas às funções executivas. Como pôde ser verificado no capítulo anterior, ao longo da história da inteligência, pelo prisma psicométrico, a inteligência fluida (Gf) tem sido um fator de destaque e, segundo Flanagan e Ortiz (2001), é o mais associado à inteligência geral. No fator da inteligência fluida o raciocínio indutivo é um componente importante, referindo-se à capacidade de analisar um conjunto de informações e estabelecer relações entre elas, criando-se novas idéias e conceitos, organizando-se sistematicamente as informações (Primi, 2002).

Outras áreas de estudo, como a neuropsicologia, têm buscado compreender as relações entre a inteligência fluida e construtos mais específicos dessas áreas. Nesse sentido, diversos artigos de abordagem neuropsicológica fazem menção à associação entre inteligência fluida e funções executivas (Blair, 2006; Salthouse, 2005; Fernandez-Duque, Baird & Posner, 2000). Visto que o raciocínio indutivo, conforme anteriormente revisado,

pode ser compreendido como um componente da inteligência fluida, pressupõe-se que o desempenho em testes que avaliem tal raciocínio, como o desenvolvido na presente tese, correlacione-se com testes que avaliam funções executivas.

As funções executivas referem-se à capacidade de o sujeito engajar-se em comportamentos orientados a objetivos, ou seja, à realização de ações voluntárias, independentes, autônomas, auto-organizadas e orientadas para metas específicas (Ardila & Ostrosky-Solís, 1996; Gazzaniga, Ivry & Mangun, 2002). Ou seja, são habilidades que, integradas, capacitam o indivíduo a tomar decisões, avaliar e adequar seus comportamentos e estratégias, buscando a resolução de um problema (Malloy-Diniz, Sedo, Fuentes e Leite, 2008). Por exemplo, em um teste de inteligência, o objetivo é resolver um problema para chegar à resposta correta, logo as funções executivas auxiliarão para que o sujeito, engajado nesse objetivo, trace uma estratégia de como resolver o problema, bem como mude ou aperfeiçoe tal estratégia caso necessário para atingir o objetivo de acertar a resposta. Dentre as principais habilidades que se relacionam com as funções executivas, nesse trabalho serão descritas cinco, quais sejam, o planejamento, a flexibilidade, memória de trabalho, controle inibitório e atenção seletiva.

O planejamento é a capacidade de estabelecer uma estratégia seqüencial direcionada a atingir um objetivo. A flexibilidade cognitiva pressupõe a capacidade do sujeito mudar ou alternar estratégias de ação ou pensamento, conforme a necessidade para a resolução do problema. Já a memória de trabalho é um depósito temporário de armazenamento de

informações que podem ser acessadas, manipuladas e reorganizadas para serem utilizadas em alguma tarefa, como por exemplo, guardar um número de telefone para logo em seguida utilizá-lo para fazer uma ligação (Malloy-Diniz & colaboradores, 2008).

Com relação ao controle inibitório, este diz respeito à filtragem e seleção de informações (Gazzaniga, Ivry & Mangun, 2002), consistindo na capacidade do indivíduo inibir respostas prepotentes ou distratoras que bloqueiam o curso de uma ação, ou ainda a interrupção de uma resposta já em curso (Barkley, 2001). Por fim, a atenção seletiva é a capacidade do sujeito de atentar a determinadas características do estímulo, ignorando aqueles que são irrelevantes à tarefa (Gazzaniga, Ivry & Mangun, 2002).

Acima foram citadas as habilidades que estão mais difundidas, porém autores como Malloy-Diniz e colaboradores (2008) acrescentam outras duas habilidades, quais sejam categorização e fluência. A primeira é um processo de agrupamento de elementos que apresentam características semelhantes. Tal processo está vinculado à formação de conceitos, raciocínio indutivo, dedutivo e abstração. Já a segunda habilidade é de emissão de comportamentos verbais e/ou não verbais numa seqüência lógica e respeitando determinadas regras.

O presente trabalho abordará mais especificamente as habilidades de controle inibitório e atenção seletiva. Analisando as definições anteriormente descritas sobre tais habilidades, pode-se perceber a dificuldade de separação desses dois construtos, já que tais mecanismos tendem a se complementar (Gazzaniga, Ivry & Mangun, 2002). Ou seja, pode

parecer natural que, ao voltar a atenção para alguns estímulos relevantes, automaticamente a atenção a alguns outros estímulos seja inibida. Porém, há uma diferença entre ambos os processos, que tem sido verificada empiricamente por meio de estudos eletrofisiológicos. Por exemplo, o trabalho de Knight e Grabowecky (1995) investigou a diferença nos potenciais evocados diante de estímulos auditivos que deveriam ser atentados ou ignorados. Observou-se que pessoas normais apresentam diferença na ativação entre os dois tipos de estímulos, com maior negatividade para estímulos que deveriam ser atentados do que para os que deveriam ser ignorados. Por outro lado, pessoas com lesão pré-frontal não apresentaram tal diferença, sendo que o estímulo a ser atentado foi ativado de forma semelhante aos sujeitos normais, porém não houve menor negatividade para os sons que deveriam ser ignorados.

Apesar de distintas, as duas habilidades ativam o sistema do cíngulo anterior. Isso foi verificado, por exemplo, por meio de estudos com neuroimagem funcional em que indivíduos respondiam a um teste de geração semântica. Nesse teste, que avalia controle inibitório, o sujeito precisa selecionar um verbo para relacioná-lo a um substantivo apresentado pelo examinador. Durante essa tarefa ocorre maior ativação do cíngulo anterior, provavelmente refletindo maior monitoramento da atenção. No entanto, quando os mesmos substantivos são apresentados repetidamente, os participantes tendem a emitir o mesmo verbo e tal processo está relacionado à diminuição da ativação no cíngulo anterior. Os autores explicaram tais achados sugerindo que, com a repetição dos substantivos e, conseqüentemente, por memorização, a repetição dos verbos gerados, a tarefa deixa de ser

nova e difícil, demandando menor controle atencional e ativando em menor intensidade o cíngulo anterior (Raichle & colaboradores, 1994).

Resultados semelhantes foram observados com o Teste de Stroop que avalia atenção seletiva. Nesse teste, o participante deve realizar diferentes tarefas: em um momento inicial, deve dizer a cor com que uma palavra está escrita, sendo a situação convergente, ou seja, a palavra escrita e a cor com que ela foi escrita são semelhantes; e, em uma situação seguinte divergente, a pessoa precisa também dizer a cor com que a palavra foi escrita, mas neste caso ela deve ignorar o conteúdo da palavra visto que a palavra escrita e cor são diferentes. Assim como no teste de geração semântica, no Teste de Stroop a ativação do cíngulo anterior tende a aumentar na segunda situação, incongruente, que necessita de um maior monitoramento da atenção (Posner, 1994 citado por Gazzaniga, Ivry & Mangun, 2002).

Como já descrito, o cíngulo anterior trabalha concomitantemente com o córtex pré-frontal para coordenarem um comportamento orientado a um objetivo. A função do cíngulo parece ser a de monitoramento da atividade seletiva do pré-frontal para que estímulos irrelevantes sejam inibidos, dessa forma há um sistema de filtragem dinâmica (Gazzaniga, Ivry & Mangun, 2002).

Didaticamente, como foi classificado acima, o Teste de Geração Semântica é proposto a avaliar controle inibitório e o Teste de Stroop, atenção seletiva. No entanto são funções relacionadas, sendo praticamente impossível separá-las nesses tipos de testes que demandam tais processos cognitivos. Ressalta-se que, no presente trabalho, adotou-se a



separação didática, porém partindo do pressuposto de que não há essa dicotomização das funções durante a execução dos testes. Essa separação apenas é possível ser percebida no campo da neurofisiologia com auxílio de recursos como Potencial Evocado e Tomografia Computadorizada. Porém, em testes psicológicos que trabalham com processos cognitivos subjacentes a essas funções, não há como detectar tal separação, mesmo porque as tarefas usualmente não são puras, sendo que vários processos podem ser utilizados para o indivíduo chegar a uma resposta. Por exemplo, tanto no Teste de Stroop quanto no Teste de Geração Semântica tendem a ocorrer tanto a seleção de informação quanto a inibição de respostas, que é justamente o sistema de filtragem dinâmica.

Essas duas habilidades, controle inibitório e atenção seletiva, foram escolhidas para serem abordadas nesse trabalho, pois a capacidade de raciocinar indutivamente requer, entre vários processos cognitivos, que o indivíduo iniba informações irrelevantes e selecione quais irão conduzi-lo a descobrir a resposta correta. Tais variáveis serão relacionadas tanto ao desempenho em um teste de raciocínio indutivo quanto ao potencial de aprendizagem nesse teste. Espera-se, visto que esse raciocínio está correlacionado com tais variáveis, que pessoas que apresentam essas funções desenvolvidas tendam a apresentar melhores desempenhos nesse potencial de aprendizagem. Tal pesquisa justifica-se especialmente porque, até o momento, não foram encontradas trabalhos destinadas a investigar essa questão.

Até esse momento da introdução foram desenvolvidas as principais variáveis desse estudo que compreende avaliação dinâmica, teste dinâmico, raciocínio indutivo e a teoria que fundamenta a construção dos itens e intervenção do raciocínio indutivo no teste dinâmico, bem como alguns critérios que serão utilizados para a verificação de evidências de validade do teste dinâmico em relação ao potencial de aprendizagem mensurado por ele, que se refere ao raciocínio indutivo. A próxima seção tratará de um aspecto mais técnico, no entanto essencial quando fala-se em construção de teste, que é independente da fundamentação teórica que embasa, pois existem certos padrões que devem ser seguidos para a construção de um teste, bem como para averiguar suas propriedades psicométricas.

## 1.5 - Construção de instrumento psicológico e seus requisitos psicométricos

A legitimidade dos testes psicológicos precisa ser respaldada em estudos que comprovem suas qualidades psicométricas, respeitando as especificações que garantam reconhecimento e credibilidade por parte da comunidade científica e dos leigos. Desde a década de 1950, no cenário internacional, houve a preocupação com a qualidade dos testes psicológicos. A Associação Psicológica Americana (APA), e depois por outras instituições, que trabalharam para estabelecer critérios para a padronização e uso dos testes (Noronha & Vendramini, 2003).

No Brasil, com a resolução nº 002\2003, também se pôde notar esse cuidado mais efetivo em relação à adequação dos testes, pois se definiu que estes devem apresentar requisitos mínimos para serem reconhecidos como tais, sendo considerada falta de ética do uso dos instrumentos que estejam fora dos padrões exigidos. Essa resolução define e regulamenta o uso, a elaboração e a comercialização de testes psicológicos, especificando que é necessário ter um psicólogo responsável por cada teste publicado, bem como foi definido um sistema contínuo de avaliação dos testes, verificando a adequação destes para o uso. Tal resolução se baseia nos conceitos, princípios e procedimentos definidos nos documentos: *International Test Commission – ITC* (2001), *American Educational Research Association\American Psychological Association\National Council on Measurement in Educational – AERA, APA, NCME –* (1999) e *Canadian Psychological Association – CPA* (1996).

Conforme descrito na Resolução nº 002/2003, para serem considerados adequados, os instrumentos de avaliação psicológica precisam apresentar alguns requisitos mínimos e obrigatórios referentes à sua fundamentação teórica, evidências empíricas de validade e precisão, propriedades psicométricas dos itens (exceto os classificados como testes projetivos), sistema de correção e interpretação dos escores, procedimentos de aplicação, correção e um manual contendo essas informações (para mais detalhes consultar o *site* [www.pol.org.br](http://www.pol.org.br)). O CFP dispõe de uma Comissão Consultiva em Avaliação Psicológica, composta por psicólogos reconhecidos na área dos testes psicológicos, para analisar e emitir pareceres sobre os instrumentos. Em suma, determinado teste só poderá vir a ser utilizado depois de obtida a aprovação do CFP, exceto quando o uso for para pesquisa, pois nesse caso pode-se utilizar testes ainda não aprovados.

A presente seção abordará a construção de um teste psicológico, pretendendo verificar suas propriedades psicométricas internas, dentre elas a precisão, bem como sua validade, sem o intuito de normatização nesse momento. No Capítulo de Método do trabalho será explanado todo o procedimento para se efetuar os estudos de construção, precisão e validade. No entanto, para se compreender melhor esse universo metodológico e os objetivos do trabalho faz-se necessário abarcar conceitualmente esses três elementos, construção, precisão e validade.

Ao pensar na construção de um instrumento psicológico, imediatamente surgem as questões de precisão e validade, bem como de padronização (procedimentos padronizados

de aplicação e correção) e normatização (interpretação dos escores), pois somente com todos esses fatores agregados é possível tornar um teste adequado para uso. De acordo com Urbina (2006) ao se construir um teste é fundamental conciliar o seu objetivo com uma fundamentação teórica adequada. Esse realmente é o primeiro passo, pois é necessário saber o construto que o teste irá se propõe a avaliar e então fundamentar esse construto, de modo a conhecê-lo e compreendê-lo, o que permitirá a elaboração de itens adequados ao teste, bem como as formas de correção e interpretação dos resultados. O presente projeto se enquadra nesse quesito, pois como se pôde perceber, nas seções anteriores, os construtos trabalhados no teste, principalmente, teste dinâmico e raciocínio indutivo foram fundamentados, o que respaldará toda a construção do instrumento. Além dessa união, objetivo e fundamentação teórica, também deve-se especificar a população que o teste pretende atingir (Urbina, 2006), pois, por exemplo, ao se construir um teste de inteligência, os itens para conseguirem mensurar a capacidade de um indivíduo, precisam estar coerentes com sua faixa etária. Nesse trabalho esse ponto também se aplica ao direcionar o teste dinâmico para crianças de 7 a 11 anos.

Urbina (2006) ainda relata passos do desenvolvimento de um teste, ou seja, a sua construção como um todo, quais sejam, criar itens com o formato adequado e determinar os procedimentos de administração e avaliação, submeter os itens para especialistas e fazer as reformulações necessárias, testar os itens com uma amostra representativa, avaliar os resultados dessa aplicação na amostra, adicionar, eliminar e/ou modificar itens se necessário a partir de análise qualitativa e quantitativa. Fazer outras aplicações em amostras

diferentes para verificar se os resultados se mantêm, padronizar e fixar o tamanho do teste e a seqüência dos itens, bem como procedimentos de aplicação e administração final, aplicar novamente numa amostra representativa da população a qual o teste se destina e desenvolver os dados normativos, a precisão e evidências de validade, publicar o teste juntamente com o manual descrevendo todos os dados de padronização, validade e precisão e revisar estudos com o passar do tempo.

Serão abordadas mais especificamente a precisão e as evidências de validade que são partes essenciais na construção e que farão parte desse estudo. A precisão é um índice que indica o quanto de erro pode estar embutido nos resultados de um teste, ou em outras palavras, o quanto um teste é preciso, é fidedigno em mensurar um determinado construto. Frente isso, é importante ressaltar que nenhum instrumento psicológico está livre de erros, então o importante é garantir menor quantidade de erro possível. Os erros podem ocorrer por diversos motivos, entre eles, conteúdo dos itens, tempo entre aplicações, desatenção e “chutes” do examinado ao fazer o teste e subjetividade do avaliador, caso o teste seja mais vulnerável a essa questão no momento da aplicação e correção.

Existem cinco delineamentos principais para obtenção da estimativa do erro (Anastasi e Urbina, 2000). O primeiro é a *Forma Teste-Reteste* que se refere quando os sujeitos são avaliados com o mesmo instrumento, mas em ocasiões diferentes e controla apenas as variâncias de conteúdo estando sujeita às variâncias de tempo. O segundo é denominado *Formas alternativas* em que sujeito pode ser avaliado no mesmo dia ou em

dias diferentes com testes paralelos. Por esse procedimento se estabelece controle sobre duas fontes de erro, a saber, as amostras de conteúdo de itens e a estabilidade temporal, conforme a escolha de seu esquema de aplicação. O terceira, *Forma de precisão do avaliador*, é um tipo de precisão que consiste em solicitar que dois ou mais avaliadores diferentes pontuem o protocolo de teste dos mesmos sujeitos. Após a correção se estabelece a correlação entre os avaliadores. Essa técnica é muito usada quando os testes dependem muito do julgamento do avaliador.

Por fim o quarto delineamento é designado de *Forma das metades ou Split-Half* que é um coeficiente de precisão, sendo conseguido em uma única aplicação de uma única forma do teste. Divide o teste em duas partes iguais e as correlaciona, nesse caso deve-se tomar o cuidado de se manter a homogeneidade ou equivalência de cada parte de modo a minimizar a variância de conteúdo. E o quinto, denominado de *Coefficientes de Kuder-Richardson e Alfa de Cronbach* representam medidas obtidas pelo grau de associação entre os itens e o teste, ou seja, o quanto um item se correlaciona com os demais para medir o referido construto. Estabelece-se uma relação entre as respostas nos itens individualmente com o escore total do teste. Esses índices também são conseguidos com uma única aplicação de uma única forma do teste. Está sujeito a variações pela homogeneidade do teste, ou seja, quanto mais homogêneo for o teste maior será o coeficiente de fidedignidade calculado por esses procedimentos. No presente trabalho será verificada a precisão pelo Coeficiente de Alfa e também por um outro meio, não muito usual, mas tão bom quanto os

descritos, que é pelos indicadores Modelo de Rasch (Wright & Stone, 2004), que verifica se os acertos aos itens estão coerentes com a habilidade do sujeito respondente.

Como parâmetro do índice de precisão, o Conselho Federal de Psicologia por meio Resolução 002/2003 adotou de padrões reconhecidos internacionalmente (ITC, 2001, APA, AERA, NCME, 1999 & CPA, 1996), nos quais a precisão mínima que um teste precisa apresentar para ser considerado adequado é 0,60. Porém autores como Cronbach (1996), Messick (1989) e Urbina (2006) e sugerem que esse índice seja no mínimo de 0,80.

Com relação a validade de acordo com a *American Educational Association*, *American Psychological Association*, *National Council on Measurement in Education* (1999): “Refere-se ao grau no qual evidência e teoria sustentam as interpretações dos escores dos testes vinculados pelo propósito do uso dos testes” (p. 9), sendo considerada fundamental para o desenvolvimento e avaliação de um teste. O processo de validação é constituído por um conjunto de evidências que possam assegurar cientificamente as interpretações dos escores do teste. Assim, nesse processo não se valida o teste em si, mas as interpretações propostas por ele; dessa forma, intenções diferentes sobre uma mesma interpretação podem ser validadas, e essas interpretações dizem respeito ao construto que o teste se propõe a mensurar. Por exemplo a interpretação de extroversão em um teste, esta pode ser associada com desempenho acadêmico ou profissional, mas é preciso comprovar empiricamente.



O CFP, na análise dos testes, entre os anos de 2003 e primeiro semestre de 2007 adotou os tipos de validade de conteúdo, critério e construto, os quais são chamados por Muniz (2004) de “santíssima trindade da validade”, por serem os primeiros tipos de validade surgidos no meio científico e muito utilizados até hoje. O primeiro tipo de validade tinha o objetivo de avaliar o conteúdo dos itens de determinado instrumento, verificando se estão adequados ao propósito de representar um domínio de comportamento a ser mensurado. Depois surgiu a validade de critério, que verifica a efetividade de um teste em prever o desempenho de um sujeito em uma situação específica. Por último, criou-se a validade de construto para averiguar a extensão que um teste mede determinado construto.

Entretanto, pode-se observar que tanto a validade de conteúdo quanto a de critério sempre apresentarão informações referentes ao construto; por isso, mais recentemente, se discutiu a amplitude do conceito de validade de construto de tal forma que se chegou à conclusão que ele é um conceito abrangente que inclui sob si todos os outros, sendo, portanto, inadequado sustentar a existência de procedimentos específicos para validade de construto, diferentes dos demais. Todos os procedimentos de se averiguar a validade de um teste têm algo específico a mostrar se um teste mede determinado construto. Nesse sentido, a terminologia “tipos de validade” foi intencionalmente mudada para “fontes de evidência” de validade, reforçando a conotação do termo validade como um conceito único, de tal forma que as diferentes formas de se estudá-la (conteúdo, correlações teste-critério, etc.) trazem dados que são usados para sustentar a argumentação a respeito da validade das interpretações dos escores do teste com uma finalidade particular, isto é, cada evidência

busca aspectos diferentes da validade, e todas elas trazem algo sobre o porquê determinado teste mede um construto. Assim, considera-se adequado mais um tipo específico de procedimento chamado validade de construto. (Anastasi & Urbina, 2000; APA AERA, NCME, 1999).

Frente a essas questões sobre o conceito validade, já no primeiro semestre de 2007, o CFP na análise dos testes incorporou uma nova nomenclatura para as evidências de validade, as quais foram propostas pelos *Standards* (APA, AERA, NCME, 1999). Essas evidências de validade são categorizadas em: (a) evidência com base no conteúdo do teste, (b) evidência com base no processo de resposta, (c) evidência com base na estrutura interna, (d) evidência com base na relação com outras variáveis e (e) evidências baseadas nas conseqüências de testagem.

Evidência com base no conteúdo informa sobre a abrangência e representatividade dos itens de um teste em representam todas as nuances do construto que o teste se propõe a medir. As pessoas que desenvolvem os testes freqüentemente trabalham para especificar o conteúdo, descrevendo-o cuidadosamente em detalhes com uma classificação das áreas e tipos de itens, por exemplo. Essa evidência pode incluir uma análise lógica ou empírica pela qual se pretende verificar se o conteúdo do teste representa adequadamente o domínio e se há relevância deste para as interpretações propostas pelos escores do teste. Evidência embasada no conteúdo também pode ser obtida por meio de especialistas que julgam a relação entre as partes do teste com as facetas do construto (APA, AERA, NCME, 1999).

Com relação à evidência com base no processo de resposta, esta pode fornecer informações detalhadas do desempenho ou a resposta emitida pelo examinando e se esses processos têm relação com o construto que o teste propõe medir. Essas evidências, geralmente, são obtidas por análises individuais de respostas, de questionamentos aos respondentes sobre suas estratégias de resolução dos itens ou respostas particulares aos itens. Inferências sobre os processos envolvidos no desempenho podem também ser desenvolvidas, analisando a relação entre as partes do teste ou o teste com outras variáveis. Essa validade também pode incluir estudos por meio de observações de juízes que verificam se os dados estão em consonância com o que pretendem medir, de acordo com a definição de construto. A evidência com base no processo de resposta pode contribuir para questões sobre as diferenças nos significados ou interpretação dos escores de um teste, cruzando subgrupos relevantes de examinandos (APA, AERA, NCME, 1999).

Já a evidência com base na estrutura interna indica o nível em que as relações entre os itens do teste e/ou seus componentes ou subtestes são coerentes com a estrutura de relações proposta pela definição teórica. A estrutura conceitual de um teste pode implicar em uma dimensão ou mais dimensões quando há multidimensionalidade, também pode pressupor uma estrutura hierárquica ou de fatores isolados. Nesses casos, essa forma de evidência verifica a coesão entre a estrutura prevista com a observada. Nessa categoria também inclui-se a análise da estrutura das relações entre os itens de um teste destinados por teoria, a avaliar um mesmo construto (APA, AERA, NCME, 1999).

No que se refere à evidência com base na relação com outras variáveis, o objetivo é a análise da relação dos escores do teste com variáveis externas. As variáveis externas podem ser classificadas como medidas de critério que é esperado por teoria possuir alguma relação com o construto medido pelo teste, assim como medidas obtidas por outros testes que supostamente meçam o mesmo construto, construtos relacionados ou ainda construtos diferentes. A evidência com base na relação com outras variáveis pode ser obtida por dois procedimentos: convergente-discriminante ou relação teste-critério.

Evidência convergente-discriminante indica relação entre escores do teste e outras medidas com a intenção de avaliar construtos similares (convergentes) ou diferentes (divergentes). Evidência na relação teste-critério busca compreender a eficácia dos escores do teste em prever determinado critério. A variável critério é uma medida de algum evento que, por si só, já é importante como acidente no trânsito, que não é necessariamente um processo mental, mas tem relação teórica com os construtos psicológicos medidos pelos testes. A escolha do critério, bem como o procedimento de mensuração utilizado para obter as medidas do critério é ponto central neste método, pois o valor do estudo depende da relevância, precisão e validade dessas medidas externa (APA, AERA, NCME, 1999).

Por fim, a evidência com base nas consequências do teste ou consequencial informa não propriamente sobre a validade do teste, mas sobre questões de política social do uso dos testes. Espera-se sempre que os resultados de um teste possam trazer algum benefício, seja para definir uma estratégia mais adequada de terapia, seja para uma organização

contratar um indivíduo que apresente um perfil próximo ao esperado, para decidir um novo conteúdo educacional que se adequou mais às necessidades dos alunos. No entanto, para se obter resultados que sejam benéficos, é necessário ter o discernimento de qual teste utilizar em determinada situação, bem como aplicá-lo e interpretá-lo coerentemente com o que já foi estipulado no manual deste teste. A má utilização de um teste pode comprometer a validade das interpretações, fazendo ruir a legitimidade dos dados obtidos. Assim, mesmo tendo instrumentos válidos, se a utilização for incorreta, a validade fica comprometida. É evidente que essa propriedade não é do teste em si, mas sim do contexto mais amplo no qual ele é usado. Portanto, as conseqüências da testagem assumem uma importância grande no que se refere a validade no sentido amplo. Os procedimentos para esse tipo de validade buscam mostrar se os benefícios antecipados pelo emprego do instrumento estão de fato ocorrendo (APA, AERA, NCME, 1999).

Ressalta-se que essa última evidência de validade não foi incluída para as análises dos testes, pois na verdade, como pode ser verificado, é algo que transcende ao teste. É um tipo de validade que pressupõem a utilização adequada do teste por parte do psicólogo. Apesar de ser um questionamento importante, pois trabalha com a ética em relação aos testes psicológicos, é difícil controlar essas situações e se torna inviável de ser julgado dentro de padrões estabelecidos criados para enquadrar o teste psicológico como adequado ou não para ser utilizado, mesmo porque essa evidência não é direta ao mesmo.

Com relação as evidências de validade no SATEPSI, recomenda-se preferencialmente que o manual do teste psicológico mostre pelo menos uma evidência de validade, mas quando única, não pode ser referente a estrutura interna, pois é preciso apresentar evidências de validade em relação a um critério externo ou comparando com outro teste, pois assim garante mais a legitimidade das interpretações do teste. Não adianta apenas saber a estrutura de um teste, ou seja, quantos fatores ele avalia, é de extrema importância ter informações adicionais para essa estrutura. Por exemplo, um teste que avalia capacidade verbal e numérica, que são dois fatores, além de saber que ele tem essa estrutura, é imprescindível que ele se correlacione com outros testes que também avalie essas duas capacidades, ou mesmo mostre predição em relação a algum desempenho como o acadêmico. Ou seja, estrutura interna apenas diz quantos fatores um teste avalia, mas para ter informações adicionais que ajudam nas interpretações e utilidade para diversos contextos de um teste, é necessário investigar outras evidências de validade.

Pensando sobre as evidências de validade expostas, o presente estudo buscará para o teste dinâmico de raciocínio indutivo evidência de validade baseada no conteúdo, na estrutura interna e na relação com outras variáveis, tanto no procedimento convergente-discriminante quanto teste-critério. A evidência baseada no conteúdo será investigada, tanto para os itens quanto para a intervenção, sendo que nesses dois casos, além de serem fundamentados de acordo com uma teoria, também haverá análise de especialistas. A estrutura interna será explorada para ver se o teste realmente avaliará somente raciocínio indutivo e se o mesmo poderá ser dividido em fatores, pois os itens apresentam conteúdos

diferenciados, alguns são pictóricos e outros geométricos. Além disso, também avaliam paradigmas do raciocínio indutivo que apesar de serem relacionados, alguns são para verificar relações entre objetos, outros averiguar apenas os atributos (cor, forma, tamanho) e ainda tem itens de analogia e complemento de séries. Pode-se pensar também em relação a complexidade dos itens, pois fatores podem vir a surgir em decorrência dos itens serem mais fáceis em um grupo e mais difíceis em outros.

Já a evidência de validade baseada na relação com outras variáveis, pelo procedimento convergente, será investigada com uma prova de raciocínio abstrato e dois testes que avaliam funções executivas, mais especificamente atenção seletiva e controle inibitório. Por fim, a evidência de validade baseada nas relações com outras variáveis pelo procedimento teste-critério será buscada por meio do desempenho acadêmico.

## **CAPÍTULO 2- METODOLOGIA, RESULTADOS E DISCUSSÃO DOS ESTUDOS 1**

### **E 2.**

O principal objetivo desse trabalho foi a construção de um teste dinâmico para verificar o potencial de aprendizagem relacionado ao raciocínio indutivo. A construção se referiu também a investigação das propriedades psicométricas em relação à estrutura interna, precisão e validade desse instrumento. No entanto para atingir esse objetivo do desenvolvimento do teste dinâmico, dividiu-se o trabalho em 2 Estudos. A seguir serão descritos separadamente os dois estudos, cada qual com seus objetivos específicos, participantes, materiais, procedimentos, resultados e discussão. No capítulo 3 serão feitas as considerações finais para os 2 Estudos.

Resumidamente, o Estudo 1 foi dividido em duas partes, a primeira teve por finalidade a construção de itens de raciocínio indutivo, conforme a Teoria de Klauer (1990), para que os itens mais adequados, conforme análise qualitativa e quantitativa, fizessem parte do pré e pós teste do instrumento dinâmico. A segunda parte do estudo objetivou a construção de um procedimento de intervenção do instrumento dinâmico para a resolução dos itens indutivos, também fundamentado na mesma teoria dos itens. Após esse estudo da elaboração do pré-teste, pós-teste e etapa de intervenção, foi possível a construção do teste dinâmico, que compreendeu a junção dessas três partes. Então o teste dinâmico foi testado e o Estudo 2 abrange a verificação das propriedades psicométricas e evidências de validade para esse instrumento.



Ressalta-se que os dois estudos utilizaram amostras virtualmente independentes. Somente no estudo dois os sujeitos que compõem a 5ª série são praticamente os mesmos do estudo piloto do Estudo 1. O tempo entre uma aplicação e outra foi de quatro meses. Escolheu-se aplicar novamente nessa amostra, pois tais indivíduos já haviam respondido os testes utilizados no estudo três como critérios. Como o tempo para a aplicação de todos os testes em uma nova amostra dessa série estava se tornando escasso decorrente o final do ano, optou-se por aplicar o teste dinâmico nesses sujeitos da 5ª série.

## **2.1 - Estudo 1. Construção de itens indutivos e elaboração da intervenção para a resolução desses itens**

O Estudo 1 teve como objetivo a construção do teste dinâmico. O trabalho foi realizado para a elaboração de itens que pudessem compor o pré-teste e o pós-teste do instrumento dinâmico, bem como desenvolvimento do procedimento de intervenção, que juntos formam o teste dinâmico. Dessa maneira o Estudo 1 foi dividido em duas partes, a primeira abrangendo a construção dos itens e a segunda abordando o desenvolvimento da intervenção.

### **2.1.1 – Estudo 1 – Parte 1**

Essa primeira parte do Estudo 1 envolveu a construção dos itens que fazem parte do pré e pós teste no instrumento dinâmico. Os objetivos específicos dessa etapa foram: a) elaborar os itens para o instrumento planejado; b) analisar a validade de conteúdo da versão prévia do instrumento, mas ainda na forma estática; c) explorar a estrutura interna do teste por meio da análise fatorial; d) verificar o coeficiente de precisão do teste na forma estática; e e) analisar, por meio da Teoria de Resposta ao Item, os parâmetros de dificuldade dos itens.

#### **2.1.1.1 – Estudo 1 – Parte 1 – Estudo Piloto**

Inicialmente foram construídos 59 itens e após análise de validade do conteúdo, por meio da teoria que os embasam, dividiu-se esses itens construídos em duas Formas, A e B.

Em seguida, foi conduzido o Estudo Piloto, com o objetivo de verificar se as Formas A e B já apresentavam, por meio de análise quantitativa, propriedades psicométricas mínimas para que fossem utilizadas com a população maior.

### **2.1.1.2 – Participantes**

Participaram do estudo piloto 71 sujeitos, cursando a 5ª série do ensino fundamental de uma escola pública. A média de idade foi de 10,92 com DP= 0,77. Do total da amostra 47,9% é do sexo feminino e 49,3% responderam a Forma A. Desses indivíduos 52,1% estudavam na 5ª série A e o restante na C. Para esse momento não houve critério de exclusão e a amostra foi por conveniência, tendo sido avaliadas todas as crianças de 5ª série dessa escola cujos responsáveis autorizaram.

### **2.1.1.3- Materiais**

#### **2.1.1.3.1- Teste de Raciocínio Indutivo para Crianças**

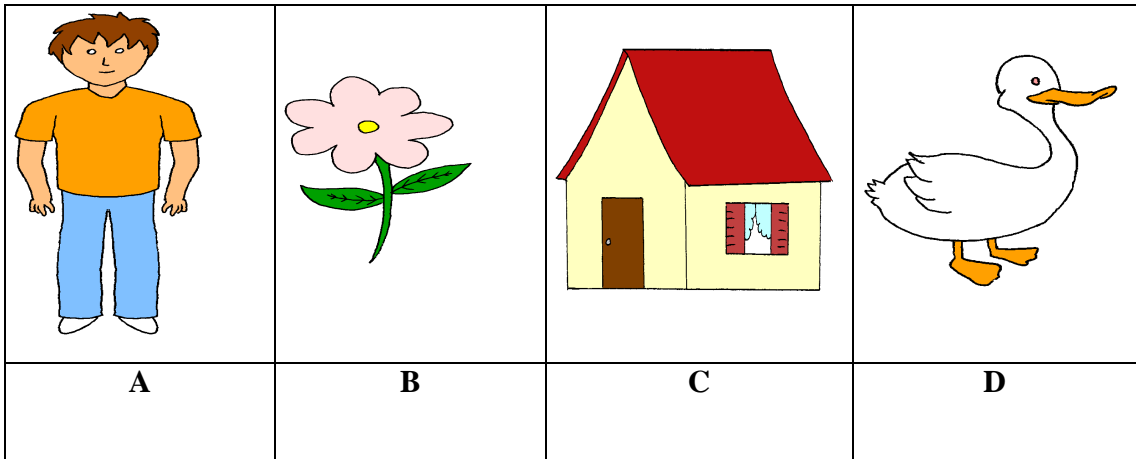
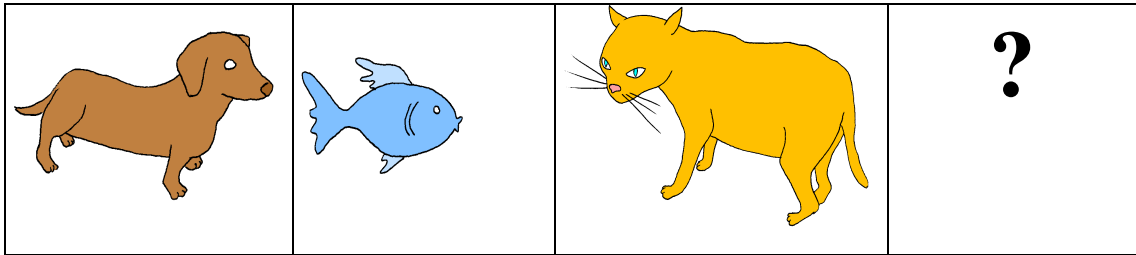
Inicialmente foram construídos 59 itens de raciocínio indutivo baseados na Teoria Prescritiva do Raciocínio Indutivo proposta por Karl Josef Klauer (1990) que visa o treinamento da aptidão, no caso, o raciocínio indutivo. Conforme já explicitado na introdução o raciocínio indutivo pode ser explicado por paradigmas extremamente relacionados, quais sejam, generalização/discriminação (SD), classificação cruzada (CC), reconhecimento/diferenciação de relações (RE) e construção de sistemas (SC). Para cada paradigma foram elaborados 12, 12, 17 e 18 itens respectivamente. Assim como também

explicado na introdução, os paradigmas, do primeiro ao último, vão se complexificando, sendo que o primeiro, generalização/discriminação, é mais fácil que o último, construção de sistemas. Cada item contém quatro alternativas de respostas, sendo apenas uma a correta. O conteúdo dos itens é pictórico, por exemplo, cachorro, borboleta, casa, carro, ou geométrico, como quadrado e triângulo. Nos itens com conteúdo pictórico, assegurou-se que todos os itens a serem representados fossem de alta familiaridade para as crianças por meio da seleção dos substantivos a partir da Lista de Frequência de Palavras (Associação Brasileira de Dislexia, 1995). Nessa lista são apresentadas, separadamente, palavras de alta, média e baixa frequência na língua portuguesa escrita para os níveis de escolaridade da 1ª à 4ª série do ensino fundamental. Para os itens geométricos esse procedimento não foi necessário, pois são considerados neutros (Winfred & Woerh, 1993).

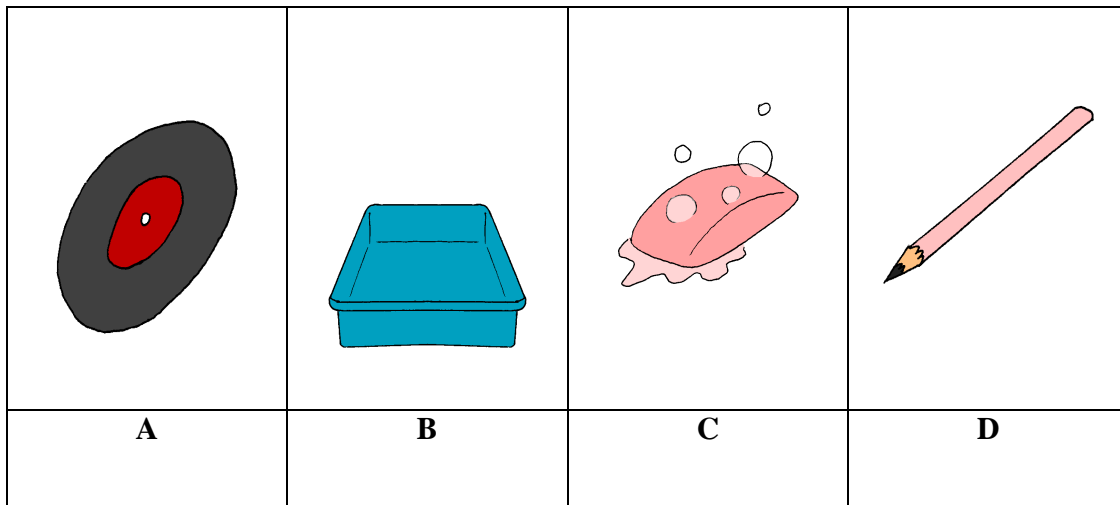
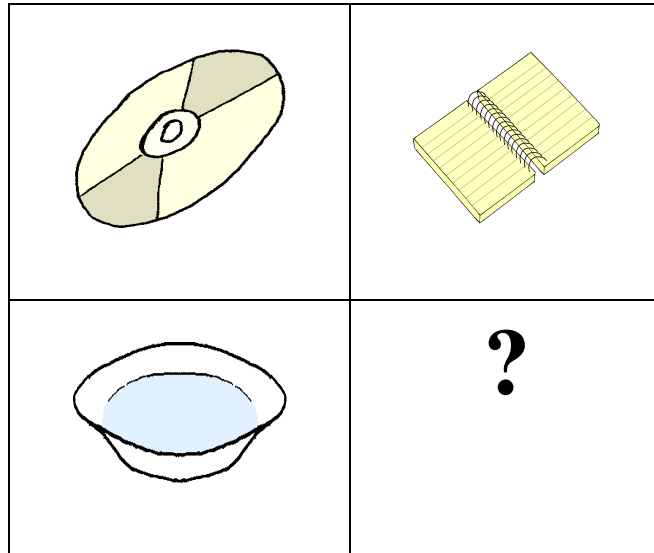
As figuras foram primeiramente desenhadas à mão, depois escaneadas para o computador e, em seguida, foram montadas as tarefas com os itens e as alternativas de respostas. Ao todo foram criados 800 desenhos para compor tanto os itens quanto as alternativas. Ressalta-se que os itens do tipo figuras geométricas foram feitas no programa Microsoft Paint.

Do total de 59 itens, oito foram usados para a ancoragem do teste, pois para esse primeiro momento o teste foi subdividido em duas partes, sendo que os itens âncora foram inseridos em ambas as partes. Esse procedimento permitiu que os dois subtestes fossem analisados em uma única escala para a análise de todos os itens. Então, os itens âncoras

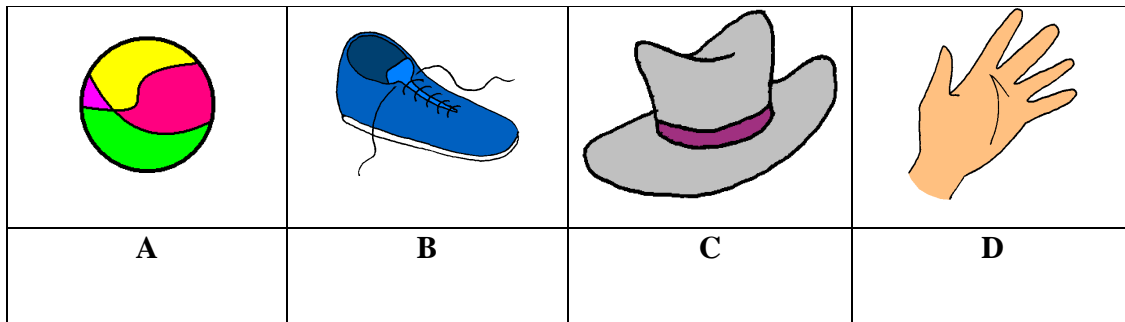
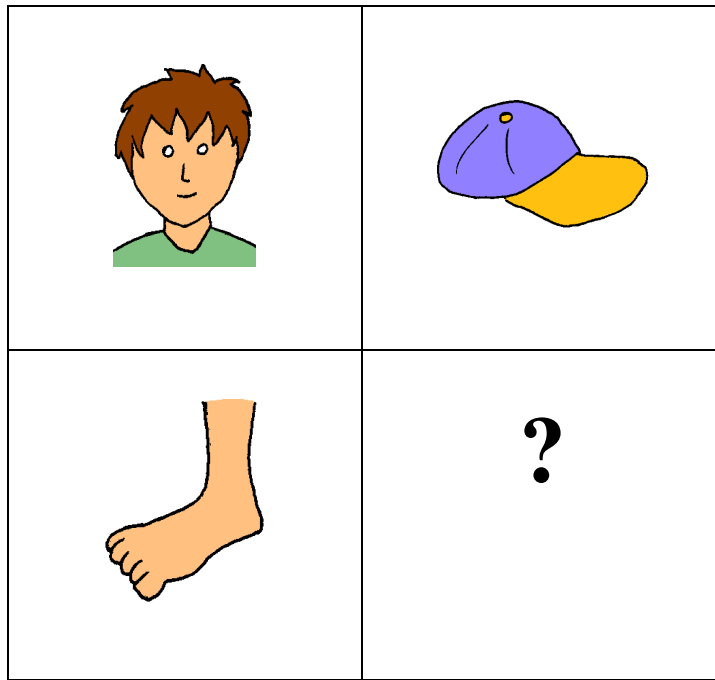
também foram elaborados de forma a abranger todos os paradigmas, bem como abarcar itens fáceis, médios e difíceis de serem resolvidos. Esse também é um pressuposto da ancoragem, ou seja, contemplar todos os níveis de complexidade dos itens. A seguir são apresentados exemplos de cada tipo de tarefa referente a cada um dos paradigmas (Figuras 13, 14, 15 e 16).



**Figura 13-** Tarefa de generalização e discriminação de atributos.

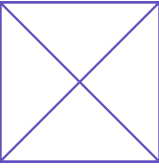
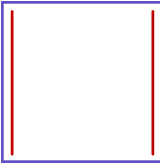


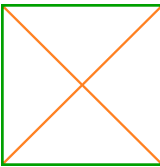





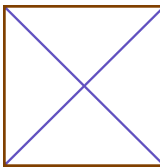

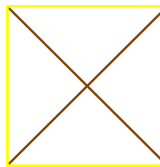
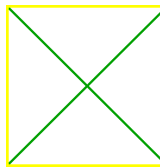
**Figura 14-** Tarefa de Classificação Cruzada.



**Figura 15-** Tarefa de Reconhecimento e Diferenciação de Relações.



		
		
		<b>?</b>

			
<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>

**Figura 16-** Tarefa de Construção de Sistemas.

Os 51 itens, não-âncoras, foram divididos em duas partes ou subtestes, a Forma A, com 26 itens, e a Forma B, com 25. Além desses itens, nas duas formas acrescentaram-se os oito itens âncoras. Assim, ao todo as Formas A e B foram compostas por 34 e 33 itens, respectivamente, distribuídos de forma equilibrada para os paradigmas. Os testes foram administrados na forma estática, em lápis e papel.

Na correção para as Formas estáticas foi atribuído 0 ponto ao errar a alternativa e 1 ponto ao acertar. Os escores emergidos são o escore total e os escores por paradigma, sendo que o escore total corresponde à soma dos acertos em todos os 34 itens Forma A e 33 Forma B e, nos escores por paradigmas, os acertos correspondem às respostas corretas em cada grupo de itens, incluindo SD, CC, RE e SC. O máximo de pontos em cada prova é 34 para a A e 33 para B. Os escores por paradigma são 7 SD, 7 CC, 10 RE e 10 SC na Forma A, e 7 SD, 7 CC, 9 RE e 10 SC na Forma B.

#### **2.1.1.4 - Procedimento**

Mediante autorização do Comitê de Ética da Universidade São Francisco (2003/11767-9), do consentimento da escola e dos pais dos voluntários, o teste foi aplicado às crianças. As Formas A e B do teste foram aplicadas de maneira aleatória dentro das classes, procurando equilibrar a quantidade de respondentes entre as Formas. Antes de iniciar a aplicação se esclareceu o objetivo da testagem. Junto com a Forma a ser respondida era entregue uma folha de resposta, e era indicado que os alunos não deveriam rabiscar os cadernos. O tempo de resposta do teste variou entre 10 e 30 minutos.

### **2.1.1.5. Resultados e discussão do Estudo-piloto**

Conforme já explicitado anteriormente, esse primeiro estudo referiu-se à construção de itens para verificar quais deveriam fazer parte do pré e pós-teste do instrumento dinâmico. Para isso foi necessário: a) elaborar itens para o instrumento planejado; b) analisar a validade de conteúdo da versão prévia do instrumento; c) verificar a estrutura interna dos itens por meio da análise fatorial; d) efetuar uma análise por meio da Teoria de Resposta ao Item para descobrir os parâmetros de dificuldade dos itens; e) verificar o coeficiente de precisão dos itens do instrumento, e f) analisar as estatísticas descritivas da amostra no teste.

Conforme anteriormente explicitado, 59 itens foram elaborados e julgados por juízes, que verificaram sua coerência de acordo com a teoria de Josef Klauer (1990). Após essa análise de conteúdo feita pelos juízes, dividiu-se os itens em dois testes com formas paralelas, Forma A e Forma B privilegiando o equilíbrio da quantidade de itens em cada paradigma do raciocínio indutivo proposto por Klauer (1990). A Forma A ficou composta por 34 itens e a B, por 33 itens. As formas continham oito itens iguais (dois de cada paradigma, quais sejam, similaridade e diferença de atributos, similaridade e diferença de relações, classificação cruzada e sistema de construção). Os oito itens iguais, conhecidos como âncoras, possibilitaram que, na análise com a Teoria de Resposta ao Item por meio do modelo de Rasch, uma análise única fosse conduzida, agrupando as formas numa mesma escala.

Após o desenvolvimento das duas formas do teste de raciocínio indutivo, optou-se pela condução de um estudo-piloto para algumas propriedades psicométricas das duas formas, antes de iniciar a aplicação com maior quantidade de sujeitos. Esse procedimento é uma etapa importante na construção de qualquer instrumento, pois possibilita verificar se a escala construída se encontra minimamente adequada para que seja dada continuidade ao estudo. Caso seja verificado, por meio do estudo-piloto, que o instrumento ainda não agrega confiabilidade suficiente, deve-se melhorá-lo quanto às suas deficiências. Resumindo, o estudo piloto nos direciona quanto ao seguimento para testagem do instrumento (Urbina, 2007).

Os dados coletados no estudo-piloto foram analisados pela Teoria de Resposta ao Item por meio do modelo de Rasch utilizando-se o programa Winsteps, com a finalidade de verificar a precisão do instrumento, a correlação item-theta e os ajustes dos itens em relação à escala. Os ajustes são representados pelos índices *Outfit* e *Infit*. Tais ajustes podem ser verificados pelos parâmetros *Outfit* e *Infit*. *Outfit* refere-se à média dos ajustes dos itens, parâmetro esse mais sensível aos *outliers*, em que o desajuste do item ou sua discrepância ocorre distante do nível de habilidade. Ou seja, os casos de *outfits* referem-se a situações em que o sujeito não acerta itens muito fáceis para sua habilidade, ou então acerta itens muito distantes da sua habilidade. *Infit* também é uma medida de desajuste do item, estando relacionado às discrepâncias próximas ao nível de habilidade do sujeito, ou seja, ele não acerta a itens próximos do seu nível de habilidade (Baker, 2001). É desejável que a precisão seja no mínimo 0,60, a correlação no mínimo 0,30 e os ajustes de no máximo 1,2, este

último conforme Writh e Stone (2004). A Tabela 2 apresenta algumas informações psicométricas do teste de raciocínio indutivo em geral.

**Tabela 2-** Propriedades psicométricas do teste de raciocínio indutivo por meio da análise de Rasch.

	Escore			
	Bruto	Dificuldade	<i>Infit</i>	<i>Outfit</i>
Média	18,0	0,26	1,00	0,97
D.P	4,3	0,78	0,18	0,41
Máximo	27,0	2,30	1,52	3,58
Mínimo	7,0	-1,71	0,70	0,54
Precisão ( <i>Real</i> )=				
0,67				

Pelos dados da Tabela 2 pode-se verificar que a precisão da escala foi de 0,67 e o índice médio de dificuldade dos itens de 0,26, indicando que o instrumento apresenta dificuldade média, sendo que a métrica da escala foi média 0 e desvio padrão 1. Os índices de ajustes, *infit* e *outfit* estão dentro do parâmetro esperado. No entanto é importante verificar as informações psicométricas relacionadas a cada item específico da escala, pois isso permite identificar pontos em que a escala pode ser melhorada. A Tabela 3 especifica essas informações. Pode-se observar que somente em alguns itens houve problemas.

Ressalta-se que os itens considerados problemáticos foram os que contemplavam concomitantemente desajuste e baixa correlação item-theta.

**Tabela 3**– Propriedades psicométricas dos itens do teste de raciocínio indutivo por meio da análise de Rasch.

Nº	Dificuldade	<i>Infit</i>	<i>Outfit</i>	<i>r</i>	Item
1	-2,82	0,87	0,52	0,38	A01SD
2	-2,36	0,81	0,51	0,48	A02SD
3	0,02	1,07	1,07	0,24	A03SD
4	-1,18	0,96	0,85	0,35	C04SD
<b>5</b>	<b>0,94</b>	<b>1,35</b>	<b>1,36</b>	<b>-0,12</b>	<b>A05SD</b>
6	-0,94	0,93	0,88	0,38	C06SD
7	-0,38	0,95	0,92	0,39	A07SD
8	1,36	0,88	0,80	0,47	A08CC
9	0,78	1,07	1,11	0,20	A09CC
<b>10</b>	<b>4,01</b>	<b>1,06</b>	<b>2,08</b>	<b>-0,11</b>	<b>A10CC</b>
<b>11</b>	<b>0,15</b>	<b>1,39</b>	<b>1,54</b>	<b>-0,21</b>	<b>A11CC</b>
12	0,26	1,03	1,04	0,27	C12CC
13	-0,81	0,93	0,88	0,39	A13CC
14	1,39	0,80	0,74	0,57	A14CC
15	-3,56	0,93	0,35	0,31	A15RE
16	-0,97	0,99	0,97	0,30	A16RE
17	0,53	0,95	1,00	0,37	A17RE
18	-2,03	0,97	0,79	0,30	A18RE
19	-1,35	1,00	0,88	0,29	C19RE
20	-0,24	0,99	0,95	0,35	A20RE
21	0,67	0,87	0,84	0,49	A21RE
22	-0,05	0,95	0,92	0,39	C22RE
23	0,53	0,87	0,84	0,50	A23RE
24	-0,51	0,92	0,86	0,43	A24RE
25	-0,86	1,05	1,12	0,19	C25SC
26	-0,66	0,94	0,86	0,40	A26SC
27	2,20	0,98	0,89	0,28	A27SC
28	0,39	0,97	0,96	0,36	C28SC
29	1,09	1,04	1,01	0,28	A29SC
30	0,20	1,04	1,05	0,27	A30SC
31	2,20	1,07	1,31	0,11	A31SC
32	1,76	0,96	0,93	0,33	A32SC

Tabela 3 (continuação) – Propriedades psicométricas dos itens do teste de raciocínio indutivo por meio da análise de Rasch.

Nº	Dificuldade	<i>Infit</i>	<i>Outfit</i>	<i>r</i>	Item
33	-0,11	1,08	1,04	0,24	A33SC
34	4,01	,93	0,38	0,29	A34SC
35	-0,62	1,10	1,03	0,16	B01SD
36	-2,43	1,02	0,94	0,15	B02SD
37	-0,42	0,96	0,91	0,36	B03SD
38	0,76	1,15	1,17	0,06	B05SD
39	-,17	0,95	0,92	0,38	B07SD
40	1,03	1,07	1,16	0,14	B08CC
41	1,01	1,12	1,15	0,09	B10CC
<b>42</b>	<b>0,63</b>	<b>1,29</b>	<b>1,40</b>	<b>-0,17</b>	<b>B11CC</b>
43	-1,83	0,92	0,74	0,37	B13CC
44	0,89	1,02	1,00	0,26	B14CC
45	-1,34	0,91	0,81	0,39	B15RE
46	0,89	0,95	0,89	0,38	B16RE
47	-0,11	0,82	0,78	0,58	B17RE
48	-1,37	1,01	1,01	0,22	B18RE
<b>49</b>	<b>3,14</b>	<b>1,14</b>	<b>3,78</b>	<b>-0,44</b>	<b>B20RE</b>
50	-1,80	0,88	0,65	0,44	B21RE
51	-1,83	0,89	0,84	0,37	B23RE
52	-0,91	0,85	0,79	0,50	B24SC
53	-1,61	1,06	1,06	0,12	B26SC
54	0,25	1,03	1,03	0,25	B27SC
55	0,89	0,99	0,99	0,30	B29SC
56	1,67	0,98	0,90	0,30	B30SC
57	-1,41	1,06	1,06	0,14	B31SC
58	0,01	1,04	1,10	0,21	B32SC
59	1,03	1,05	1,02	0,22	B33SC

Os resultados da Tabela 3 demonstraram que cinco itens apresentaram desajustes e correlação item-theta nula ou negativa. Ao analisar quantitativa e qualitativamente tais



itens, observou-se que dois deles, os itens 11 e 49, poderiam ter mais que uma resposta correta. Diante disso decidiu-se excluir esses itens. Os três outros itens, de números 5, 10 e 42, apresentaram problemas na escolha da alternativa correta, ou seja, a maioria das crianças, dentre elas as mais habilidosas, escolhiam outra alternativa como sendo a correta. Para esses três itens a resposta correta foi recodificada visto que, de fato, a resposta apontada pelas crianças também atendia aos critérios de correção. A recodificação consistiu na mudança da alternativa correta, ou seja, na nova análise foram consideradas corretas as respostas efetuadas pelas crianças. Em seguida, fez-se a análise novamente, excluindo os dois itens e recodificando a resposta dos outros cinco. Ressalta-se que o item 31 apresentou índice Outfit um pouco elevado e correlação baixa, porém ao analisá-lo qualitativamente, constatou-se que não apresentava problemas semelhantes aos outros cinco itens citados, então decidiu-se mantê-lo como estava. A Tabela 4 mostra os novos parâmetros para esses itens.

**Tabela 4**– Propriedades psicométricas dos itens do teste de raciocínio indutivo, considerando os recodificados e os excluídos, por meio da análise de Rasch.

Nº	Dificuldade	Infit	Outfit	r	Item
1	-2,75	0,91	0,61	0,32	A01SD
2	-2,29	0,83	0,63	0,42	A02SD
3	0,15	1,09	1,11	0,27	A03SD
4	-1,08	0,98	0,84	0,36	C04SD
<b>5</b>	<b>1,94</b>	<b>1,06</b>	<b>0,95</b>	<b>0,26</b>	<b>A05SD</b>
6	-0,83	0,91	0,86	0,42	C06SD
7	-0,26	0,88	0,85	0,49	A07SD
8	1,53	0,89	0,84	0,46	A08CC
9	0,94	1,11	1,18	0,21	A09CC
<b>10</b>	<b>-0,55</b>	<b>1,09</b>	<b>1,37</b>	<b>0,20</b>	<b>A10CC</b>
<b>11</b>	<b>EXCLUÍDO</b>				<b>A11CC</b>
12	0,41	1,08	1,09	0,27	C12CC
13	-0,70	0,95	0,93	0,38	A13CC
14	1,56	0,81	0,79	0,54	A14CC
15	-3,49	0,91	0,30	0,33	A15RE
16	-0,86	1,05	1,03	0,27	A16RE
17	0,68	0,98	1,05	0,37	A17RE
18	-1,95	0,95	0,76	0,34	A18RE
19	-1,26	1,02	0,93	0,28	C19RE
20	-0,12	1,01	0,99	0,35	A20RE
21	0,82	0,87	0,82	0,51	A21RE
22	0,09	0,98	0,96	0,38	C22RE
23	0,68	0,87	0,83	0,52	A23RE
24	-0,40	0,95	0,86	0,43	A24RE
25	-0,75	1,10	1,25	0,17	C25SC
26	-0,55	0,93	0,82	0,45	A26SC
27	2,39	1,02	1,02	0,24	A27SC
28	0,54	1,03	1,01	0,33	C28SC
29	1,25	1,05	1,02	0,30	A29SC
30	0,33	1,06	1,08	0,30	A30SC
31	2,39	1,12	1,57	0,07	A31SC

Tabela 4 (continuação) – Propriedades psicométricas dos itens do teste de raciocínio indutivo, considerando os recodificados e os excluídos, por meio da análise de Rasch.

Nº	Dificuldade	<i>Infit</i>	<i>Outfit</i>	<i>r</i>	Item
32	1,94	1,02	1,03	0,27	A32SC
33	0,02	1,09	1,05	0,28	A33SC
34	4,21	0,94	0,38	0,27	A34SC
35	-0,50	1,18	1,09	0,13	B01SD
36	-2,35	1,07	1,07	0,09	B02SD
37	-0,27	0,92	0,85	0,43	B03SD
38	1,21	1,01	1,18	0,30	B05SD
39	-0,03	0,96	0,90	0,40	B07SD
40	1,21	1,10	1,16	0,21	B08CC
41	1,35	1,09	1,06	0,24	B10CC
<b>42</b>	<b>-0,23</b>	<b>0,94</b>	<b>0,87</b>	<b>0,42</b>	<b>B11CC</b>
43	-1,74	0,96	0,80	0,30	B13CC
44	1,06	1,08	1,04	0,27	B14CC
45	-1,21	0,95	0,86	0,33	B15RE
46	1,06	1,04	0,97	0,32	B16RE
47	0,02	0,85	0,79	0,52	B17RE
48	-1,27	1,00	0,91	0,28	B18RE
<b>49</b>	<b>EXCLUÍDO</b>				<b>B20RE</b>
50	-1,71	0,83	0,55	0,48	B21RE
51	-1,74	0,82	0,82	0,42	B23RE
52	-0,80	0,82	0,73	0,52	B24SC
53	-1,51	1,10	1,12	0,13	B26SC
54	0,40	1,10	1,12	0,23	B27SC
55	1,06	1,03	1,06	0,30	B29SC
56	1,89	0,95	0,80	0,41	B30SC
57	-1,31	1,06	0,90	0,24	B31SC
58	0,15	1,15	1,32	0,12	B32SC
59	1,21	1,14	1,10	0,20	B33SC

Como pode ser observado na Tabela 4, os novos resultados mostraram melhores ajustes e maior correlação item-theta dos três itens recodificados. Também é importante destacar que a precisão da escala aumentou para 0,71. Frente a esses dados que sustentam adequadas propriedades psicométricas para as Formas A e B, continuou-se o estudo alterando as alternativas dos itens recodificados para que somente uma resposta fosse a correta. Essa alteração foi realizada modificando o desenho da alternativa que era a correta inicialmente. Foram também acrescentados nove novos itens para que o teste tivesse mais opções de tarefas e também para as Formas ficarem equiparadas na quantidade de itens. Então acrescentou-se itens quatro na Forma A (1SD, 1RE e 2CC) e cinco na Forma B (2SD, 2CC e 1 RE). Assim, cada forma ficou composta por 38 itens.

### **2.1.2 – Estudo 1 – Parte 1 – Aplicação principal. Seleção da amostra de itens para a composição do pré e pós-teste do instrumento dinâmico**

Após o Estudo Piloto, o Teste de Raciocínio Indutivo, nas Formas A e B, com 38 itens cada, foi aplicado em uma amostra com maior número de sujeitos. Tendo como objetivos os já citados, quais sejam explorar a estrutura interna dos itens por meio da análise fatorial, verificar o coeficiente de precisão do instrumento na forma estática e analisar, por meio da Teoria de Resposta ao Item, os parâmetros de dificuldade dos itens. Pois por meio dessa parte do estudo foram escolhidos os 16 itens que compuseram o pré e os 16 que formaram o pós.

#### **2.1.2.1 – Participantes**

Participaram da Parte 1 do presente estudo 417 crianças, amostra por conveniência, sendo que 219 responderam a Forma A. Desse total da amostra, 167 eram de escola pública e 250, de escola particular. Na amostra da escola pública, 87 sujeitos responderam a Forma A e 80, a B. Na amostra da escola particular, 132 responderam a Forma A e 117, a B.

As crianças cursavam da 1ª à 5ª série, sendo que a aplicação na 5ª série ocorreu somente na escola particular. Entre as séries, na escola pública foram avaliados 23, 48, 42 e 54 das séries 1ª, 2ª, 3ª e 4ª, respectivamente. Na escola particular foram 58, 71, 49, 33 e 39 nas séries 1ª, 2ª, 3ª, 4ª e 5ª. A média de idade para cada série pode ser visualizada pela Tabela 5.

**Tabela 5** – Média e desvio padrão das idades por série e tipo de escola.

Série	Média	Desvio Padrão
<b>Escola particular</b>		
1 <sup>a</sup>	6,62	0,49
2 <sup>a</sup>	7,49	0,53
3 <sup>a</sup>	8,71	0,68
4 <sup>a</sup>	9,58	0,56
5 <sup>a</sup>	10,54	0,50
<b>Escola pública</b>		
1 <sup>a</sup>	6,78	0,52
2 <sup>a</sup>	7,81	0,45
3 <sup>a</sup>	9,0	0,70
4 <sup>a</sup>	10,15	1,00

Antes de iniciar a aplicação, perguntava-se para a professora sobre possíveis alunos que pudessem apresentar algum problema que interferisse negativamente no resultado do teste, como, por exemplo, deficiência mental ou distúrbios de aprendizagem. Dentro da amostra, foram identificadas três crianças disléxicas conforme informação dada pelas professoras. Os protocolos desses alunos foram marcados e, depois, excluídos do banco de dados, especialmente porque não conseguiram resolver as tarefas, tendo sido ajudados pelas professoras.

### **2.1.2.2 - Materiais**

#### **2.1.2.2.1- Teste de Raciocínio Indutivo para Crianças**

Após algumas modificações necessárias frente às informações obtidas no estudo piloto, a versão do Teste de Raciocínio Indutivo para Crianças ficou composta por 68 itens, e os paradigmas ficaram representados por 17 itens SD, 18 CC, 21 RE e 20 SC. Os 8 itens âncoras continuaram e foram somados na representação dos paradigmas. As Formas A e B ficaram compostas por 38 itens cada já com os âncoras incorporados, mantendo o equilíbrio da quantidade de itens por paradigma. Os testes foram administrados na forma estática, em lápis e papel. A forma dinâmica e computadorizada do teste foi apenas posteriormente criada, a partir dos melhores itens verificados nessa forma estática.

A correção para as Formas estáticas compreendeu em 0 ponto ao errar a alternativa correta e 1 ponto ao acertar. Os escores emergidos são total e por paradigma, no total se soma os acertos dos 38 itens e por paradigmas os acertos correspondentes aos itens SD, CC, RE e SC. O máximo de pontos em cada prova é 38 e por paradigma 8 SD, 11 RE, 9 CC e 10 SC na Forma A e na Forma B são 9 SD, 10 RE, 9 CC e 10 SC.

#### **2.1.2.4 - Procedimento**

Após a autorização do Comitê de Ética da Universidade São Francisco (2003/11767-9), do consentimento das escolas e dos pais dos voluntários, os testes foram aplicados em amostras diferentes, porém equivalentes em suas características como série escolar e sexo. A aplicação foi coletiva, em grupos de aproximadamente 30 alunos de cada classe, realizada por uma psicóloga treinada e pela autora do projeto. Antes do início da aplicação explicou-se o objetivo do projeto, bem como se pediu para que as crianças

indicassem caso houvesse algum desenho desconhecido ou que elas não sabiam identificar no teste.

Após esse primeiro contato, foram entregues os cadernos de teste para cada criança, sendo que metade da classe respondeu a Forma A e a outra metade a Forma B, essa divisão foi aleatória. Uma das aplicadoras leu junto com as crianças as instruções do teste, que são as mesmas para as duas formas. Tanto na Forma A quanto na B, contém três exemplos para ser resolvidos junto com as crianças. Após a compreensão da tarefa pelos sujeitos, iniciou-se o teste e as respostas foram marcadas em uma folha de resposta. As aplicadoras permaneceram na sala para esclarecimento de eventuais dúvidas e para verificar se realmente compreenderam a tarefa. O tempo de aplicação variou de 10 a 40 minutos, sendo que as séries iniciais demoravam mais para responder.

#### **2.1.2.5 - Resultados e discussão do Estudo 1, Parte 1- Aplicação principal**

As análises do Estudo 1, aplicação principal, foram conduzidas de maneira mais detalhada em relação ao estudo-piloto. Isso porque as análises tiveram como objetivo permitir a seleção dos itens para compor o pré e pós-teste do instrumento dinâmico no Estudo 2, a partir de uma amostra maior e mais diversificada em relação às séries.

Inicialmente encontram-se apresentadas as estatísticas descritivas e inferenciais da amostra no teste em função de série. Em seguida, os resultados da análise fatorial que procurou investigar a dimensionalidade da escala. Posteriormente é descrita a precisão da



escala e, por fim, novamente utilizando-se da Teoria de Resposta ao Item por meio do modelo de Rasch, são descritos os parâmetros de dificuldade dos itens, juntamente com correlação item-theta e índices de ajustes. Essa última análise foi a principal fonte na escolha dos itens. Ressalta-se que pela Teoria de Resposta ao Item, por meio dos itens âncoras que permaneceram nas duas formas, a análise foi realizada juntando as duas formas.

#### **2.1.2.5.1. Estatísticas descritivas e inferenciais no Teste de Raciocínio Indutivo em função de série.**

Com relação às estatísticas descritivas foi feito um escore total para Forma A e outro para a Forma B, já que cada metade da amostra respondeu a um tipo de forma. Também foram criados escores de cada paradigma, em cada forma, somando-se os itens respectivos a SD, CC, RE e SC. A Tabela 6 mostra as estatísticas descritivas referentes aos escores mencionados. Vale destacar que a pontuação máxima para cada um dos escores da Forma A total, Forma B total, SD\_A total, SD\_B total, CC\_A total, CC\_B total, RE\_A total, RE\_B total, SC\_A total e SC\_B total é respectivamente 38, 38, 8, 9, 9, 9, 11, 10, 10 e 10.

**Tabela 6** – Estatísticas descritivas dos escores das Formas A e B.

Variáveis	N	Mínimo	Máximo	Média	Desvio Padrão
A_TT	219	3	35	17,69	6,81
B_TT	197	5	30	17,27	6,25
T_A_SD	219	0	8	4,86	1,88
T_B_SD	198	0	9	4,34	2,28
T_A_CC	219	0	8	3,40	1,98
T_B_CC	198	1	8	3,66	1,54
T_A_RE	219	0	11	5,81	2,55
T_B_RE	197	1	9	4,91	1,89
T_A_SC	219	0	9	3,62	2,07
T_B_SC	198	0	10	4,35	2,39

*Nota: A\_TT (escore total Forma A), B\_TT (escore total Forma B), T\_A\_SD (paradigma SD da Forma A), T\_B\_SD (paradigma SD da Forma B), T\_A\_CC (paradigma CC da Forma A), T\_B\_CC (paradigma CC da Forma B), T\_A\_RE (paradigma RE da Forma A), T\_B\_RE (paradigma RE da Forma B), T\_A\_SC (paradigma SC da Forma A), T\_B\_SC (paradigma SC da Forma B).*

Observando a Tabela 6, nota-se que, ao comparar a amostra que respondeu a Forma A com a que respondeu a Forma B, as médias em todos os escores totais e por paradigma são bem próximas. No entanto, como a amostra contempla uma diversidade de séries, é importante ter a descrição dessas estatísticas por série, dessa forma também é possível identificar se as médias nos escores continuam equiparadas como quando verifica-se a amostra como um todo. Visto que as séries abrangem uma faixa etária dos 6 aos 11 anos, espera-se que as médias aumentem conforme maior a escolaridade, pois o teste aborda um construto ligado à inteligência e esta tende a aumentar com o passar dos anos nessa faixa etária indicada. A Tabela 7 apresenta as estatísticas descritivas dos escores de acordo com cada série.

**Tabela 7**– Estatísticas descritivas dos escores das Formas A e B separando por série.

Variáveis	N	Mínimo	Máximo	Média	D.P
<b>1ª Série</b>					
A_TT	45	6	25	13,20	4,40
B_TT	36	7	26	12,53	4,01
T_A_SD	45	1	7	3,98	1,57
T_B_SD	36	0	8	2,94	1,80
T_A_CC	45	0	7	2,56	1,42
T_B_CC	36	1	7	3,31	1,19
T_A_RE	45	0	9	4,29	2,13
T_B_RE	36	1	7	3,69	1,67
T_A_SC	45	0	5	2,38	1,33
T_B_SC	36	0	6	2,58	1,44
<b>2ª Série</b>					
A_TT	61	3	31	16,59	6,31
B_TT	57	5	27	15,25	5,69
T_A_SD	61	0	8	4,46	1,87
T_B_SD	58	0	8	3,66	2,24
T_A_CC	61	0	8	3,25	1,97
T_B_CC	58	1	7	3,17	1,40
T_A_RE	61	1	10	5,54	2,23
T_B_RE	57	1	9	4,61	1,84
T_A_SC	61	0	9	3,34	2,01
T_B_SC	58	0	9	3,76	2,19
<b>3ª Série</b>					
A_TT	47	6	30	19,00	6,75
B_TT	44	7	30	19,73	5,96
T_A_SD	47	1	8	5,36	1,84
T_B_SD	44	1	9	5,39	2,10
T_A_CC	47	0	8	3,66	2,07
T_B_CC	44	1	7	4,18	1,68
T_A_RE	47	2	10	6,09	2,64
T_B_RE	44	2	8	5,09	1,65
T_A_SC	47	0	9	3,89	2,06
T_B_SC	44	1	10	5,07	2,34

Tabela 7 (continuação) – Estatísticas descritivas dos escores das Formas A e B separando por série.

Variáveis	N	Mínimo	Máximo	Média	D.P
<b>4ª Série</b>					
A_TT	44	8	31	19,00	6,40
B_TT	43	9	29	19,47	5,63
T_A_SD	44	1	8	5,18	1,85
T_B_SD	43	1	9	4,74	2,12
T_A_CC	44	0	8	3,64	1,90
T_B_CC	43	1	7	3,84	1,46
T_A_RE	44	1	10	6,14	2,34
T_B_RE	43	2	9	5,58	1,81
T_A_SC	44	1	9	4,05	1,96
T_B_SC	43	1	10	5,30	2,36
<b>5ª Série</b>					
A_TT	22	12	35	24,55	6,27
B_TT	17	13	30	22,24	5,35
T_A_SD	22	3	8	6,09	1,60
T_B_SD	17	3	9	5,88	1,76
T_A_CC	22	0	8	4,50	2,32
T_B_CC	17	2	8	4,24	1,88
T_A_RE	22	4	11	8,45	2,06
T_B_RE	17	2	9	6,29	1,72
T_A_SC	22	2	9	5,50	2,04
T_B_SC	17	2	9	5,82	2,18

*Nota: A\_TT (escore total Forma A), B\_TT (escore total Forma B), T\_A\_SD (paradigma SD da Forma A), T\_B\_SD (paradigma SD da Forma B), T\_A\_CC (paradigma CC da Forma A), T\_B\_CC (paradigma CC da Forma B), T\_A\_RE (paradigma RE da Forma A), T\_B\_RE (paradigma RE da Forma B), T\_A\_SC (paradigma SC da Forma A), T\_B\_SC (paradigma SC da Forma B).*

Analisando essa Tabela 7, pode-se notar que há, como esperado, um aumento nas médias das pontuações dos escores em relação às séries. Somente entre a 3ª e 4ª series a media se manteve, não sendo possível identificar esse aumento. Para a verificar se essas

diferenças observadas entre as séries são significativas, efetuou-se uma ANOVA. Os resultados encontram-se sumariados na Tabela 8.

**Tabela 8** – Resultados da ANOVA da diferença entre séries nos escores das Formas A e B do teste de raciocínio indutivo.

Variáveis	<i>gl</i>	<i>F</i>	<i>p</i>
A_TT	4	14,63	0,00
B_TT	4	16,27	0,00
T_A_SD	4	7,44	0,00
T_B_SD	4	11,19	0,00
T_A_CC	4	4,45	0,00
T_B_CC	4	4,16	0,00
T_A_RE	4	12,60	0,00
T_B_RE	4	9,06	0,00
T_A_SC	4	11,31	0,00
T_B_SC	4	12,45	0,00

*Nota: A\_TT (escore total Forma A), B\_TT (escore total Forma B), T\_A\_SD (paradigma SD da Forma A), T\_B\_SD (paradigma SD da Forma B), T\_A\_CC (paradigma CC da Forma A), T\_B\_CC (paradigma CC da Forma B), T\_A\_RE (paradigma RE da Forma A), T\_B\_RE (paradigma RE da Forma B), T\_A\_SC (paradigma SC da Forma A), T\_B\_SC (paradigma SC da Forma B).*

Conforme a Tabela 8, houve diferença significativa entre os grupos de séries analisadas, quais sejam, 1ª, 2ª, 3ª, 4ª e 5ª nos escores das Formas A e B. Para identificar entre quais séries houve diferenças significativas, foram conduzidas análises de Tukey. Os resultados podem ser visualizados nas Tabelas 9, 10, 11, 12, 13, 14, 15, 16, 17 e 18.

**Tabela 9**– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore total da Formas A.

Variável dependente	(I) Série2	(J) Série2	Diferença de média	<i>p</i>
A_TT	1	2	-3,39*	0,04
		3	-5,80*	0,00
		4	-5,80*	0,00
		5	-11,34*	0,00
	2	1	3,39*	0,04
		3	-2,41	0,25
		4	-2,41	0,26
		5	-7,95*	0,00
	3	1	5,80*	0,00
		2	2,41	0,25
		4	0,00	1,00
		5	-5,54*	0,00
	4	1	5,80*	0,00
		2	2,41	0,26
		3	0,00	1,00
		5	-5,54*	0,00
	5	1	11,34*	0,00
		2	7,95*	0,00
		3	5,54*	0,00
		4	5,54*	0,00

\* $p < 0,05$

Observando a Tabela 9 constata-se que para o escore total da Forma A houve diferença significativa entre séries, exceto entre as séries 2<sup>a</sup> e 3<sup>a</sup>, 2<sup>a</sup> e 4<sup>a</sup>, 3<sup>a</sup> e 4<sup>a</sup>. Tal resultado sugere que 2<sup>a</sup>, 3<sup>a</sup>, 4<sup>a</sup> série tendem a obter desempenhos semelhantes no escore total da Forma A. Os dados obtido pelas séries na Forma B podem ser visualizados na Tabela 10.

**Tabela 10**– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore total da Formas B.

Variável dependente	(I) Série2	(J) Série2	Diferença de média	<i>p</i>
B_TT	1	2	-2,71	0,13
		3	-7,20*	0,00
		4	-6,93*	0,00
		5	-9,70*	0,00
	2	1	2,71	0,13
		3	-4,48*	0,00
		4	-4,22*	0,00
		5	-6,99*	0,00
	3	1	7,20*	0,00
		2	4,48*	0,00
		4	0,26	0,99
		5	-2,50	0,49
	4	1	6,93*	0,00
		2	4,22*	0,00
		3	-0,26	0,99
		5	-2,77	0,39
	5	1	9,70*	0,00
		2	6,99*	0,00
		3	2,50	0,49
		4	2,77	0,39

\* $p < 0,05$

A Tabela 10 possibilita perceber que os resultados da ANOVA com as séries para a Forma B se diferiram dos dados verificados com a Forma A. Na Forma B a média da 5ª série não se diferenciou significativamente com as séries 3ª e 4ª. O mesmo foi observado entre a 1ª e 2ª série. Também notou-se que na Forma B 3ª e 4ª apresentaram medias semelhantes, porem se diferenciaram significativamente da 2ª série. A próxima Tabela traz as informações da ANOVA em relação ao paradigma SD na Forma A.

**Tabela 11**– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Forma A considerando o paradigma SD.

Variável dependente	(I) Série2	(J) Série2	Diferença de média	<i>p</i>
T_A_SD	1	2	0,-48	0,64
		3	-1,38*	0,00
		4	-1,20*	0,01
		5	-2,11*	0,00
	2	1	0,48	0,64
		3	-0,90	0,07
		4	-0,72	0,24
		5	-1,63*	0,00
	3	1	1,38*	0,00
		2	0,90	0,07
		4	0,18	0,98
		5	-0,72	0,50
	4	1	1,20*	0,01
		2	0,72	0,24
		3	-0,18	0,98
		5	-0,90	0,29
	5	1	2,11*	0,00
		2	1,63*	0,00
		3	0,72	0,50
		4	0,90	0,29

\* $p < 0,05$

Com o escore do paradigma SD na Forma A nota-se que as diferenças de média significativas ocorreram somente entre a 5ª série em relação a 1ª e 2ª e entre a 1ª com a maioria das séries, exceto com a 2ª. Pela Forma B os resultados com o SD estão contidos na Tabela 12.



**Tabela 12**– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Formas B considerando o paradigma SD.

Variável dependente	(I) Série2	(J) Série2	Diferença de média	<i>p</i>
T_B_SD	1	2	-0,71	0,49
		3	-2,44*	0,00
		4	-1,80*	0,00
		5	-2,93*	0,00
	2	1	0,71	0,49
		3	-1,73*	0,00
		4	-1,08	0,07
		5	-2,22*	0,00
	3	1	2,44*	0,00
		2	1,73*	0,00
		4	0,64	0,60
		5	-0,49	0,91
	4	1	1,80*	0,00
		2	1,08	0,07
		3	-0,64	0,60
		5	-1,13	0,31
	5	1	2,93*	0,00
		2	2,22*	0,00
		3	0,49	0,91
		4	1,13	0,31

\* $p < 0,05$

No paradigma SD pela Forma B observa-se que apenas um resultado se diferiu quando visto esse paradigma na Forma A. Pela Tabela 12 constata-se a diferença de média significativa entre a 2<sup>a</sup> e 3<sup>a</sup> séries. A Tabela 13 apresenta os resultados da ANOVA decorrente do paradigma CC na Forma A.

**Tabela 13**– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Forma A considerando o paradigma CC.

Variável dependente	(I) Série2	(J) Série2	Diferença de média	<i>p</i>
T_A_CC	1	2	-0,69	0,36
		3	-1,10*	0,05
		4	-1,08	0,06
		5	-1,94*	0,00
	2	1	0,69	0,36
		3	-0,41	0,80
		4	-0,39	0,84
		5	-1,25	0,07
	3	1	1,10*	0,05
		2	0,41	0,80
		4	0,02	1,00
		5	-0,84	0,44
	4	1	1,08	0,06
		2	0,39	0,84
		3	-0,02	1,00
		5	-0,86	0,42
	5	1	1,944*	0,00
		2	1,25	0,07
		3	0,84	0,44
		4	0,86	0,42

\* $p < 0,05$

Na Tabela 13 se verifica diferença significativa de média apenas entre a 1<sup>a</sup> com a 3<sup>a</sup> e 5<sup>a</sup> série. A Tabela 14 traz os resultados para o paradigma CC na Forma B.

**Tabela 14**– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore total da Forma B considerando o paradigma CC.

Variável dependente	(I) Série2	(J) Série2	Diferença de média	<i>p</i>
T_B_CC	1	2	0,13	0,99
		3	-0,87	0,07
		4	-0,53	0,51
		5	-0,93	0,21
		2	-0,13	0,99
	2	3	-1,00*	0,00
		4	-0,66	0,18
		5	-1,06	0,07
		1	0,87	0,07
		2	1,00*	0,00
	3	4	0,34	0,81
		5	-0,05	1,00
		1	0,53	0,51
		2	0,66	0,18
		3	-0,34	0,81
	4	5	-0,39	0,88
		1	0,93	0,21
		2	1,06	0,07
		3	0,05	1,00
		4	0,39	0,88

\* $p < 0,05$

Nota-se que com o paradigma CC na Forma B houve apenas uma diferença de média significativa que incluiu as séries 1ª e 3ª. As diferenças das séries para o escore do paradigma RE na Forma A podem ser visualizadas na Tabela 15.

**Tabela 15**– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Forma A considerando o paradigma RE.

Variável dependente	(I) Série2	(J) Série2	Diferença de média	<i>p</i>
T_A_RE	1	2	-1,25	0,05
		3	-1,79*	0,00
		4	-1,84*	0,00
		5	-4,16*	0,00
	2	1	1,25	0,05
		3	-0,54	0,74
		4	-0,59	0,69
		5	-2,91*	0,00
	3	1	1,79*	0,00
		2	0,54	0,74
		4	-0,05	1,00
		5	-2,36*	0,00
	4	1	1,84*	0,00
		2	0,59	0,69
		3	0,05	1,00
		5	-2,31*	0,00
	5	1	4,16*	0,00
		2	2,91*	0,00
		3	2,36*	0,00
		4	2,31*	0,00

\* $p < 0,05$

Observa-se pela Tabela 15 que a 1ª série apresentou média significativamente diferente com a maioria das série, exceto com a 2ª. A 5ª série também apresentou medias significativamente diferentes em relação as outras séries. Na Forma B os resultados em função do paradigma RE constam na Tabela 16.

**Tabela 16**– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Forma B considerando o paradigma RE.

Variável dependente	(I) Série2	(J) Série2	Diferença de média	<i>p</i>
T_B_RE	1	2	-0,92	0,10
		3	-1,39*	0,00
		4	-1,88*	0,00
		5	-2,60*	0,00
		2	1	0,92
	2	3	-0,47	0,65
		4	-0,96	0,05
		5	-1,68*	0,00
		1	1,39*	0,00
		2	0,47	0,65
	3	4	-0,49	0,68
		5	-1,20	0,11
		1	1,88*	0,00
		2	0,96	0,53
		3	0,49	0,68
	4	5	-0,71	0,61
		1	2,60*	0,00
		2	1,68*	0,00
		3	1,20	0,11
		4	0,71	0,61

\* $p < 0,05$

Constata-se na Tabela 16 que as diferenças de média significativas ocorreram entre a 1ª com a maioria das séries, exceto a 2ª, sendo que esta última apresentou diferença significativa com 5ª. Esses dados se diferem dos encontrados na Forma A, uma vez que a 5ª série se diferenciou significativamente apenas com a 1ª e 2ª nesta Forma B. Por fim os dados da ANOVA com o paradigma SC na Forma A estão presentes na Tabela 17.

**Tabela 17**– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Formas A considerando o paradigma SC.

Variável dependente	(I) Série2	(J) Série2	Diferença de média	<i>p</i>
T_A_SC	1	2	-,96	0,07
		3	-1,51*	0,00
		4	-1,66*	0,00
		5	-3,12*	0,00
		2	,96	0,07
	2	3	-,54	0,57
		4	-,70	0,33
		5	-2,15*	0,00
		1	1,51*	0,00
		2	,54	0,57
	3	4	-,15	0,99
		5	-1,60*	0,01
		1	1,66*	0,00
		2	,70	0,33
		3	,15	0,99
	4	5	-1,45*	0,03
		1	3,12*	0,00
		2	2,15*	0,00
		3	1,60*	0,01
		4	1,45*	0,03

\* $p < 0,05$

Com o paradigma SC as séries 1<sup>a</sup> e 5<sup>a</sup> novamente se diferiram significativamente na média em relação as outras séries. A exceção desse dado se verifica entre a 1<sup>a</sup> e 2<sup>a</sup> que não se diferem significativamente entre si. Na Tabela 18 se encontram os dados da ANOVA de séries com o mesmo paradigma, mas na Forma B.

**Tabela 18**– Diferenças de médias verificada pela análise de Tukey entre cada série avaliada para o escore da Forma B considerando o paradigma SC.

Variável dependente	(I) Série2	(J) Série2	Diferença de média	<i>p</i>
T_B_SC	1	2	-1,17	0,08
		3	-2,48*	0,00
		4	-2,71*	0,00
		5	-3,24*	0,00
		2	1,17	0,08
	2	3	-1,31*	0,02
		4	-1,54*	0,00
		5	-2,06*	0,00
		1	2,48*	0,00
		2	1,31*	0,02
	3	4	-0,23	0,98
		5	-0,75	0,73
		1	2,71*	0,00
		2	1,54*	0,00
		3	0,23	0,98
	4	5	-0,52	0,91
		1	3,24*	0,00
		2	2,06*	0,00
		3	0,75	0,73
		4	0,52	0,91

\* $p < 0,05$

Na Tabela 18 se pode verificar resultados diferentes dessa ANOVA considerando o paradigma SC na Forma A. Aqui a 2ª série mostra diferença significativa de média com a 3ª, 4ª e 5ª. Já a 5ª se diferiu somente da 1ª e 2ª séries.

Observa-se, pelas Tabelas 9 a 18, que várias diferenças significativas ocorreram e entre diversidades de séries. A diferença mais constante surgiu entre 1ª e 2ª séries, de um lado, e as demais séries de outro lado. Porém, de forma geral, 1ª e 2ª série não tenderam a

se diferenciar. A 5ª série também se diferenciou da maioria das outras séries. O escore em que houve menos diferenças entre séries foi o B\_T\_CC, onde foi verificada somente uma diferença de média, entre 2ª e 3ª série. No escore A\_T\_CC também ocorreram poucas diferenças significativas, somente entre a 1ª e 3ª e 1ª e 5ª série. Os itens CC pressupõem um raciocínio diferenciado do que geralmente utiliza-se em testes de inteligência, como já explicado na introdução. Esses itens CC, de fato, foram os menos sensíveis a diferenças entre séries.

A análise correlacional entre série e os escores também permite investigar se, ao aumentar a série cursada, o indivíduo tende a aumentar seu desempenho nas Formas A e B. A Tabela 19 apresenta os dados obtidos ao efetuar a análise correlacional das variáveis série e escores.



**Tabela 19**– Análise de correlação entre os escores das Formas A e B com a variável série.

Escores Forma A e B	Séries
A_TT	0,53**
B_TT	0,56**
T_A_SD	0,43**
T_B_SD	0,45**
T_A_CC	0,30**
T_B_CC	0,27**
T_A_RE	0,48**
T_B_RE	0,47**
T_A_SC	0,47**
T_B_SC	0,48**

\*  $p < 0,01$

A Tabela 19 mostrou correlações positivas e moderadas entre série e o desempenho nos escores das Formas A e B, exceto em dois casos, envolvendo as variáveis do paradigma CC, em que houve correlações baixas. Tais informações sugerem que, ao avançar nas séries, o desempenho nas Formas tende a aumentar. Essas informações são evidência de validade baseada nas relações com outras variáveis, mais especificamente com a série.

#### **2.1.2.5.2. Análise fatorial do Teste de Raciocínio Indutivo**

Essa seção teve como objetivo compreender a estrutura interna da escala. Para tanto foi conduzida uma análise fatorial, de modo a verificar se a estrutura do Teste de Raciocínio Indutivo é semelhante à proposta de Klauer, com seis paradigmas. De acordo com a Teoria de Klauer (1990), são propostos seis paradigmas do raciocínio indutivo. Na presente tese foram agrupados dois paradigmas, logo os itens foram construídos baseados

em quatro paradigmas (tal redução encontra-se justificada na introdução). Objetivou-se, portanto, verificar se o teste, frente os itens elaborados, se apresentaria em uni ou multidimensional. Para isso efetuou-se a análise fatorial oblíqua por meio do programa *Microfact* que utiliza os coeficientes de correlação tetracórica, pois são adequados para itens com respostas dicotômicas (Primi & Almeida, 1998). Ressalta-se que essas análises foram realizadas separadamente para as Formas A e B.

Procurando investigar uma possível multidimensionalidade do construto raciocínio indutivo estruturado em quatro fatores, conforme os quatro paradigmas, se pediu inicialmente quatro fatores na análise, tanto para a Forma A quanto para B. Os procedimentos de rotação foram feitos sempre pela PROMAX assumindo associação entre os fatores. No entanto, após rotar as Formas, verificou-se que a estrutura com quatro fatores, compatível com a teoria, não ocorreu para nenhuma das formas. Ressalta-se que pela literatura resgatada não foram encontrados dados informando se os paradigmas se agrupam ou não em fatores separados. Klauer (1990) propôs a teoria, mas a análise fatorial psicométrica para testar estatisticamente a teorização dos paradigmas, não foi constatada diante da literatura resgatada nessa presente pesquisa. É importante considerar, porém, que a ausência da separação em fatores não invalida a distinção teórica a respeito dos diferentes tipos de processamento envolvidos na solução dos problemas apresentados. Deve ser considerado que esses paradigmas são utilizados para formular atividades do treinamento do raciocínio indutivo, ou seja, a cada paradigma a criança aprender a raciocinar com mais informações e utilizando até mesmo procedimentos diferentes.

Analisando os quatro fatores obtidos, tanto na Forma A quanto na B, também não foi possível constatar uma coerência teórica entre os itens que compôs cada fator, porém todos se correlacionavam entre si de forma positiva e moderada ou baixa. Ressalta-se que o primeiro fator apresentou eigenvalue de 9,84 (Forma A) e 8,71 (Forma B) enquanto os outros fatores demonstram aproximadamente 2,5. Além disso, a quantidade de itens nos dois primeiros fatores foi bem maior. Diante desse resultado a análise fatorial das duas formas foi refeita, mas pedindo três fatores. Novamente foram verificadas as mesmas questões da rotação com quatro fatores, a saber. A correlação encontrada entre os fatores 1 e 3, nas duas formas, foi de 0,67, considerada alta, e portanto mais uma vez se refez a análise, pedindo apenas dois fatores, nas duas formas. A Tabela 20 apresenta a solução obtida após a rotação PROMAX das duas formas. Para as análises fatoriais anteriores não foram descritas as estruturas, pois decidiu-se descrever apenas a estrutura que se mostrou mais coerente.

**Tabela 20-** Fatores extraídos da análise fatorial por meio do programa Microfact pedindo dois fatores para as Formas A e B.

	Item			Item Forma			
	Forma A	Fator1	Fator2	B	Fator1	Fator2	
1	A01SD	<b>0,92</b>	-0,12	1	C04SD_20	0,28	0,17
2	A02SD	<b>0,90</b>	-0,21	2	C06SD_02	<b>0,60</b>	0,20
3	A03SD_01	<b>0,46</b>	-0,04	3	C09CC	0,12	-0,17
4	C04SD_20	<b>0,54</b>	0,14	4	C12CC	0,11	0,15
5	A05SD	0,04	0,30	5	C19RE	0,19	0,37
6	C06SD_02	<b>0,50</b>	0,04	6	C22RE_27	<b>0,87</b>	<b>-0,33</b>
7	A07SD	-0,01	<b>0,37</b>	7	C25SC	<b>0,78</b>	-0,06
8	A08CC_05	-0,01	<b>0,54</b>	8	C28SC	<b>0,39</b>	0,08
9	C09CC	-0,29	<b>0,85</b>	9	B01SD_04	<b>0,41</b>	0,19
10	A10CC	0,12	-0,20	10	B02SD	0,25	<b>0,50</b>
11	A11CC_07	<b>0,31</b>	0,25	11	B03SD_19	<b>0,50</b>	0,16
12	C12CC	-0,20	<b>0,68</b>	12	B05SD	0,12	0,23
13	A13CC_24	<b>0,38</b>	0,26	13	B07SD_17	<b>0,38</b>	0,24
14	A14CC_23	-0,00	<b>0,58</b>	14	B08CC_22	0,23	-0,26
15	A15RE	<b>0,60</b>	0,09	15	B10CC	0,04	-0,03
16	A16RE	<b>0,59</b>	0,15	16	B11CC	0,15	-0,27
17	A17RE_10	<b>0,70</b>	-0,34	17	B13CC	0,25	<b>0,54</b>
18	A18RE	<b>0,55</b>	0,16	18	B14CC_08	<b>0,35</b>	0,07
19	C19RE	<b>0,43</b>	0,16	19	B15RE	<b>0,66</b>	-0,08
20	A20RE_11	<b>0,64</b>	-0,05	20	B16RE_09	0,25	-0,07
21	A21RE_12	0,24	<b>0,37</b>	21	B17RE_28	<b>0,56</b>	0,15
22	C22RE_27	<b>0,48</b>	0,24	22	B18RE	0,06	<b>0,34</b>
23	A23RE	<b>0,46</b>	-0,19	23	B20RE_25	0,03	-0,33
24	A24RE	<b>0,50</b>	-0,02	24	B21RE	0,24	<b>0,52</b>
25	C25SC	<b>0,65</b>	-0,05	25	B23RE	<b>0,48</b>	0,18
26	A26SC_16	<b>0,75</b>	-0,04	26	B24SC_32	<b>0,58</b>	-0,05
27	A27SC	-0,10	0,10	27	B26SC	<b>0,59</b>	0,08
28	C28SC	0,29	0,10	28	B27SC_14	<b>0,40</b>	-0,13
29	A29SC_29	<b>0,37</b>	-0,05	29	B29SC	<b>0,65</b>	-0,25
30	A30SC_31	0,24	0,24	30	B30SC_13	<b>0,34</b>	0,01
31	A31SC	<b>0,32</b>	0,14	31	B31SC	<b>0,68</b>	<b>0,33</b>

Tabela 20 (continuação)- Fatores extraídos da análise fatorial por meio do programa Microfact pedindo dois fatores para as Formas A e B.

Item		Fator1		Fator2		Item Forma B		Fator1		Fator2	
32	A32SC_15	<b>0,39</b>	0,10	32	B32SC	<b>0,29</b>	0,00				
33	A33SC	0,18	0,01	33	B33SC	0,26	0,12				
34	A34SC_30	<b>0,55</b>	-0,07	34	B34SD_18	0,12	<b>0,54</b>				
35	A35SD	<b>0,61</b>	0,11	35	B35SD_03	<b>0,30</b>	0,14				
36	A36CC_21	<b>0,41</b>	-0,20	36	B36CC	-0,03	0,10				
37	A37CC_06	<b>0,75</b>	-0,16	37	B37CC	-0,10	<b>0,48</b>				
38	A38RE_26	-0,02	0,13	38	B38RE	0,13	-0,16				
Eigenvalue		9,84	2,52			8,71	2,56				
Var.Explicada		24,34	4,80			19,71	4,60				
Precisão		0,84	0,62			0,85	0,57				

*Nota: A primeira letra do item possibilita identificar a que forma ele pertence: se a letra for A, o item pertence à Forma A; se a letra for B, o item pertence à Forma B; e se a letra for C, o item é comum às duas Formas. O número dos itens no teste encontra-se após o tipo de paradigma; no caso dos itens comuns a ambas as formas, o primeiro número refere-se à posição na Forma A e o segundo número\_ à posição na Forma B. O tipo de paradigma é ilustrado pela composição de letras, quais seja SD (paradigma de Similaridade e Diferenciação), CC (paradigma Classificação Cruzada), RE (paradigma de Relações entre objetos) e SC (paradigma de Sistema de Construção).*

Observando a estrutura de dois fatores na Forma A, identificou-se um primeiro fator mais robusto, com *eigenvalue* de 9,84, variância explicada de 24,34 e composto por 25 itens com carga fatorial acima de 0,30, alcançando uma precisão de 0,84. Esse fator é composto por itens representantes de todos os paradigmas. Já o fator dois é menos expressivo quando comparado com o 1, apresentando *eigenvalue* de 2,52, variância explicada de 4,80 e composto por seis itens, obtendo precisão de 0,62. Nota-se que esse segundo fator é prioritariamente composto por itens do paradigma CC. Apesar de verificar

esse segundo fator constatou-se que a correlação entre os dois fatores foi de 0,57, considerada moderadamente alta.

Com relação à Forma B, conforme descrito na Tabela 20, a análise com dois fatores seguiu um padrão semelhante ao da Forma A. Identificou-se um primeiro fator mais robusto, com *eigenvalue* de 8,71, variância explicada de 19,71 e composto por 19 itens acima de 0,30, atingindo uma precisão de 0,85. A configuração da estrutura também abrangeu itens de todos os paradigmas. No que diz respeito ao fator 2, esse foi menos expressivo frente o fator 1, apresentando *eigenvalue* de 2,56, variância explicada de 4,60 e formado por oito itens que obtêm precisão de 0,57. Dos itens que compuseram esse segundo fator, quatro eram representativos do paradigma RE, dois SD e dois CC. Não há um sentido teórico que justifique esse fator com tal estruturação. A correlação entre os dois fatores foi de 0,38, podendo ser considerada baixa, porém quase moderada.

Observou-se, portanto, que em ambas as formas houve correlação moderada a alta entre os dois fatores e que as estruturas obtidas com dois fatores não são teoricamente esperadas. Por isso, aplicou-se nova análise fatorial pedindo apenas um fator. Os resultados encontram-se descritos na Tabela 21.

**Tabela 21-** Fator extraído da análise fatorial por meio do programa Microfact pedindo 1

Fator para as Formas A e B.

	Item Forma A	Fator		Item Forma B	Fator
1	A01SD	<b>0,82</b>	1	C04SD_20	<b>0,38</b>
2	A02SD	<b>0,74</b>	2	C06SD_02	<b>0,70</b>
3	A03SD_01	<b>0,43</b>	3	C09CC	0,01
4	C04SD_20	<b>0,62</b>	4	C12CC	0,20
5	A05SD	0,24	5	C19RE	<b>0,41</b>
6	C06SD_02	<b>0,53</b>	6	C22RE_27	<b>0,63</b>
7	A07SD	0,24	7	C25SC	<b>0,71</b>
8	A08CC_05	<b>0,36</b>	8	C28SC	<b>0,42</b>
9	C09CC	0,21	9	B01SD_04	<b>0,52</b>
10	A10CC	-0,02	10	B02SD	<b>0,55</b>
11	A11CC_07	<b>0,48</b>	11	B03SD_19	<b>0,59</b>
12	C12CC	0,26	12	B05SD	0,27
13	A13CC_24	<b>0,56</b>	13	B07SD_17	<b>0,52</b>
14	A14CC_23	<b>0,39</b>	14	B08CC_22	0,07
15	A15RE	<b>0,66</b>	15	B10CC	0,02
16	A16RE	<b>0,69</b>	16	B11CC	-0,01
17	A17RE_10	<b>0,46</b>	17	B13CC	<b>0,60</b>
18	A18RE	<b>0,66</b>	18	B14CC_08	<b>0,38</b>
19	C19RE	<b>0,53</b>	19	B15RE	<b>0,58</b>
20	A20RE_11	<b>0,60</b>	20	B16RE_09	0,20
21	A21RE_12	<b>0,49</b>	21	B17RE_28	<b>0,63</b>
22	C22RE_27	<b>0,64</b>	22	B18RE	0,27
23	A23RE	<b>0,32</b>	23	B20RE_25	-0,18
24	A24RE	<b>0,49</b>	24	B21RE	<b>0,57</b>
25	C25SC	<b>0,61</b>	25	B23RE	<b>0,58</b>
26	A26SC_16	<b>0,72</b>	26	B24SC_32	<b>0,53</b>
27	A27SC	-0,02	27	B26SC	<b>0,61</b>
28	C28SC	<b>0,36</b>	28	B27SC_14	<b>0,32</b>
29	A29SC_29	<b>0,33</b>	29	B29SC	<b>0,48</b>
30	A30SC_31	<b>0,40</b>	30	B30SC_13	<b>0,35</b>
31	A31SC	<b>0,41</b>	31	B31SC	<b>0,87</b>
32	A32SC_15	<b>0,46</b>	32	B32SC	0,29

Tabela 21 (continuação)- Fator extraído da análise fatorial por meio do programa Microfact pedindo 1 Fator para as Formas A e B.

	Item Forma A	Fator		Item Forma B	Fator
33	A33SC	0,18	33	B33SC	<b>0,32</b>
34	A34SC_30	<b>0,50</b>	34	B34SD_18	<b>0,44</b>
35	A35SD	<b>0,68</b>	35	B35SD_03	<b>0,37</b>
36	A36CC_21	0,26	36	B36CC	0,02
37	A37CC_06	<b>0,63</b>	37	B37CC	0,19
38	A38RE_26	0,07	38	B38RE	0,02

Para a Forma A, nessa nova análise o *eigenvalue* e a variância explicada referente ao primeiro fator permaneceram os mesmos da análise anterior, porém o fator aumentou sua composição para 30 itens com cargas acima 0,30 e a precisão melhorou para 0,87. Com relação aos itens, dos oito que não atingiram carga fatorial satisfatória para serem considerados pertencentes à estrutura, três fazem parte do fator dois da análise anterior. Esses dados indicam que a Forma A tende a apresentar unidimensionalidade.

Para a Forma B, conforme a Tabela 21, o *eigenvalue* e a variância explicada do fator 1 também praticamente não se modificaram em relação aos resultados da análise anterior, mas a quantidade de itens no fator 1 com cargas acima de 0,30 aumentou para 25, e a precisão continuou estável em 0,87. Essa estrutura fatorial incorporou cinco dos oito itens que faziam parte do fator dois na análise anterior, sendo que tais itens apresentaram cargas fatoriais de moderadas a altas. Assim como para a Forma A, a Forma B também foi mais bem explicada com a estrutura unidimensional.



Diante dos dados concebeu-se as Formas A e B como unidimensionais visto que, apesar de possuírem itens diversificados de acordo com os paradigmas propostos por Klauer (1990), a estrutura multidimensional composta por fatores correspondentes a esses paradigmas não foi verificada. Essa parece ter sido a primeira vez que se procurou investigar a estrutura fatorial advinda da teoria de Klauer. O resultado encontrado apontando a unidimensionalidade não deixa de ser coerente com a teoria citada, pois todos os itens são elaborados para serem resolvidos com o raciocínio indutivo, sendo que a diferença principal entre os itens é a complexidade do pensar indutivamente, mas sempre partido do ponto de semelhanças entre as informações. Ressalta-se que, apesar da análise fatorial ter apontado a unidimensionalidade para as Formas A e B, as estatísticas descritivas e a correlação efetuadas com escores separando cada paradigma deve ser mantida, pois para cada paradigma o processo de raciocinar indutivamente tem sua particularidade, como já explicado na introdução. Logo é interessante verificar como os indivíduos se comportam frente aos conjuntos de itens de cada paradigma. Além disso, o presente estudo tratou-se de uma pesquisa inicial, devendo-se considerar suas limitações e a necessidade de exploração mais detalhada dos resultados encontrados.

#### **2.1.2.5.3. Análise da precisão do Teste de Raciocínio Indutivo**

Alguns dados de precisão referente a cada Forma do teste já foram referidos na seção anterior, tendo sido, de forma geral, altas. No entanto, como a amostra é composta por diferentes séries, é importante verificar a precisão separada por série. Tais informações

tendem a esclarecer para quais séries as Formas podem ser mais confiáveis quando aplicadas. Na Tabela 22 a seguir podem ser conferidas essas precisões, obtidas por meio da análise de Alfa de Cronbach.

**Tabela 22**– Índices de precisão para as Forma A e B em cada série avaliada.

Forma	Série	Precisão
A	1 <sup>a</sup>	0,48
A	2 <sup>a</sup>	0,67
A	3 <sup>a</sup>	0,80
A	4 <sup>a</sup>	0,84
A	5 <sup>a</sup>	0,75
B	1 <sup>a</sup>	NÃO HOUVE
B	2 <sup>a</sup>	0,58
B	3 <sup>a</sup>	0,76
B	4 <sup>a</sup>	0,77
B	5 <sup>a</sup>	0,77

As precisões encontradas apontam índices satisfatórios (acima de 0,60) e muito satisfatórios (acima de 0,80) dependendo da série, exceto para a 1<sup>a</sup> série nas duas Formas, principalmente na B, em que não houve precisão, e para a 2<sup>a</sup> série também na Forma B. É importante esse conhecimento de quais séries são mais precisamente avaliadas, pois isso possibilita compreender melhor os resultados de um teste.

#### **2.1.2.5.4. Análise do Teste de Raciocínio Indutivo pela Teoria de Resposta ao Item**

A presente análise foi a última conduzida na parte 1 do Estudo 1, bem como a que principalmente direcionou para a seleção dos itens que compuseram o pré e pós-teste do instrumento dinâmico. A escolha dos itens seguiu o padrão de 16 itens para cada momento do teste (pré e pós), sendo quatro para cada paradigma, e esses deveriam variar em termos de grau de dificuldade, sendo um fácil, dois médios e um difícil, dentro das possibilidades existentes.

Por exemplo, na Teoria de Resposta ao Item pelo modelo de Rasch, que foi a análise realizada, a dificuldade dos itens é associada ao nível de habilidade dos sujeitos em que possuem 50% de chance de acerto, sendo que pode variar entre -4 e +4. Logo valores próximos a zero são considerados de média dificuldade; muito abaixo, são considerados fáceis; e muito acima, difíceis. No entanto, nem sempre foi possível seguir esse padrão, pois além do critério da dificuldade, era preciso que os itens fossem ajustados e com uma correlação item-theta adequada. Por fim, não houve uma variabilidade de níveis de dificuldade dentro de todos os paradigmas. Todas essas questões dificultaram a seleção de itens a partir do critério de dificuldade variada, o que possibilitaria construir um teste que identifique sujeitos com habilidades diversificadas.

Como já especificado, a análise foi realizada por meio da Teoria de Resposta ao Item com o modelo de Rasch. As duas Formas foram rodadas como uma única escala, pois os itens em comum, itens âncoras, possibilitam que essa ação de equalização seja efetuada no programa Winsteps. A Tabela 23 apresenta os itens escolhidos para o pré (itens

sombreados) e pós-teste (itens não sombreados), bem como suas propriedades psicométricas relevantes na seleção.

**Tabela 23**– Propriedades psicométrica por meio da análise Rasch para dos itens escolhidos para compor o instrumento dinâmico.

Nº	Dificuldade	<i>Infit</i>	<i>Outfit</i>	r	Item	Paradigma
4	-0,58	0,93	0,92	0,45	C04SD	SD
6	-0,52	0,9	0,87	0,49	C06SD	SD
26	-0,41	0,83	0,78	0,57	A26SC	SC
11	-0,36	0,96	0,95	0,44	A11CC	CC
56	-0,32	0,92	0,92	0,45	B24SC	SC
41	-0,32	0,85	0,83	0,54	B03SD	SD
39	-0,3	0,95	0,94	0,43	B01SD	SD
51	-0,24	0,85	0,82	0,54	B17RE	RE
21	-0,22	0,96	0,93	0,45	A21RE	RE
13	-0,16	0,93	0,89	0,47	A13CC	CC
3	-0,07	1,03	1,01	0,38	A03SD	SD
64	-0,06	0,97	0,95	0,41	B34SD	SD
37	0,01	0,91	0,87	0,50	A37CC	CC
14	0,03	1,03	1,04	0,37	A14CC	CC
20	0,16	0,91	0,89	0,49	A20RE	RE
22	0,22	0,84	0,82	0,55	C22RE	RE
65	0,28	1,01	1,01	0,35	B35SD	SD
43	0,3	0,9	0,88	0,47	B07SD	SD
58	0,41	1,05	1,07	0,29	B27SC	SC
44	0,43	1,17	1,26	0,15	B08CC	CC
30	0,43	1,02	1,1	0,36	A30SC	SC
17	0,45	1	1,04	0,39	A17RE	RE
48	0,62	1	1,03	0,34	B14CC	CC
38	0,7	1,27	1,31	0,13	A38RE	RE
32	0,83	0,98	0,95	0,4	A32SC	SC
50	0,86	1,08	1,2	0,22	B16RE	RE
34	0,87	0,96	1,04	0,41	A34SC	SC
8	0,98	1,02	1,08	0,35	A08CC	CC
60	1,01	1,01	1,07	0,30	B30SC	SC
29	1,11	1,06	1,35	0,27	A29SC	SC
36	1,44	1,14	1,28	0,2	A36CC	CC
53	2,36	1,12	1,97	-0,05	B20RE	RE

É importante destacar que, pela Teoria de Resposta ao Item por meio do modelo de Rasch, a escala atingiu a precisão (*Real*) de 0,80 e foi considerada de dificuldade média (0), sendo que a habilidade média da amostra foi abaixo desse dado (-0,21). Essas informações indicam que a dificuldade da escala é maior que a habilidade dos sujeitos da amostra, porém a diferença não é significativa, pois está ao redor de 0.

Retomando a Tabela 23, nota-se que os itens escolhidos estão na maioria dentro dos parâmetros de ajustes (menor que 1,2) e que as correlações item-theta estão adequadas. Somente três itens apresentaram desajustes e correlação não apropriada, quais sejam 29, 38 e 53, mas somente esse último é mais problemático, pois além de ter correlação nula, o índice *outfit* foi 1,97. Mas decidiu-se mantê-lo por causa do nível alto de dificuldade para compor o conjunto de quatro itens do paradigma RE no pós-teste. Inclusive o desajuste desse item pode ter sido em função do nível de habilidade.

Observando a Tabela 23, verifica-se que realmente não foi possível seguir os três níveis de dificuldade de acordo com o modelo de Rasch. Porém, dentro de cada paradigma, procurou-se contemplar um item que mostrava menor nível de dificuldade, um com maior dificuldade e dois com dificuldade intermediária. Por exemplo, no conjunto dos itens do paradigma SD que compõem o pré-teste, considerados os mais fáceis entre os paradigmas, o item com maior nível de dificuldade foi o 65 (B35), com 0,28, que pode ser considerado um item com nível médio de dificuldade de acordo com o modelo de Rasch, mas, nesse

caso, foi considerado o mais difícil dentre os outros três. Isso foi necessário pois não havia outro item do mesmo paradigma com nível de dificuldade mais alta.

Outro dado que pode ser constatado na Tabela 23 é que alguns itens, se fossem verificados somente pela análise fatorial, não comporiam o pré e o pós-teste. Tais itens são o 08, 16, 20, 36 e 38. Porém, diante do conjunto das outras informações e devido ao critério dos níveis de dificuldade, decidiu-se que eles fariam parte do instrumento dinâmico. Se faz necessário ter ciência dessas informações, pois tais itens foram investigados no Estudo 3 e foram mais bem compreendidos a partir desse histórico.

A seguir serão descritos e explicados cada item eleito para a composição do pré e pós-teste para que se possa entender a natureza de cada item, os processos e estratégias cognitivas utilizadas para a resolução da tarefa. Ressalta-se que a natureza, os processos e as estratégias descritas foram baseadas na teoria prescritiva do raciocínio indutivo de Josef Klauer (1990) eleita para embasar os itens e a intervenção do teste dinâmico construído.

A descrição referida será efetuada tanto para os itens do pré quanto para os itens do pós. Essas duas fases são constituídas de itens semelhantes e com níveis de dificuldade equivalentes. Todos os itens foram formulados para que a estratégia utilizada fosse do raciocínio indutivo, ou seja, comparar regularidades entre os objetos, sendo que os processos cognitivos envolvidos são basicamente generalização e discriminação. Então, para que a criança atinja a resposta correta, ela precisa identificar o que há de semelhante e diferente observando um padrão de regularidades, verificando qual alternativa é a resposta

correta. No entanto, apesar desses serem os processos e estratégias básicas, cada item, ou cada tarefa, apresenta particularidades de acordo com os paradigmas do raciocínio indutivo proposto por Klauer (1990). O detalhamento abrange os 32 itens, separados por paradigmas com a numeração correspondente no pré e pós e com suas descrições.

Os quatro primeiros itens no pré (1, 2, 3 e 4) e pós (17, 18, 19 e 20) fazem parte do paradigma Generalização/Discriminação de atributos-SD. Generalização é o estabelecimento de similaridades de atributos para diferentes objetos formando um grupo, pois apesar de haver diferenciações, partes desses objetos são comuns. Para a resolução desses problemas a criança procede analiticamente, comparando as propriedades dos objetos, procurando descobrir algo em comum entre eles. A criança pode fazer isso sistematicamente, ou seja, ir comparando um a um, da esquerda para a direita, por exemplo. Durante a análise o indivíduo utiliza estratégias de monitoramento, checando quais atributos dos itens excluídos não possuem, bem como aprendendo a distinguir esse monitoramento da solução do problema. A discriminação (DI) é o processo de averiguar diferenças entre objetos a respeito de atributos em que também há a verificação do que é comum entre os objetos, para conseqüentemente identificar a diferença. Ao se fazer a generalização, automaticamente se faz a discriminação, pois é necessária para a resolução da tarefa de generalização e vice-versa, então nesse trabalho optou-se por agrupar os paradigmas. Na Tabela 24 encontra-se a descrição do paradigma SD.



**Tabela 24**– Descrição dos itens do pré e pós-teste referentes ao paradigma SD

<b>Item</b>	<b>Formato do problema</b>	<b>Processo</b>	<b>Estratégia</b>	<b>O que identificar?</b>
SD_1	Encontrar atributos comuns	Similaridade e Diferença em atributos	Indutiva	Cor igual em objetos diferentes.
SD_2	Encontrar atributos comuns	Similaridade e Diferença em atributos	Indutiva	Movimentos simétricos entre a orelha e a pata do gato; Cada gato apresenta um tipo de movimento.
SD_3	Expansão de classe	Similaridade e Diferença em atributos	Indutiva	Motorizados; Tamanho aumentando; Levam maior quantidade de pessoas.
SD_4	Encontrar atributos comuns	Similaridade e Diferença em atributos	Indutiva	Todos voam e têm asas.
SD_17	Encontrar atributos comuns	Similaridade e Diferença em atributos	Indutiva	Mesmo tamanho e cor de fundo; Cor do interior da asa modifica simetricamente, encima e embaixo em todas as borboletas com cores diferentes.
SD_18	Formação de classe	Similaridade e Diferença em atributos	Indutiva	Frutas
SD_19	Encontrar atributos comuns	Similaridade e Diferença em atributos	Indutiva	Tamanho igual; Cores e objetos diferentes.
SD_20	Encontrar atributos comuns	Similaridade e Diferença em atributos	Indutiva	Objetos que contêm fogo

Os itens 5, 6, 7, 8 (pré) e 21, 22, 23 e 24 (pós) foram elaborados de acordo com o paradigma Classificação Cruzada-CC. Esse tipo de tarefa é organizada em esquemas de atributos em matrizes de quatro partes, ou quatro objetos, dois em cima e dois embaixo, onde pelo menos dois atributos são considerados ao mesmo tempo. Por exemplo, o objeto do primeiro quadrante precisa ter no mínimo um atributo igual ao objeto do segundo quadrante, ao seu lado, e outro atributo igual ao terceiro quadrante, que está abaixo dele. Isso deve ocorrer para cada objeto de um quadrante. A classificação cruzada requer a determinação de atributos diferentes e comuns comparando os objetos, mas cruzando as informações entre todos os objetos para saber, por exemplo, onde uma determinada figura melhor se encaixa. As estratégias de monitoramento também são importantes. Visualiza-se na Tabela 25 A descrição dos itens do paradigma CC.

**Tabela 25**– Descrição dos itens do pré e pós-teste referentes ao paradigma CC.

<b>Item</b>	<b>Formato do problema</b>	<b>Processo</b>	<b>Estratégia</b>	<b>O que identificar?</b>
CC_5	4 quadrículos	Similaridade e Diferença em atributos cruzando informações	Indutiva	Formato parecido com a figura acima e utilidade com a figura ao lado
CC_6	4 quadrículos	Similaridade e Diferença em atributos cruzando informações	Indutiva	Categoria igual a figura acima e formato parecido com a figura ao lado
CC_7	4 quadrículos	Similaridade e Diferença em atributos cruzando informações	Indutiva	Formato parecido com a figura acima e mesma categoria com a figura ao lado
CC_8	4 quadrículos	Similaridade e Diferença em atributos cruzando informações	Indutiva	Formato parecido com a figura acima e categoria igual com a do lado
CC_21	4 quadrículos	Similaridade e Diferença em atributos cruzando informações	Indutiva	Categoria de transporte com a figura acima e de algo terrestre com a figura ao lado
CC_22	4 quadrículos	Similaridade e Diferença em atributos cruzando informações	Indutiva	Formato parecido com a figura acima e mesma categoria com a figura ao lado
CC_23	4 quadrículos	Similaridade e Diferença em atributos cruzando informações	Indutiva	Formato parecido com a figura acima e categoria com a figura ao lado
CC_24	4 quadrículos	Similaridade e Diferença em atributos cruzando informações	Indutiva	Cor igual a figura acima e categoria com a figura ao lado

O paradigma Reconhecimento e Diferenciação de Relações-RE é representado pelo itens 9, 10, 11 e 12 no pré e 25, 26, 27 e 28 no pós. Também são dois paradigmas, mas que foram agrupados em apenas um, seguindo o mesmo raciocínio do paradigma SD, ou seja, para fazer o reconhecimento é preciso identificar diferenciações e vice-versa.

O paradigma RE é possível quando pelo menos dois objetos estão presentes e a inferência de relação decorre de um procedimento de comparação entre pares, mas, nesse caso, é preciso encontrar a similaridade entre relações e não mais atributos, o que torna a resolução do item mais complexa. Essas relações podem ser de acordo com a função de objetos, ou causa e efeito, ou algo único de um conjunto de problemas. Por exemplo, em problemas com analogia é preciso determinar a relação específica entre um par de objetos que será o padrão de comparação. Várias relações podem ser possíveis, então a estratégia de solução consiste em mapear a relação de um par de objetos e estabelecer, em um par incompleto, este mesmo tipo de relação. Nesse processo de mapear uma relação para aplicar em outra situação, o monitoramento metacognitivo para verificar se a escolha é a correta sempre deve estar presente. A descrição dos itens desse paradigma RE consta na Tabela 26.

**Tabela 26**– Descrição dos itens do pré e pós-teste referentes ao paradigma RE.

<b>Item</b>	<b>Formato do problema</b>	<b>Processo</b>	<b>Estratégia</b>	<b>O que identificar?</b>
RE_9	Analogia A-B/C-D	Similaridade e diferenciação de relações	Indutiva	Relação de proteger algo Então: guarda-chuva protege da chuva e o boné do sol.
RE_10	Analogia A-B/C-D	Similaridade e diferenciação de relações	Indutiva	Relação de produção. Então: vaca produz leite e árvore produz fruta (laranja)
RE_11	Complemento de série (transformação)	Similaridade e diferenciação de relações	Indutiva	A figura vai perdendo partes; Cor permanece.
RE_12	Complemento de série (transformação)	Similaridade e diferenciação de relações	Indutiva	O círculo aumenta; O triângulo e a cor não mudam
RE_25	Complemento de série (transformação)	Similaridade e diferenciação de relações	Indutiva	O círculo aumenta, mas cor permanece. O retângulo diminui e muda de cor a cada fase.
RE_26	Analogia A-B/C-D	Similaridade e diferenciação de relações	Indutiva	Relação de entrada. Então: A porta é entrada da casa e a orelha do corpo.
RE_27	Complemento de série (continuidade/ transformação)	Similaridade e diferenciação de relações	Indutiva	Aparecimento do Sol entre as montanhas.
RE_28	Complemento de série (Sequência/ transformação)	Similaridade e diferenciação de relações	Indutiva	Alternância de triângulo e círculo, mas esse último aumenta de tamanho.

Por fim os últimos quatro itens do pré e pós, 13, 14, 15, 16 e 29, 30, 31, 32, respectivamente, são englobados no paradigma Sistema de Construção-SC. Envolve um raciocínio semelhante ao usado para resolver problemas de classificação cruzada, mas no caso do SC existem pelo menos duas relações em que similaridade e diferenciação podem ser verificadas. Para a solução desses tipos de item é preciso reconhecer onde cada relação é operacionalizada e onde existe similaridade e diferença. A estratégia metacognitiva nesses casos é uma checagem sistemática para similaridade de relações entre linhas e colunas, sempre iniciando na célula esquerda acima de uma matriz. Os itens referentes ao paradigma SC estão descritos na Tabela 27.

**Tabela 27**– Descrição dos itens do pré e pós-teste referentes ao paradigma SC.

<b>Item</b>	<b>Formato do problema</b>	<b>Processo</b>	<b>Estratégia</b>	<b>O que identificar?</b>
SC_13	Matrizes (Três colunas e três linhas)	Similaridade e diferenciação de relações, cruzando informações	Indutiva	Cada linha um tipo de objeto; Na linha o objeto rota (mas essa informação é irrelevante); A cor dos objetos na terceira coluna é igual a da primeira coluna.
SC_14	Matrizes (três colunas e três linhas)	Similaridade e diferenciação de relações, cruzando informações	Indutiva	Cada linha um tipo de objeto; Cores diferentes para cada tipo de objeto; Cada coluna uma rotação do objeto; As cores permanecem as mesmas em cada linha.
SC_15	Matrizes (três colunas e três linhas)	Similaridade e diferenciação de relações, cruzando informações	Indutiva	Cada linha um tipo de objeto; Na linha objetos aumentam; A cor do objeto na última coluna é a mesma da primeira coluna.
SC_16	Matrizes (três colunas e duas linhas)	Similaridade e diferenciação de relações, cruzando informações	Indutiva	Cada linha um tipo de objeto; O objeto muda de cor nas colunas.As cores dos tipos de objetos mudam simetricamente; Na linha objetos aumentam de tamanho.

Tabela 27 (continuação)- Descrição dos itens do pré e pós-teste referentes ao paradigma SC.

<b>Item</b>	<b>Formato do problema</b>	<b>Processo</b>	<b>Estratégia</b>	<b>O que identificar?</b>
SC_29	Matrizes (Três linhas e três colunas)	Similaridade e diferenciação de relações, cruzando informações	Indutiva	Cada linha um tipo de objeto; Há mudanças de cores em cada tipo de objeto; As mudanças ocorrem simetricamente para cada tipo de objeto nas linhas; Única cor na primeira coluna; Duas cores (uma na parte superior e outra na inferior do objeto) na segunda coluna; Na terceira coluna apenas inverte o local das cores.
SC_30	Matrizes (duas linhas e duas colunas)	Similaridade e diferenciação de relações, cruzando informações	Indutiva	Cor do objeto em linha permanece; Preenchimento na coluna desaparece; Na segunda linha e primeira coluna o objeto é um triângulo, mas as respostas são apenas círculos (é um distrator)



Tabela 27 (continuação)- Descrição dos itens do pré e pós-teste referentes ao paradigma SC.

<b>Item</b>	<b>Formato do problema</b>	<b>Processo</b>	<b>Estratégia</b>	<b>O que identificar?</b>
SC_31	Matrizes (Duas linhas e três colunas)	Similaridade e diferenciação de relações, cruzando informações	Indutiva	Cada linha dois objetos diferentes na mesma coluna (um dentro do outro); As cores permanecem de cada tipo de objeto; Há rotação para a esquerda do objeto de dentro e para a direita do objeto de fora
SC_32	Matrizes (Três linhas e duas colunas)	Similaridade e diferenciação de relações, cruzando informações	Indutiva	Um tipo de objeto em cada linha; Dentro de cada um deles tem uma bolinha; As cores dos objetos permanecem; A cor da bolinha muda em cada coluna, mas igual com a linha abaixo; Os objetos fazem rotação para a direita.

Além dessa descrição dos itens a Tabela 28 apresenta as alternativas corretas para cada item.

**Tabela 28-** Alternativa correta de cada item.

<b>ITENS TIPO SIMILARIDADE E DISCRIMINAÇÃO-SD</b>
<p style="text-align: center;"><b>ITEM 1</b></p> <p>Resposta A: sem indicadores de semelhanças para a tarefa Resposta B: atributo incompleto (cor diferente) <b>Resposta C: correta</b> Resposta D: atributo incompleto (cor diferente)</p>
<p style="text-align: center;"><b>ITEM 2</b></p> <p>Resposta A: correspondência de pata e orelha, mas não entre essas partes Resposta B: não há correspondência <b>Resposta C: correta</b> Resposta D: não há correspondência</p>
<p style="text-align: center;"><b>ITEM 3</b></p> <p><b>Resposta A: correta</b> Resposta B: atributo incompleto (não carrega maior quantidade de pessoas) Resposta C: atributo incompleto (não é terrestre) Resposta D: apesar de permitir locomoção, não é compatível com os dois atributos</p>
<p style="text-align: center;"><b>ITEM 4</b></p> <p><b>Resposta A: correta</b> Resposta B: atributo incompleto (não tem asas) Resposta C: atributo incompleto (tem “asas”, mas não voa) Resposta D: atributo incompleto (voa, mas não tem asas)</p>
<p style="text-align: center;"><b>ITEM 17</b></p> <p>Resposta A: atributo incompleto (mudança da cor embaixo não é simétrico) Resposta B: atributo incompleto (tamanho não corresponde) <b>Resposta C: correta</b> Resposta D: atributo incompleto (cor da borboleta não corresponde)</p>
<p style="text-align: center;"><b>ITEM 18</b></p> <p>Resposta A: atributo incorreto (não é fruta) Resposta B: atributo incorreto (não é fruta) Resposta C: atributo incorreto (não é fruta) <b>Resposta D: correta</b></p>
<p style="text-align: center;"><b>ITEM 19</b></p> <p>Resposta A: atributo incompleto (tamanho errado e objeto repetido) Resposta B: atributo incompleto (tamanho errado e cor repetida) Resposta C: atributo incompleto (tamanho errado e cor repetida) <b>Resposta D: correta</b></p>
<p style="text-align: center;"><b>ITEM 20</b></p> <p>Resposta A: não há atributo equivalente <b>Resposta B: correta</b> Resposta C: atributo equivocado (a resposta é com fogo e não fumaça) Resposta D: não há atributo equivalente</p>

Tabela 28 (continuação)- Alternativa correta de cada item.

<b>ITENS TIPO CLASSIFICAÇÃO CRUZADA-CC</b>
<b>ITEM 5</b>
Resposta A: possui 1 atributo, mas com o objeto errado
<b>Resposta B: correta</b>
Resposta C: não possui atributos com os objetos
Resposta D: não possui atributos com os objetos
<b>ITEM 6</b>
Resposta A: possui 1 atributo, mas com o objeto errado
<b>Resposta B: correta (23%)</b>
Resposta C: possui 1 atributo, mas com o objeto errado
Resposta D: não possui os atributos adequados para a resposta correta
<b>ITEM 7</b>
Resposta A: possui apenas 1 atributo
Resposta B: não possui atributo
<b>Resposta C: correta</b>
Resposta D: não possui atributo
<b>ITEM 8</b>
<b>Resposta A: correta</b>
Resposta B: possui apenas 1 atributo
Resposta C: possui 1 atributo, mas equivocado para a resposta
Resposta D: possui os dois atributos, mas com o mesmo objeto
<b>ITEM 21</b>
<b>Resposta A: correta</b>
Resposta B: possui apenas 1 atributo correto
Resposta C: possui apenas 1 atributo correto
Resposta D: possui apenas 1 atributo correto
<b>ITEM 22</b>
<b>Resposta A: correta</b>
Resposta B: possui apenas 1 atributo (doce)
Resposta C: possui apenas 1 atributo (comprido)
Resposta D: possui 1 atributo, mas irrelevante para a resposta
<b>ITEM 23</b>
Resposta A: não há atributos
<b>Resposta B: correta</b>
Resposta C: possui 1 atributo, mas com a figura errada
Resposta D: possui 1 atributo irrelevante para a resposta
<b>ITEM 24</b>
Resposta A: possui 1 atributo irrelevante para a resposta (mas brilha como a estrela)
Resposta B: possui 1 atributo irrelevante para a resposta ( é de comer)
Resposta C: Até poderia ser, mas modificando os atributos
<b>Resposta D: correta</b>

Tabela 28 (continuação)- Alternativa correta de cada item.

<b>ITENS TIPO SIMILARIDADE E DISCRIMINAÇÃO ENTRE OBJETOS-RE</b>
<p style="text-align: center;"><b>ITEM 9</b></p> <p>Resposta A: Também é de proteção, mas não a proteção apresentada na analogia <b>Resposta B: correta</b> Resposta C: Figura igual ao item de relação e não há correspondência Resposta D: não há correspondência com as figuras de relação, apenas com a C</p>
<p style="text-align: center;"><b>ITEM 10</b></p> <p><b>Resposta A: correta</b> Resposta B: não há relação com a analogia, apenas com a figura C Resposta C: não há relações Resposta D: não há relações</p>
<p style="text-align: center;"><b>ITEM 11</b></p> <p>Resposta A: é a etapa correta, mas a cor está errada Resposta B: não é a etapa correta, e a cor também difere <b>Resposta C: correta</b> Resposta D: etapa incorreta</p>
<p style="text-align: center;"><b>ITEM 12</b></p> <p>Resposta A: transformação errada, objeto constante errado <b>Resposta B: correta</b> Resposta C: transformação correta, mas o objeto constante muda Resposta D: transformação correta, mas não há o objeto constante</p>
<p style="text-align: center;"><b>ITEM 25</b></p> <p>Resposta A: duas transformações erradas, apenas as cores corretas <b>Resposta B: correta</b> Resposta C: apenas a cor de 1 objeto está errada Resposta D: apenas a cor de 1 objeto está correta</p>
<p style="text-align: center;"><b>ITEM 26</b></p> <p>Resposta A: relação correta com a figura C, mas analogia errada com A-B Resposta B: resposta errada, possui apenas atributo com a figura C, mas não relação <b>Resposta C: correta</b> Resposta D: relação correta com a figura C, mas analogia errada com A-B</p>
<p style="text-align: center;"><b>ITEM 27</b></p> <p><b>Resposta A: correta</b> Resposta B: resposta errada ( transformação igual a anterior) Resposta C: resposta parcialmente correta, pois não consta a montanha Resposta D: resposta errada (pôr-do-sol)</p>
<p style="text-align: center;"><b>ITEM 28</b></p> <p>Resposta A: seqüência errada (prevalece a anterior) Resposta B: transformação errada Resposta C: transformação errada (junção dos objetos) <b>Resposta D: correta</b></p>

Tabela 28 (continuação)- Alternativa correta de cada item.

<b>ITENS TIPO SISTEMA DE CONSTRUÇÃO-SC</b>	
<b>ITEM 13</b>	
Resposta A: cor correta, mas com borda amarela	
Resposta B: cor correta, mas com borda amarela	
<b>Resposta C: correta</b>	
Resposta D: cor incorreta	
<b>ITEM 14</b>	
<b>Resposta A: correta</b>	
Resposta B: parcialmente correta, cor diferente	
Resposta C: rotação errada, posterior a correta	
Resposta D: resposta irrelevante	
<b>ITEM 15</b>	
Resposta A: transformação incorreta, cor correta	
Resposta B: transformação correta, cor incorreta	
Resposta C: transformação correta, cor incorreta	
<b>Resposta D: correta</b>	
<b>ITEM 16</b>	
Resposta A: transformação correta, mas cor diferente	
<b>Resposta B: correta</b>	
Resposta C: transformação incorreta, cor correta	
Resposta D: transformação e cor incorreta e acréscimo de 1 elemento	
<b>ITEM 29</b>	
Resposta A: resposta errada, única cor	
Resposta B: resposta errada, não intercala	
Resposta C: está intercalado correto, mas há cor a mais	
<b>Resposta D: correta</b>	
<b>ITEM 30</b>	
<b>Resposta A: correta</b>	
Resposta B: parcialmente correta, a cor está certa, mas não há preenchimento	
Resposta C: cor incorreta	
Resposta D: cor incorreta	
<b>ITEM 31</b>	
Resposta A: parcialmente correta, rotação do objeto de fora errada	
Resposta B: parcialmente correta, rotação do objeto de dentro errada	
<b>Resposta C: correta</b>	
Resposta D: resposta errada, cores e rotação erradas	
<b>ITEM 32</b>	
Resposta A: parcialmente correta, rotação correta, mas cor da bolinha errada	
Resposta B: resposta errada, rotação e cor da bolinha erradas	
Resposta C: resposta errada, rotação errada e sem a bolinha	
<b>Resposta D: correta</b>	

Até o momento foram descritos e discutidos os resultados da primeira parte do Estudo 1, que compreendeu a escolha dos itens para a composição do pré e pós-teste do instrumento dinâmico, bem como se fez um detalhamento dos itens selecionados. A segunda parte desse Estudo 1, descrita a seguir, buscou desenvolver a intervenção desenvolvida para o instrumento dinâmico e analisar os resultados da avaliação de juízes.

### **2.1.3 – Estudo 1 – Parte 2. Elaboração e avaliação por juízes da intervenção para resolução de itens indutivos**

Na segunda parte do Estudo 1 o objetivo se concentrou na construção da etapa de intervenção do raciocínio indutivo para compor o teste dinâmico e teve como finalidade: a) elaborar os procedimentos da intervenção; b) verificar a adequação do conteúdo com a teoria que o fundamenta e c) averiguar a clareza de compreensão da intervenção.

#### **2.1.3.1 – Participantes**

Para essa fase do estudo, participaram dois professores doutores e capacitados em intervenções de processos cognitivos, bem como 15 alunos entre mestrandos e doutorandos em Avaliação Psicológica. Dez crianças, duas de cada idade que o teste propõe abranger (ou seja, 6, 7, 8, 9 e 10 anos) também compuseram a amostra para esse momento do estudo. Cada idade foi representada pela série correspondente a idade escolar padrão, ou seja 6 e 7 na 1ª, 8 na 2ª, 9 na 3ª e 10 na 4ª.

#### **2.1.3.2 – Materiais**

##### **2.1.3.2.1 – Desenvolvimento da intervenção para raciocínio indutivo**

A construção dos procedimentos para a intervenção também foi embasada na Teoria Prescritiva do Raciocínio Indutivo de Klauer (1990). Três fases foram propostas visando ao desenvolvimento da competência do raciocínio indutivo, quais sejam, conhecimento

declarativo, procedural e estratégico. Resumidamente, primeiro é ensinado para a criança sobre os seis paradigmas do raciocínio indutivo, ou seja, que entre dois ou mais objetos sempre há características parecidas, diferentes, bem como pode haver uma relação entre eles. Posteriormente é ensinado como a criança precisa proceder para encontrar as regularidades entre atributos ou relações. Por fim, é ensinado a aplicar essa aprendizagem em outros contextos, que no caso do teste dinâmico será transferir essa aprendizagem para solucionar as tarefas do pós-teste.

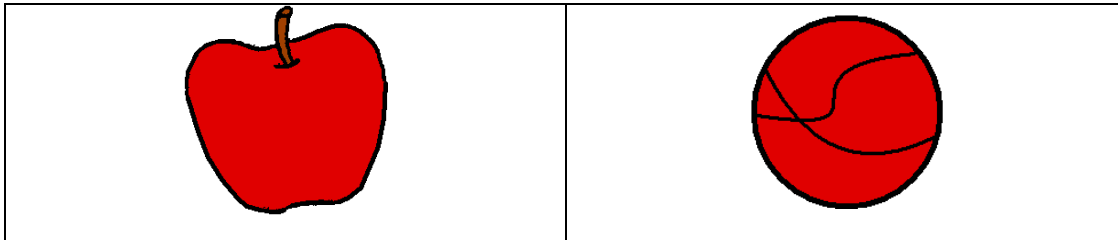
Abaixo é exemplificada parte do procedimento da intervenção, a qual corresponde ao paradigma de similaridade/diferenciação de atributos na fase do conhecimento declarativo. Logo após encontra-se outro exemplo, também de similaridade/diferenciação de atributos na fase do conhecimento procedural. No entanto, não há exemplificação para o conhecimento estratégico, pois é a aplicação da aprendizagem na intervenção em outras situações, no caso o pós-teste.

Exemplo 1:

*Todas as coisas podem ter algo em comum, ou seja, algo entre elas que combina. Então, figuras, animais, pessoas, brinquedos, comidas podem ter características comuns, ou seja, podem ter algo que é igual. Por exemplo, pode ser tamanho igual, forma igual, cor igual. Por exemplo, veja a bola e a maçã.*

*(bola vermelha e maçã vermelha)*





*O que a bola e a maçã têm igual? No que elas combinam?*

*Vamos olhar para as figuras e comparar para descobrir o que elas têm que combina.*

*Bom, a bola tem o formato redondo (ilumina<sup>4</sup> a bola) e a maçã também tem um formato redondo, um formato parecido com o da bola (ilumina a maçã). Então as formas da bola e da maçã são iguais, então a bola e a maçã combinam na forma.*

*Nesses desenhos, a bola e a maçã são da cor vermelha (ilumina os dois), então a cor também é igual. Então a bola e a maçã também combinam na cor.*

*Agora vamos fazer juntos algumas tarefas do jogo. Sempre lembrando de investigar o porquê as coisas são parecidas, vendo no que elas são parecidas. Então sempre vamos comparar as figuras e tentar descobrir no que elas são iguais, no que elas são parecidas.*

*Vamos começar?*

---

<sup>4</sup> A palavra iluminar indica que no momento da intervenção, que será gravado, o desenho correspondente a instrução para a criança será iluminado com uma moldura para ficar mais didático e ajudar a criança a atentar ao desenho. Esse recurso é uma ferramenta do software que será elaborado para o teste dinâmico.

Exemplo 2:

*(Mostrar como itens, sendo de cores diferentes, um quadrado, um tabuleiro de damas, uma caixa e um ponto de interrogação para ser preenchido pela alternativa correta. Logo abaixo dos itens mostrar um triângulo, um iô-iô e um boné e uma televisão também de cores diferentes)*

*Aqui o jogo pede para que a gente escolha o desenho certo para preencher o espaço com ponto de interrogação (ilumina esse espaço).*

*Então vamos ter que analisar qual das alternativas (ilumina as alternativas) é o desenho que é parecido com as figuras de cima. Qual desenho combina com as figuras de cima (ilumina os itens).*

*Vamos começar investigando o que as figuras de cima têm que seja parecido entre elas, no que elas combinam (ilumina os itens). Vamos comparando uma com outra para descobrir o que há de igual entre elas.*

*Vamos pensar, será que é a cor? O quadrado (ilumina o item) tem cor azul, o tabuleiro (ilumina o item) tem cor amarela e preta, e a caixa (ilumina o item) tem cor marrom. Todos têm cores diferentes, então não é a cor.*

*Vamos ver se é a função, ou seja, se eles servem pra fazer a mesma coisa. Bom, o quadrado (ilumina o item) é uma figura geométrica, a tabuleiro (ilumina o item) serve pra gente*

*brincar e a caixa (ilumina o item) serve pra gente guardar coisas. É, eles não servem para a mesma coisa.*

*Então vamos pensar se eles são de um mesmo grupo. Bem, o quadrado (ilumina o item) como já vimos é uma figura geométrica, o tabuleiro (ilumina o item) é um brinquedo e a caixa (ilumina o item) é de guardar coisas, então eles também não são de um mesmo grupo.*

*Ah, então será que é a forma? Bom, o quadrado (ilumina o item) tem a forma quadrada, o tabuleiro (ilumina o item) também tem a forma quadrada e a caixa (ilumina o item) também é quadrada. Então eles são parecidos, eles se combinam porque têm a forma igual, que é de quadrado. Então descobrimos que a figura correta também tem que ser da forma quadrada.*

*Agora que sabemos disso, ou seja, que as figuras de cima se parecem, se combinam porque são quadradas, então vamos pensar qual é a alternativa correta (ilumina todas as alternativas). Novamente vamos comparar as figuras.*

*Vamos começar pela primeira que é o triângulo (ilumina o triângulo). Bom, o triângulo não é um quadrado, então não é essa alternativa.*

*Vamos para a segunda, o iô-iô (ilumina a alternativa). Bom, o iô-iô tem a forma redonda, como uma bola, então não é parecido com a forma do quadrado que é a alternativa correta nessa tarefa.*

*A próxima alternativa que temos que investigar é um boné (ilumina a alternativa). Como podemos ver o boné também não tem a forma quadrada, então não é essa alternativa também.*

*Então vamos pensar na última alternativa que é a televisão (ilumina a alternativa). É, a televisão tem a forma quadrada parecida com as figuras de cima (ilumina os itens). Então essa é a figura correta. Vamos clicar e colocar ela ao lado das outras (coloca essa figura no lugar do ponto de interrogação). Pronto. Olha, realmente todas elas tem a forma parecida, então elas se combinam.*

*Viu só como não é difícil fazer as tarefas do jogo. Você só precisar ver o que os itens têm de igual, o que têm parecido, no que eles se combinam e depois pensar, comparar as figuras e ir vendo qual é a melhor figura para se encaixar no espaço que está faltando.*

### **2.1.3.3 – Procedimento**

Após o procedimento de intervenção inicialmente construído, pediu-se para que dois especialistas com experiência em psicometria e conhecedores da Teoria Prescritiva do Raciocínio Indutivo de Klauer, lessem e averiguassem se o procedimento da intervenção estava de acordo com a teoria que o embasa, bem como sugerissem melhorias tanto para o conteúdo teórico, quanto para o semântico. Após isso entregou-se para cada um dos mestrandos e doutorandos uma cópia dos procedimentos de intervenção, para que eles fizessem, individualmente, uma análise semântica e didática do conteúdo e das fases da

intervenção, bem como da compreensão deste. Por último, já com as reformulações sugeridas dos especialistas, doutorandos e mestrandos, aplicou-se individualmente os procedimentos da intervenção nas dez crianças com a finalidade de verificar a compreensão delas nesse procedimento. As crianças participantes foram selecionadas por conveniência em uma escola pública. Ressalta-se que nesse estudo os procedimentos de intervenção não foi computadorizada, assim o aplicador leu o conteúdo descrito no papel de maneira padronizada, ou seja, sem mudar ou interferir com mais informações.

Cada criança descreveu o que havia entendido sobre o que era para fazer, bem como foram feitas perguntas para averiguar melhor essa compreensão. Exemplo de pergunta: “A gente pode ver que as coisas são iguais ou parecidas olhando o quê?” ou “Para eu ver se duas coisas são iguais como é que eu faço para descobrir isso?”. No software computadorizado, a intervenção foi idêntica à forma no papel, ressaltando possíveis melhorias e retirando essa parte das perguntas. As instruções da intervenção computadorizada são gravadas com a fala de toda a etapa de intervenção, ou seja, a criança não ser alfabetizada.

#### **2.1.3.4 - Resultados e discussão do Estudo 1, Parte 2**

Nessa parte dos resultados não houve análises quantitativas, exceto frequência de compreensão do procedimento da intervenção e adequação do conteúdo. Aqui pautou-se mais nas análises qualitativas, que são essenciais para esse tipo de trabalho que se embasa na compreensão do indivíduo ao ler, ou ao ouvir (no caso das crianças), um texto corrido e

explicativo. Abaixo são apresentadas as quatro questões que os juízes tiveram que responder após lerem a intervenção.

- a) *A intervenção está compreensível para você? ( ) Sim ( ) Não. Justifique a sua resposta.*
- b) *A intervenção está compreensível para crianças de 7 a 11 anos? ( ) Sim ( ) Não. Justifique a sua resposta.*
- c) *A intervenção está coerente e com as fases de conhecimento declarativo e procedural? ( ) Sim ( ) Não. Justifique a sua resposta.*
- d) *Você tem sugestões ou críticas em relação à intervenção do ADRI? ( ) Sim ( ) Não. Justifique a sua resposta.*

Com relação à primeira pergunta, 100% dos juízes responderam que compreenderam a intervenção. As justificativas foram bastante semelhantes, sendo que uma resposta para exemplificar o conteúdo dessas justificativas é “está claro, pois está muito bem explicado, passo-a-passo e as figuras ajudam bastante a fixar o conteúdo”.

Na segunda pergunta houve 100% de resposta sim, ou seja os juízes concordaram que a intervenção estava compreensível para crianças, porém 80% dos 15 juízes comentaram que a intervenção estava demasiadamente longa. Todos os juízes responderam de forma bastante semelhante à pergunta anterior, dizendo que a intervenção estava compreensível para a faixa etária, mas ao mesmo tempo 80 % dos avaliadores também questionaram a duração da intervenção, que parecia estar muito longa. É importante ressaltar que, nesse

momento do estudo, o conteúdo da intervenção estava descrito no papel e ainda não havia sido gravado, tendo sido estimado que teria de 20 a 30 minutos de duração.

Já na terceira pergunta os juízes opinaram 100% pela resposta afirmativa, justificando que a intervenção abordava os pressupostos do conhecimento declarativo e procedural. Por fim, na questão 4, 80% dos juízes, os mesmo da questão 2, sugeriram, se possível, a diminuição do tempo da intervenção, pois estava muito longo e as crianças poderiam se cansar e não mais fixar a atenção na tarefa proposta. Além disso, 100% dos juízes emitiram outras sugestões, mas todas referentes a alguma incoerência da figura mostrada com o que estava escrito sobre ela. Também foram aferidas sugestões de palavras mais adequadas e correções gramaticais.

Por meio dessa análise dos juízes pôde-se verificar que o procedimento da intervenção estava coerente com a teoria que o embasava, estava compreensível, mas o ponto crítico foi a duração do teste, que estava longa e a criança poderia se cansar e não se atentar para os elementos importantes da tarefa de intervenção. Procurou-se acatar todas as sugestões, porém não foi possível reduzir o tempo da intervenção. Após analisar o que poderia ser diminuído, concluiu-se que a redução retiraria informações úteis; apenas o último exemplo, que referia ao paradigma SC, foi retirado com a justificativa de que, se o indivíduo compreendesse os outros paradigmas, poderia transferir as informações transmitidas para resolver o SC. Ressalta-se que a questão da duração longa da etapa de intervenção já era do conhecimento da autora do presente estudo, mas para ensinar a

resolução de itens indutivos abrangendo todos os paradigmas e os dois conhecimentos (declarativo e procedural) nesses paradigmas, a extensão fez-se necessária e, durante todo o desenvolvimento do projeto, foi ciente que a variável tempo na intervenção poderia interferir negativamente.

Após a elaboração dos procedimentos de intervenção e a análise pelos juízes do procedimento construído, aplicou-se a etapa de intervenção completa, mas não computadorizada em 10 crianças, de 1ª à 4ª série, idade entre 6 a 10 anos, respeitando a idade padrão de cada série. Foi analisado o relato das crianças sobre o que haviam entendido do procedimento da intervenção e algumas questões foram feitas para verificar se realmente haviam compreendido como raciocinar indutivamente. Esse momento da coleta não foi padronizado e também não foi registrado, visto que o objetivo não era uma análise qualitativa estruturada. Por meio desse contato com as crianças observou-se que o procedimento de intervenção estava adequado em relação à linguagem e às explicações, ressaltando que não foi exigido da criança que ela explicasse a forma de como resolver cada paradigma, somente foi focado o fundamento básico que é saber identificar diferenças e semelhanças entre objetos.

Para poder verificar uma possível aprendizagem momentânea das informações da intervenção, foram aplicadas duas tarefas, uma de SD e a outra de RE por analogia. A maioria das crianças acertou a SD, somente duas não conseguiram, uma da 1ª e a outra da 2ª série, mas na RE somente as crianças da 3ª e 4ª acertaram. Por meio dessa aplicação não



há como garantir que os acertos foram decorrentes da intervenção, pois não foi feita anteriormente uma linha de base da capacidade, mas são indicativos, por exemplo, que a tarefa RE tende a ser acertada por crianças mais velhas.

Com o pré-teste, pós-teste e intervenção desenvolvidos, houve a junção dessas três partes para a formação do teste dinâmico no formato computadorizado e padronizado, de forma que as crianças pudessem receber as mesmas informações. Essa versão pressupõe somente a criança interagindo com o computador. Em seguida, foi então conduzido o Estudo 2, descrito a seguir.

## **2.2 - Estudo 2- Análises psicométricas da estrutura interna e externa do instrumento dinâmico.**

Refere-se ao processo de precisão e evidências de validade do instrumento na forma dinâmica de testagem. Para isso foram verificadas a precisão do instrumento, os índices de ajuste por meio da análise de Rasch, a estrutura fatorial pelo programa Microfact, as evidências de validade convergente-discriminante com os testes Prova de Raciocínio Abstrato da BPR-5 infantil, Geração Semântica e Stroop, as evidências de validade teste-critério com o desempenho acadêmico, bem como a validade incremental do instrumento dinâmico em prever o desempenho acadêmico. Porém, antes de iniciar as evidências de validade com outros testes, investigou-se a capacidade do teste dinâmico em mensurar o potencial de aprendizagem.

### **2.2.1 - Participantes**

Participaram desse segundo estudo 317 crianças que cursavam da 1ª à 5ª série. A divisão da amostra por série se configurou em 71 crianças na 1ª série, 87 na 2ª série, 65 na 3ª série, 51 na 4ª série e 44 na 5ª série. Com relação ao sexo, o feminino compôs 52,5% da amostra. A média de idade e desvio padrão dos participantes separados por série são 1ª série  $M=6,83$  e  $DP=0,41$ , 2ª série  $M=8,06$  e  $DP=0,57$ , 3ª série  $M=9,23$  e  $DP=0,74$ , 4ª série  $M=10,49$  e  $DP=0,83$  e 5ª série  $M=11,39$  e  $DP=0,62$ . A amostra coletada foi por conveniência, sendo que o critério de exclusão foi deficiência auditiva, visual, intelectual ou algum problema que pudesse impossibilitar o sujeito realizar os testes. Como na escola

de 1ª à 4ª série os alunos com algum tipo de deficiência estudavam em sala separada, tal classe não foi incluída no estudo, facilitando o emprego dos critérios de exclusão. Na escola da 5ª série não foi constatado aluno que entrasse no critério de exclusão. Além disso, as professoras também indicavam os alunos que se enquadravam nesses critérios de exclusão.

## **2.2.2 - Materiais**

### **2.2.2.1 - Teste Dinâmico Informatizado de Raciocínio Indutivo**

A partir do Estudo 1, foi desenvolvido o Teste Dinâmico Informatizado de Raciocínio Indutivo para avaliar o potencial de aprendizagem da criança para resolver tarefas que exijam raciocínio indutivo. Construiu-se o pré e pós-teste, cada um com dezesseis itens escolhidos a partir do primeiro estudo. O procedimento de intervenção desenvolvido no segundo estudo compôs a parte intermediária, o que caracteriza o teste como dinâmico.

O teste como um todo foi computadorizado. Para isso desenvolveu-se um software na linguagem Delphi que faz a aplicação com três fases, uma seguida da outra, de forma automática, conforme a criança resolva as tarefas. Essas fases compreendem o pré-teste, a etapa da intervenção e o pós-teste. Com relação ao pré e o pós teste, as tarefas foram apresentadas ao indivíduo, uma a uma. A passagem de um item para outro ocorre imediatamente após a escolha da resposta pelo indivíduo, ou seja, após clicar sobre a alternativa que julgar ser a correta, não havendo a possibilidade de mudança de opção. Essa

informação de escolha única é esclarecida para o indivíduo no teste. As tarefas sempre são compostas por 3 ou 4 figuras e um espaço em branco com um ponto de interrogação que deve ser completado por uma alternativa. Sempre há quatro alternativas de resposta, com apenas uma correta. O indivíduo escolhe a resposta que julgar correta clicando apenas uma vez sobre a alternativa. Automaticamente a alternativa preenche o espaço em branco.

As instruções do pré e pós teste, além de escritas na tela do computador, foram gravadas para que a criança possa ouvi-las, ou seja, não é necessário a criança ter a aquisição da leitura. Ressalta-se que, antes de iniciar os 16 itens, há três tarefas para exemplificar o que é para fazer no pré-teste. Para o pós-teste não é necessário, pois ele se inicia após o término da intervenção que, ao final, faz a ligação com o pós-teste: *“Bom, agora você irá jogar sozinho novamente. Lembre-se do que aprendemos. Verifique sempre o que as coisas têm que são parecidos, como cor, tamanho, formato, função, ou seja, se servem para a mesma coisa, ou mesmo se a tarefa descreve um começo, meio e fim. Mas não se esqueça que também as coisas podem ter relação, assim como vimos nessa última tarefa e na tarefa da luva que coloca na mão e do boné que coloca na cabeça. Vamos lá, agora é com você”*.

Com relação a etapa de intervenção, esta se inicia após a criança ter respondido o item 16 do pré teste. Nessa parte a criança apenas ouve e/ou lê as instruções da intervenção, sem precisar efetuar nenhuma ação. A etapa da intervenção é feita apenas uma vez, sem a opção de refazê-la completamente ou em parte. Para a intervenção o software foi

programado para iluminar os desenhos que são citados ao longo da fala. Ex: “*A primeira figura das alternativas é o menininho dormindo (iluminar a alternativa). Essa figura é igual à primeira figura dos desenhos acima (iluminar esse item de cima), então não é, mesmo porque geralmente depois que tomamos o café da manhã, a gente não volta a dormir, a gente vai fazer outra coisa, como ir pra escola, ou assistir televisão, ou ir brincar ou fazer tarefa de casa, não é mesmo?*”

Para testes informatizados, como o construído no presente estudo, a interferência do aplicador é praticamente nula, sendo que a interação da criança ocorre com o computador. E isso é explicado para a criança antes de começar a responder ao teste. Na aplicação a criança utiliza um fone de ouvido para ouvir as instruções dos exemplos antes do pré-teste e o procedimento da intervenção, que é a fase do meio do teste. O aplicador apenas interfere no item de exemplo, assegurando que a criança compreendeu o que é para ser feito no pré-teste.

A correção do teste dinâmico para verificar o potencial de aprendizagem é realizada pela subtração entre a pontuação do pós e do pré-teste. No entanto, até o momento não há uma normalização em relação à pontuação. Então maiores informações sobre a efetivação dos escores no teste dinâmico se encontram na seção de resultados.

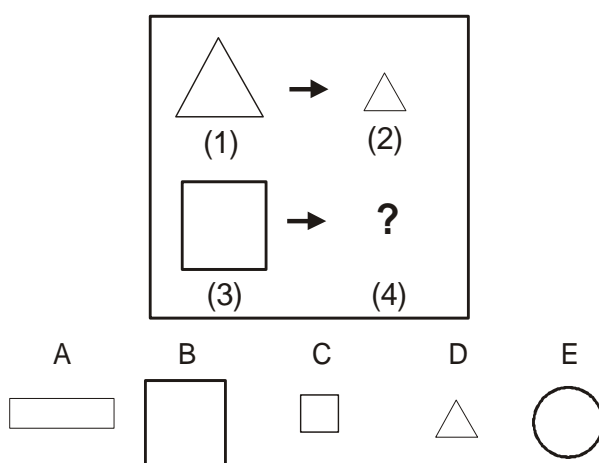
**2.2.2.2- Raciocínio Abstrato da Bateria de Provas de Raciocínio-BPR5i (Primi, Almeida & Cruz, em construção).**

A Bateria de Provas de Raciocínio Infantil (BPR-5i) foi criada no sentido de avaliar a realização cognitiva dos estudantes que cursam de 1<sup>a</sup> a 5<sup>a</sup> séries do ensino fundamental, no que diz respeito à apreensão de relações entre elementos (raciocínio indutivo) e aplicação das relações inferidas a novas situações (raciocínio dedutivo) (Lemos, Almeida & Guisande, 2006). Esta bateria é formada por quatro provas que se diferem no conteúdo em que os itens são formulados, mas que têm em comum, a avaliação do raciocínio como uma função cognitiva dominante, isto é, o fator geral (fator *g*), no entanto tais provas avaliam mais especificamente a inteligência fluida. As provas que compõem são Raciocínio Abstrato, Raciocínio Verbal, Raciocínio Numérico e Raciocínio Prático.

No presente estudo foi utilizada somente a prova de Raciocínio Abstrato. Essa prova foi eleita para o estudo por mensurar a inteligência fluida de forma mais pura em relação as outras provas. O Raciocínio Abstrato apresenta correlação positiva e significativa de 0,69 com as Matrizes Progressivas Coloridas de Raven. Este último teste avalia a inteligência fluida e a correlação encontrada foi alta, o que permite afirmar convergência entre os dois testes, ou seja, estão avaliando o mesmo construto. Ressalta-se ainda que a prova RA apresenta um índice de precisão de 0,91 por meio do procedimento Alfa de Cronbach, considerado muito alto.

A Prova RA é composta por trinta itens de conteúdo abstrato. Nesta prova são apresentadas analogias figurativas, que o sujeito deverá completar. Para tanto, é necessário que o indivíduo aprenda a relação entre os dois primeiros elementos (1) e (2) e descubra

uma quarta figura (?) que venha a repetir essa relação inferida com o terceiro elemento (3) indicado, conforme se pode verificar no exemplo que se segue (Figura 26).



**Figura 17**– Exemplo de um item da Prova RA da BPR5i.

A relação entre (1) e (2) deve ser aplicada no segundo par de figuras a constituir: (3) e (4). Neste caso, o triângulo (1) transformou-se ficando menor em (2). Transformando-se do mesmo modo, o quadrado (3) ficaria menor como na alternativa C.

A resposta do indivíduo consiste em assinalar, dentre cinco alternativas de resposta, aquela que considera correta para completar as relações da analogia. Não há tempo limite para a realização desta prova. Os resultados dos sujeitos na prova correspondem ao número de itens respondidos corretamente.

### 2.2.2.3 – Teste de Geração Semântica (Capovilla & Macedo, no prelo)

Para avaliar o controle inibitório, que é a filtragem e seleção de informações, utilizou-se a tarefa de geração semântica informatizada. Nela são apresentadas 120 figuras ilustrando substantivos e o sujeito deve dizer uma ação (i.e., um verbo) semanticamente relacionada a cada substantivo. Os itens pertencem a duas categorias, alta e baixa seleção. Na categoria de alta seleção os substantivos, no caso as figuras, podem ser associadas a várias palavras; já na de baixa seleção, os substantivos são mais facilmente associados a apenas uma palavra.

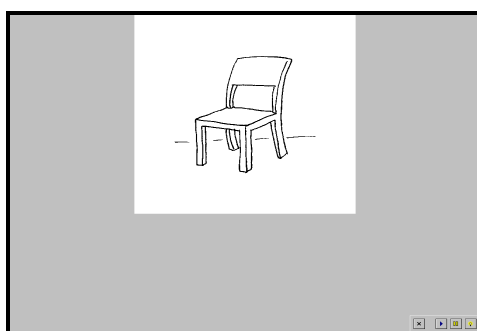
A tarefa de geração semântica foi computadorizada de forma a facilitar a aplicação e o registro das respostas, e foi executada em notebook. A instrução inicial é “Você verá alguns desenhos e deverá dizer um verbo, ou seja, uma ação que você considere relacionada a cada desenho. Por exemplo, diante do desenho de “lápiz”, você poderá dizer “escrever”. Se aparecer o desenho de “pente”, você poderá dizer “pentear”. Vamos começar”.

Cada figura é apresentada, uma a uma, sendo registrada a vocalização do sujeito em cada figura, bem como o tempo de reação. O examinador pode operar o software, selecionando os botões “Parar”, “Ir para o próximo” ou “Pausa” que aparecem no canto inferior direito da tela do computador. A Figura 18 ilustra uma das pranchas do Teste de Geração Semântica com o desenho de cadeira.

A versão computadorizada permite reexaminar auditivamente as locuções do sujeito, o que possibilita ao examinador fazer análises qualitativas e quantitativas ulteriores. Foram analisados diferentes desempenhos, como escore nos itens de baixa seleção, escore



nos itens de alta seleção, tempo de reação nos itens de baixa seleção e tempo de reação nos itens de alta seleção. Esse teste foi criado e tem sido utilizado com diversas faixas etárias que abrange crianças com cinco anos de idade até idosos com 80 anos.



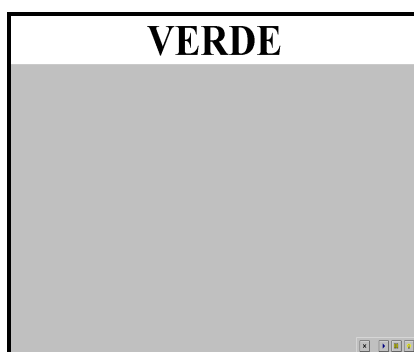
**Figura 18-** Layout da tela para a figura de “cadeira” do Teste de Geração Semântica.

#### **2.2.2.4 – Teste de Stroop Computadorizado (Capovilla, Montiel, Macedo & Charin, no prelo)**

O Teste de Stroop Computadorizado foi aplicado para avaliar a atenção seletiva, ou seja, a capacidade do sujeito de atentar a determinadas características do estímulo, ignorando características irrelevantes à tarefa. O teste é composto em três partes, cada uma contendo 24 estímulos.

Esta versão computadorizada apresenta um estímulo por vez na tela do computador, sendo que, em cada parte, a tela inicial corresponde às instruções escritas descrevendo a tarefa do participante. Tais instruções, além de escritas na tela, também são lidas em voz

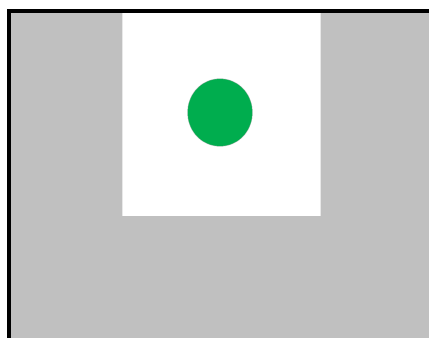
alta pelo aplicador ao sujeito. A primeira parte do teste apresenta os nomes de quatro cores (amarelo, azul, verde e vermelho), com palavras escritas em letras maiúsculas, fonte Times New Roman, tamanho 72, em tinta preta. Cada nome aparece seis vezes, em ordem pseudo-randômica, de forma que uma mesma palavra não apareça duas vezes seguidas. Como o objetivo desta parte do teste é apenas verificar se o sujeito tem leitura correta dos itens, cada palavra fica exposta por tempo indeterminado na tela, sendo a tarefa do sujeito ler as palavras escritas tão rapidamente quanto possível. A Figura 19 ilustra uma tela da primeira parte do Teste de Stroop Computadorizado.



**Figura 19** - Layout da tela para a parte 1 do Teste de Stroop Computadorizado.

A parte 2 apresenta 24 círculos coloridos, sendo seis círculos para cada uma das quatro cores, distribuídos pseudo-randomicamente de modo que uma mesma cor não apareça duas vezes seguidas. Cada círculo fica exposto por 40 ms, sendo a tarefa do sujeito dizer a cor dos círculos tão rapidamente quanto possível. O objetivo desta parte é servir como linha de base para a análise de acertos e tempo de reação da terceira parte, em que o

sujeito também deve dizer a cor de estímulos, porém tais estímulos são palavras escritas correspondentes a nomes de cores em situação divergente (i.e., em que o nome de cor escrito não corresponde à cor da letra). A Figura 20 ilustra uma tela da segunda parte do Teste de Stroop Computadorizado.



**Figura 20-** Layout de uma tela da parte 2 do Teste de Stroop Computadorizado.

Na terceira parte, semelhante à segunda parte, os círculos são substituídos por nomes escritos de cores, não havendo concordância entre o nome escrito e a cor da tinta. Os desempenhos no Teste de Stroop correspondem ao efeito de interferência cor-palavra para escore e tempo de reação, ou seja, o número de acertos na parte 3 subtraído do número de acertos na parte 2, e o tempo de reação médio da parte 3 subtraído do tempo de reação médio da parte 2. Tais desempenhos devem ser computados apenas para sujeitos que

tenham pelo menos 80% de acerto na parte 1 (Regard, 1981). A Figura 21 ilustra uma tela da terceira parte do Teste de Stroop Computadorizado. Ressalta-se que esse teste, também foi criado e tem sido utilizado com diversas faixas etárias que abrange crianças com cinco anos de idade até idosos com 80 anos.



**Figura 21-** Layout da tela para a parte 3 do Teste de Stroop Computadorizado.

#### **2.2.2.5 – Desempenho Acadêmico**

Para a obtenção das notas acadêmicas dos alunos, coletaram-se, com a permissão da direção das escolas, as notas finais dos alunos em cada disciplina. Além desse critério, também se criou uma escala com os nomes dos alunos para que as respectivas professoras, de cada série, pontuassem os cinco alunos com maior e os cinco com menor facilidade de aprender de acordo com as observações no decorrer do ano. Pediu-se para que não considerasse comportamento, somente a facilidade de aprender quando são ensinados os conteúdos. Ressalta-se que esse critério de nota atribuído pela percepção das professoras

não foi possível ser realizado com a amostra da 5ª série, pois como são diversos professores com menor quantidade de contato, pois são menos aulas por semana, a nota atribuída poderia refletir de maneira mais frágil a percepção sobre os alunos com maior facilidade e dificuldade para aprender.

### **2.2.3 – Procedimento**

Após o consentimento dos pais por meio do Termo de Consentimento Livre e Esclarecido (0092.0.142.000-07) para que as crianças participassem da pesquisa, assim como o consentimento destas, iniciou-se a coleta dos dados. Para tanto, o Teste de Geração Semântica e o Teste de Stroop foram aplicados de forma individual nos participantes da pesquisa, em duas etapas com intervalo de uma a duas semanas entre cada etapa. O Teste Dinâmico de Raciocínio Indutivo foi aplicado em grupos de 7 crianças no máximo. Já a prova do Raciocínio Abstrato foi aplicada coletivamente nos grupos de cada série. O tempo de aplicação do Teste Dinâmico foi aproximadamente de 45 minutos, para a aplicação do Geração Semântica 25 minutos, o Stroop 15 minutos cada e a RA 30 minutos.

As aplicações coletivas do Teste Dinâmico ocorreram no laboratório de informática das escolas. O Teste Dinâmico foi realizado em apenas uma sessão em que a criança era submetida as três fases, pré-teste, intervenção e pós-teste. Já a prova RA foi realizada dentro das salas de aula de cada série. A coleta individual do Geração e do Stroop foi realizada em uma sala de aula vazia que continha uma mesa e uma cadeira de tamanhos adequados para as crianças. Ressalta-se que além da autora da pesquisa, outras três pessoas

devidamente treinadas colaboraram na coleta dos dados individuais. Para a aplicação dos testes que ocorrem fora da sala de aula, as crianças foram retiradas das suas salas pelo aplicador que a conduziu até a sala de aplicação da tarefa. Nesse momento explicou-se a finalidade do teste que iriam realizar. Por fim o Desempenho Acadêmico foi coletado no final do ano na secretaria da escola com a permissão da direção e também foi nesse período que as professoras responderam à escala de facilidade de aprender. Para melhor compreensão das aplicações, elaborou-se a Tabela 29 com informações de qual teste aplicado, tipo de aplicação, quantidade de sujeitos por aplicação e total dos respondente, bem como tempo de aplicação para cada sessão.

**Tabela 29**– Informações sobre as sessões das aplicações dos testes utilizados no Estudo 2

Teste	Sujeitos por sessão	Total de respondentes	Tipo de aplicação	Tempo da aplicação (aproximadamente)
Dinâmico	Máximo 7	318	Coletiva	45 minutos
RA	Máximo 30	303	Coletiva	25 minutos
Stroop	1 sujeito	172	Individual	15 minutos
Geração	1 sujeito	172	Individual	25 minutos

#### 2.2.4 Resultados e Discussão do Estudo 2

Para alcançar os objetivos do Estudo 2, ou seja, verificar as propriedades psicométricas do teste dinâmico construído no Estudo 1, foram verificadas a precisão do instrumento, os índices de ajuste de seus itens por meio da análise efetuada pelo modelo de Rasch, a estrutura fatorial oblíqua pelo programa Microfact, as evidências de validade convergente-discriminante com os testes Prova de Raciocínio Abstrato da BPR-5 infantil,

Geração Semântica e Stroop, as evidências de validade teste-critério com o desempenho acadêmico, bem como a validade incremental do instrumento dinâmico em prever o desempenho acadêmico. Além dessas análises também são apresentadas as estatísticas descritivas dos instrumentos que possibilitam uma melhor compreensão dos dados encontrados. Foi efetuada, ainda, uma análise dos erros no instrumento dinâmico com a finalidade de discutir melhor a efetividade da intervenção desenvolvida, bem como explorar essas informações para entender o que levou ao erro.

É importante destacar que nesse tipo de teste dinâmico há três principais escores, quais sejam, total do pré-teste, total do pós-teste e total do ganho, ou seja, quanto o indivíduo conseguiu melhorar no pós em relação ao pré-teste. Portanto as análises descritivas do instrumento dinâmico e as evidências de validade foram realizadas para esses três escores. Isso foi feito porque, apesar do escore mais importante em um instrumento dinâmico ser o ganho, os escores no pré e no pós-teste também podem fornecer informações relevantes para a compreensão do próprio ganho e também para o instrumento como um todo. Além disso, algumas vezes, quando pertinente, também foram utilizados os escores separados de cada paradigma no pré e pós.

Deve ser destacado, também, que as análises das propriedades psicométricas internas, como precisão e índices de ajuste, bem como a análise da dimensionalidade, foram efetuadas dispendo pré e pós separadamente. Isso foi feito porque, apesar do instrumento ser único, compreende-se que, por causa da etapa de intervenção, pode ocorrer uma

mudança, mesmo que temporária, do entendimento do indivíduo em relação a como resolver os itens no pós-teste. Pode-se dizer que os sujeitos do pré são diferentes dos sujeitos do pós, apesar de serem os mesmos indivíduos. Logo, apesar do pré e pós-teste serem formas paralelas do mesmo teste, os indivíduos podem ter mudado, o que caracteriza a necessidade de analisar as partes separadamente, bem como são formas que podem ter propriedades bem diferentes.

#### **2.2.4.1 Análise do Teste Dinâmico de Raciocínio Indutivo pela Teoria de Resposta ao Item**

Inicialmente, nessa seção são pontuados alguns índices para o pré e o pós-teste a partir da Teoria de Resposta ao Item com o modelo de Rasch, quais sejam, índices de ajuste, correlação item-theta, dificuldade e precisão (*Real*). Tais informações encontram-se descritas na Tabela 30.



**Tabela 30**– Índices psicométricos do pré e pós-teste por meio Teoria de Resposta ao Item com o modelo de Rasch.

Item	Dificuldade	<i>Infit</i>	<i>Outfit</i>	Correlação
PRÉ-TESTE/Precisão 0,42				
A03SD_01	-2,21	1,04	1,09	0,31
C06SD_02	-0,52A	0,94	0,90	0,44
B35SD_03	1,05	1,10	1,12	0,20
<b>B01SD_04</b>	<b>0,98</b>	<b>1,16</b>	<b>1,18</b>	<b>0,15</b>
<b>A08CC_05</b>	<b>0,98A</b>	<b>1,24</b>	<b>1,20</b>	<b>0,35</b>
A37CC_06	0,71	0,91	0,86	0,42
A11CC_07	-0,85	1,01	1,00	0,38
B14CC_08	0,62A	1,07	1,17	0,26
B16RE_09	0,86A	1,10	1,19	0,32
A17RE_10	0,45A	0,92	0,87	0,35
A20RE_11	0,16A	0,93	0,86	0,41
A21RE_12	-0,22A	0,99	1,00	0,33
B30SC_13	1,01A	1,20	1,19	0,31
B27SC_14	0,87	0,98	0,97	0,32
<b>A32SC_15</b>	<b>1,44</b>	<b>1,03</b>	<b>1,56</b>	<b>0,19</b>
A26SC_16	-0,41A	0,90	0,91	0,46
PÓS-TESTE/Precisão 0,45				
B07SD_17	-0,97	1,08	1,08	0,28
B34DS_18	-0,72	0,95	0,92	0,42
B03SD_19	-0,32A	0,94	0,92	0,44
C04SD_20	-0,58A	0,96	0,92	0,41
<b>A36CC_21</b>	<b>2,53</b>	<b>1,11</b>	<b>1,42</b>	<b>0,09</b>
B08CC_22	1,27	1,07	1,14	0,24
A14CC_23	1,03	1,06	1,16	0,26
A13CC_24	-0,16A	0,90	0,88	0,46
<b>B20RE_25</b>	<b>2,36A</b>	<b>1,14</b>	<b>1,48</b>	<b>0,06</b>
<b>A38RE_26</b>	<b>0,70A</b>	<b>1,22</b>	<b>1,28</b>	<b>0,17</b>
C22RE_27	1,08	0,97	1,04	0,34
B17RE_28	0,15A	0,90	0,86	0,50
A29SC_29	2,20	0,99	1,04	0,25
A34SC_30	1,64	0,92	0,79	0,39
A30SC_31	0,43A	1,11	1,13	0,27
B24SC_32	0,39	0,87	0,82	0,49

*Nota: A primeira letra designa a Forma, sendo que a C representa os itens âncoras. Os itens em negrito foram os que apresentaram os índices psicométricos, ajuste e correlação, não adequados.*

A Tabela 30 permite verificar algumas propriedades psicométricas do pré e do pós-teste analisadas pela Teoria de Resposta ao Item com o modelo de Rasch. Na coluna do índice de dificuldade, os itens ancorados na análise das Formas A e B estão destacados pela letra A, pois mostraram-se adequados com a medida obtida anteriormente. Porém os outros itens âncoras, como apresentaram índice de dificuldade que variaram de forma não aceitável em relação a medida anterior, não foram usados para a ancoragem, sendo deixados livres na análise. Observou-se que os padrões de dificuldade do pré e do pós-teste foram semelhantes às dificuldades iniciais verificadas nas Formas A e B, porém alguns itens apresentaram maiores variâncias. Para verificar se a variação é aceitável, se recorre ao *displacement*, que é uma medida do quanto a dificuldade do item variou, entre duas aplicações. O valor aceitável é no máximo 0,50. Essa variância dos parâmetros pode ter ocorrido em função da forma de aplicação, ou seja, estimou-se os parâmetros de dificuldade dos itens inicialmente com as Formas A e B, que foram aplicadas na forma lápis e papel. A estimação nesse segundo momento, ancorada na inicial, foi realizada com os itens aplicados em forma computadorizada.

Em relação aos índices de ajuste, *infit e outfit*, e à correlação *item-theta*, somente três itens em cada parte apresentaram inadequações. Tais itens estão em negrito na Tabela 30. Apenas o item A38RE\_26 demonstrou problema nos três índices. As precisões observadas podem ser consideradas baixas, no entanto deve-se considerar a pouca heterogeneidade da amostra, composta por sujeitos com pouca habilidade em relação ao teste. que será mostrada nas análises seguintes. Ressalta-se que, apesar da observância de

desajuste em alguns itens, nas análises seguintes foram considerados todos os itens, pois estas foram análises iniciais de um teste ainda em construção e com poucos itens, portanto não se justifica a eliminação de itens nesse momento da análise.

#### **2.2.4.2 - Análise das estatísticas descritivas e inferenciais no Teste Dinâmico de Raciocínio Indutivo em função de tipo de paradigma, momento da testagem e série**

Nessa seção são expostas as estatísticas descritivas. Nessas análises, além dos escores totais no pré e no pós-teste, também foram considerados os escores específicos em cada paradigma. Isso foi feito porque, mesmo tendo sido verificada a unidimensionalidade do pré e do pós-teste, como ocorreu com as Formas A e B do Estudo 1, é importante analisar os escores dos paradigmas, pois se espera que os sujeitos acertem mais itens SD e menos SC, independente de ser pré ou pós-teste, conforme já explicado anteriormente.

Ressalta-se, ainda, que essa análise foi realizada utilizando a dificuldade dos paradigmas e a habilidade das pessoas calculados pelo modelo de Rasch, e não o escore bruto comumente analisado. Isso porque, para que se possa comparar o desempenho do indivíduo no pré e no pós-teste e em cada paradigma, é necessário que os itens possuam um ponto de referência em relação à dificuldade. Assim, tanto os itens do pré quanto do pós-teste foram equalizados a partir dos índices de dificuldade obtidos no Estudo 1. Essa equalização se faz necessária, pois garante saber os índices de dificuldade previamente definidos, mantendo a comparabilidade no pré e pós-teste, colocando-os numa escala que propicia um mesmo sentido em relação aos valores, permitindo assim uma comparação

entre esses dois momentos. Essa equalização auxilia, por exemplo, a solução do problema dos itens terem ficado mais fáceis no pós-teste.

É importante destacar que na equalização foram ancoradas (utilizados os índices de dificuldade já estimados no Estudo 1) os índices de dificuldade de nove itens no pré e sete no pós-teste. A princípio ancoraram-se todos os itens, mas o índice *displacement*, que indica o quanto a medida de dificuldade poderia variar se ficasse livre, apontou que alguns itens mudariam mais e optou-se por não ancorá-los. A Tabela 31 apresenta a média de dificuldade do pré e pós-teste gerais, seus respectivos paradigmas e as habilidades das pessoas, o que possibilita comparar a dificuldade com a habilidade, além de se verificar qual paradigma os sujeitos apresentam melhor desempenho.

**Tabela 31**– Índices de dificuldade do teste e habilidade dos sujeitos.

Variáveis	Dificuldade	Habilidade
Pré-teste geral	0,31	-0,69
SD_pré	0,16	-0,40
CC_pré	0,51	-0,38
RE_pré	0,47	-0,56
SC_pré	0,94	-0,33
Pós-teste geral	0,69	-0,25
SD_pós	-0,68	-0,32
CC_pós	1,03	-0,38
RE_pós	1,03	-0,13
SC_pós	1,03	-0,26

Por meio da Tabela 31 constata-se que o pré e o pós-teste foram difíceis de serem resolvidos perante as habilidades dos sujeitos. Um nível de dificuldade ou habilidade médio

seria 0. A maioria das diferenças entre dificuldade e habilidade foram maiores que a distância de 1 ponto, exceto com o pós-teste geral e SD\_pós. Esse último, inclusive, foi o único em que a média de habilidade dos sujeitos superou a dificuldade dos itens. Esses dados refletem que o teste dinâmico para essa amostra foi difícil, apesar dos índices de dificuldade não serem altos.

Como a amostra foi composta por alunos de diferentes anos escolares (1<sup>a</sup> à 5<sup>a</sup>), espera-se que indivíduos de séries mais avançadas apresentem maior habilidade do que alunos de séries iniciais. A Tabela 32 mostra os índices de habilidade dos sujeitos separando-os por série. A média de dificuldade do teste (pré, pós, paradigma) é a mesma.

**Tabela 32-** Índices de habilidade dos sujeitos separando-os por série.

Variáveis	N	Mínimo	Máximo	Média	D.P
<b>1ª série</b>					
SD_pré	72	-2,97	1,63	-0,85	1,40
CC_pré	72	-2,08	1,67	-0,53	1,14
RE_pré	72	-2,04	1,59	-0,73	1,09
SC_pré	72	-1,69	2,13	-0,55	1,03
SD_pós	72	-3,19	1,82	-0,83	1,13
CC_pós	72	-1,74	2,31	-0,42	1,23
RE_pós	72	-1,67	2,28	-0,35	1,05
SC_pós	72	-1,64	1,03	-0,67	0,99
Pré_geral	72	-10,47	4,58	-3,20	3,24
Pós_geral	72	-6,85	4,88	-2,26	2,36
<b>2ª série</b>					
SD_pré	88	-2,97	3,15	-0,56	1,42
CC_pré	88	-2,08	3,08	-0,66	1,14
RE_pré	88	-2,04	1,59	-0,70	1,06
SC_pré	88	-1,69	2,13	-0,69	0,96
SD_pós	88	-3,19	1,82	-0,74	1,17
CC_pós	88	-1,74	2,31	-0,53	1,18
RE_pós	88	-1,67	2,28	-0,21	1,00
SC_pós	88	-1,64	2,26	-0,46	1,15
Pré_geral	88	-9,08	6,59	-3,30	3,07
Pós_geral	88	-8,24	3,49	-2,00	2,59
<b>3ª série</b>					
SD_pré	65	-2,97	1,63	-0,26	1,22
CC_pré	65	-2,08	3,08	-0,49	1,26
RE_pré	65	-2,04	1,59	-0,79	1,04
SC_pré	65	-1,69	2,13	-0,19	1,14
SD_pós	65	-1,80	1,82	-0,01	1,23
CC_pós	65	-1,74	2,31	-0,10	1,26
RE_pós	65	-1,67	2,28	-0,26	1,09
SC_pós	65	-1,64	3,73	0,15	1,21
Pré_geral	65	-10,47	5,81	-1,91	3,64
Pós_geral	65	-5,41	8,66	-0,22	2,86

Tabela 32 (continuação)- Índices de habilidade dos sujeitos separando-os por série.

Variáveis	N	Mínimo	Máximo	Média	D.P
<b>4ª série</b>					
SD_pré	51	-2,97	3,15	-0,27	1,23
CC_pré	51	-2,08	1,67	-0,45	1,02
RE_pré	51	-2,04	2,98	-0,49	1,30
SC_pré	51	-1,69	3,56	-0,31	1,35
SD_pós	51	-1,80	1,82	0,14	1,21
CC_pós	51	-1,74	3,83	-0,60	1,38
RE_pós	51	-1,67	3,81	-0,18	1,17
SC_pós	51	-1,64	2,26	-0,38	1,01
Pré_geral	51	-9,03	9,45	-1,82	4,51
Pós_geral	51	-5,41	8,97	-1,02	3,15
<b>5ª série</b>					
SD_pré	44	-2,97	1,63	0,31	0,10
CC_pré	44	-2,08	3,08	0,65	1,30
RE_pré	44	-2,04	2,98	0,25	1,15
SC_pré	44	-1,69	3,56	0,50	1,25
SD_pós	44	-3,19	1,82	0,35	1,26
CC_pós	44	-1,74	2,31	-0,19	1,17
RE_pós	44	-1,67	2,28	0,68	1,25
SC_pós	44	-1,64	3,73	0,40	1,33
Pré_geral	44	-5,01	9,16	2,21	3,52
Pós_geral	44	-5,34	8,85	1,25	3,31

Com os dados contidos na Tabela 32, nota-se que as séries demonstram indivíduos mais habilidosos no pós-teste geral em comparação ao pré-teste geral, exceto para a 5ª série. Porém não foi constatada a diferença de desempenho esperada entre os paradigmas, ou seja, melhor desempenho no paradigma SD, depois CC, em seguida RE e por último SC, conforme seria esperado a partir da teoria (Klauer, 1990). Para verificar se há diferença significativa entre as médias de habilidade dos sujeitos para cada série aplicou-se a análise de variância (ANOVA), conforme representado na Tabela 33.

**Tabela 33**– Diferenças de média entre as séries para as habilidades dos sujeitos geral e em cada paradigma

Variáveis	<i>gl</i>	<i>F</i>	<i>p</i>
Pré_geral	4	20,51	0,00
Pós_geral	4	14,51	0,00
SD_pré	4	6,11	0,00
CC_pré	4	10,27	0,00
RE_pré	4	7,27	0,00
SC_pré	4	9,23	0,00
SD_pós	4	12,58	0,00
CC_pós	4	1,88	0,11
RE_pós	4	7,20	0,00
SC_pós	4	8,88	0,00

A Tabela 33 mostra que a média de habilidade dos sujeitos no teste dinâmico, exceto a variável ganho que não está sendo verificada no momento, diferiu significativamente entre as séries. Para investigar entre quais séries ocorreu essa diferença, foi conduzida análise de Tukey para cada variável conforme as Tabelas 34, 35, 36, 37, 38, 39, 40, 41, 42 e 43.



**Tabela 34-** Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o escore geral do Pré-teste.

Variável Dependente	(I) Serie	(J) Serie	Diferença de média (I-J)	Erro padrão	<i>p</i>
Pré_geral	1	2	0,09	0,564	1,00
		3	-1,29	0,607	0,21
		4	-1,38	0,649	0,21
		5	-5,41*	0,679	0,00
	2	1	-0,09	0,564	1,00
		3	-1,39	0,580	0,12
		4	-1,47	0,624	0,13
		5	-5,50*	0,655	0,00
	3	1	1,29	0,607	0,21
		2	1,39	0,580	0,12
		4	-0,09	0,663	1,00
		5	-4,11*	0,692	0,00
	4	1	1,38	0,649	0,21
		2	1,47	0,624	0,13
		3	0,09	0,663	1,00
		5	-4,03*	0,730	0,00
	5	1	5,41*	0,679	0,00
		2	5,50*	0,655	0,00
		3	4,12*	0,692	0,00
		4	4,03*	0,730	0,00

\* $p < 0,05$

Por meio da Tabela 34 constata-se que as diferenças significativas de média em relação a habilidade dos sujeitos no escore geral do Pré-teste ocorreram somente entre a 5ª série com as demais séries. No escore geral do Pós-teste os dados podem ser visualizados na Tabela 35.

**Tabela 35-** Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o escore geral do Pós-teste.

Variável Dependente	(I) Serie	(J) Serie	Diferença de média (I-J)	Erro padrão	<i>p</i>
Pós_geral	1	2	-0,31	0,445	0,95
		3	-2,04*	0,479	0,00
		4	-1,24	0,512	0,11
		5	-3,51*	0,535	0,00
	2	1	0,31	0,445	0,96
		3	-1,74*	0,458	0,00
		4	-0,93	0,492	0,32
		5	-3,20*	0,517	0,00
	3	1	2,04*	0,479	0,00
		2	1,74*	0,458	0,00
		4	0,80	0,523	0,54
		5	-1,46	0,546	0,06
	4	1	1,24	0,512	0,11
		2	0,93	0,492	0,32
		3	-0,80	0,523	0,54
		5	-2,26*	0,576	0,00
	5	1	3,51*	0,535	0,00
		2	3,20*	0,517	0,00
		3	1,46	0,546	0,06
		4	2,26*	0,576	0,00

\* $p < 0,05$

A diferença da média de habilidade dos sujeitos, por séries no escore geral do Pós-teste, foi significativa entre 1ª série e 3ª, 2ª e 3ª e novamente na 5ª com as demais séries.

Em relação aos paradigmas A Tabela 36 mostra os dados em função do SD.

**Tabela 36-** Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma SD no pré-teste.

Variável Dependente	(I) Serie	(J) Serie	Diferença de média (I-J)	Erro padrão	<i>p</i>
SD_pré	1	2	-0,29	0,205	0,63
		3	-0,59	0,221	0,06
		4	-0,58	0,237	0,10
		5	-1,15*	0,247	0,00
	2	1	0,29	0,205	0,63
		3	-0,30	0,211	0,62
		4	-0,30	0,227	0,69
		5	-,87*	0,239	0,00
	3	1	0,59	0,221	0,06
		2	0,30	0,211	0,62
		4	0,00	0,242	1,00
		5	-0,57	0,252	0,16
	4	1	0,58	0,237	0,10
		2	0,30	0,227	0,69
		3	-0,00	0,242	1,00
		5	-0,57	0,266	0,20
	5	1	1,15*	0,247	0,00
		2	0,87*	0,239	0,00
		3	0,57	0,252	0,16
		4	0,57	0,266	0,20

\* $p < 0,05$

Para o paradigma SD a média de habilidade das séries se diferiu significativamente apenas com 1<sup>a</sup> e 5<sup>a</sup>, 2<sup>a</sup> e 5<sup>a</sup> séries. A diferença de habilidade nesse mesmo paradigma, mas no Pós-teste encontra-se descrita na Tabela 37.

**Tabela 37-** Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma SD no pós-teste.

Variável Dependente	(I) Serie	(J) Serie	Diferença de média (I-J)	Erro padrão	<i>p</i>
SD_pós	1	2	-0,09	0,190	0,99
		3	-0,82*	0,204	0,00
		4	-0,97*	0,218	0,00
		5	-1,18*	0,228	0,00
	2	1	0,09	0,190	0,99
		3	-0,73*	0,195	0,00
		4	-0,88*	0,210	0,00
		5	-1,10*	0,220	0,00
	3	1	0,82*	0,204	0,00
		2	0,73*	0,195	0,00
		4	-0,15	0,223	0,96
		5	-0,37	0,232	0,52
	4	1	0,97*	0,218	0,00
		2	0,88*	0,210	0,00
		3	0,15	0,223	0,96
		5	-0,21	0,245	0,91
	5	1	1,18*	0,228	0,00
		2	1,10*	0,220	0,00
		3	0,37	0,232	0,52
		4	0,21	0,245	0,91

\* $p < 0,05$

No Pós-teste, nota-se por meio da Tabela 37 que as diferenças de média significativas da habilidade no paradigma SD, entre as séries, se concentrou na 1ª e 2ª série com as demais. Tal resultado se difere do Pré-teste, demonstrando mais diferenças. A Tabela 38 apresenta os resultados da análise considerando o paradigma CC.

**Tabela 38-** Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma CC no pré-teste.

Variável Dependente	(I) Serie	(J) Serie	Diferença de média (I-J)	Erro padrão	<i>p</i>
CC_pré	1	2	0,13	0,186	0,96
		3	-0,04	0,200	1,00
		4	-0,08	0,214	1,00
		5	-1,18*	0,224	0,00
	2	1	-0,13	0,186	0,96
		3	-0,17	0,191	0,91
		4	-0,21	0,206	0,84
		5	-1,31*	0,216	0,00
	3	1	0,04	0,200	1,00
		2	0,17	0,191	0,91
		4	-0,05	0,219	1,00
		5	-1,14*	0,228	0,00
	4	1	0,08	0,214	1,00
		2	0,21	0,206	0,84
		3	0,05	0,219	1,00
		5	-1,10*	0,240	0,00
	5	1	1,18*	0,223	0,00
		2	1,31*	0,216	0,00
		3	1,14*	0,228	0,00
		4	1,10*	0,240	0,00

\* $p < 0,05$

A média de habilidade dos sujeitos, no paradigma CC, se diferiu significativamente somente entre a 5ª série com as demais. Na Tabela 39 se observa os dados com relação ao CC no Pós-teste.

**Tabela 39-** Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma CC no pós-teste.

Variável Dependente	(I) Serie	(J) Serie	Diferença de média (I-J)	Erro padrão	<i>p</i>	
CC_pós	1	2	0,12	0,200	0,97	
		3	-0,32	0,212	0,55	
		4	0,18	0,227	0,93	
		5	-0,23	0,237	0,87	
		2	1	-0,12	0,197	0,97
	2	3	-0,44	0,203	0,19	
		4	0,06	0,218	1,00	
		5	-0,35	0,229	0,55	
		3	1	0,32	0,212	0,55
		2	0,44	0,203	0,19	
	3	4	0,50	0,232	0,20	
		5	0,09	0,242	1,00	
		4	1	-0,18	0,227	0,93
		2	-0,06	0,218	1,00	
		3	-0,50	0,232	0,20	
	4	5	-0,41	0,255	0,50	
		1	0,23	0,237	0,87	
		2	0,35	0,229	0,55	
		3	-0,09	0,242	1,00	
		4	0,41	0,255	0,50	

\* $p < 0,05$

Verificando a Tabela 39 constata-se que a média de habilidade dos sujeitos no paradigma CC no Pós-teste não se diferenciou por série, divergindo do resultado encontrado no Pré-teste com esse paradigma. A Tabela 40 traz os dados da análise com o paradigma RE.

**Tabela 40-** Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma RE no pré-teste.

Variável Dependente	(I) Serie	(J) Serie	Diferença de média (I-J)	Erro padrão	<i>p</i>
RE_pré	1	2	-0,02	0,177	1,00
		3	0,06	0,191	1,00
		4	-0,23	0,204	0,78
		5	-0,97*	0,213	0,00
		2	1	0,02	0,177
	2	3	0,08	0,182	0,99
		4	-0,21	0,200	0,82
		5	-0,95*	0,206	0,00
		3	1	-0,06	0,191
	3	2	-0,08	0,182	0,99
		4	-0,30	0,208	0,62
		5	-1,04*	0,217	0,00
		4	1	0,23	0,204
	4	2	0,21	0,200	0,82
		3	0,30	0,208	0,62
		5	-0,74*	0,230	0,01
		5	1	0,97*	0,213
	5	2	0,95*	0,206	0,00
		3	1,04*	0,217	0,00
		4	0,74*	0,229	0,01

\* $p < 0,05$

Com o paradigma RE, foram identificadas diferenças significativas de médias de habilidade somente entre a 5ª série com as demais. Na Tabela 41 é possível verificar se esse padrão permanece no pós-teste.

**Tabela 41-** Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma RE no pós-teste.

Variável Dependente	(I) Serie	(J) Serie	Diferença de média (I-J)	Erro padrão	<i>p</i>	
RE_pós	1	2	-0,14	0,174	0,93	
		3	-0,09	0,187	0,99	
		4	-0,17	0,200	0,92	
		5	-1,04*	0,210	0,00	
		2	1	0,14	0,174	0,93
	2	3	0,05	0,179	1,00	
		4	-0,03	0,193	1,00	
		5	-0,90*	0,202	0,00	
		3	1	0,09	0,187	0,99
		2	-0,05	0,179	1,00	
	3	4	-0,08	0,205	1,00	
		5	-0,94*	0,214	0,00	
		4	1	0,17	0,200	0,92
		2	0,03	0,193	1,00	
		3	0,08	0,205	1,00	
	4	5	-0,87*	0,225	0,00	
		1	1,04*	0,210	0,00	
		2	0,90*	0,202	0,00	
		3	0,94*	0,213	0,00	
		4	0,87*	0,225	0,00	

\* $p < 0,05$

Observando os dados da Tabela 41 constata-se que o padrão das séries que apresentam diferença de média significativa no paradigma RE se manteve igual aos resultados da Tabela anterior. Já para o SC a análise pode ser visualizada na Tabela 42.



**Tabela 42-** Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma SC no pré-teste.

Variável Dependente	(I) Serie	(J) Serie	Diferença de média (I-J)	Erro padrão	<i>p</i>
SC_pré	1	2	0,14	0,178	0,94
		3	-0,36	0,192	0,32
		4	-0,24	0,205	0,77
		5	-1,05*	0,215	0,00
	2	1	-0,14	0,178	0,94
		3	-0,50	0,183	0,05
		4	-0,38	0,197	0,31
		5	-1,19*	0,207	0,00
	3	1	0,36	0,192	0,32
		2	0,50	0,183	0,05
		4	0,12	0,210	0,98
		5	-0,69*	0,219	0,02
	4	1	0,24	0,205	0,77
		2	0,38	0,197	0,31
		3	-0,12	0,210	0,98
		5	-0,81*	0,231	0,01
	5	1	1,05*	0,215	0,00
		2	1,19*	0,207	0,00
		3	0,69*	0,219	0,02
		4	0,82*	0,231	0,01

\* $p < 0,05$

A média de habilidade dos sujeitos no paradigma SC, considerando o Pré-teste, somente foi significativa entre a 5<sup>a</sup> série com as demais. A análise do SC no Pós-teste encontra-se descrita na Tabela 43.

**Tabela 43-** Análise de Tukey sobre as diferenças de média de habilidade para as variáveis do teste dinâmico entre cada série considerando o paradigma SC no pós-teste.

Variável Dependente	(I) Serie	(J) Serie	Diferença de média (I-J)	Erro padrão	<i>p</i>
SC_pós	1	2	-0,20	0,181	0,80
		3	-0,81*	0,195	0,00
		4	-0,28	0,208	0,65
		5	-1,06*	0,218	0,00
	2	1	0,20	0,181	0,80
		3	-0,61*	0,186	0,01
		4	-0,08	0,200	0,99
		5	-0,86*	0,210	0,00
	3	1	0,81*	0,200	0,00
		2	0,61*	0,186	0,01
		4	0,53	0,213	0,10
		5	-0,25	0,222	0,80
	4	1	0,29	0,208	0,65
		2	0,08	0,200	0,99
		3	-0,53	0,213	0,10
		5	-0,77*	0,234	0,01
	5	1	1,06*	0,218	0,00
		2	0,86*	0,210	0,00
		3	0,25	0,222	0,80
		4	0,77*	0,234	0,01

\* $p < 0,05$

Como pode-se notar as diferenças significativas da média de habilidade dos sujeitos entre as séries no paradigma SC se diferiram dos resultados da Tabela anterior. No Pós-teste as diferenças significativas ocorreram na 1ª e 2ª séries com 3ª e 5ª séries. A diferença da 5ª série com as demais permaneceu, exceto com a 3ª série.

Pode-se observar nas Tabelas 34 a 43 que a maioria das diferenças significativas entre as habilidades ocorreu entre a 5ª série em relação às demais séries, exceto com a 3ª, nas variáveis pós-teste geral, SD\_pré, SD\_Pós, SC\_pós, e com a 4ª série em SD\_pré e

SD\_pós. Para a 3ª série também houve diferença significativa em relação a 1ª e 2ª séries nas variáveis pós-teste geral, SD\_pós e SC\_pós. Por fim, a 4ª série superou a 1ª e 2ª série na variável SD\_pós. Diante disso nota-se que a 5ª e 3ª série foram as que maiores habilidades demonstraram na resolução das variáveis do teste dinâmico. Esses resultados apontam que o teste não se mostrou sensível para detectar diferenças de habilidade entre todas as séries, caracterizando pouca evidência de validade teste-critério com série. Porém esse resultado pode ter sido afetado pela pouca quantidade de itens em cada paradigma, seria interessante um número maior de itens, e isso em relação a todas as análises com paradigmas. A variável mais sensível ao efeito da série foi SD\_pós. No entanto, retomando a Tabela 32, percebe-se que no paradigma RE houve maior aumento da habilidade dos sujeitos do pré para o pós-teste, independente da série.

Até o momento foram analisados os escores derivados do pré e do pós-teste, gerais e em cada paradigma. Agora serão descritas as análises do ganho, medida esta que pode ser obtida somente com testes dinâmicos. Tal variável é a diferença entre o pré e o pós-teste, verificando se, no pós-teste, o indivíduo apresentou maior habilidade.

Observando as análises já descritas, o pré e o pós-teste separadamente, pode-se verificar alguns dados. Por exemplo, conforme a Tabela 31, de estatísticas descritivas, identifica-se que os sujeitos foram, de forma geral, mais habilidosos no pós-teste. Esse dado parece indicar que os procedimentos de intervenção colaborou para que os indivíduos tivessem maior capacidade de resolver os itens no pós-teste. Tal resultado é desejável nesse tipo de instrumento dinâmico. No entanto é necessário verificar se a diferença positiva que

ocorre no pós em relação ao pré é significativa. Para isso aplicou-se uma análise de medidas repetidas, visto que o mesmo sujeito foi avaliado em dois momentos diferentes. Considerando cada um dos quatro paradigmas, foram incluídas oito medidas. Inicialmente procedeu-se à análise para a diferença geral no teste entre pré e pós. Por meio da análise constatou-se diferença significativa ( $F= 43,60$ ,  $gl=1$  e  $p<0,001$ ) entre os escores do pré e pós-teste, indicando que houve ganho de habilidade no pós teste. Também se verifica que esse perfil não interagiu com série ( $F=2,74$ ,  $gl= 40$  e  $p= 0,74$ ) ou sexo ( $F= 0,06$ ,  $gl= 1$  e  $p=0,94$ ) apontando que os indivíduos independente da série ou sexo apresentaram o ganho no pós-teste. Esse foi um resultado positivo para o teste dinâmico sustentando o propósito de identificar potenciais de aprendizagem.

Na Tabela 44, a seguir, apresenta os resultados da mesma análise de pré e pós entre os indivíduos, mas aqui a questão não é o ganho, somente verifica se há diferença significativa entre os sujeitos em função de série e sexo. Como pode ser visualizado, há diferença significativa de média entre séries. Tais análises confirmam o efeito de série observado em análises prévias.

**Tabela 44-** Resultados da diferença entre grupos emergidos da Análise de Medidas Repetidas.

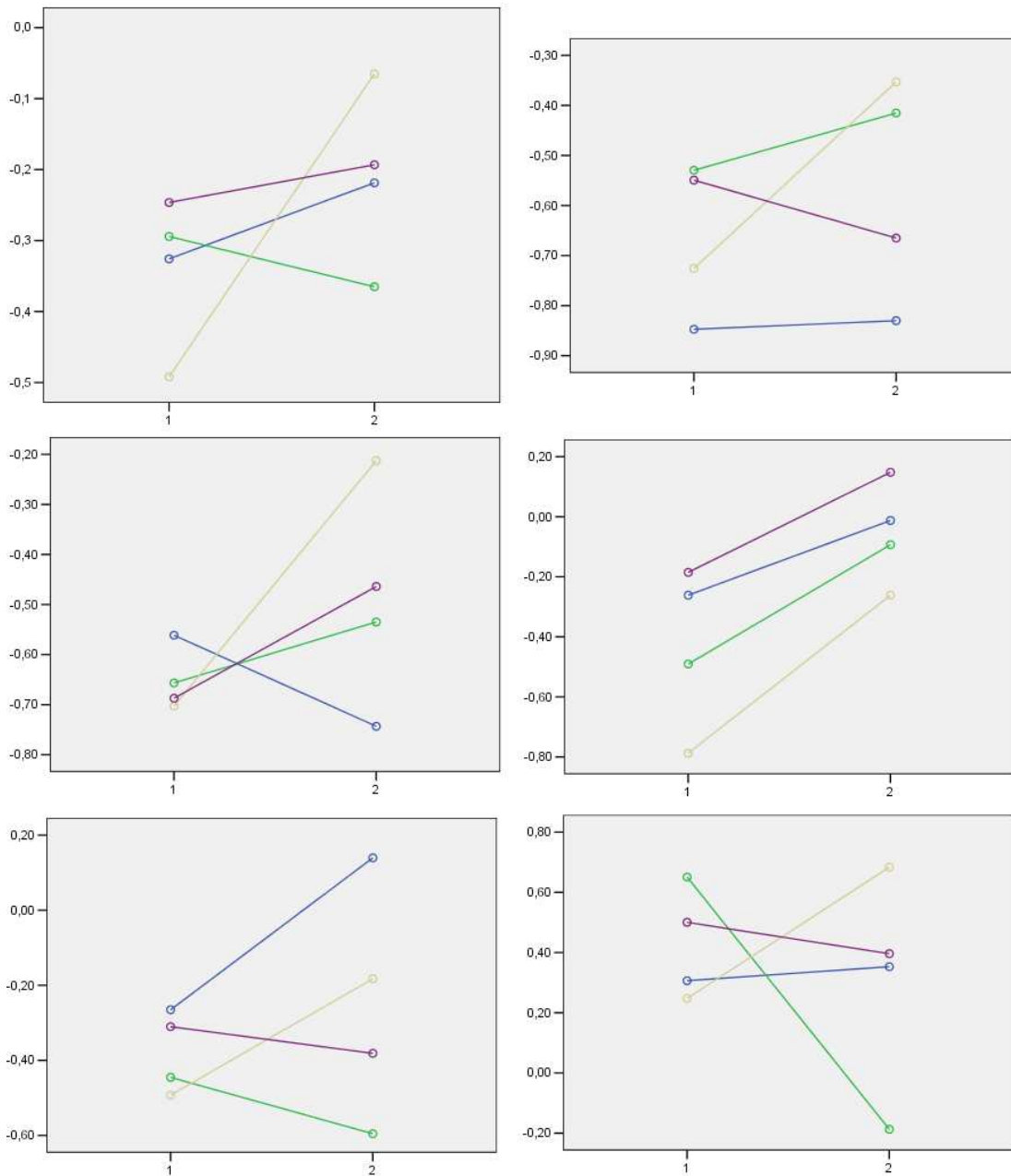
Força da variância	Soma das Quadrados	gl	MQ	F	p	Eta <sup>2</sup>
Entre-grupos	97,09	1	97,09	112,24	0,00	0,27
Serie	82,45	4	20,61	23,83	0,00	0,24
Sexo	0,18	1	0,18	0,21	0,65	0,00
Serie * Sexo	7,57	4	1,892	2,19	0,07	0,03
Erro	268,16	310	0,87			

A Tabela 44 forneceu a informação que há diferença significativa entre o pré e o pós, indicando ganho dos indivíduos. Agora a próxima análise, também utilizando o procedimento de medidas repetidas, verificou se esse perfil de ganho é alterado em relação aos paradigmas, bem como modificado considerando paradigma e série (Tabela 45).

**Tabela 45-** Resultados da Análise de Medidas Repetidas investigando o ganho de habilidade por meio da diferença significativa entre pré e pós teste considerando os paradigmas.

Efeito	Valor	<i>F</i>	<i>p</i>
prepos	0,98	7,32	0,01
prepos * Serie	0,97	2,43	0,05
paradigma	0,99	1,02	0,39
paradigma * Serie	0,90	2,74	0,00
prepos * paradigma	0,94	6,37	0,00
prepos * paradigma * Serie	0,94	1,57	0,09

A Tabela 45 descreve que o ganho difere significativamente entre os paradigmas, bem como esse perfil dos paradigmas é alterado em função da série. Os gráficos abaixo permitem a visualização desses resultados. O primeiro gráfico se refere a diferença de ganho entre os paradigmas e as outras é a interação desse ganho com cada série (Figura 22).



**Figura 22** – Diferença dos ganhos considerando tipo (paradigma) e série

*Nota: Da esquerda para a direita pode ser visualizado o ganho no geral sem separação por série e depois apresentado por cada série seqüencialmente da 1ª à 5ª. O número 1*

*designa o resultado do pré-teste e o número dois representa o resultado no pós-teste e as cores representam os paradigmas, então, SD é azul, CC verde, RE amarela e SC roxa.*

Por meio da Figura 22, que considera a amostra geral, nota-se que no paradigma RE ocorreu maior ganho da habilidade, seguido pelo SD, e depois SC. No caso do CC houve um decréscimo, ou seja, eles demonstraram maior dificuldade para resolver esse tipo de item no pós. Na Figura 22 observa-se que o perfil dos ganhos diferiu entre as séries. Houve, entretanto, um padrão com o paradigma RE, pois em todas as séries o ganho foi maior nesse tipo de item. Outro dado é que somente a 3ª série apresentou ganho em todos os paradigmas, sugerindo que essa série conseguiu ter um maior aproveitamento da etapa de intervenção como um todo.

Resumindo a presente seção, observou-se que o instrumento dinâmico parece ser adequado para investigar o potencial de aprendizagem, o que pôde ser conferido por meio do ganho significativo entre o pré e pós teste. É importante notar que a variável RE foi a que apresentou maior melhora do desempenho; assim, talvez o procedimento da intervenção tenha auxiliado especificamente a resposta a esse tipo de item. Outro dado importante é que, pelas informações obtidas até o momento, o teste dinâmico se mostra de difícil resolução. Procurando investigar melhor essas questões fez-se uma análise mais detalhada das repostas dos sujeitos aos itens do pré e pós teste, descrita no tópico a seguir.



### **2.2.4.3 - Análise das respostas aos itens no pré e pós-teste dinâmico e verificação da eficácia da intervenção em função do tipo de paradigma explorando os erros cometidos.**

A partir da descrição dos itens do teste de raciocínio indutivo, apresentada na Parte 1 do Estudo 1, foi possível analisar a escolha feita pelos participantes em cada item. A seguir, na Tabela 46, é analisada a frequência de escolha, em cada item, da alternativa correta e das alternativas erradas, bem como são apresentadas as características de cada alternativa. Dessa maneira é possível compreender melhor os erros dos sujeitos, ou seja, o que eles deixarem de identificar.

**Tabela 46-** Alternativa correta e porcentagem de escolha as alternativas de cada item.

<b>ITENS TIPO SIMILARIDADE E DISCRIMINAÇÃO-SD</b>
<b>ITEM 1</b> Resposta A: sem indicadores de semelhanças para a tarefa (12%) Resposta B: atributo incompleto (cor diferente) (5%) <b>Resposta C: correta (79%)</b> Resposta D: atributo incompleto (cor diferente) (4%)
<b>ITEM 2</b> Resposta A: correspondência de pata e orelha, mas não entre essas partes (9%) Resposta B: não há correspondência (24%) <b>Resposta C: correta (45%)</b> Resposta D: não há correspondência (22%)
<b>ITEM 3</b> <b>Resposta A: correta (18%)</b> Resposta B: atributo incompleto (não carrega maior quantidade de pessoas) (52%) Resposta C: atributo incompleto (não é terrestre) (14%) Resposta D: apesar de permitir locomoção, não é compatível com os dois atributos (16%)
<b>ITEM 4</b> <b>Resposta A: correta (19%)</b> Resposta B: atributo incompleto (não tem asas) (36%) Resposta C: atributo incompleto (tem “asas”, mas não voa) (34%) Resposta D: atributo incompleto (voa, mas não tem asas) (10%)
<b>ITEM 17</b> Resposta A: atributo incompleto (mudança da cor embaixo não é simétrico) (19%) Resposta B: atributo incompleto (tamanho não corresponde) (6%) <b>Resposta C: correta (65%)</b> Resposta D: atributo incompleto (cor da borboleta não corresponde) (9%)
<b>ITEM 18</b> Resposta A: atributo incorreto (não é fruta) (13%) Resposta B: atributo incorreto (não é fruta) (10%) Resposta C: atributo incorreto (não é fruta) (17%) <b>Resposta D: correta (60%)</b>
<b>ITEM 19</b> Resposta A: atributo incompleto (tamanho errado e objeto repetido) (11%) Resposta B: atributo incompleto (tamanho errado e cor repetida) (3%) Resposta C: atributo incompleto (tamanho errado e cor repetida) (39%) <b>Resposta D: correta (47%)</b>
<b>ITEM 20</b> Resposta A: não há atributo equivalente (7%) <b>Resposta B: correta (58%)</b> Resposta C: atributo equivocado (a resposta é com fogo e não fumaça) (27%) Resposta D: não há atributo equivalente (1%)

Tabela 46 (continuação)- Alternativa correta e porcentagem de escolha as alternativas de cada item.

ITENS TIPO CLASSIFICAÇÃO CRUZADA-CC
<b>ITEM 5</b> Resposta A: possui 1 atributo, mas com o objeto errado (26%) <b>Resposta B: correta (26%)</b> Resposta C: não possui atributos com os objetos (32%) Resposta D: não possui atributos com os objetos (17%)
<b>ITEM 6</b> Resposta A: possui 1 atributo, mas com o objeto errado (28%) <b>Resposta B: correta (23%)</b> Resposta C: possui 1 atributo, mas com o objeto errado (21%) Resposta D: não possui os atributos adequados para a resposta correta (28%)
<b>ITEM 7</b> Resposta A: possui apenas 1 atributo (21%) Resposta B: não possui atributo (20%) <b>Resposta C: correta (54%)</b> Resposta D: não possui atributo (6%)
<b>ITEM 8</b> <b>Resposta A: correta (25%)</b> Resposta B: possui apenas 1 atributo (22%) Resposta C: possui 1 atributo, mas equivocado para a resposta (28%) Resposta D: possui os dois atributos, mas com o mesmo objeto (25%)
<b>ITEM 21</b> <b>Resposta A: correta (8%)</b> Resposta B: possui apenas 1 atributo correto (61%) Resposta C: possui apenas 1 atributo correto (11%) Resposta D: possui apenas 1 atributo correto (20%)
<b>ITEM 22</b> <b>Resposta A: correta (21%)</b> Resposta B: possui apenas 1 atributo (doce) (65%) Resposta C: possui apenas 1 atributo (comprido) (6%) Resposta D: possui 1 atributo, mas irrelevante para a resposta (8%)
<b>ITEM 23</b> Resposta A: não há atributos (2%) <b>Resposta B: correta (25%)</b> Resposta C: possui 1 atributo, mas com a figura errada (58%) Resposta D: possui 1 atributo irrelevante para a resposta (16%)
<b>ITEM 24</b> Resposta A: possui 1 atributo irrelevante para a resposta (mas brilha como a estrela) (1%) Resposta B: possui 1 atributo irrelevante para a resposta ( é de comer) (40%) Resposta C: Até poderia ser, mas modificando os atributos (17%) <b>Resposta D: correta (42%)</b>

Tabela 46 (continuação)- Alternativa correta e porcentagem de escolha as alternativas de cada item.

<b>ITENS TIPO SIMILARIDADE E DISCRIMINAÇÃO ENTRE OBJETOS-RE</b>
<b>ITEM 9</b> Resposta A: Também é de proteção, mas não a proteção apresentada na analogia (40%) <b>Resposta B: correta (24%)</b> Resposta C: Figura igual ao item de relação e não há correspondência (32%) Resposta D: não há correspondência com as figuras de relação, apenas com a C (4%)
<b>ITEM 10</b> <b>Resposta A: correta (24%)</b> Resposta B: não há relação com a analogia, apenas com a figura C (43%) Resposta C: não há relações (17%) Resposta D: não há relações (16%)
<b>ITEM 11</b> Resposta A: é a etapa correta, mas a cor está errada (8%) Resposta B: não é a etapa correta, e a cor também difere (16%) <b>Resposta C: correta (31%)</b> Resposta D: etapa incorreta (45%)
<b>ITEM 12</b> Resposta A: transformação errada, objeto constante errado (13%) <b>Resposta B: correta (35%)</b> Resposta C: transformação correta, mas o objeto constante muda (29%) Resposta D: transformação correta, mas não há o objeto constante (23%)
<b>ITEM 25</b> Resposta A: duas transformações erradas, apenas as cores corretas (46%) <b>Resposta B: correta (9%)</b> Resposta C: apenas a cor de 1 objeto está errada (15%) Resposta D: apenas a cor de 1 objeto está correta (30%)
<b>ITEM 26</b> Resposta A: relação correta com a figura C, mas analogia errada com A-B (7%) Resposta B: resposta errada, possui apenas atributo com a figura C, mas não relação (19%) <b>Resposta C: correta (33%)</b> Resposta D: relação correta com a figura C, mas analogia errada com A-B (42%)
<b>ITEM 27</b> <b>Resposta A: correta (24%)</b> Resposta B: resposta errada ( transformação igual a anterior) (8%) Resposta C: resposta parcialmente correta, pois não consta a montanha (57%) Resposta D: resposta errada (pôr-do-sol) (11%)
<b>ITEM 28</b> Resposta A: seqüência errada (prevalece a anterior) (10%) Resposta B: transformação errada (20%) Resposta C: transformação errada (junção dos objetos) (24%) <b>Resposta D: correta (46%)</b>

Tabela 46 (continuação)- Alternativa correta e porcentagem de escolha as alternativas de cada item.

ITENS TIPO SISTEMA DE CONSTRUÇÃO-SC
<b>ITEM 13</b> Resposta A: cor correta, mas com borda amarela (8%) Resposta B: cor correta, mas com borda amarela (47%) <b>Resposta C: correta (24%)</b> Resposta D: cor incorreta (22%)
<b>ITEM 14</b> <b>Resposta A: correta (21%)</b> Resposta B: parcialmente correta, cor diferente (12%) Resposta C: rotação errada, posterior a correta (61%) Resposta D: resposta irrelevante (6%)
<b>ITEM 15</b> Resposta A: transformação incorreta, cor correta (12%) Resposta B: transformação correta, cor incorreta (39%) Resposta C: transformação correta, cor incorreta (36%) <b>Resposta D: correta (13%)</b>
<b>ITEM 16</b> Resposta A: transformação correta, mas cor diferente (13%) <b>Resposta B: correta (42%)</b> Resposta C: transformação incorreta, cor correta (19%) Resposta D: transformação e cor incorreta e acréscimo de 1 elemento (26%)
<b>ITEM 29</b> Resposta A: resposta errada, única cor (7%) Resposta B: resposta errada, não intercala (62%) Resposta C: está intercalado correto, mas há cor a mais (21%) <b>Resposta D: correta (10%)</b>
<b>ITEM 30</b> <b>Resposta A: correta (16%)</b> Resposta B: parcialmente correta, a cor está certa, mas não há preenchimento (64%) Resposta C: cor incorreta (17%) Resposta D: cor incorreta (3%)
<b>ITEM 31</b> Resposta A: parcialmente correta, rotação do objeto de fora errada (14%) Resposta B: parcialmente correta, rotação do objeto de dentro errada (31%) <b>Resposta C: correta (38%)</b> Resposta D: resposta errada, cores e rotação erradas (18%)
<b>ITEM 32</b> Resposta A: parcialmente correta, rotação correta, mas cor da bolinha errada (15%) Resposta B: resposta errada, rotação e cor da bolinha erradas (26%) Resposta C: resposta errada, rotação errada e sem a bolinha (23%) <b>Resposta D: correta (36%)</b>

Uma vez que já foi descrita a porcentagem de escolha para cada alternativa, a seguir encontra-se descrita a porcentagem de indivíduos que erraram os itens do pré e pós teste. Para facilitar essa análise dividiu-se o teste em quatro grupos, quais sejam, itens errados por 25% dos sujeitos, itens errados de 26% a 50% dos sujeitos, itens errados de 50% a 74% dos sujeitos e itens errados por 75% ou mais dos indivíduos. Essa informação pode ser verificada de forma geral e por série nas Tabelas 47, 48, 49, 50, 51 e 52.

**Tabela 47-** Porcentagem de erros e índices de dificuldade dos itens do pré e pós teste.

<b>PRÉ-TESTE GERAL</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
SD_1 (-2,21)	SD_2 (-0,52)	RE_11 (0,16)	SD_3 (1,05)
	CC_7 (-0,85)	RE_12 (-0,22)	SD_4 (0,98)
		SC_16 (-0,41)	CC_05 (0,98)
			CC_06 (0,71)
			CC_08 (0,62)
			RE_9 (0,86)
			RE_10 (0,45)
			SC_13 (1,01)
			SC_14 (0,87)
			SC_15 (1,44)
<b>PÓS-TESTE GERAL</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
	SD_17 (-0,97)	CC_24 (-0,16)	CC_21 (2,53)
	SD_18 (-0,72)	RE_26 (0,70)	CC_22 (1,27)
	SD_19 (-0,32)	RE_28 (0,15)	CC_23 (1,03)
	SD_20 (-0,58)	SC_31 (0,43)	RE_25 (2,36)
		SC_32 (0,39)	RE_27 (1,08)
			SC_29 (2,20)
			SC_30 (1,64)

Verificando a Tabela 47 e comparando os erros nos itens entre o pré e pós-teste, constata-se maior porcentagem de acertos no pós-teste, principalmente relacionados aos itens SD. As próximas Tabelas mostram esses dados para a 1ª série.

**Tabela 48-** Porcentagem de erros dos itens do pré e pós teste – 1ª série.

<b>PRÉ-TESTE</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
	SD_1	SD_2	SD_3
		CC_7	SD_4
		CC_8	CC_5
		RE_11	CC_6
		RE_12	RE_9
			RE_10
			SC_13
			SC_14
			SC_15
			SC_16
<b>PÓS-TESTE</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
	SD_17	SD_19	CC_21
	SD_18	SD_20	CC_22
		CC_23	RE_25
		CC_24	RE_27
		RE_26	SC_29
		RE_28	SC_30
		SC_31	SC_32

Na Tabela 48 é possível averiguar a porcentagem de erros e acertos nos itens do pré e pós dos alunos da 1ª série. Nota-se uma concentração de itens errados por no mínimo 75% da amostra, porém no pós essa quantidade de erro é menor e os itens SD são acertados por mais indivíduos. Na Tabela 49 visualiza-se a porcentagem de erros cometida pelos alunos da 2ª série.

**Tabela 49-** Porcentagem de erros dos itens do pré e pós teste – 2ª série.

<b>PRÉ-TESTE</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
SD_1		SD_2 CC_7 RE_12 SC_16	SD_3 SD_4 CC_5 CC_6 CC_8 RE_9 RE_10 RE_11 SC_13 SC_14 SC_15
<b>PÓS-TESTE</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
	SD_17 SD_18 SD_20	SD_19 CC_24 RE_26 RE_28 SC_31	CC_21 CC_22 CC_23 RE_25 RE_27 SC_29 SC_30 SC_32

Para a 2ª série, a porcentagem de sujeitos que cometeram erros também foi menor no pós-teste e acertando mais os itens SD. No entanto, tanto a 1ª quanto a 2ª série, grande porcentagem de sujeitos tendem ao erro na maioria dos itens. O dados para a terceira série podem ser conferidos na Tabela 50.



**Tabela 50-** Porcentagem de erros dos itens do pré e pós teste – 3ª série.

<b>PRÉ-TESTE</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
SD_1	CC_7 SC_16	SD_2 RE_10 SC_13	SD_3 SD_4 CC_5 CC_6 CC_8 RE_9 RE_11 RE_12 SC_14 SC_15
<b>PÓS-TESTE</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
	SD_17 SD_18 SD_19 SD_20 RE_28 SC_32	CC_22 CC_24 RE_27 SC_31	CC_21 CC_23 RE_25 RE_26 SC_29 SC_30

Na 3ª série nota-se que no pós-teste os itens SD tendem a ser acertados por uma grande porcentagem de sujeitos, bem como itens RE começam a ser mais acertados pelos sujeitos. A Tabela 51 apresenta os dados para a 4ª série.

**Tabela 51-** Porcentagem de erros dos itens do pré e pós teste – 4ª série.

<b>PRÉ-TESTE</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
SD_1	SD_2 CC_7	CC_6 RE_9 RE_10 RE_11 RE_12 SC_16	SD_3 SD_4 CC_5 CC_8 SC_13 SC_14 SC_15
<b>PÓS-TESTE</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
	SD_17 SD_18 SD_19 SD_20	CC_24 RE_27 RE_28 SC_31 SC_32	CC_21 CC_23 CC_23 RE_25 RE_26 SC_29 SC_30

Na 4ª série continua o padrão de maior quantidade de acertos nos itens SD no pós-teste. Por fim a Tabela 52 mostra os resultados dos alunos da 5ª série.

**Tabela 52-** Porcentagem de erros dos itens do pré e pós teste – 5ª série.

<b>PRÉ-TESTE</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
SD_1	CC_6	CC_5	SD_3
SD_2	CC_7	CC_8	SD_4
	SC_16	RE_9	SC_15
		RE_10	
		RE_11	
		RE_12	
		SC_13	
		SC_14	
<b>PÓS-TESTE</b>			
<b>25%</b>	<b>26% - 50%</b>	<b>51% - 74%</b>	<b>75%</b>
RE_28	SD_17	CC_23	CC_21
	SD_18	RE_26	CC_22
	SD_19	RE_27	RE_25
	SD_20	SC_31	SC_29
	CC_24	SC_30	
	SC_32		

Observando a Tabela 52 constata-se que na 5ª série os acertos aos itens começam a ser mais freqüentes, tanto no pré quanto no pós e em todos os paradigmas. Por meio das Tabelas 47 a 52, verifica-se que uma maior porcentagem de os participantes tenderam a acertar mais os itens SD, tanto no pré quanto no pós, sendo que nesse último, a porcentagem foi ainda maior. No entanto, isso não significa que o maior ganho de habilidade foi no paradigma SD, pois como já explicitado, para fazer essa comparação é preciso que os itens estejam equalizados, e os escores utilizados para tais identificações são a dificuldade do teste e a habilidade do sujeito. O ganho em RE parece ser maior se comparado ao ganho em SD, pois os itens pressupõem maior dificuldade, ao contrário dos

SD que são itens mais fáceis, exigindo menos habilidade para a resolução. Assim a quantidade de sujeitos que acertam os SD já era maior no pré teste em relação ao acerto em RE, pois tais itens tendem a ser mais pareados com a habilidade do sujeito, sendo que menores ganhos na habilidade podem resultar em maiores acertos em SD do que em RE. O principal objetivo dessa análise de erro não foi verificar o ganho, e sim entender os erros para também compreender melhor as deficiências dos indivíduos, bem como as falhas no procedimento da intervenção. Então a seguir as análises qualitativas têm essa finalidade.

Os itens com acerto menor que 25% no Pós, considerando a amostra geral, foram CC\_21, CC\_22, RE\_25 e SC\_29. Se forem retirados os indivíduos da 5ª série observa-se que os itens mais difíceis foram CC\_21, CC\_22 (exceto 3ª), CC\_23 (exceto 1ª), RE\_25, SC\_29, e SC\_30. Ressalta-se que os itens CC\_21, CC\_22, CC\_25, SC\_29 e SC\_30 são considerados muito difíceis (esse dados de dificuldade podem ser visualizados na Tabela 47). Já o CC\_22 teve dificuldade média alta. Os itens CC\_21 e RE\_25 apresentam índices de desajustados, de forma semelhante à análise das dificuldades. Outro dado interessante em relação a esses itens é que a resposta da maioria da amostra concentrou-se em apenas uma alternativa que não a correta, exceto o RE\_25 que se concentrou em 2 alternativas, também não corretas. Ao verificar o resultado geral, o item RE\_27 também foi incorporado na porcentagem de 75% de erros, mas isso ocorreu devido a uma alta concentração de erros nesse item na primeira e segunda série apenas, sendo que também nesse item a resposta se concentrou em uma alternativa que não a correta.

Os itens 21, 22 e 23 fazem parte do paradigma CC, dentre os quatro existentes no pós, que exige do sujeito cruzamento informações no momento de raciocinar em relação a atributos, verificando semelhança com a figura ao lado esquerdo e a figura da direita acima. Esse tipo de raciocínio não é comum, pois como o formato do item é quadriculado, normalmente se faz associação somente com a figura ao lado e não com a debaixo ou acima. Por meio das respostas mais emitidas se percebe justamente essa questão, elegendo a alternativa mais semelhante, que possua mais atributos com a figura ao lado.

Já o item 25 é RE, ou seja, de regularidade de relações e não de atributos. Nesse item o indivíduo precisa se atentar que a bola de dentro do quadrado aumenta e não muda de cor e o quadrado diminui e muda de cor. Porém, as alternativas mais escolhidas foram justamente as que não contemplam essas transformações dos objetos. Em uma alternativa apenas a cor do quadrado é correta e, na outra, somente a diminuição do quadrado. Essas respostas ignoram os outros elementos.

O item 29 pertence ao paradigma Sistema de Construção-SC. Nesse item o indivíduo precisa verificar o que ocorre com as figuras nas duas primeiras linhas horizontais acima e nas três verticais. Essa tarefa apresenta somente a relação de mudança de cor da parte superior e inferior dos objetos. A alternativa mais escolhida não contemplou esse aspecto, ou seja, os sujeitos novamente ignoraram todas as informações relevantes para a resposta correta. O 30 também é SC, mas nesse caso o indivíduo precisa perceber que a cor da figura permanece e o preenchimento desaparece. Porém, a não atenção sobre a

relação entre os elementos acima parece ter ocorrido novamente, pois pela alternativa escolhida pela maioria que não foi a correta, pois possui característica (atributo) com a figura ao lado, mas não possuía relação com os elementos na linha acima e com a coluna, o que seria essencial na resposta correta a itens do tipo SC.

Por fim, o item 27 também é de relações RE, mas é de continuidade sobre a seqüência de um evento, no caso, nascer do sol. Então é preciso identificar que o sol está nascendo por trás da montanha. A resposta mais emitida também fazia parte da seqüência correta, mas correspondia a uma fase posterior da resposta correta, e não contém a montanha. Esse item foi menos acertado pelos alunos da primeira e segunda série.

Ao se fazer essa mesma análise para os itens do pré teste constata-se maior número de itens com porcentagem de erro maior que 74% ao se retirar a quinta série. Permanecendo essa série, haveria apenas três itens, quais sejam SD\_3, SD\_4 e SC\_15; excluindo-a, obtém-se nove itens considerando pelo menos três séries, como na análise do pós, o que é mais compatível com o resultado geral em que se encontram 10 itens. Tais itens são SD\_3, SD\_4, CC\_5, CC\_6, CC\_8 (exceto 1ª), RE\_9 (exceto 4ª), SC\_13 (exceto 3ª), SC\_14 e SC\_15. A maioria desses itens podem ser considerados de dificuldade média alta, exceto o SC\_15 que já pode ser considerado difícil. No entanto não ocorreu o comportamento de concentração de resposta conforme observado nos itens do pós-teste, variando entre três respostas, incluindo a correta, sendo que, ao se somar as porcentagens das alternativas não-corretas escolhidas, atinge-se os 75%.

Nos itens CC\_5, CC\_6 e CC\_8 foi onde ocorreu variabilidade mais equilibrada, pois para cada alternativa a porcentagem foi próxima a 25%. Essa informação parece indicar que os sujeitos não apresentavam habilidade para responder a esses itens. Pode-se supor que os sujeitos “chutaram” as alternativas desses itens. Esse comportamento já era esperado, pois a forma de raciocinar nesses itens difere da comumente utilizada em tarefa com formato de quatro quadrículos, dois em cima e dois embaixo. Normalmente raciocina-se de forma linear, porém nos itens do tipo CC o raciocínio é cruzado, verificando atributo similar com a figura do quadrado acima e ao lado do local que a resposta deve ser completada. Uma outra possível explicação para essa variabilidade é que as alternativas dos itens apresentam bons distratores para resolução do item, dificultando a solução correta desses itens.

Os itens SD\_3 e SD\_4 compõem o paradigma similaridade e diferenciação entre atributos. No primeiro é preciso identificar que a similaridade de todos os objetos é a questão do meio de transporte, bem como do primeiro ao último a quantidade de pessoas transportada é maior, sendo que posterior as aplicações e já nessas análises, constata-se que esse item é RE, o que contribuiu para o erro, já que são itens mais difíceis que SD no geral. Referente ao segundo, a similaridade é que todos voam e possuem asas. As alternativas mais escolhidas, tanto no item 3 quanto no 4, contemplaram apenas um atributo, ou seja, os indivíduos não juntaram as duas informações de regularidades.

Já o RE\_9 faz parte do paradigma relações e no caso é uma tarefa de analogia. O indivíduo precisa identificar que o boné é uma proteção contra o sol, assim como o guarda-chuva é em relação a chuva. A primeira alternativa mais escolhida, com 40% das respostas, refere-se a que o boné protege, no caso a cabeça do ser humano; e a segunda resposta mais emitida, com 32%, refere-se a um objeto com a mesma função do boné, que é de proteção. As duas respostas estão erradas e ignoram a analogia, pois a primeira tem relação, mas equivocada diante da tarefa, enquanto a segunda se atenta somente para atributo.

Por fim os itens SC\_13, SC\_14 e SC\_15 referem-se ao paradigma Sistema de Construção-SC, que pressupõe a identificação de similaridade e diferença de relações entre linhas e colunas. As três tarefas são formadas por nove quadrados que correspondem três linhas e três colunas e o indivíduo precisa se atentar ao que acontece na linha e nas colunas. No item SC\_13 é necessário verificar que na linha o objeto muda de cor e de posição, mas pela coluna pode-se perceber que a cor da última figura é sempre igual a primeira da linha. A resposta mais emitida, e errada, foi a alternativa 2 com 47% das escolhas, em que a opção une as duas cores da linha que deve ser completada. Ou seja, os indivíduos ignoraram a relação nas linhas acima e nas colunas.

O SC\_14 possui os mesmos nove quadrados, porém o que ocorre nas linhas é a rotação do objeto, e pela coluna pode ser observado que a cor não muda. A alternativa mais escolhida e errada (61%) corresponde a uma figura com rotação posterior a resposta



correta. Nesse caso parece que os indivíduos compreenderam a proposta, mas a capacidade de rotar figuras indica que deve ser melhorada. E o SC\_15 apresenta o mesmo princípio da SC\_13 porém acrescenta-se o aumento da figura na linha. As respostas mais emitidas foram a dois (39%) e a três (36%). Na similaridade do aumento da figura estão corretas, mas as cores são erradas. Dessa forma, percebe-se que os sujeitos se atentaram apenas para uma regra o que não permitiu chegar à resposta correta.

Essa análise dos erros permitiu averiguar algumas deficiências dos sujeitos. O que mais se pode constatar é que as maiores dificuldades são de abstração e capacidade de integrar diferentes informações. No entanto nota-se que o princípio básico para resolver os itens de raciocínio indutivo, que é o de verificar similaridades e diferença em atributos, tendeu a melhorar no pós-teste. O enriquecimento desse princípio básico pôde ser verificado, por exemplo, quando nos três itens CC, do pré-teste, os indivíduos não souberam qual era a alternativa correta e “chutaram”, de forma que as respostas às alternativas ficam equilibradas entre todas. Já nos três itens CC do pós-teste, apesar dos indivíduos tenderem a errar, muito provavelmente pelo formato do item, a maioria dos indivíduos tendeu a escolher a alternativa com atributo mais parecido com a figura ao lado. Porém apenas esse princípio básico não é suficiente para a resolução dos paradigmas, principalmente CC, RE e SC. Mesmo em SD que é o princípio de similaridade e diferença, às vezes não é suficiente, pois pressupõe junção de várias informações e os indivíduos tendem a se atentar em apenas uma informação.

Por meio das porcentagens de indivíduos que acertaram os itens, observa-se que poucos itens foram acertados por 25% da amostra e a metade dos itens também não foi acertada por 50% da amostra, bem como se nota que os itens SD são os que contemplam maior porcentagem de acertos. Esses dados também corroboram com as análises por meio da Teoria de Resposta ao Item por meio do Rasch, em que se identificou que o pré e pós teste, bem como os paradigmas, exceto o SD\_pós, possuem maior nível de dificuldade que a habilidade dos sujeitos. Mesmo com os resultados mostrando que o teste dinâmico propiciou ganhos aos indivíduos, pode-se perceber que a melhora não ocorreu em todos os itens, principalmente os mais complexos envolvendo os paradigmas CC e SC. Essa falta de ganho de aprendizagem “momentânea” em alguns itens pode ter ocorrido por vários fatores, entre eles características específicas do procedimento da intervenção realizada, sendo a variável duração uma das principais questões. O procedimento de intervenção abrange o ensinamento de dos paradigmas, exceto SC, porém o primeiro a ser ensinado corresponde ao SD, e antes de iniciar as informações sobre os paradigmas, há uma explicação detalhada que enfoca mais o princípio básico de similaridade e diferença entre atributos. Como a etapa da intervenção parece ter se mostrado longo, os ensinamentos posteriores ao SD, ou até mesmo do SD, podem já não ter surtido tanto efeito em decorrência da diminuição da atenção às instruções, que pode ter sido maior no início.

A duração demasiadamente extensa da etapa da intervenção foi enfatizada por todos os juízes que o avaliaram, porém, para que as instruções permanecessem na maioria dos paradigmas, não foi possível reduzir o tempo. Frente a essa realidade se decidiu assumir

essa possível variável de interferência, o que parece ter se confirmado, não somente pelas análises ao teste, mas durante as aplicações, onde se notava o cansaço de maneira mais nítida em alguns alunos.

Uma outra possibilidade explicativa para os ganhos terem sido modestos refere-se ao formato do teste, pois foi usado um teste dinâmico auto-aplicável, por meio do computador, o que pode não ser a melhor forma de aplicação. No artigo publicado por Tzuriel e Shamir (2002) sobre a aplicação de um teste dinâmico também computadorizado, mas com a intervenção de um examinador, esses autores concluíram que essa aplicação mútua, computador e examinador, foi mais eficaz do que somente com o examinador, porém não houve um estudo usando apenas o computador. No entanto tais autores também são contra a aplicação somente computadorizada.

Sumariamente, até o momento foram enfatizadas análises que possibilitaram compreender de maneira mais detalhada o propósito do teste, que é o de verificar o potencial de aprendizagem do indivíduo e que pelas análises pode-se constatar que em geral as crianças melhoraram no pós-teste, após a etapa de intervenção. As próximas análises enfatizarão a dimensionalidade do teste, procurando identificar a estrutura do pré e pós-teste, e os estudos de evidências de validade com outros instrumentos. É preciso pontuar que a amostra do Estudo 2 apresentou habilidades baixas, logo o ganho menor pode ter sido consequência dos poucos recursos cognitivos que as crianças apresentaram. A capacidade de aprender foi verificada, porém em menor intensidade, o que pode ser reflexo de um

aparato cognitivo reduzido para o tipo de questão apresentada, o que também leva a refletir sobre uma possível maior eficácia caso o teste dinâmico seja conduzido uma pessoa (Tzuriel & Shamir, 2002), ou caso seja ajustado à real necessidade do indivíduo (Tzuriel & Shamir, 2002, Sternberg & Grigorenko, 2002, Linhares, 1996, Simões, 1995, entre outros).

#### 2.2.4.4 - Análise da evidência de validade baseada na estrutura interna

Apesar das análises fatoriais aplicadas nos itens das Formas A e B terem sugerido unidimensionalidade para essas formas, e os itens do pré e pós-teste terem serem retirados dessas formas, objetivou-se investigar a estrutura que os testes pré e pós apresentariam, visto que, dependendo da amostra, a estrutura pode variar. Dessa forma, procurando seguir novamente o que propõe a teoria de Klauer (1990), pré e pós-teste foram explorados inicialmente gerando quatro fatores para averiguar se os itens referentes a cada paradigma se agrupavam. As análises foram rodadas pelo programa Microfact, do mesmo modo que as Formas A e B do Estudo 1.

A estrutura solicitada de quatro fatores, tanto para o pré quanto para o pós-teste, não apresentou a separação dos paradigmas, bem como não foi encontrada uma lógica teórica para os quatro fatores, sendo que todos os fatores se correlacionaram entre si. Diante disso aplicou-se novamente a análise fatorial, mas sugerindo três fatores tanto para o pré quanto para o pós-teste. Os resultados dessas análises indicaram os mesmos problemas das análises com quatro fatores, com isso partiu-se para outra análise para pré e pós-teste pedindo dois fatores.

Ao se analisar os dois fatores extraídos para o pré teste e buscando um sentido teórico, não foi possível identificar alguma explicação plausível para a separação dos itens. O primeiro fator ficou composto por quatro itens, apresentando *eigenvalue* de 2,92, variância explicada de 2,22, e obtendo uma precisão de 0,49. O segundo fator englobou 7

itens, *eigenvalue* de 1,65, variância explicada de 0,91 e precisão de 0,54. A correlação entre os dois fatores atingiu 0,47. Diante dessas informações, optou-se por realizar a análise com apenas um fator, sendo que essa extração foi a mais condizente e pode ser conferida na Tabela 53.

**Tabela 53** – Fator extraído para o pré-teste por meio da análise tetracórica.

Itens	Fator1
A03SD_01	0,25
C06SD_02	0,45
B35SD_03	0,06
B01SD_04	-0,15
A08CC_05	0,38
A37CC_06	0,64
A11CC_07	0,38
B14CC_08	0,15
B16RE_09	0,39
A17RE_10	0,30
A20RE_11	0,44
A21RE_12	0,30
B30SC_13	0,22
B27SC_14	0,32
A32SC_15	0,17
A26SC_16	0,53

A estrutura com um fator para o pré teste ficou composta pela maioria dos itens, exceto B35SD\_03, B01SD\_04, B14CC\_08, B30SC\_13 e A32SC\_15, que apresentaram cargas fatoriais fracas, abaixo de 0,30. Essa estrutura com 10 itens que abrange todos os paradigmas possui um *eigenvalue* de 2,92, variância explicada de 12,55 e precisão atingindo 0,57. A maior diferença entre essa estrutura e a anterior com dois fatores é notada

pela variância explicada que aumentou substancialmente. Além disso, essa estrutura é teoricamente mais coerente.

Por outro lado, na análise do pós-teste solicitando dois fatores, observou-se uma possível explicação para a separação dos itens, identificando-se estruturas separadas em decorrência do conteúdo dos itens. O primeiro fator abrangeu itens com conteúdos estruturais ou formais, em que a criança precisa identificar similaridades entre os objetos utilizando-se de informações presentes nos itens, como cor, transformação de tamanho, rotação, entre outros. Já o segundo fator, discriminado como semântico, exigia um conhecimento sobre o objeto, ou seja, sobre o conceito subjacente à figura, além das questões estruturais. Por exemplo, a criança precisava saber que sorvete é um doce, camelo é um animal terrestre ou que fogão produz fogo. O primeiro fator foi formado por 7 itens com *eigenvalue* igual a 3,30, variância explicada de 2,68 e precisão de 0,65. Já o segundo fator apresentou 4 itens, *eigenvalue* igual a 2, variância explicada de 1,29 com precisão de 0,44. Nesse fator, três itens são do tipo CC. Como não houve correlação entre os fatores, mesmo a diferença sendo somente o conteúdo, optou-se por deixar os fatores separados.

É interessante observar que, apesar de terem emergido dois fatores diferentes em relação ao conteúdo dos itens, ambos os fatores referem-se ao raciocínio indutivo. Em outros trabalhos, como por exemplo o realizado por Wilhelm em 2000 (citado por Wilhelm, 2005), o autor buscou verificar estruturas do raciocínio, investigando a separação de raciocínio indutivo e dedutivo e, para isso, utilizou diversos testes nomeados desses

raciocínios e com diferentes conteúdos. Por meio de análises fatoriais confirmatórias, a melhor solução encontrada para a estrutura foi a que separou os testes somente por meio do conteúdo da tarefa, chegando a solução de três fatores denominados de verbal, quantitativo e figural-espacial, mas todos contendo tarefas de raciocínio indutivo e dedutivo. Ou seja, três fatores referentes aos mesmos tipos de raciocínios, com conteúdos diferentes.

Conforme a Tabela 54 a seguir, ao todo 10 itens não apresentaram cargas fatoriais acima de 0,30, sendo cinco itens em cada teste, bem como não foi encontrada uma estrutura unidimensional para o pós-teste. Como são dados iniciais e de uma amostra que apresentou bastante dificuldade em resolver os itens propostos, seria interessante aplicar em uma amostra mais heterogênea em relação a habilidade e ver se as estruturas se mantêm.



**Tabela 54**– Fatores extraídos para o pós-teste por meio da análise tetracórica.

Itens	Fator 1	Fator 2
B07SD_17	0,13	-0,01
B34SD_18	0,49	0,13
B03SD_19	0,64	0,08
C04SD_20	0,17	-0,43
A36CC_21	-0,18	-0,42
B08CC_22	-0,10	-0,44
A14CC_23	-0,22	-0,63
A13CC_24	0,47	-0,14
B20RE_25	-0,06	-0,04
A38RE_26	-0,16	-0,09
C22RE_27	0,22	-0,27
B17RE_28	0,55	-0,17
A29SC_29	0,65	0,25
A34SC_30	0,37	-0,43
A30SC_31	-0,01	-0,14
B24SC_32	0,73	-0,00

Em função dessas ressalvas, as análises descritas anteriormente utilizando TRI e todos os itens continuam válidas e importantes para a análise do teste dinâmico, em que a principal finalidade é verificar o potencial de aprendizagem. Todas essas questões psicométricas são extremamente relevantes, porém não se pode perder o sentido maior do teste dinâmico e sua singularidade que o difere dos testes tradicionais de inteligência. Essa idiosincrasia dos testes dinâmicos reflete-se nas pesquisas que os utilizam, em que não se encontram muitos dados sobre as questões psicométricas, mas sim uma ênfase na eficácia em avaliar o potencial das crianças (Grigorenko & Sternberg, 2002 e Sternberg & Grigorenko, 2000). Supondo que os dois extremos da questão são equivocados (ou seja,

sem qualquer análise psicométrica ou somente com tal análise), o ideal seria a junção do conhecimento teórico e psicométrico, como também defendido por Wilhelm (2005) e Primi (2002), este último tendo construído um teste para inteligência geral seguindo esse princípio da união psicometria e teoria.

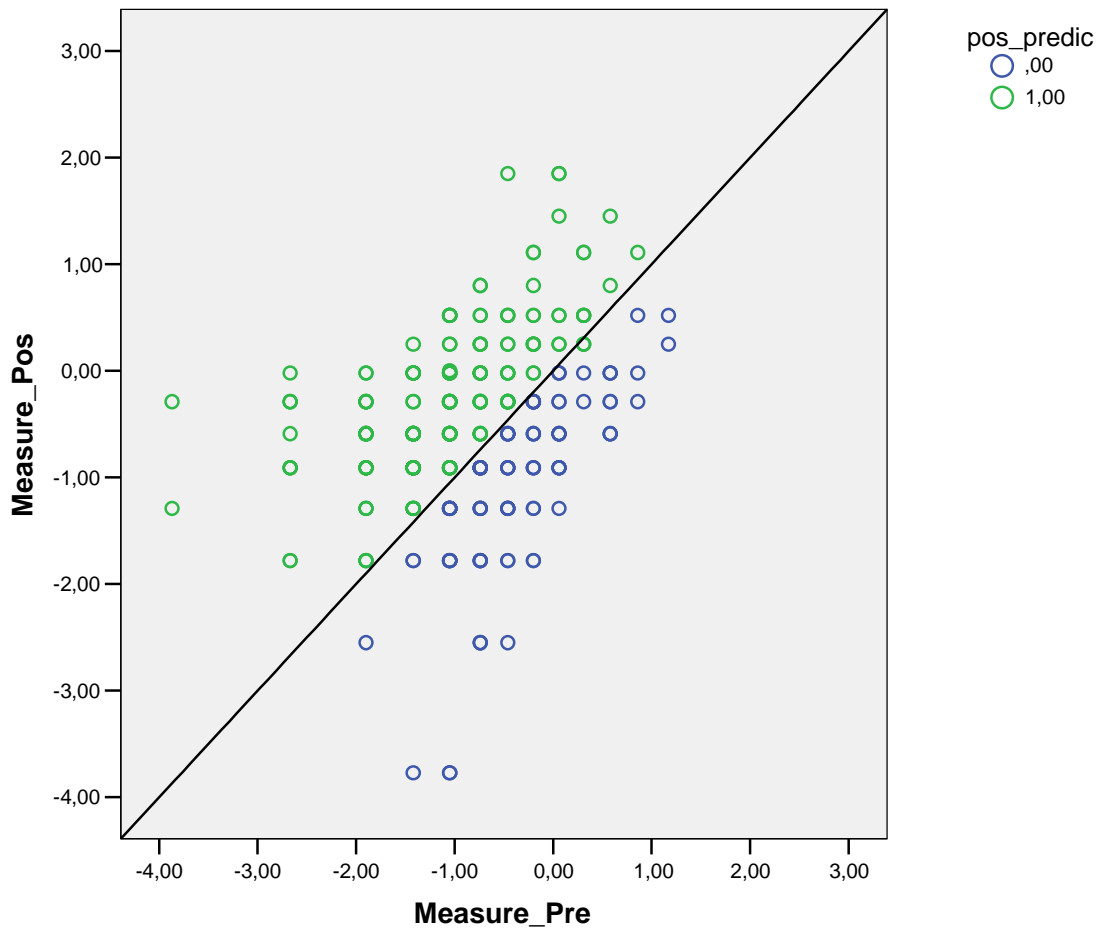
#### **2.2.4.5 - Análise das evidências de validade baseada nas relações com outras variáveis**

Os estudos de validade baseada nas relações com outras variáveis foram desenvolvidos utilizando-se, além do Teste Dinâmico de Raciocínio Indutivo, um outro teste psicológico que avalia raciocínio indutivo, dois testes neuropsicológicos, um mensurando controle inibitório e o outro atenção seletiva, bem como se fez uso do critério desempenho escolar. Os estudos envolvendo os testes psicológico e neuropsicológicos são classificados como procedimentos convergente-discriminante, e o estudo com o desempenho escolar como teste-critério. Essa diferenciação de procedimentos faz parte das evidências de validade baseadas nas relações com outras variáveis (*American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1999*). Essas investigações serão citadas na seguinte seqüência, primeiro validade convergente-discriminante do teste dinâmico com o teste Prova de Raciocínio Abstrato da BPR-5i que avalia raciocínio indutivo; segundo, com o teste neuropsicológico Geração Semântica que mensura controle inibitório; terceiro, com o teste neuropsicológico de Stroop que avalia atenção seletiva; e, por último, a validade teste-critério com o desempenho escolar. Ressalta-se que as análises correlacionais foram

realizadas com os thetas (habilidades) dos indivíduos no teste dinâmico e com os escores clássicos nos outros testes.

Para as análises correlacionais, nos estudo de validade, decidiu-se criar duas novas variáveis derivadas do escore ganho (Pós\_Pré) no teste dinâmico. A primeira variável, Pós\_Pré2, refere-se ao escore ganho considerando aumento zero para os sujeitos que não obtiveram ganho, ou seja, indivíduos com diferença zero ou negativa entre o pré e o pós. Já a segunda variável diz respeito à normalização dos ganhos da variável Pós\_Pré2 para tornar a possibilidade de ganho equiparada entre os indivíduos com baixa nota no pré com os indivíduos com notas mais altas no pré. Isso porque, sendo o pré-teste uma linha de base, os sujeitos com baixo desempenho no pré têm a possibilidade de apresentar maior quantidade de ganho, ao contrário das pessoas que já começam com a linha de base mais elevada. Por exemplo, visto que a habilidade de um sujeito pode variar de -4 a +4, se a linha de base desse sujeito for +2 no pré-teste, no pós-teste esse indivíduo terá a possibilidade de atingir habilidade +4, ou seja, apenas dois pontos a mais do que já apresentava. Mas se a linha de base for -2, essa pessoa também terá a chance de alcançar a habilidade +4, porém aqui o ganho será de 6 pontos em relação à linha de base. Então a variabilidade do ganho é maior para os indivíduos com linha de base menor.

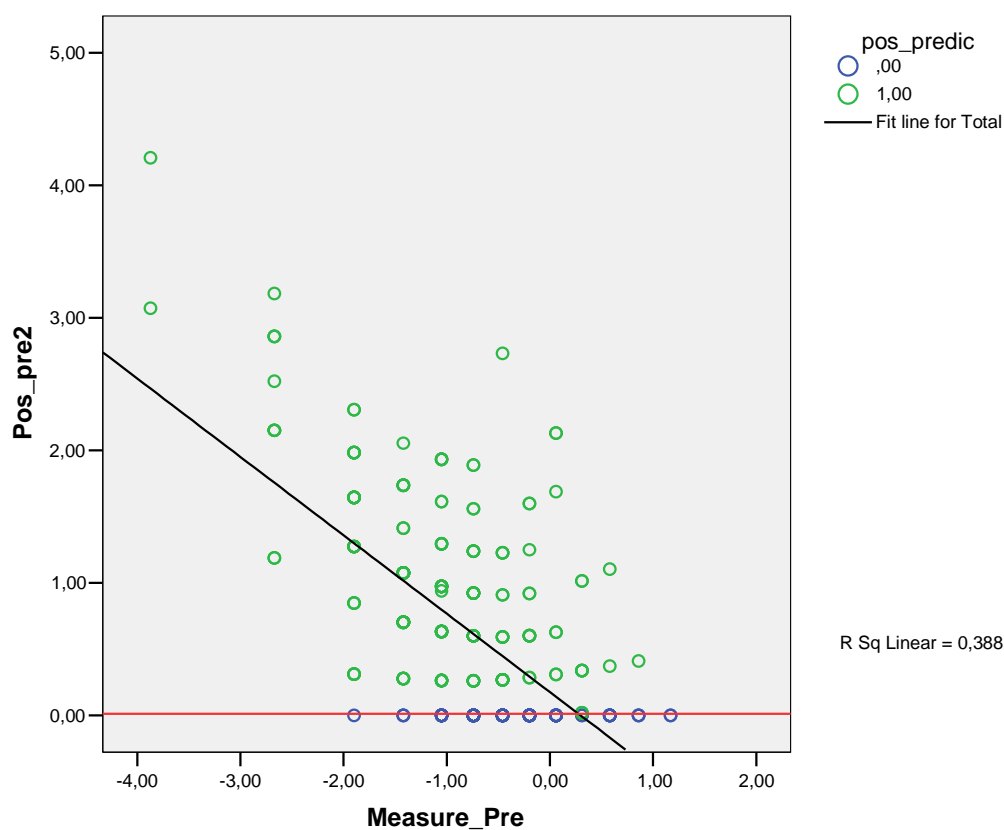
A Figura 23 mostra um gráfico representando os indivíduos que ganharam habilidade (círculos verdes) e os que não ganharam (círculos azuis). A Figura foi feita utilizando a nota da habilidade do sujeito no pré e no pós-teste.



**Figura 23** – Indivíduos que obtiveram ganhos e perdas de habilidade.

Por meio da Figura 23 observam-se vários indivíduos sem ganho de habilidade, mas como já comentado decidiu-se zerar o escore criando uma nova nota no teste dinâmico, que foi a Pós-Pré2, sendo que para fazer o gráfico da Figura 24 se fez uso dessa nova nota. Então a Figura 24 tem a finalidade de possibilitar a visualização das diversas possibilidades

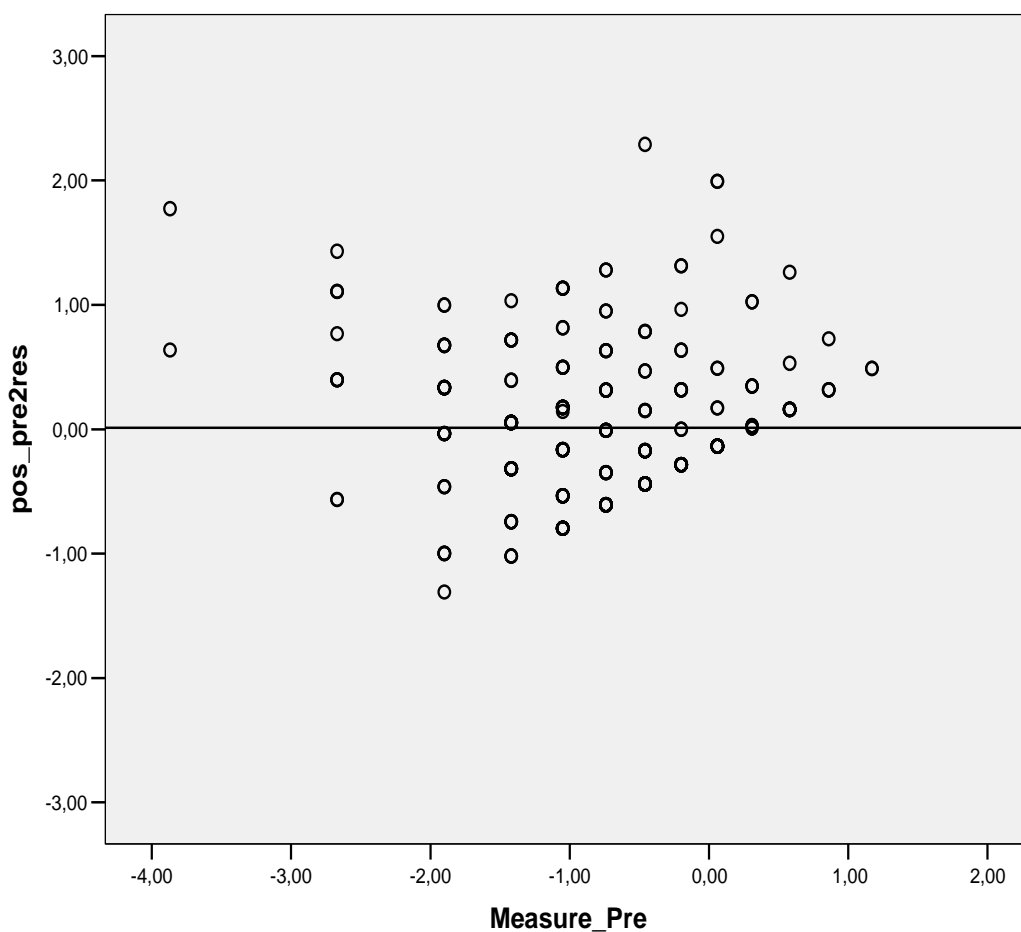
de aumento no ganho dependente da nota na linha de base que é o pré-teste. Nessa Figura a nota ganho zerada, para os indivíduos que não obtiveram ganho, pode ser notada por meio dos círculos azuis situadas no traçado vermelho.



**Figura 24-** Diferentes possibilidades de quantidade de ganho dependente da linha de base (pré-teste).

Verificando a Figura 24, constata-se justamente o que já foi discutido sobre os sujeitos com menor desempenho na linha de base, ou seja, pré-teste, terem a possibilidade de maior aumento no pós-teste e, conseqüentemente, apresentarem maior ganho do que os

indivíduos que já possuem uma linha de base com melhor desempenho. Por isso criou-se a nota Pós\_Pré2res para diminuir essa variação dependente da linha de base. O gráfico da Figura 25 mostra o resultado da normalização, deixando as variações mais equiparadas.



**Figura 25-** Normalização das variações de ganho entre sujeitos com linha de base baixa e alta.

Ressalta-se que antes de iniciar as análises correlacionais verificou-se como seriam as correlações, com a nota Pós\_Pré2res, caso tivesse mantido as notas dos indivíduos que

não apresentaram ganho. Essa exploração foi feita com a BPR-5i e constatou-se que as correlações foram semelhantes. Então decidiu-se manter a Pós\_Pré2res como se havia criado inicialmente. Antes de iniciar a descrição dos resultados correlacionais, serão apresentadas as estatísticas descritivas da amostra na Prova de Raciocínio Abstrato da BPR-5i. Na Tabela 55 estão sumariadas essas informações para a amostra geral e por série.

**Tabela 55**– Estatística descritiva da amostra para na Prova de Raciocínio Abstrato.

Variável	N	Mínimo	Máximo	Média	D.P
RA_TT	303	2	27	13,82	6,34
RA_TT(1ª)	63	2	22	8,62	4,45
RA_TT(2ª)	74	2	23	11,89	5,22
RA_TT(3ª)	59	4	25	15,25	5,73
RA_TT(4ª)	46	5	27	18,26	5,08
RA_TT(5ª)	40	4	26	17,98	5,94

*Nota: RA\_TT = score geral na prova de Raciocínio Abstrato na BPR-5i*

O máximo de pontos na prova RA é 30 e o mínimo 0. Por meio do desvio padrão nota-se que houve uma boa variabilidade nos escores obtidos pelos sujeitos. As análises correlacionais de Pearson também foram separadas por amostra total e série. A Tabela 56 apresenta os resultados obtidos na análise correlacional com a amostra geral. Para os escores do RA criaram-se três escores, o total geral, total com amostra abaixo do percentil 25 no RA e total da amostra com percentil acima de 75 no RA, esses percentis foram baseados na amostra da pesquisa. Adotou-se esse procedimento para explorar melhor se ocorreriam correlações diferenciadas em amostras com alto e baixo desempenho no RA.

**Tabela 56**– Correlações obtidas entre os escores do teste dinâmico com as notas no RA.

Variáveis	RA_total	RA_baixo	RA_alto
SD_pré	0,07	-0,01	0,18
CC_pré	0,16**	0,02	0,21
RE_pré	0,18**	0,17	0,08
SC_pré	0,11	-0,02	0,22*
SD_pós	0,12	0,13	0,28**
CC_pós	0,07	-0,05	0,00
RE_pós	0,18**	0,06	0,12
SC_pos	0,16**	0,05	0,25*
Measure_Pré	0,18**	0,06	0,25*
Measure_Pós	0,19**	0,09	0,27**
Pós_Pré	0,02	0,03	0,02
Pós_Pré2	-0,02	0,01	-0,05
Pós_Pré2res	0,12*	0,05	0,16

\*\* p<0,01, \* p<0,05

Houve sete correlações significativas encontradas entre o teste dinâmico e a prova RA, especificamente com as variáveis CC\_pré, RE\_pré, RE\_pós, SC\_pós, Measure\_Pré, Measure\_Pós e o resíduo do ganho (Pós\_Pré2res). Com o RA da amostra de baixo desempenho houve correlação significativa apenas com RE\_pré. Todos os resultados foram positivos, porém baixos. Esse foi um dado instigante, pois esperavam-se maiores correlações entre o teste dinâmico e a Prova RA, já que essa prova foi construída para avaliar o raciocínio indutivo, assim como o teste dinâmico. No entanto os resultados indicam pouca associação entre os dois testes, bem como as amostras de baixo e alto desempenho não apresentaram correlações, exceto uma com alto RA.

Com o escore ganho, a ausência de correlação pode ser explicada pela finalidade diferente dos dois testes, ou seja, apesar da capacidade mensurada ser o raciocínio indutivo, no RA o objetivo é verificar o quanto que o sujeito já possui dessa capacidade, e no teste



dinâmico é investigado o quanto ele pode vir a desenvolver. Assim, tais diferentes avaliações podem ter levado a essa ausência de correlação significativa. Essa lógica das baixas correlações encontradas decorrentes do objetivo da mensuração, segue a colocação de Sternberg e Grigorenko (2002) quando descrevem que baixas correlações entre teste dinâmico e desempenho escolar podem, também, ser em virtude do teste dinâmico não ter a finalidade de verificar o desempenho já adquirido, como o acadêmico.

Porém quando foram analisados separadamente os escores pré-teste e pós-teste em cada paradigma, pode-se considerá-los testes estáticos, de forma semelhante ao RA, ou seja, mensurando uma capacidade já existente. Porém, mesmo nesses casos não houve correlações significativas. Logo, uma possível explicação para as baixas correlações refere-se aos conteúdos dos testes. O RA é composto somente por itens geométricos, sendo todos de analogia (A/B assim como C/D), o que pressupõe um esquema de pensamento igual. Já o teste dinâmico abrange conteúdos na maioria figurais e com formatos de tarefa diferenciados, implicando esquemas de pensamentos distintos. Essas desigualdades podem ter levado à não associação, mesmo sendo o mesmo construto, da mesma forma como ocorre com o construto inteligência fluida que em análises fatoriais há uma separação por conteúdo dos itens (Carroll, 1993; Wilhelm, 2005).

Análises correlacionais entre o teste dinâmico e o RA também foram aplicadas separando a amostra por série. As correlações significativas encontradas na 1ª série foram SD\_pre com RA\_total ( $r = -0,37$   $p < 0,05$ ) e Measure\_pré com RA\_baixo ( $r = 0,25$   $p < 0,05$ ), na 2ª série SD\_pós e Measure\_pós com RA\_alto ( $r = -0,24$   $p < 0,05$ ), na 3ª série SD\_pré com

RA\_baixo ( $r = 0,26$   $p < 0,05$ ), CC\_pré com RA\_total e RA\_alto ( $r = 0,26$   $p < 0,05$  e  $r = 0,36$   $p < 0,05$  respectivamente), SD\_pós com RA\_baixo ( $r = 0,37$   $p < 0,05$ ), na 4ª série CC\_pós com RA\_baixo e RA\_total ( $r = -0,32$   $p < 0,05$  e  $r = 0,32$   $p < 0,05$  respectivamente), com a 5ª série CC\_pós com RA\_total e RA\_alto ( $r = -0,34$   $p < 0,05$  e  $r = -0,43$   $p < 0,05$  respectivamente).

Novamente foram observadas poucas correlações, duas por série, exceto a 3ª que apresentou quatro correlações, todas positivas. Entre as correlações negativas, as que mais chamam a atenção são justamente as duas correlações da 5ª série, em que até os indivíduos com alta habilidade em RA tendem a apresentar baixo desempenho no pós-teste. Essa série, mesmo em análises anteriores, demonstrou resultados pouco condizentes com o esperado, porém isso pode ser atribuído ao menos parcialmente à falta de seriedade no momento da realização do teste dinâmico, pois foi observado que a maioria dos indivíduos brincavam e não se atentavam à atividade.

A seguir, encontra-se descrito o estudo de evidência de validade baseado nas relações com outras variáveis realizado com o Teste de Stroop, que avalia a atenção seletiva. A amostra que compôs esse estudo não teve a participação dos alunos de 1ª e 2ª série, pois o Stroop necessita que o indivíduo saiba ler, e as crianças dessas séries não conseguiram fazer o Stroop adequadamente por falta da habilidade de leitura. Inicialmente serão apresentadas as estatísticas descritivas do Stroop e a comparação de média nesse instrumento dos alunos por série. Por meio da Tabela 57 pode-se visualizar as estatísticas descritivas, nesse teste, das séries juntas e separadas.

**Tabela 57**– Estatística descritiva do desempenho da amostra geral e por série no teste Stroop.

Variáveis	N	Mínimo	Máximo	Média	D. P
Escore total	172	48	72	69,63	3,27
MTR_p2	172	0,41	1,26	0,68	0,16
MTR_p3	172	0,76	2,85	1,27	0,31
MTR_Interf	172	0,09	2,04	0,59	0,26
MTRa_p2	172	0,41	1,26	0,67	0,15
MTRa_p3	171	0,78	2,96	1,28	0,34
MTRa_Interf	171	0,09	2,15	0,60	0,28
<b>3ª Série</b>					
Escore total	61	57	72	69,49	3,31
MTR_p2	61	0,46	1,20	0,72	0,16
MTR_p3	61	0,84	2,85	1,40	0,36
MTR_Interf	61	0,22	2,04	0,67	0,30
MTRa_p2	61	0,46	1,20	0,72	0,16
MTRa_p3	61	0,90	2,96	1,41	0,35
MTRa_Interf	61	0,21	2,15	0,69	0,31
<b>4ª Série</b>					
Escore total	47	48	72	69,98	3,91
MTR_p2	47	0,46	1,10	0,66	0,15
MTR_p3	47	0,78	1,64	1,14	0,23
MTR_Interf	47	0,15	0,99	0,48	0,19
MTRa_p2	47	0,46	1,10	0,66	0,15
MTRa_p3	46	0,79	1,64	1,12	0,22
MTRa_Interf	46	0,20	0,99	0,47	0,18
<b>5ª Série</b>					
Escore total	43	65	72	69,70	2,20
MTR_p2	43	0,43	1,26	0,66	0,16
MTR_p3	43	0,76	2,20	1,27	0,30
MTR_Interf	43	0,09	1,37	0,61	0,23
MTRa_p2	43	0,43	1,26	0,66	0,16
MTRa_p3	43	0,78	2,25	1,28	0,29
MTRa_Interf	43	0,09	1,28	0,62	0,22

*Nota: Escore Total (escore total no teste Stroop), MTRp2 (tempo de reação na segunda parte do Stroop em que palavra e cor são congruentes), MTRp3 (tempo de reação no Stroop na terceira parte em que palavra e cor são incongruentes), MTR\_Interf (subtração do MTRp3-MTRp2, para verificar a questão da interferência na parte em que palavra e cor são incongruentes, ou seja, leitura interferindo na codificação da cor). MTRa\_p2, MTRa\_p3 e MTRa\_Interf são os mesmos escores de tempo de reação, porém somente com acerto.*

A principal medida no teste Stroop é o MTRa\_Interf. Esse dado mostra se o tempo do indivíduo, que acertou o item, aumentou com a incongruência do item, ou seja, qual foi a diferença no tempo de reação da parte 3 (em que há interferência, pois deve-se dizer a cor das letras e não ler a palavra) subtraindo o tempo de reação da parte 2 (em que não há interferência, pois a tarefa é dizer a cor de círculos). Na parte 3 deve-se, por exemplo, dizer a cor da palavra amarelo que se encontra escrita em vermelho. É nesse momento que ocorre a avaliação da atenção seletiva, ou seja, o indivíduo precisa prestar atenção apenas na cor da palavra e não na escrita, mas quando se tem a leitura automatizada, é necessário inibir essa leitura.

As outras medidas de tempo de reação, separadamente, avaliam a velocidade ao responder o Teste de Stroop, o que também é interessante verificar, mas somente com as pessoas que acertaram o item, pois se a pessoa erra, não tem sentido verificar a velocidade de processamento. Já o escore total é a soma de quantos itens o participante acertou. As discussões pautaram-se no escore MTRa\_Interf, mas todas as medidas serão informadas, uma vez que pode ser de interesse do leitor.

Observando a Tabela 57 constata-se que, para o tempo de reação de interferência nos itens acertados (MTRa\_Interf), a 4ª série foi a série menos prejudicada pela situação de incongruência (ou seja, que apresentou menor aumento no tempo de reação da parte 2 para a parte 3), seguida da 5ª e por último a 3ª série. Para verificar se essa diferença de média é significativa fez-se a análise multivariada ANOVA. A Tabela 58 apresenta os resultados

obtidos, já com a análise de Tukey para identificar entre quais séries há diferença significativa, caso ocorra.

**Tabela 58**– Diferenças de média entre as séries no teste Stroop verificadas pela ANOVA.

Variáveis	(I) Série	(J) Série	Diferença de média (I-J)	Erro padrão	<i>p</i>
Escore total	3	4	-0,49	0,631	0,72
		5	-0,21	0,647	0,95
	4	3	0,49	0,631	0,72
		5	0,28	0,686	0,91
	5	3	0,21	0,647	0,95
		4	-0,28	0,686	0,91
MTR_p2	3	4	0,06	0,030	0,10
		5	0,07	0,031	0,08
	4	3	-0,02	0,030	0,10
		5	0,01	0,033	0,98
	5	3	-0,07	0,031	0,08
		4	-0,01	0,033	0,98
MTR_p3	3	4	0,26*	0,059	0,00
		5	0,13	0,061	0,08
	4	3	-0,26*	0,059	0,00
		5	-0,13	0,064	0,12
	5	3	-0,13	0,061	0,08
		4	0,13	0,064	0,12
MTR_Interf	3	4	0,20*	0,049	0,00
		5	0,06	0,050	0,41
	4	3	-0,20*	0,049	0,00
		5	-0,13*	0,053	0,04
	5	3	-0,06	0,050	0,41
		4	0,13*	0,053	0,04
MTRa_p2	3	4	0,06	0,030	0,10
		5	0,06	0,031	0,10
	4	3	-0,06	0,030	0,10
		5	0,00	0,033	1,00
	5	3	-0,06	0,031	0,10
		4	-0,00	0,032	1,00
MTRa_p3	3	4	0,30*	0,057	0,00
		5	0,13	0,060	0,07
	4	3	-0,30*	0,059	0,00
		5	-0,16*	0,064	0,03
	5	3	-0,13	0,060	0,07
		4	0,16*	0,064	0,03

Tabela 58 (continuação)- Diferenças de média entre as séries no teste Stroop verificadas pela ANOVA.

Variáveis	(I) Série	(J) Série	Diferença de média (I-J)	Erro padrão	<i>p</i>
MTRa_Interf	3	4	0,23*	0,049	0,00
		5	0,07	0,050	0,34
	4	3	-0,23*	0,049	0,00
		5	-0,16*	0,053	0,01
	5	3	-0,07	0,050	0,34
		4	0,16*	0,053	0,01

A Tabela 58 mostra que a diferença de média entre a 4<sup>a</sup> e a 3<sup>a</sup> e a 4<sup>a</sup> e a 5<sup>a</sup> série realmente são significativas, mas isso não ocorre entre a 3<sup>a</sup> e a 5<sup>a</sup> série. Estudo prévio havia revelado diferenças significativas no tempo de reação de interferência entre crianças de 5<sup>a</sup> série e de 7<sup>a</sup> e 8<sup>a</sup> séries (Menezes, 2008). Com relação às associações entre o Teste de Stroop e os escores do Teste Dinâmico, a Tabela 59 apresenta as correlações encontradas sem separação por série.

**Tabela 59-** Correlação entre os escores do Teste Dinâmico e o Teste de Stroop sem separação por série

Variáveis	Total	MTR_p2	MTR_p3	MTR_Interf	MTRa_p2	MTRa_p3	MTRa_Interf
Measure_Pré	0,04	-0,20*	-0,19*	-0,10	-0,19*	018*	-0,10
Measure_Pós	0,13	-0,23**	-0,23**	-0,14	-0,21**	-0,23**	-0,15
Pós_Pré	0,09	-0,01	-0,02	-0,02	-0,01	-0,04	-0,04
Pós_Pré2	0,09	-0,01	-0,01	-0,02	-0,02	-0,03	-0,02
Pós_Pré2res	0,13	-0,18*	-0,17*	-0,10	-0,18*	-0,17*	-0,11

\* $p < 0,05$ , \*\* $p < 0,01$

Verificando a Tabela 59, constata-se que não ocorreram correlações significativas entre os escores do Teste Dinâmico e a principal variável do Teste de Stroop que é a

MTRa\_Interf, sugerindo que não há relação entre a atenção seletiva, avaliada por esse teste, e o desempenho nos escores do teste dinâmico. O resultado esperado seria a correlação significativa, pois pressupõe-se que, durante a resolução dos itens de raciocínio indutivo, bem como no decorrer da etapa de intervenção para a solução desses itens, o indivíduo necessite utilizar o processo da atenção seletiva, o que permite que se atente para as informações relevantes na tarefa e no treino (Gazzaniga, Ivry & Mangun, 2002), conseqüentemente possibilitando um melhor desempenho.

No entanto, sabe-se que o processo da atenção seletiva já está presente no indivíduo desde os 2 ou 3 anos de idade, mas de maneira pouco desenvolvida, sendo que a maturação ocorre ao longo dos anos e tende a se estabilizar até a vida adulta. Por exemplo, conforme citado por Menezes (2008), Adleman e cols. (2002) e Leung, Skudlarski, Gatenby, Peterson e Gore (2000) revelaram melhora do desempenho no Teste de Stroop no início da fase adulta, período em que foi possível verificar uma ativação cerebral continuada da região do CPF responsável pela atenção seletiva. Como a presente amostra foi composta por sujeitos entre 9 e 11 anos de idade, e nesse período da vida o processo da atenção seletiva já se encontra mais bem desenvolvido, porém, pode ser que não muito maduro a ponto de contribuir significativamente para o desempenho em testes como esse construído no presente estudo. A análise com os paradigmas sustenta o mesmo resultado, ou seja, indicando o processo da atenção seletiva como não correlacionado ao desempenho nos itens referentes aos paradigmas. Essa informação pode ser identificada na Tabela 60.



**Tabela 60**– Correlação entre os escores dos paradigmas com o teste Stroop.

Variáveis	Total	MTR_p2	MTR_p3	M_Interf	MTRa_p2	MTRa_p3	MTRa_Interf
SD_pré	-0,09	0,00	-0,13	-0,16	-0,00	-0,08	-0,09
CC_pré	-0,04	-0,16*	-0,07	0,00	-0,15	-0,09	-0,00
RE_pré	0,03	-0,20*	-0,17*	-0,08	-0,20*	-0,16*	-0,07
SC_pré	0,11	-0,10	-0,08	-0,03	-0,08	-0,09	-0,05
SD_pós	0,10	-0,20*	-0,21**	-0,14	-0,20*	-0,22**	-0,14
CC_pós	0,01	-0,14	-0,16*	-0,12	-0,12	-0,13	-0,08
RE_pós	0,15	-0,12	-0,12	-0,07	-0,10	-0,16	-0,13
SC_pós	0,07	-0,13	-0,05	0,02	-0,12	-0,04	0,03

Considerando, também, os fatores idade e maturação do processo da atenção seletiva, efetuou-se a análise correlacional entre o Teste de Stroop e os escores do Teste Dinâmico e dos paradigmas separando por série. Na Tabela 61 se averigua tais dados.

**Tabela 61**– Correlações por série entre os escores do teste dinâmico, incluindo os paradigmas, com o teste Stroop.

Variáveis	Total	MTR_p2	MTR_p3	M_Interf	MTRa_p2	MTRa_p3	MTRa_Interf
<b>3ª série</b>							
Measure_Pré	0,14	-0,18	-0,22	-0,16	-0,17	-0,22	-0,16
Measure_Pós	0,17	-0,18	-0,24	-0,18	-0,16	-0,24	-0,19
Pós_Pré	-0,00	0,03	0,02	0,01	0,03	0,00	-0,01
Pós_Pré2	-0,05	-0,01	-0,01	-0,01	-0,02	-0,04	-0,03
Pós_Pré2res	0,05	-0,17	-0,20	-0,14	-0,17	-0,22	-0,16
<b>4ª série</b>							
Measure_Pré	-0,14	-0,11	-0,23	-0,19	-0,09	-0,26	-0,20
Measure_Pós	0,02	-0,23	-0,12	0,04	-0,23	-0,13	0,06
Pós_Pré	0,16	-0,11	0,12	0,23	-0,12	0,15	0,26
Pós_Pré2	0,24	-0,14	0,08	0,21	-0,15	0,13	0,25
Pós_Pré2res	0,14	-0,24	-0,09	0,08	-0,24	-0,06	0,12
<b>5ª série</b>							
Measure_Pré	0,22	-0,21	-0,07	0,05	-0,21	-0,08	0,05
Measure_Pós	0,35*	-0,28	-0,46**	-0,40**	-0,26	-0,52**	-0,50**
Pós_Pré	0,21	-0,14	-0,43**	-0,45**	-0,13	-0,48**	-0,55**
Pós_Pré2	0,21	-0,05	-0,34*	-0,41**	-0,05	-0,37*	-0,46**
Pós_Pré2res	0,33*	-0,16	-0,38*	-0,38*	-0,16	-0,41**	-0,43**
<b>3ª série</b>							
SD_pré	-0,09	0,12	-0,09	-0,17	0,12	-0,06	-0,13
CC_pré	-0,02	-0,18	-0,08	-0,00	-0,18	-0,07	0,01
RE_pré	0,07	-0,17	-,17	-0,11	-0,18	-0,17	-0,11
SC_pré	0,25	-0,09	-0,12	-0,09	-0,08	-0,15	-0,13
SD_pós	0,18	-0,30*	-0,21	-0,08	-0,30*	-0,24	-0,12
CC_pós	0,02	-0,03	-0,19	-0,20	-0,02	-0,13	-0,13
RE_pós	0,10	-0,02	-0,05	-0,04	-0,00	-0,07	-0,08
SC_pós	0,06	-0,05	-0,07	-0,06	-0,03	-0,08	-0,07
<b>4ª série</b>							
SD_pré	-0,06	-0,07	-0,15	-0,14	-0,08	-0,15	-0,12
CC_pré	-0,27	-0,14	-0,08	0,01	-0,11	-0,17	-0,04
RE_pré	-0,01	-0,06	-0,10	-0,08	-0,06	-0,11	-0,07
SC_pré	-0,08	-0,03	-0,27	-0,30*	-0,00	-0,28	-0,30*
SD_pós	-0,01	-0,09	-0,05	0,02	-0,08	0,00	0,08
CC_pós	-0,02	-0,25	-0,19	-0,04	-0,23	-0,23	-0,06
RE_pós	0,09	-0,16	0,00	0,13	-0,17	-0,05	0,09
SC_pós	-0,01	-0,14	-0,05	0,04	-0,16	-0,06	0,10

Tabela 61 (continuação)- Correlações por série entre os escores do teste dinâmico, incluindo os paradigmas, com o teste Stroop.

Variáveis	Total	MTR_p2	MTR_p3	M_Interf	MTRa_p2	MTRa_p3	MTRa_Interf
<b>5ªsérie</b>							
SD_pré	-0,20	-0,00	-0,16	-0,21	-0,01	-0,03	-0,03
CC_pré	0,29	-0,04	-0,04	-0,02	-0,04	-0,11	-0,10
RE_pré	0,03	-0,27	-0,12	0,04	-0,26	-0,10	0,05
SC_pré	0,24	-0,10	0,13	0,23	-0,11	0,08	0,19
SD_pós	0,13	-0,11	-0,33*	-0,36*	-0,10	-0,34*	-0,37*
CC_pós	0,09	-0,27	-0,36*	-0,28	-0,27	-0,37*	-0,30
RE_pós	0,42**	-0,08	-0,30	-0,33*	-0,05	-0,40**	-0,49**
SC_pós	0,31*	-0,28	-0,21	-0,08	-0,27	-0,24	-0,13

\*p<0,05, \*\*p<0,01

Nota-se, pela Tabela 61, que somente a 5ª série mostrou correlações significativas entre a variável de tempo de reação de interferência em acertos (MTRa\_Interf) e os escores do teste dinâmico referentes ao desempenho no pós-teste, medidas do ganho e os paradigmas SD\_pós e RE\_pós, com correlações negativas e moderadas, exceto o paradigma SD em que a correlação foi baixa. Somente uma correlação foi verificada na 4ª série, entre a interferência no acerto (MTRa\_Interf) e o paradigma SC\_pré. Tais resultados sugerem que o processo da atenção seletiva parece apresentar maior relevância para o desempenho nos escores do teste dinâmico, bem como de outros semelhantes, em crianças mais velhas.

Esse dado parece corroborar a hipótese de que a atenção seletiva está correlacionada com o potencial de aprendizagem somente em crianças mais velhas. Diferentes razões podem ser elencadas como explicativas desse fenômeno, incluindo desde uma possível dificuldade dos testes aplicados para a população mais jovem (o que poderia gerar efeito de piso e, conseqüentemente, ausência de correlação significativa) até a ausência de correlação

propriamente dita entre os construtos avaliados em crianças mais jovens (ou seja, talvez devido ao desenvolvimento ainda incipiente da atenção seletiva nessas crianças jovens, sua participação e auxílio na solução de problemas indutivos seria reduzida).

Pensando nesses processos cognitivos que são subjacentes no momento da resolução de tarefas e da etapa da intervenção contidas no teste dinâmico, ou seja, atenção seletiva e controle inibitório, outro teste foi aplicado, mas mais voltado à avaliação do controle inibitório, ou seja, a capacidade de inibir informações irrelevantes. Ressalta-se que controle inibitório e atenção seletiva, apesar de didaticamente separados, parecem ocorrer concomitantemente, sendo que mais recentemente têm sido tratados em alguns estudos como uma função única (Miyake, Friedman, Emerson, Witzki & Howerter, 2000). No entanto, como foi utilizado um teste nomeadamente construído para mensurar o controle inibitório, o Geração Semântica, e por ser recente a discussão da junção dessas duas funções, optou-se por tratar de maneira separada o controle inibitório, porém considerando a nova discussão. Assim como realizado com os outros testes, inicialmente será apresentada a estatística descritiva do teste Geração Semântica. A Tabela 62 mostra essas informações referentes ao grupo total e separado por série.

**Tabela 62-** Estatística descritiva do tempo de reação no Teste de Geração Semântica da amostra geral e separada por série.

Variáveis	N	Mínimo	Máximo	Média	D.P
<b>Todas as séries</b>					
Md_Alta	172	1,52	11,90	3,66	1,74
Md_Bx	172	1,23	7,00	2,55	0,97
Interf	172	-0,65	8,07	1,11	1,13
MdBxTRa	172	1,19	10,33	2,49	1,07
MdAlTRa	172	1,51	9,24	3,43	1,55
InterfTRa1	172	-5,55	3,44	-0,94	1,08
<b>3ª Série</b>					
Md_Alta	61	1,69	8,73	3,63	1,44
Md_Bx	61	1,34	5,62	2,60	0,87
Interf	61	-0,65	3,91	1,03	0,84
MdBxTRa	61	1,32	5,61	2,52	0,87
MdAlTRa	61	1,65	8,02	3,37	1,33
InterfTRa1	61	-3,77	1,19	0,84	0,85
<b>4ª Série</b>					
Md_Alta	49	1,52	11,90	3,60	2,20
Md_Bx	49	1,32	5,41	2,43	0,96
Interf	49	-0,27	8,07	1,17	1,53
MdBxTRa	49	1,35	10,33	2,44	1,39
MdAlTRa	49	1,51	9,24	3,30	1,80
InterfTRa1	49	-5,55	3,44	-0,86	1,33
<b>5ª Série</b>					
Md_Alta	41	1,94	8,56	3,90	1,61
Md_Bx	41	1,31	5,53	2,68	0,95
Interf	41	-0,45	4,73	1,22	1,12
MdBxTRa	41	1,38	5,45	2,60	0,93
MdAlTRa	41	1,80	8,64	3,77	1,54
InterfTRa1	41	-4,87	0,56	-1,18	1,16

*Nota: Md\_Alta (média do tempo de reação nos itens com alta seleção de resposta), Md\_Bx (média do tempo de reação nos itens com baixa seleção de resposta), Interf (subtração do Md\_Alta-Md\_Bx, para verificar a questão da interferência da quantidade de possibilidade de respostas no tempo de reação), MdBxTRa, MdAlTRa e InterfTRa são os mesmos escores de tempo de reação, porém considerando somente acertos.*

Assim como no Teste de Stroop, a principal medida a ser verificada no Teste de Geração Semântica é o InterfTRa, pois é essa nota que mostra se o tempo de reação para responder é maior nos itens de alta seleção de resposta do que nos de baixa, como é pressuposto pela teoria. Ou seja, tende a haver maior demanda de controle inibitório quando o sujeito se depara com vários estímulos para serem inibidos; logo, nos itens com alta seleção de respostas, que há maior número de informações, o sujeito levaria um tempo mais elevado para conseguir emitir uma resposta. O IntefTRa é o índice de interferência que considera somente as respostas corretas, pois é incoerente computar respostas erradas, uma vez que o ideal é a junção da velocidade do processamento com o acerto. Os outros índices separadamente verificam velocidade de processamento da resposta, exceto o Interf, que é o mesmo que o IntefTRa, mas computa acerto e erro juntos.

Por meio da Tabela 62 nota-se que o tempo de reação de interferência para acertos (InterfTRa) apresentou uma média maior nos indivíduos da 5ª série se comparada com as médias da 3ª e 4ª série, sugerindo que as crianças, na 5ª série, começam a ser influenciadas pela demanda de controle inibitório, apresentando maior tempo de reação em itens de alta seleção quando comparados aos de baixa seleção. Porém, para verificar se essa diferença é significativa foi aplicada a análise multivariada ANOVA. Na Tabela 63 estão sumariados os resultados encontrados.

**Tabela 63**– Diferença de média no teste Geração Semântica verificada pela ANOVA.

Variáveis	<i>gl</i>	<i>F</i>	<i>p</i>
Md_Alta	2	0,38	0,69
Md_Bx	2	0,88	0,42
Interf	2	0,38	0,69
MdBxTRa	2	0,23	0,80
MdAlTRa	2	1,21	0,30
InterfTRa	2	1,30	0,28

Na Tabela 63 constata-se que não houve diferença significativa entre as séries no escore InterfTRa do Geração Semântica, bem como também não houve tais diferenças nos outros escores. Tais resultados corroboram parcialmente com a pesquisa realizada por Tortella, Dias Capovilla, Trevisan, Hipólito e Berberian (2008) com amostra de 1ª à 4ª série do ensino fundamental público. Um dos objetivos da pesquisa foi verificar a validade desenvolvimental do Teste Geração Semântica a partir da escolaridade. Os resultados apontaram diferença de média significativa no escore InterfTRa somente a partir da 4ª série, tendo essa série melhor desempenho do que a 1ª, 2ª e 3ª série.

Para investigar possíveis associações entre o Teste de Geração Semântica e os escores do teste dinâmico fez-se uma análise correlacional de Pearson. Os resultados encontrados podem ser visualizados na Tabela 64. Ressalta-se que essa análise foi considerado o grupo geral de 3ª, 4ª e 5ª série.

**Tabela 64-** Análise correlacional entre o Teste de Geração Semântica e os escores no teste dinâmico.

Variáveis	Measure_Pré	Measure_Pós	Pós_Pré	Pos_pre2	Pós_Pré2res
Md_Alta	0,09	0,02	-0,07	-0,09	-0,03
Md_Bx	-0,01	-0,05	-0,04	-0,05	-0,06
Interf	0,14	0,07	-0,08	-0,10	0,01
MdBxTRa	-0,02	-0,11	-0,09	-0,09	-0,12
MdAlTRa	0,11	0,03	-0,09	-0,11	-0,03
InterfTRa1	-0,17*	-0,14	0,05	0,06	-0,07

\* $p < 0,05$

Como pode ser observado na Tabela 64, houve somente uma correlação significativa, entre InterfTRa e Measure\_Pré ( $r = -0,17$   $p < 0,05$ ), porém a magnitude é muito baixa. Tal resultado sugere que indivíduos com maior controle inibitório tende a apresentar melhor desempenho no pré-teste do teste dinâmico. Esse dado é coerente, uma vez que se pressupõe a necessidade de inibir informações irrelevantes durante a resolução das tarefas no pré-teste para que o indivíduo consiga se atentar aos estímulos corretos que o conduzirão a resposta correta. No entanto, a correlação encontrada não permite considerações consistentes, especialmente porque se esperava mais correlações com os outros escores do teste dinâmico. Dessa forma também foi explorada a correlação do teste Geração Semântica com os paradigmas contidos no pré e pós-teste. A Tabela 65 apresenta os dados emergidos.



**Tabela 65**– Análise correlacional entre o teste Geração Semântica e os escores dos paradigmas.

	Md_Alta	Md_Bx	Interf	MdBxTRa	MdAlTRa	InterfTRa1
SD_pré	0,13	-0,00	0,20*	-0,10	0,16	-0,32**
CC_pré	0,02	0,01	0,03	0,05	0,04	-0,01
RE_pré	0,04	0,01	0,05	0,02	0,05	-0,05
SC_pré	-0,01	-0,08	0,05	-0,08	0,01	-0,10
SD_pós	0,03	-0,07	0,10	-0,11	0,02	-0,14
CC_pós	0,01	-0,03	0,03	-0,07	0,01	-0,09
RE_pós	-0,00	0,00	-0,01	-0,04	0,02	-0,07
SC_pós	0,02	-0,02	0,05	-0,02	0,01	-0,04

\* $p < 0,05$ , \*\* $p < 0,01$

A Tabela 65 informa que somente o paradigma SD no pré-teste mostrou correlação significativa com o Geração Semântica, das duas correlações que ocorreram, a relevante é a InterTRa ( $r = -0,32$   $p < 0,01$ ), que é considerada uma baixa correlação. O resultado sugere que as pessoas com maior controle inibitório tende a apresentar melhor desempenho no paradigma SD\_pré. No entanto esse é o paradigma com menor número de informações e menos complexo. Esperavam-se mais correlações com outros paradigmas.

Retomando os resultados obtidos no Teste de Stroop, em que somente foram verificadas correlações a partir da 5ª série, ou seja, sujeitos com idade maior na amostra, bem como considerando que o controle inibitório pode ocorrer concomitantemente com a atenção seletiva, foram conduzidas análises correlacionais separando por série. No entanto não ocorreram correlações significativas entre o Geração Semântica e os escores do teste dinâmico, nem mesmo com os paradigmas.

Essa ausência de correlação significativa entre os desempenhos no Teste de Geração Semântica e o teste dinâmico pode ser devida a diferentes fatores. É possível que os construtos avaliados de fato não se correlacionem, o que não é esperado a partir da literatura que sugere relação entre funções executivas e inteligência fluida, uma vez que raciocínio indutivo é um fator robusto dessa inteligência (Blair, 2006; Salthouse, 2005 entre outros). Outra possibilidade é que essa correlação somente se verifique em crianças mais velhas ou adultos ou, ainda, que os instrumentos usados, especialmente o Teste de Geração Semântica, não sejam válidos para avaliar os construtos nessa faixa etária.

Os três estudos de evidências de validade conduzidos até o momento são classificados como convergente-discriminante. O quarto estudo, descrito a seguir, refere-se à evidência de validade baseada nas relações com outras variáveis, mas pelo procedimento teste-critério. O desempenho escolar foi o critério e foi obtido com base na média final por matéria. Como há matérias diferentes entre a primeira etapa do ensino fundamental (1ª à 4ª série) e a segunda etapa (5ª série), as análises são apresentadas separadamente. Além disso, na 1ª à 4ª série foi considerada uma outra avaliação de desempenho escolar, além da nota nas matérias. Essa outra avaliação foi constituída da observação da professora em sala de aula, ou seja, quem ela considerava com maior e menor facilidade para aprender os conteúdos ensinados. Essa classificação foi obtida perguntado-se à professora quem ela considerava ser os cinco alunos com maior e os cinco com menor facilidade de aprender. Assim, as análises com essa nota foram baseadas numa amostra reduzida, de 10 alunos por sala. Primeiramente são apresentadas as análises com as séries de 1ª à 4ª. As estatísticas

descritivas do desempenho acadêmico serão inicialmente colocados para que se possa compreender a amostra nesse critério (Tabela 66).

**Tabela 66**– Estatística descritiva da amostra total e por série em relação as matérias.

Matérias	N	Mínimo	Máximo	Média	D.P
<b>Amostra total</b>	287	3	10	6,63	1,72
Português					
Matemática	287	2	10	6,60	1,86
Ciências	287	1	10	6,72	1,90
História	287	1	10	6,72	1,85
Geografia	286	1	10	6,73	1,86
Ed.Artística	287	4	10	7,52	1,31
Ed.Física	287	5	10	8,43	1,22
Informática	287	7	10	8,29	0,70
Inglês	126	5	9	5,94	1,35
<b>1ª série</b>					
Português	70	3	10	7,13	1,91
Matemática	70	3	10	7,24	2,08
Ciências	70	3	10	7,49	2,09
História	70	3	10	7,49	2,09
Geografia	70	3	10	7,49	2,09
Ed.Artística	70	4	10	7,74	1,16
Ed.Física	70	6	10	8,46	0,96
Informática	70	7	10	8,19	0,74
<b>2ª série</b>					
Português	84	3	10	6,64	1,80
Matemática	84	2	10	6,71	1,95
Ciências	84	3	10	6,82	1,85
História	84	4	10	6,99	1,76
Geografia	84	4	10	6,99	1,76
Ed.Artística	84	5	10	8,04	0,78
Ed.Física	84	5	10	8,37	1,22
Informática	84	7	10	8,45	0,48
<b>3ª série</b>					
Português	65	3	9	6,49	1,49
Matemática	65	3	10	6,46	1,56
Ciências	65	3	9	6,45	1,49
História	65	3	9	6,36	1,39
Geografia	64	3	10	6,34	1,45
Ed.Artística	65	5	10	6,95	1,30
Ed.Física	65	5	10	8,40	1,36
Informática	65	7	10	8,08	0,69
Inglês	65	5	9	6,14	1,49

Tabela 66 (continuação) – Estatística descritiva da amostra total e por série em relação as matérias.

Matérias	N	Mínimo	Máximo	Média	D.P
<b>4ª série</b>					
Português	49	3	9	6,14	1,51
Matemática	49	3	9	5,80	1,38
Ciências	49	1	10	5,96	1,71
História	49	1	9	5,74	1,51
Geografia	49	1	10	5,91	1,62
Ed.Artística	49	5	10	7,24	1,69
Ed.Física	49	5	10	8,80	1,12
Informática	49	7	10	8,41	0,80
Inglês	49	5	9	5,69	1,12

As notas podem variar de 0 a 10, sendo a média mínima 5 para a aprovação. Por meio da Tabela 66, pode-se notar que as maiores médias se concentram nas matérias de Educação Física, Educação Artística e Informática e a menor média é vista na matéria de Inglês, exceto para a 1ª e 2ª série, pois não têm essa matéria. Essa menor média no Inglês pode ser devida a dificuldade em aprender uma outra língua, pois é um conteúdo mais distante da realidade do dia-a-dia, especialmente em escolas públicas como foi o caso da presente amostra, sendo que a maioria das crianças de escolas públicas tende a ter contato com a língua inglesa apenas na escola e a partir da 1ª série. Além disso, segundo Carrol (1993), a aprendizagem de língua estrangeira é o conteúdo que demanda maior capacidade da inteligência fluida. Com relação às correlações entre as matérias (desempenho acadêmico) e as variáveis do teste dinâmico, a Tabela 67 mostra os resultados obtidos com as variáveis totais e ganho.

**Tabela 67**– Correlações obtidas entre o desempenho acadêmico e escores do teste dinâmico.

Variáveis	Measure_Pré	Measure_Pós	Pós_Pré	Pós_Pré2	Pós_Pré2_res
Português	0,16**	0,27**	0,08	0,08	0,23**
Matemática	0,17**	0,24**	0,06	0,07	0,22**
Ciências	0,13*	0,23**	0,07	0,08	0,21**
História	0,104	0,20**	0,08	0,09	0,19**
Geografia	0,13*	0,22**	0,07	0,08	0,20**
Ed.Artística	0,07	0,10	0,02	0,04	0,11
Ed.Física	0,10	0,15*	0,04	0,05	0,15*
Informática	0,12	0,12	0,00	-0,01	0,08
Inglês	0,33**	0,48**	0,09	0,10	0,39**

\*\* p<0,01, \* p<0,05

Observa-se, na Tabela 67, que ocorreram várias correlações do Measure\_Pré, Measure\_Pós e Pós\_Pré2res com o desempenho acadêmico em diversas matérias, exceto Informática e Educação Artística. No entanto a maioria das correlações, apesar de positivas e significativas, foram baixas. As maiores ocorreram entre esses escores do teste dinâmico e a matéria de Inglês, que é justamente o conteúdo que demanda maior capacidade da inteligência fluida que, por sua vez tem, o raciocínio indutivo como um fator robusto, como foi verificado por Carrol (1993). Um dado interessante é que das variáveis ganho, somente a Pós\_Pré2res, que é a normalização das possibilidades de ganho, apresentou correlações com o desempenho. Esse resultado mostra a importância, nos testes dinâmicos, dessa normalização.

As baixas correlações entre o desempenho acadêmico e o teste dinâmico, especificamente com a variável ganho, são esperadas, pois o raciocínio indutivo não é

trabalhado diretamente nas matérias escolares. O conteúdo do ensino no sistema educacional brasileiro é voltado para a inteligência cristalizada, ou seja, adquirir informações, o que sem dúvida é importante, porém não enfatiza o raciocínio indutivo. A pesquisa de Primi (2005) também encontrou magnitudes de correlação baixas entre desempenho acadêmico e potencial de aprendizagem, explorando também raciocínio indutivo, mas com amostra universitária.

Tais resultados corroboram a observação de Sternberg e Grigorenko (2002), referindo a pouca predição do potencial de aprendizagem em relação ao desempenho acadêmico, justificado pela composição dos escores no teste dinâmico ser diferente das notas escolares, pelo fato dos testes dinâmicos avaliarem o potencial de aprender, e não o que a criança já aprendeu e, por fim, pelo fato do construto potencial de aprendizagem de raciocínio indutivo poder ser diferente do que normalmente é trabalhado na escola. Na presente pesquisa esse fator construto parece ser relevante, uma vez que raciocínio indutivo não é diretamente trabalhado no ensino brasileiro. Já no que se refere aos escores separados do teste dinâmico Measure\_Pré e Measure\_Pós, considerados desempenhos advindos de formas estáticas, as associações baixas tendem a sustentar o que já foi encontrado em pesquisas como dos autores Cruz (2008), Santos e Primi (2005), Primi, Santos e Vendramini (2002), Primi e Almeida (1998).

O raciocínio indutivo é um componente da inteligência fluida, e esta tende a ter seu auge na adolescência (Horn & Noll, 1997). Como a amostra é composta por crianças de 6 a

11 anos, é importante olhar esse dado por série. Para tanto, tais informações são detalhadas após a discussão dos resultados da Tabela 68, que se refere às correlações efetuadas com os paradigmas.

**Tabela 68-** Correlações obtidas entre o desempenho acadêmico e os paradigmas.

Variáveis	SD_pré	CC_pré	RE_pré	SC_pré	SD_pós	CC_pós	RE_pós	SC_pós
Português	0,17**	0,07	0,06	0,14*	0,28**	0,15*	0,08	0,13*
Matemática	0,17**	0,07	0,06	0,15*	0,23**	0,15*	0,10	0,12
Ciências	0,13*	0,05	0,06	0,14*	0,22**	0,13*	0,10	0,11
História	0,11	0,05	0,04	0,10	0,19**	0,11	0,09	0,10
Geografia	0,11	0,06	0,05	0,14*	0,21**	0,13*	0,09	0,11
Ed.Artístico	0,07	0,03	0,03	0,02	0,11	0,10	-0,01	0,06
Ed.Física	0,06	0,09	0,06	0,12*	0,23**	0,07	-0,03	0,03
Informática	0,10	0,02	0,06	0,10	0,20**	0,03	0,00	0,03
Inglês	0,08	0,23*	0,10	0,45**	0,43**	0,31**	0,15	0,31**

\*\* p<0,01, \*p<0,05

Os paradigmas que mostraram maior quantidade de correlações significativas com as notas foram SC\_pré e SD\_pós. No entanto, apesar de positivas, as correlações foram baixas, exceto com o desempenho na matéria de Inglês, com que houve correlações positivas e moderadas. Essa matéria também se correlacionou com os paradigmas CC\_pré, CC\_pós e SC\_pós, sendo que as correlações com Inglês foram as maiores encontradas. A explicação anterior, quando a matéria de Inglês se correlacionou com as variáveis da Tabela 67, pode ser aplicada nesses resultados. Todas essas correlações foram efetuadas com a amostra geral e, a seguir, serão expostos os dados com amostra por série (Tabela 69).



**Tabela 69-** Correlações entre desempenho acadêmico e os escores totais e ganho do teste dinâmico separando por série.

Variáveis	Measure_Pré	Measure_Pós	Pós_Pré	Pós_Pré2	Pós_Pré2_res
<b>1ª série</b>					
Português	0,00	0,14	0,09	0,03	0,05
Matemática	-0,04	0,10	0,09	0,05	0,05
Ciências	-0,02	0,13	0,09	0,04	0,05
História	-0,02	0,13	0,09	0,04	0,05
Geografia	-0,02	0,13	0,09	0,04	0,05
Ed.Artística	-0,03	0,10	0,08	0,08	0,11
Ed.Física	0,05	-0,01	-0,04	0,03	0,08
Informática	-0,03	0,09	0,08	0,07	0,08
<b>2ª série</b>					
Português	0,18	0,35**	0,19	0,20	0,38**
Matemática	0,18	0,38**	0,21	0,24*	0,43**
Ciências	0,20	0,37**	0,19	0,23*	0,43**
História	0,20	0,37**	0,19	0,21	0,40**
Geografia	0,20	0,37**	0,19	0,21	0,40**
Ed.Artística	0,14	0,01	-0,09	-0,06	0,01
Ed.Física	0,05	0,16	0,11	0,07	0,12
Informática	0,14	0,22*	0,11	0,10	0,22*
<b>3ª série</b>					
Português	0,33**	0,41**	0,02	0,01	0,29*
Matemática	0,44**	0,41**	-0,09	-0,09	0,25*
Ciências	0,28*	0,33**	0,00	-0,01	0,23
História	0,29*	0,39**	0,04	0,05	0,31*
Geografia	0,28*	0,45**	0,10	0,09	0,35**
Ed.Artística	0,07	0,31*	0,19	0,16	0,27*
Ed.Física	0,06	0,15	0,08	0,12	0,20
Informática	0,30*	0,32*	-0,03	-0,06	0,17
Inglês	0,38**	0,54**	0,08	0,04	0,38**

Tabela 69 (continuação)- Correlações entre desempenho acadêmico e os escores totais e ganho do teste dinâmico separando por série.

Variáveis	Measure_Pré	Measure_Pós	Pós_Pré	Pós_Pré2	Pós_Pré2_res
<b>4ª série</b>					
Português	0,40**	0,48**	0,03	0,10	0,43**
Matemática	0,50**	0,50**	-0,06	0,02	0,42**
Ciências	0,41**	0,50**	0,03	0,09	0,43**
História	0,35*	0,44**	0,04	0,11	0,40**
Geografia	0,45**	0,45**	-0,06	0,02	0,38**
Ed.Artística	0,30*	0,42**	0,08	0,09	0,33*
Ed.Física	0,26	0,28*	-0,02	-0,02	0,19
Informática	0,12	0,01	-0,12	-0,08	0,00
Inglês	0,28*	0,38**	0,06	0,15	0,39**

\*\* p<0,01, \*p<0,05

Efetuada as correlações separando por série constata-se um aumento de correlações e na magnitude das correlações ao ir crescendo as séries. Dessa forma notam-se muitas correlações moderadas. Como colocado anteriormente, o raciocínio indutivo tende a se apresentar mais desenvolvido com o aumento da idade, tendo seu auge na adolescência. Mesmo não sendo trabalhado diretamente nos conteúdos escolares, indiretamente a aprendizagem escolar também contribui para o desenvolvimento do raciocínio indutivo. Os resultados encontrados indicam tais pressupostos. Novamente se verifica a importância da variável que representa o ganho normalizado, pois por meio dela observou-se as correlações com desempenho acadêmico na maioria das séries, exceto a primeira que não teve correlações as variáveis do estudo. Com a separação por série, a matéria Inglês

continuou se correlacionando, ora de maneira baixa, ora moderada, mas não manteve o padrão das maiores correlações entre todas que surgiram.

Os resultados de várias correlações positivas moderadas, ou seja, acima de 0,40 entre desempenho acadêmico e o escore ganho pela nota Pós\_Pré2res, não corroboram com a observação de Sternberg e Grigorenko (2002) da pouca predição do potencial de aprendizagem em relação ao desempenho acadêmico. Tais resultados sugerem uma posição mais positiva dessa relação potencial e desempenho acadêmico. No entanto, são necessárias mais investigações procurando identificar se esse dado mais animador ganhará maior solidez. Também foram realizadas análises correlacionais por paradigma, na Tabela 70 se visualiza os resultados encontrados.

**Tabela 70-** Correlações obtidas entre o desempenho escolar e os paradigmas do raciocínio indutivo.

Variáveis	SD_pré	CC_pré	RE_pré	SC_pré	SD_pós	CC_pós	RE_pós	SC_pós
<b>1ª série</b>								
Português	0,17	-0,02	-0,12	-0,06	0,20	0,16	-0,14	0,08
Matemática	0,17	-0,03	-0,16	-0,08	0,15	0,16	-0,13	0,05
Ciências	0,15	-0,03	-0,14	-0,04	0,17	0,11	-0,09	0,11
História	0,15	-0,03	-0,14	-0,04	0,17	0,11	-0,09	0,11
Geografia	0,15	-0,03	-0,14	-0,04	0,17	0,11	-0,09	0,11
Ed.Artística	0,15	-0,05	-0,06	-0,19	0,22	0,06	-0,12	0,12
Ed.Física	0,12	0,04	0,01	0,18	-0,00	0,09	-0,04	-0,05
Informática	-0,03	0,01	0,03	-0,00	0,15	-0,04	0,01	0,08
<b>2ª série</b>								
Português	0,33**	0,01	0,08	0,05	0,41**	0,14	0,13	0,13
Matemática	0,34**	0,01	0,07	0,06	0,43**	0,15	0,14	0,17
Ciências	0,32**	-0,09	0,10	0,11	0,39**	0,15	0,15	0,21
História	0,30**	0,03	0,10	0,08	0,42**	0,14	0,12	0,18
Geografia	0,30**	0,03	0,10	0,08	0,42**	0,14	0,12	0,18
Ed.Artística	0,15	0,05	0,14	0,02	0,10	-0,01	-0,03	-0,13
Ed.Física	0,09	0,04	-0,06	0,05	0,22*	0,05	-0,20	0,17
Informática	0,24*	-0,10	-0,04	0,28*	0,35**	0,05	-0,07	0,17
<b>3ª série</b>								
Português	0,02	0,21	0,26*	0,32*	0,42**	0,02	0,20	0,31*
Matemática	0,09	0,29*	0,31*	0,41**	0,42**	0,05	0,21	0,30*
Ciências	-0,03	0,15	0,28*	0,30*	0,34**	0,01	0,18	0,24
História	-0,01	0,22	0,25*	0,29*	0,39**	0,02	0,20	0,30*
Geografia	-0,05	0,19	0,25*	0,32*	0,45**	0,07	0,17	0,36**
Ed.Artística	-0,05	0,01	0,01	0,21	0,28*	0,14	-0,04	0,35**
Ed.Física	-0,08	0,16	0,12	0,01	0,29*	0,04	0,09	-0,03
Informática	0,17	0,03	0,19	0,19	0,27*	0,19	0,15	0,06
Inglês	0,06	0,25*	0,18	0,50**	0,47**	0,26*	0,20	0,36**

Tabela 70 (continuação)- Correlações obtidas entre o desempenho escolar e os paradigmas do raciocínio indutivo.

Variáveis	SD_pré	CC_pré	RE_pré	SC_pré	SD_pós	CC_pós	RE_pós	SC_pós
<b>4ª série</b>								
Português	0,31*	0,17	0,19	0,50**	0,43**	0,32*	0,31*	0,19
Matemática	0,30*	0,20	0,30	0,61**	0,41**	0,34*	0,40**	0,20
Ciências	0,23	0,24	0,22	0,46**	0,49**	0,33*	0,36*	0,13
História	0,20	0,15	0,18	0,46**	0,46**	0,29*	0,34*	0,07
Geografia	0,24	0,22	0,22	0,58**	0,43**	0,32*	0,35*	0,07
Ed.Artística	0,27	0,21	0,08	0,27	0,35*	0,35*	0,13	0,24
Ed.Física	0,10	0,12	0,19	0,34*	0,36*	0,20	0,08	-0,01
Informática	0,07	0,20	0,00	0,11	0,18	0,01	-0,15	-0,05

\*\* p<0,01, \*p<0,05

Por meio da Tabela 70, nota-se que a primeira série novamente não apresentou correlações, mas dessa vez foi entre o desempenho acadêmico e os paradigmas. Todas as correlações emergidas são positivas e baixas ou moderadas. Os paradigmas que mostraram maior quantidade de correlação diferiram dependendo da série, porém o SD\_pós é o único que se correlacionou, na maioria das vezes moderadamente, com a maioria das séries, exceto 1ª, e com as matérias, exceto Ed.Artística na 3ª e Informática na 4ª série. A variável SD\_pós são itens do paradigma Similaridade e Diferenciação para serem resolvidos após a etapa da intervenção. Esse maior número de correlação pode ser devido ao aumento da variabilidade, ou seja, antes da etapa da intervenção a maioria dos indivíduos não acertavam e após vários indivíduos obtiveram melhores desempenhos. Retomando a análise dos erros, foi constatado que realmente essa variável foi a que os indivíduos, exceto os da 5ª, conseguiram acertar mais após a etapa da intervenção. Esse maior acerto também é

devido a facilidade dos itens desse paradigma se comparado com os itens de outros paradigmas.

Como já relatado anteriormente, além dessas notas de desempenho por matéria, foi coletada uma nota de desempenho de acordo com a percepção das professoras. Esse cuidado se fez necessário, pois notas escolares, principalmente de 1ª à 4ª série, em que a repetência é rara devido ao sistema de progressão continuada, é contaminada por diversos fatores, entre eles comportamento da criança, trabalhos em grupo, participação nas aulas, dentre outros. Na literatura há uma discussão sobre esse critério desempenho escolar ser confiável mediante a média na matéria (Depresbiteris, 1997). Somente relembrando, essa nota é constituída pela atribuição da professora, a partir da observação em sala de aula, dos cinco alunos com maior e dos cinco alunos com menor facilidade de aprender os conteúdos ensinados em sala de aula. A Tabela 71 sumariza as estatísticas descritivas das duas amostras derivadas a partir desse procedimento, isto é, com maior e menor facilidade para aprender.

**Tabela 71-** Estatísticas descritivas das amostras com maior e menor facilidade para aprender.

Variáveis	N	Mínimo	Máximo	Média	D.P
<b>Maior facilidade</b>					
Measure_Pré	61	-2,67	0,86	-0,64	0,73
Measure_Pós	61	-2,55	1,85	-0,25	0,80
Pós_Pré	61	-1,89	2,73	0,55	1,08
Pós_Pré2	61	2,00	2,73	0,79	0,79
Pós_Pré2res	61	-1,02	2,30	0,23	0,64
SD_pré	61	-2,97	3,15	0,02	1,40
CC_pré	61	-2,08	3,08	-0,40	1,10
RE_pré	61	-2,04	1,59	-0,43	1,15
SC_pré	61	-1,69	3,56	-0,06	1,34
SD_pós	61	-3,19	1,82	0,08	1,19
CC_pós	61	-1,74	3,83	-0,01	1,34
RE_pós	61	-1,67	3,81	0,04	1,22
SC_pós	61	-1,64	2,26	0,01	1,29
<b>Menor facilidade</b>					
Measure_Pré	59	-3,87	0,31	-1,17	0,77
Measure_Pós	59	-3,77	0,52	-0,88	0,78
Pós_Pré	59	-2,80	3,18	0,45	1,24
Pós_Pré2	59	0,00	3,18	0,74	0,88
Pós_Pré2res	59	-0,80	1,43	-0,13	0,57
SD_pré	59	-2,97	1,63	-0,88	1,30
CC_pré	59	-2,08	1,67	-0,55	1,22
RE_pré	59	-2,04	1,59	-0,88	1,04
SC_pré	59	-1,69	0,95	-0,68	0,88
SD_pós	59	-3,19	1,82	-0,86	1,17
CC_pós	59	-1,74	2,31	-0,55	1,13
RE_pós	59	-1,67	2,28	-0,37	1,05
SC_pós	59	-1,64	1,03	-0,51	1,02

Comparando as médias das duas amostras constata-se que o grupo percebido pelas professoras como o que apresenta maior facilidade para aprender realmente mostrou médias mais elevadas no teste dinâmico inteiro, ou seja, considerando todos os escores. Porém é

necessário averiguar se essa diferença de médias entre os grupos é significativa, para isso utilizou-se o teste *t de Student*. A Tabela 72 mostra os dados encontrados.

**Tabela 72-** Diferença de média entre os grupos com maior e menor facilidade de aprender.

Variáveis	<i>t</i>	<i>Gl</i>	<i>P</i>
Measure_Pré	-3,85	118	0,00
Measure_Pós	-4,37	118	0,00
Pós_Pré	-0,48	118	0,63
Pós_Pré2	-0,35	118	0,73
Pós_Pré2res	-3,29	118	0,00
SD_pré	-3,68	118	0,00
CC_pré	-0,71	118	0,48
RE_pré	-2,25	118	0,03
SC_pré	-3,01	118	0,00
SD_pós	-4,38	118	0,00
CC_pós	-2,42	118	0,02
RE_pós	-1,98	118	0,05
SC_pós	-2,42	118	0,02

Por meio da Tabela 72 pode-se verificar que houve diferença de média entre a maior parte das variáveis do teste dinâmico, exceto as de ganho Pós\_Pré, Pós\_Pré2 e o paradigma CC\_pré. Novamente observa-se que a variável Pós\_Pré2res é o escore de ganho que é mais sensível nas análises para detecção do resultado de correlações e, no caso, diferenças de médias, confirmando mais uma vez a importância da normalização dos ganhos. Já o paradigma CC apresentou diferença significativa no pós e não no pré-teste, resultado que pode ser explicado pela maior capacidade que o grupo com maior facilidade de aprender demonstrou perante o procedimento da intervenção. Para essa amostra também foram realizadas análises correlacionais. Na Tabela 73 são mostrados os resultados para a amostra geral com o teste dinâmico.



**Tabela 73**– Correlações obtidas entre o teste dinâmico e a nota da professora por meio da percepção em sala de aula

Teste dinâmico	Nota Professora
Measure_Pré	0,33**
Measure_Pós	0,37**
Pós_Pré	0,04
Pós_Pré2	0,03
Pós_Pré2res	0,29**
SD_pré	0,32**
CC_pré	0,07
RE_pré	0,20*
SC_pré	0,27**
SD_pós	0,37**
CC_pós	0,22*
RE_pós	0,18*
SC_pós	0,22*

\*\*  $p < 0,01$ , \* $p < 0,05$

Com essa amostra também não foi verificada correlação significativa com as variáveis de ganho Pós\_Pré e Pós\_Pré2, somente com Pós\_Pré2res. O CC\_pré foi o único paradigma que não apresentou correlação significativa, porém apesar dos outros paradigmas demonstrarem correlação significativa, bem como os escores totais e ganho, todas foram baixas. A próxima Tabela (74) elenca a mesma análise correlacional, mas separando por série.

**Tabela 74-** Correlações obtidas entre o teste dinâmico e a nota da professora por meio da percepção em sala de aula separada por série.

Teste Dinâmico	NProf (1 <sup>a</sup> )	NProf (2 <sup>a</sup> )	NProf(3 <sup>a</sup> )	NProf(4 <sup>a</sup> )
Measure_Pré	0,11	0,27	0,44*	0,56**
Measure_Pós	0,30	0,48**	0,44*	0,45*
Pós_Pré	0,13	0,22	-0,15	-0,09
Pós_Pré2	0,01	0,25	-0,12	0,06
Pós_Pré2res	0,08	0,52**	0,23	0,47*
SD_pré	0,24	0,40*	0,10	0,61**
CC_pré	-0,05	0,00	0,26	0,09
RE_pré	-0,03	0,18	0,31	0,33
SC_pré	0,00	0,12	0,57**	0,51**
SD_pós	0,30	0,62**	0,46*	0,37
CC_pós	0,39*	0,25	-0,10	0,38
RE_pós	-0,08	0,16	0,25	0,43*
SC_pós	0,11	0,21	0,37*	0,25

Separando as análises por série detectam-se correlações significativas positivas e moderadas, exceto duas que foram baixas, mas quase moderadas (Nprof1<sup>a</sup> com CC\_pós  $r = 0,39$   $p < 0,05$  e Nprof3<sup>a</sup> com SC\_pós  $r = 0,37$   $p < 0,05$ ). Dependendo da série o perfil de correlações se modificou, mas a variável Measure\_Pós foi a única que se correlacionou significativamente com as séries, exceto na 1<sup>a</sup> série, que também demonstrou o menor número de correlações, apenas uma.

A seguir serão descritas as análises correlacionais com a amostra da 5<sup>a</sup> série. A Tabela 75 contém as informações referentes ao desempenho acadêmico com as variáveis totais e ganho do teste dinâmico.

**Tabela 75**– Correlações obtidas entre o desempenho acadêmico e escores do teste dinâmico para a 5ª série.

Variáveis	Measure_Pré	Measure_Pós	Pós_Pré	Pós_Pré2	Pós_Pré2_res
Português	0,42**	0,22	-0,08	-0,13	0,11
História	0,40**	0,11	-0,19	-0,22	0,01
Geografia	0,44**	0,27	-0,05	-0,08	0,17
Ciências	0,57**	0,36*	-0,05	-0,14	0,20
Matemática	0,47**	0,26	-0,09	-0,19	0,09
Ed. Física	0,38*	0,25	-0,03	0,02	0,24
Ed.Artística	0,42**	0,04	-0,27	-0,39*	-0,14
Inglês	0,39**	0,39**	0,11	-0,01	0,22

\*\* p<0,01, \* p<0,05

As correlações encontradas entre desempenho acadêmico e os escores totais e ganho do teste dinâmico foram em menor quantidade ao se comparar com as 3ª e 4ª série, e as associações concentraram-se no escore total do pré-teste. No escore ganho houve somente uma correlação, que foi negativa com a matéria Educação Artística. Possíveis explicações para essa correlação negativa é que o raciocínio indutivo e a matéria de Educação Artística de fato envolvam diferentes habilidades que não têm relação entre si. Para essa amostra também foi verificada as correlações com os paradigmas (Tabela 76).

**Tabela 76-** Correlações obtidas entre o desempenho acadêmico e escores dos paradigmas do teste dinâmico para a 5ª série.

Variáveis	SD_pré	CC_pré	RE_pré	SC_pré	SD_pós	CC_pós	RE_pós	SC_pós
Português	-0,02	0,26	0,14	0,49**	0,09	0,06	0,15	0,30
História	0,06	0,14	0,30	0,34*	-0,15	0,14	0,11	0,17
Geografia	0,01	0,24	0,32*	0,37*	-0,04	0,16	0,27	0,32*
Ciências	0,15	0,37*	0,24	0,49**	0,13	0,30*	0,29	0,29
Matemática	0,05	0,29	0,21	0,45**	0,05	0,18	0,25	0,23
Ed. Física	0,19	0,13	0,27	0,26	-0,07	0,30	0,18	0,31*
Ed. Artística	-0,03	0,47**	0,08	0,35*	-0,07	-0,05	0,13	0,08
Inglês	-0,07	0,31*	0,12	0,45**	0,23	0,25	0,30*	0,26

\*\* p<0,01, \* p<0,05

A análise correlacional com os paradigmas mostrou correlações baixas e moderadas, mas poucas, sendo o maior número de associação entre o paradigma SC\_pré e o desempenho em nas matérias, exceto Educação Física. Aliás Educação Física, Educação Artística e Informática, no caso das séries de 1ª à 4ª, são as que apresentam menor número de correlações com o teste dinâmico. Essa constatação parece adequada, pois são matérias que envolvem outras habilidades e que mesmo assim são pouco exigidas na escola. Concluindo essa parte das notas escolares, salvas algumas diferenças, pode-se constatar que tanto o critério desempenho acadêmico, construído por meio das notas escolares tradicionais, quanto o critério desempenho composto pela percepção das professoras mostraram resultados positivos de associação entre potencial de aprendizagem do raciocínio indutivo e desempenho escolar, bem como entre a habilidade desse raciocínio já adquirida e esse mesmo critério desempenho.

Apesar da literatura apontar para pouca ou nenhuma evidência da predição do potencial de aprendizagem em relação ao desempenho acadêmico (Sternberg & Grigorenko), no presente trabalho foi notada uma relação mais robusta entre essas variáveis. Os testes estáticos de inteligência tendem a apresentar predição em relação ao desempenho acadêmico e seria importante que os testes dinâmicos também demonstrassem alguma contribuição para a predição desse desempenho, o que tornaria esse tipo de testagem ainda mais relevante.

Diante disso procurou-se verificar a validade incremental do teste dinâmico em comparação ao teste estático Prova de Raciocínio Abstrato da BPR-5i. A validade incremental propicia a informação, nesse caso, de se o teste dinâmico prediz alguma informação única além da já predita pelo teste estático RA. Para isso aplicou-se uma análise de regressão hierárquica, colocando no primeiro passo o RA como preditor e em um segundo variáveis Measure\_Pré e Pós\_Pré2res do teste dinâmico. Somente foram elencadas essas duas variáveis do teste dinâmico, pois a Measure\_Pré é considerada uma nota como um teste estático, sem a interferência da etapa da intervenção que ocorre na Measure\_Pós e a Pós\_Pré2res é justamente o escore ganho, ou seja, por ela que se verifica o potencial de aprendizagem. A Tabela 77 mostra os resultados encontrados para cada série.

**Tabela 77**– Análise de regressão hierárquica com o teste estático e o dinâmico para a predição do desempenho acadêmico.

Modelo	Variáveis preditoras	B	Erro Padrão	<i>b</i>	<i>t</i>	<i>p</i>	<i>F</i>	R <sup>2</sup>	
<b>1ª Série</b>									
1	(Constant)	8,03	0,403			19,95	0,00	2,17	0,35
	E_RA_TT	-0,06	0,043	-0,19	-1,47	0,15			
2	(Constant)	8,03	0,446			18,02	0,00	0,74	0,37
	E_RA_TT	-0,06	0,045	-0,18	-1,36	0,18			
	Measure_Pré	0,04	0,248	0,02	0,14	0,89			
	Pós_Pré2res	-0,08	0,357	-0,03	-0,24	0,81			
<b>2ª Série</b>									
1	(Constant)	7,80	0,357			21,85	0,00	2,23	0,03
	E_RA_TT	-0,04	0,027	-0,18	-1,50	0,14			
2	(Constant)	8,24	0,368			22,40	0,00	9,10	0,29
	E_RA_TT	-0,04	0,024	-0,16	-1,57	0,12			
	Measure_Pré	0,37	0,185	0,21	1,98	0,05			
	Pós_Pré2res	1,11	0,235	0,49	4,72	0,00			
<b>3ª Série</b>									
1	(Constant)	6,90	0,401			17,20	0,00	0,04	0,00
	E_RA_TT	-0,01	0,025	-0,03	-0,20	0,85			
2	(Constant)	7,06	0,379			18,63	0,00	5,48	0,23
	E_RA_TT	-0,01	0,022	-0,03	-0,22	0,83			
	Measure_Pré	0,40	0,153	0,31	2,63	0,01			
	Pós_Pré2res	0,67	0,221	0,36	3,02	0,00			
<b>4ª Série</b>									
1	(Constant)	5,86	0,557			10,53	0,00	2,11	0,04
	E_RA_TT	0,05	0,031	0,22	1,45	0,15			
2	(Constant)	6,62	0,492			13,44	0,00	8,04	0,37
	E_RA_TT	0,02	0,027	0,09	0,67	0,51			
	Measure_Pré	0,42	0,140	0,38	3,00	0,01			
	Pós_Pré2res	0,66	0,234	0,37	2,81	0,01			
<b>5ª Série</b>									
1	(Constant)	6,89	0,529			13,04	0,00	0,00	0,00
	E_RA_TT	-0,00	0,027	-0,01	-0,05	0,96			
2	(Constant)	6,44	0,499			12,91	0,00	4,24	0,27
	E_RA_TT	0,01	0,025	0,04	0,29	0,78			
	Measure_Pré	0,75	0,237	0,47	3,18	0,00			
	Pós_Pré2res	0,23	0,281	0,12	0,81	0,42			

Como pode ser observado na Tabela 77 variável RA, em todas as séries, não se mostrou preditora do desempenho acadêmico, já a variável ganho, que se refere ao potencial de aprendizagem, foi preditora do desempenho acadêmico nas séries 2<sup>a</sup>, 3<sup>a</sup> e 4<sup>a</sup>. A variável Measure\_Pré, considerado um teste estático dentro do teste dinâmico também apresentou predição para o desempenho acadêmico e na maioria das séries, exceto a 1<sup>a</sup>. Tais resultados apontam evidência de validade incremental, ou seja a contribuição única de um teste em prever algum critério, no caso o teste dinâmico, sugerindo que o potencial de aprendizagem também tende a propiciar predição do desempenho acadêmico.

### **CAPÍTULO 3- CONSIDERAÇÕES FINAIS**

O principal objetivo desse trabalho foi a construção de um teste dinâmico e informatizado do raciocínio indutivo para crianças de 7 a 11 anos de idade. Para se atingir esse objetivo foi necessário o desenvolvimento de várias etapas anteriores, sendo a primeira a busca por uma fundamentação teórica. Conseqüentemente, todas as outras etapas foram embasadas na Teoria Prescritiva do Raciocínio Indutivo de Josef Klauer (1990). Essa teoria, além de definir e operacionalizar o raciocínio indutivo, abrange um treinamento sobre esse construto. Com isso as etapas de construção dos itens para o pré e pós-teste e intervenção do teste dinâmico foram fundamentadas pela Teoria de Klauer (1990).

Por meio do Estudo 1 foi possível investigar, utilizando-se de análises qualitativas e quantitativas, a adequação dos itens e do procedimento de intervenção, bem como efetuar as mudanças necessárias para a melhora dessas duas partes que compuseram o teste dinâmico. A partir dos bons resultados obtidos, tanto dos itens quanto do procedimento de intervenção, foi possível a construção do teste dinâmico no seu formato pré-teste, intervenção e pós-teste. No caso da intervenção, somente não foi possível responder a sugestão para a diminuição do tempo, pois implicaria retirar dessa parte instruções relevantes para que o indivíduo compreendesse como, de acordo com a Teoria de Klauer (1990), poderia resolver tarefas utilizando o raciocínio indutivo.

A Teoria de Klauer (1990) sugere que inicialmente seja ensinado para a criança o conhecimento declarativo, ou seja, o conhecimento das características de objetos e das



relações entre objetos, por exemplo. Após é necessário ensinar o conhecimento procedural, que é como resolver as tarefas com o raciocínio indutivo, para isso é preciso ensinar como resolver, por meio dos paradigmas do raciocínio indutivo, que são seis, mas na presente tese uniram-se dois e utilizaram-se quatro. Diante disso a etapa da intervenção ficou extensa, mas essa limitação foi assumida no trabalho e realmente ao aplicar o teste dinâmico constatou-se nitidamente o cansaço de várias crianças ao fazerem o teste. Sendo que o tempo longo da etapa da intervenção pode ter sido uma variável que interferiu negativamente para verificar o potencial de aprendizagem das crianças de maneira mais apurada, pois os ganhos foram significativos, mas poderiam ter sido maiores.

Frente essas observações sugere-se que o teste dinâmico seja separado em quatro partes, avaliando de forma separada e em momento diferentes, um em cada dia, por exemplo, os paradigmas SD, CC, RE e SC. Claro que poderá ocorrer uma aprendizagem em cascata, ou seja, quando mensurar o CC a criança ter aprendido algo anterior no SD e não ser possível verificar o potencial de aprendizagem em cada paradigma de forma pura. No entanto, essa mensuração pura não é o objetivo, pois a Teoria de Klauer (1990) propõem justamente essa adição de aprendizagem entre os paradigmas, ou seja, para resolver CC, RE e SC é preciso ter aprendido como resolver o SD.

Outra sugestão para futuras pesquisas é utilizar concomitantemente o teste dinâmico na forma informatizada, mas com um examinador transmitindo as instruções, bem como aplicar somente sem o examinador e verificar os resultados. Na pesquisa de Tzuriel e

Shamir (2002) eles relatam que a aplicação computadorizada junto com o examinador é a mais adequada, porém a pesquisa comparou esse formato apenas com a aplicação sem computador e não investigaram a forma utilizando somente o computador. Porém, é importante destacar que para uma melhor compreensão das dificuldades do indivíduo, bem como para se obter uma análise mais clínica, o teste dinâmico aplicado com um tutor pode ser a forma mais apropriada, pois permite o acesso a informações que não podem ser obtidas somente com a aplicação padronizada informatizada.

Como descrito anteriormente, apesar da durabilidade da etapa da intervenção ter sido considerada longa para as crianças, optou-se por proceder ao Estudo 2, que abarcou as propriedades psicométricas, precisão e estudos de evidências de validade, para o teste dinâmico. Para isso foram verificadas a precisão do instrumento, os índices de ajuste por meio da análise no Winsteps utilizando o modelo de Rasch, a estrutura fatorial oblíqua pelo programa Microfact, as evidências de validade convergente-discriminante com os testes Prova de Raciocínio Abstrato da BPR-5 infantil, Geração Semântica e Stroop, as evidências de validade teste-critério com o desempenho acadêmico, bem como a validade incremental do instrumento dinâmico em prever o desempenho acadêmico. Antes de iniciar as evidências de validade com critério externo, BPR-5i, Geração Semântica, Stroop e Desempenho Acadêmico, foi investigado se o teste dinâmico possibilitava identificar ganho, ou seja, potencial de aprendizagem.

Com relação à precisão, índice de ajuste e análise fatorial, foi preciso explorar pré-teste e pós-teste separadamente, pois como há uma etapa de intervenção no meio, são considerados testes diferentes, apesar de serem equivalentes. As precisões foram baixas, 0,42 para o pré-teste e 0,45 para o pós-teste, esse resultado pode ter ocorrido em função da pouca quantidade de itens somada a diversidade de conteúdo e forma de raciocinar indutivamente decorrente dos paradigmas. Tanto para o pré quanto para o pós foram somente quatro itens avaliando cada paradigma que propõem raciocínios diferentes, apesar de todos serem indutivos.

Ao que se refere os índices de ajuste *infit*, *outfit* e também correlação *item-theta*, apenas três itens do pré e três do pós apresentaram algum desajuste ou correlação baixa. Tais dados sugerem que a maioria dos itens do pré e pós foi adequada para comporem o teste dinâmico. Já a análise fatorial apontou uma estrutura unidimensional para o pré-teste e multidimensional, com dois fatores, para o pós-teste e que foram interpretados como sendo um estrutural e outro semântico. Tais resultados apontam que apesar da existência dos paradigmas, o raciocínio indutivo é o construto que prevalece conferindo evidência de validade da estrutura interna para o pré e pós-teste. Ainda em relação ao pré e pós-teste explorados separadamente, foi encontrada evidência de validade teste-critério com série de forma parcial, pois não foram observados aumento dos desempenhos em função das séries. Os únicos dados identificados foram o melhor desempenho da 5ª série ao comparar com as outras e da 3ª série confrontando com a 1ª e 2ª.

A informação mais relevante que um teste dinâmico precisa fornecer é a identificação do potencial de aprendizagem, no caso da presente tese, em relação a resolução de tarefas com raciocínio indutivo. O teste construído nesse trabalho apresentou essa característica, ou seja, conseguiu detectar potencial de aprendizagem, por meio dos ganhos de habilidade do pré para o pós-teste. Esse ganho pode ser visualizado independente de série e sexo, no entanto o ganho nos paradigmas é diferenciado dependendo da série. Porém independente da série o ganho maior foi em relação ao paradigma RE, e a 3ª série foi a única que mostrou ganho em todos os paradigmas. Uma das razões para a 3ª série ter apresentado melhor desempenho no ganho do que a 4ª e 5ª pode ser em virtude da seriedade no momento de realizar o teste. A 5ª série durante as aplicações do teste dinâmico se mostrou demasiadamente indisciplinada, por exemplo, os alunos brincavam, falavam alto e atrapalhavam os outros examinados. Os dados obtidos dessa série poderiam ter sido bem melhores.

Aliás, a questão comportamental para a realização de qualquer teste é bastante relevante para a compreensão dos resultados obtidos. Por exemplo, por meio dos dados encontrados nessa pesquisa, os alunos da 5ª série obtiveram desempenhos menores, em relação ao potencial de aprendizagem, comparado com a 3ª série. No entanto, esse resultado pode ter ocorrido em função dos comportamentos de desmotivação e displicência dos alunos da 5ª série. Portanto pode-se questionar o quanto os alunos da 5ª série poderiam ter se beneficiado da intervenção se tivessem apresentado um comportamento mais pró-ativo,

mais motivador para fazer o teste dinâmico. Quando se avalia a parte cognitiva é preciso considerar essas variáveis comportamentais. De acordo com Simões (1995), fatores como ansiedade, falta de motivação, impulsividade e outros também podem afetar a capacidade que o sujeito já possui sobre o construto a ser mensurado por meio de testes dinâmicos. Tais fatores não foram analisados no presente estudo e sugere-se que pesquisas futuras incluam sua avaliação, de forma a analisar o potencial de aprendizagem em testes dinâmicos, medido pelo ganho do pré para o pós-teste, também em relação a tais aspectos motivacionais e comportamentais.

Com o potencial de aprendizagem sendo verificado pelo teste dinâmico pode-se investigar as evidências de validade com outras variáveis para o escore ganho (potencial). A primeira evidência buscada foi com a Prova de Raciocínio Abstrato-RA da BPR-5i que também avalia raciocínio indutivo. Porém ao contrário do esperado, as correlações significativas, apesar de positivas, foram baixas. A explicação para tal fato recorre aos conteúdos dos itens bem diferentes em cada teste e ao processo de raciocínio indutivo que deve ser utilizado para a resolução. No RA é exigido somente analogia do tipo A está para B assim como C está para D e nos itens do teste dinâmico, além dessa analogia, há, por exemplo, complemento de série, raciocínios cruzados, entre outros. Esses resultados apontam pouca evidência de validade convergente entre os dois testes. Seria interessante, em estudo futuros aplicar testes estáticos que fossem mais compatíveis aos conteúdos e a forma de raciocinar indutivamente do teste dinâmico.

Na introdução se relatou que elementos relacionados às funções executivas, especialmente as habilidades de controle inibitório e atenção seletiva, poderiam estar relacionadas à resolução adequada de tarefas de raciocínio indutivo, bem como ajudariam no momento das instruções do procedimento da intervenção, colaborando para o ganho de habilidade. A evidência de validade, no caso, se referiu a teste-critério e foi explorada com o teste Geração Semântica, para controle inibitório, e teste de Stroop, para atenção seletiva. No entanto, somente foi constatada correlação significativa entre o escore ganho e atenção seletiva para alunos da 5ª série. A importância da atenção seletiva e controle inibitório para a resolução de testes de raciocínio e para a aprendizagem de instruções é coerente, no entanto esse dado encontrado se depara com a questão que as funções executivas percorrem um processo maturacional, estando melhores desenvolvidas a partir da adolescência (Huizinga, Dolan & Molen, 2006). Então o resultado encontrado da atenção seletiva foi adequado, mas tanto com essa função quanto com controle inibitório, novos estudos devem ser realizados, pois esses iniciais não atestaram evidência de validade teste-critério com controle inibitório e atenção seletiva, com essa última ocorreu um dado mais positivo.

Por fim verificou-se a predição do potencial de aprendizagem para o raciocínio indutivo em relação ao desempenho acadêmico. A literatura sempre aponta pouca ou nenhuma correlação entre essas variáveis (Primi, 2005, Sternberg & Grigorenko, 2002 e Escolano & Linhares, 2001). No entanto os resultados da presente tese mostraram correlações moderadas entre potencial de aprendizagem e desempenho acadêmico indicando evidência de validade teste-critério para o teste dinâmico. Esse dado se fortaleceu

com a predição única do potencial de aprendizagem em relação ao desempenho acadêmico, quando comparado com o teste estático RA e pré e pós-teste. Diante disso, o teste dinâmico também apresentou evidência de validade incremental. Mesmo frente a esses resultados positivos entre desempenho acadêmico e potencial de aprendizagem, teria sido interessante o maior controle das notas de desempenho acadêmico por meio de provas mais objetivas. Como já discutido na introdução, as notas escolares são contaminados por diversas variáveis, portanto sugere-se que, em futuras pesquisas, seja elaborada uma prova sobre os conhecimentos acadêmicos dos alunos, que pode ser formulada com a ajuda da professora. Essa metodologia pode fortalecer os resultados encontrados na presente pesquisa, tornando a avaliação de desempenho mais objetiva e, portanto, tornando mais válida a investigação da sua relação com o potencial de aprendizagem.

Por meio dos achados do Estudo 2 constata-se que o teste dinâmico, desenvolvido na presente tese, apresentou resultados favoráveis e encorajadores em relação as suas propriedades psicométricas a sua capacidade de identificar potencial de aprendizagem. Esse trabalho pode ser considerado audacioso, uma vez que não seguiu o esquema tradicionalmente utilizado para a aplicação de teste dinâmico. O formato informatizado, padronizado e sem a ajuda de um examinador para o teste dinâmico, foram características pela primeira construídas para o teste ser aplicado em crianças, pois para adolescentes e adultos já existe o teste dos autores Primi, Sternberg e Grigorenko (1997).

Os testes dinâmicos, geralmente são utilizados para avaliação de crianças, tendo o caráter do examinador propiciar ajuda de acordo com a identificação das dificuldades do indivíduo, ou seja, é uma aplicação mais idiossincrática. O presente teste dinâmico construído não possibilitou essa idiossincrasia, bem como não precisou de um examinador, pois a criança interagia somente com o computador. Essa forma de aplicação pode ser contestada e não aprovada por diversos teóricos da área da testagem dinâmica, justamente por considerarem que cada criança tem suas dificuldades e a ajuda necessita ser de acordo com essas deficiências. Com certeza essa aplicação mais idiossincrática, mais clínica tem um papel muito relevante na testagem dinâmica e não deve ser abandonada. Porém é preciso considerar e desenvolver novos tipos de aplicação, como o que foi desenvolvido na presente tese, que são mais estruturados e possibilitam uma aplicação coletiva, bem como propiciar estudar os testes dinâmicos por meio de suas propriedades psicométricas e propor uma normatização, pois até o momento não foram encontrados testes dinâmicos que oferecesse uma norma de comparação entre o escore do indivíduo e o grupo. A normatização não foi um objetivo desse trabalho, porém com um maior número de aplicações em uma amostra mais heterogênea, esse objetivo poderá ser realizado.

Outra limitação do Estudo 2 foi a amostra. Como observado pelos resultados os participantes apresentaram baixa habilidade para resolverem tanto o pré quanto o pós-teste. No entanto esse dado precisa ser compreendido também dentro do contexto que essas crianças estão inseridas. Os participantes advêm de famílias com nível socioeconômico bem baixo, com menos acesso a recursos e estimulação do desenvolvimento cognitivo.



O projeto que envolveu a elaboração da tese abrangia a aplicação do teste dinâmico em escola particular, justamente para que se obtivesse maior variabilidade dos dados, pois se sabe que nessas escolas as crianças tendem a apresentar melhores desempenhos. Prova disso são alguns manuais de testes de inteligência que trazem tabelas separadas para indivíduos do ensino público e particular, por exemplo a Bateria de Provas de Raciocínio-BPR5 (Primi & Almeida, 1998) e Matrizes Progressivas Coloridas de Raven (Angelini, Alves, Custódio, Duarte, & Duarte, 1999). Então se essas crianças parecem possuir um aparato cognitivo mais desenvolvido, acredita-se que também poderiam vir a demonstrarem mais ganhos perante as instruções no procedimento da intervenção. Mas como a amostra foi por conveniência e as escolas particulares que foram contatadas não se mostraram interessadas, a coleta dos dados se restringiu a apenas uma escola pública de 1<sup>a</sup> à 4<sup>a</sup> série e uma de 5<sup>a</sup> série. Frente essas questões, verifica-se a necessidade de investigar o teste dinâmico construído com uma amostra mais heterogênea que abarque o ensino particular e também crianças mais velhas, pois não houve efeito teto no pré e pós-teste.

A comparação entre escolas públicas e particulares também é relevante ao se pensar nas questões comportamentais e motivacionais. Conforme observações heurísticas, inclusive as efetuadas no presente estudo, pode-se perceber que as crianças de escolas particulares tendem a demonstrar maior disposição e motivação para a resolução desses tipos de testes, enquanto crianças de escolas públicas tendem a apresentar menor interesse em participar. Seria importante desenvolver uma pesquisa comparando escolas públicas e particulares para investigar aspectos de inteligência e potencial de aprendizagem, porém

controlando a variável comportamental e motivacional que podem interferir nos resultados. Como já colocado, é interessante que se faça uma investigação com alunos do ensino particular, porém é importante destacar a relevância de pesquisas com alunos da escola pública, uma vez que a maioria da população está inserida no ensino público.

Concluindo, apesar das limitações apontadas e da ousadia presente do teste dinâmico construído, considera-se esse instrumento promissor na área da testagem dinâmica, pois mostrou, além das evidências de validade, principalmente a preditiva em relação ao desempenho acadêmico, que é capaz de mensurar potenciais de aprendizagem para resolução de tarefas com raciocínio indutivo, sendo que essa última característica é que identifica se um teste dinâmico está coerente com seu propósito, ou seja, mensurar potenciais de aprendizagem.

#### 4. REFERÊNCIAS

- Adleman, N. E., Menon, V., Blasey, C. M., White, C. D., Warsofsky, I. S., Glover, G. H. & Reiss, A. L. (2002). *A developmental fMRI study of the Stroop Color-Word Task. Neuroimage, 16* (1), 61-75.
- Almeida, L. S. (1988). *Teorias da Inteligência*. Porto: Edições Jornal de Psicologia.
- Almeida, L. S. (1994) *Inteligência: Definição e Medida*. Aveiro: CIDInE.
- Almeida, L. S. & Morais, M. F. (1997). *Programa promoção cognitiva*. Barcelos: Didálvi.
- Almeida, L. S. & Primi, R. (1998). *Bateria de Provas de Raciocínio (BPR-5): Manual Técnico*. São Paulo: Casa do Psicólogo.
- Almeida, L. S., Primi, R. & Cruz, M. B. Z. (em construção). *Bateria de Provas de Raciocínio para Crianças (BPR-5i)*.
- Angelini, A.L., Alves, I.C.B., Custódio, E.M., Duarte, W.F., & Duarte, J.L.M. (1999). *Matrizes progressivas coloridas de Raven. Manual*. São Paulo, SP: Centro Editor de Testes e Pesquisas em Psicologia.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Disponível On-line em: <http://www.intestcom.org>.

- Anastasi, A. & Urbina, S. (2000). *Testagem Psicológica*. Porto Alegre: Artmed.
- Ardila, A., & Ostrosky-Solís, F. (1996). *Diagnóstico del daño cerebral: enfoque neuropsicológico*. Mexico: Editorial Trillas.
- Associação Brasileira de Dislexia (1995). *Lista de frequência de palavras para crianças*. São Paulo, SP: ABD.
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, 121(1), 65-94.
- Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Science*, 29, 109-160.
- Bowman, D. B. Markham, P. M. Roberts, R. D. (2002). Expanding the frontier of human cognitive abilities: so much more than (plain) g! *Learning and Individual Differences*, 13, 127-158.
- Canadian Psychological Association (1996). Guidelines for educational and psychological testing. Ontário, CA: CPA. Disponível on-line em: <http://www.cpa.ca/guide9.html>
- Capovilla, A. G. S. & Macedo, E. C. (no prelo). *Teste de Geração Semântica*.
- Capovilla, A. G. S., Montiel, J. M., Macedo, E. C. & Charin, S. (no prelo). *Teste de Stroop Computadorizado*.

- Carrol, J. B. (1993). *Human Cognitive Abilities: A Survey of factor-Analytic Estudos*. New York: Cambrid University Press.
- Carrol, J. B. (1997). The Three-Stratum Theory of Cognitive Abilities. Em: D. P. Flanagan, J. L. Gensheft, P. L. Harrison, (Eds). *Contemporany Intelectual Assesment: theories, tests and issues*. (pp.122-130). New York, London: The Guilford Press.
- Chan, L. L. & Osguthorpe, R. T. (1990). The effects of computerized picture-word processing on kindergarten' language development. *Journal of Research Childhood Education*, 5, 73-83.
- Conselho Federal de Psicologia – CFP (2003). *Resolução n.º 002/2003 [Online]*.Disponível: <http://www.pol.org.br>.
- Cronbach, L. J. (1996). *Fundamentos da Testagem Psicológica*. Porto Alegre: Artes Médicas.
- Cruz, M. B. Z. (2008). *Estudo de validade e precisão da Bateria de Provas Infantis-BPR-5i*. Dissertação de Mestrado não publicada. Programa de Pós-Graduação Sticto Sensu em Psicologia da Universidade São Francisco, Itatiba S.P.
- Depresbiteris, L. (1997). Avaliação da aprendizagem: Revendo conceitos e posições. Em C. P. Sousa (Org.), *Avaliação do rendimento escolar* (6ª ed.; pp. 51-79). Campinas: Papirus.

- Embretson, S. E. (1983). Construct Validity: Construct Representation Versus Nomothetic Span. *Psychological Bulletin*, 93(1), 179-197.
- Embretson, S. E. (1985). Studying Intelligence With Test Theory Models. Em: D. K. Detterman (Ed.). *Current topics in human intelligence*. Norwood, NJ: Ablex.
- Embretson, S. E. (1994). Applications of Cognitive Design Systems to Test Development. Em: C. R. Reynolds, (org.). *Cognitive assessment: a multidisciplinary perspective*. New York: Plenum Press.
- Embretson, S. E. (1999). IRT Applications in Cognitive and Developmental Assessment. Em: S. E. Embretson, S. L. Hershberger (Eds.). *The New Rules of Measurement: what every psychologist and educator should know*. Mahwah, New Jersey, London: Lawrence Erlbaum Associates.
- Embretson, S. E. & Reise, S. P. (2000). *Item Response Theory for Psychologists*. New Jersey: Lawrence Erlbaum.
- Enumo, S. R. F. & Batista, C. G. (2006). Jogo de perguntas de busca com figuras geométricas para crianças com deficiência visual (PBFV-DV). Em M. B. M. Linhares, A. C. M. Escolano & S. R. F. Enumo (Eds.), *Avaliação Assistida: Fundamentos, Procedimentos e Aplicabilidade* (pp.69-78). São Paulo, Casa do Psicólogo.

- Enumo, S. R. F., Ferrão, E. S., Motta, A. B., Moraes, E. O. & Linhares, M. B. M. (2006). *Um panorama das pesquisas sobre avaliação assistida no Brasil. Psicologia e Educação, 2*, 712-723.
- Escolano, A. C. M. & Linhares, M. B. M. (2006). Avaliação cognitiva assistida em situação de resolução de problema na predição do desempenho escolar de crianças de primeira série do primeiro grau. Em M. B. M. Linhares, A. C. M. Escolano & S. R. F. Enumo (Eds.), *Avaliação Assistida: Fundamentos, Procedimentos e Aplicabilidade* (pp.69-78). São Paulo, Casa do Psicólogo.
- Fernandez-Duque, D., Baird, J. A. & Posner, M. I (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition, 9*, 288-307.
- Feuerstein, R., Rand, Y., Hoffman, M. B. & Miller, R. (1980). *Instrumental Enrichment: An Intervention Program for Cognitive Modifiability*. Glenview, Illinois: Scott, Foresman and Company.
- Flanagan, D. P., McGrew, K. S. & Ortiz, S. O. (1999). *The Wechsler Intelligence Scales and Gf-Gc Theory*. Needham Heights, MA: Allyn and Bacon.
- Flanagan, D. P. & Ortiz, S. O. (2001). *Essentials of cross-battery assessment*. New York: John Wiley & Sons, Inc.

- Ferrioli, S. H. T., Linhares, M. B. M., Loureiro, S. R. & Marturano, E. M. (2001). Indicadores de potencial de aprendizagem obtidos através da avaliação assistida. *Psicologia: Reflexão e Crítica*, 14(1), 35-43.
- Gazzaniga, M. S., Ivry, R. B. & Mangun, G. R. (2002). *Cognitive Neuroscience: The Biology of the Mind*. New York, NY: Norton & Company.
- Gera, A. A. S. & Linhares, M. B. M. (1998). Estratégias de perguntas de busca de informações na resolução de problemas em criança com e sem queixa de dificuldade de aprendizagem (Resumo). Em: Sociedade Brasileira de Psicologia (Org.), *Resumo de Comunicações da XXVIII Reunião Anual de Psicologia* (p. 126). Ribeirão Preto: SBP.
- Goel, V. (2005). Cognitive neuroscience of deductive reasoning. In: K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 475-492). New York: Cambridge.
- Grigorenko, E. L. & Sternberg, R. J. (1998). Dynamic Testing. *Psychological Bulletin*, 124, (1), 75-111.
- Hambleton, H. K. & Swaminathan, H. (1985). *Item response theory: principles and applications*. Boston: Kluwer.
- Haywood, C., Brooks, P. & Burns, S. (1992). *Bright start: cognitive curriculum for young children*. Watertown-MA: Charlesbridge Publishing



- Haywood, H. C. & Tzuriel, D. (2002). Applications and challenges in dynamic assessment. *Peabody Journal of Education*, 77(2), 40-63.
- Holyoak, K. J & Morrison, R. G. (2005). Thinking and reasoning: a reader's guide. In: K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp.1-9). New York: Cambridge.
- Horn, J. L. & Cattell, R. B. (1966) Refinement and test of the theory of fluid and crystalized general intelligences. *Journal of Educational Psychology*, 57(5), 253-270.
- Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, 75(3), 242-259.
- Horn, J. L. & Noll, J. (1997). Human cognitive capabilities: Gf-Gc Theory. In: D. P. Flanagan, J. L. Genshaft & P. L. Harrison (Eds.). *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp.53-91). New York: The Guilford Publications.
- Huizinga, M, Dolan, C. V. & Molen, M. W. Van der. (2006). Age-related in executive function: developmental trends and a latent variable analysis. *Neuropsychologia*, 44, 2017-2036.
- Hunt, E. (1999). Intelligence and human resources: past, present and future. In: P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual*

*differences: process, trait and content determinants*. Washington, DC: American Psychological Association.

International Test Commission (2000). *ITC Guidelines on adapting tests*. International Test Commission. Disponível On-line em: <http://www.intestcom.org>.

Johnson-Laird, P. N. (1983). *Mental Models*. Cambridge: Harvard University Press.

Johnson-Laird, P. N. (2005). Mental models and thought. In: K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp.185-208). New York: Cambridge.

Kyllonen, P. C., & Cristal, R. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence*, 14, 389-434.

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

Klauer, J. K. & Phye, G. D. (1995). *Cognitive training for children: a developmental program of inductive reasoning and problem solving*. Toronto: Hogrefe & Huber Publishers.

Klauer, J. K., Willmes, K. & Phye, G. D. (2002). Inducing inductive reasoning: does it transfer to fluid intelligence? *Contemporary Educational Psychology*, 27, 1-25.

- Knight, R. T., & Grabowecky, M. (1995). Escape from linear time: Prefrontal cortex and conscious experience. In Gazzaniga, M. S. (Ed.), *The cognitive neurosciences* (pp. 1357-1371). Cambridge, MA: MIT Press.
- Lemos, G., Almeida, J. S. & Guisande, M. A. (2006). Bateria de Provas de Raciocínio: suas versões, validação e normalização. Em Machado, C. Almeida, L. Guisande, M. A., Gonçalves, M. & Ramalho, V. (Orgs.), *Actas da XI Conferência Internacional Avaliação Psicológica: Formas e Contextos* (pp. 73-80). Braga: Psiquilibrios.
- Leung, H. C., Skudlarski, P., Gatenby, J. C., Peterson, B. S. & Gore, J. C. (2000). An event-related functional MRI study of the Stroop color word interference task. *Cerebral Cortex* 10, 552-560.
- Lezak, M.D. (1995). *Neuropsychological assessment*. New York, NY: Oxford Univ Press.
- Linhares, M. B. M. (1995). Avaliação Assistida: Fundamentos, Definição, Características e Implicações para a Avaliação Psicológica. *Psicologia: Teoria e Pesquisa*, 11 (1) 23-31.
- Linhares, M. B. M. (1996). Avaliação assistida de crianças com queixa de dificuldade de aprendizagem. *Temas em Psicologia*, 4(1), 17-32.
- Linhares, M. B. M., Escolano, A. C. M & Enumo, S. R. F. (2006). *Avaliação assistida: fundamentos, procedimentos e aplicabilidade*. São Paulo: Casa do Psicólogo.

- Linhares, M. B. M., Santa Maria, M. R., Escolano, A. C. M., Gera, A. A. S. (1998). Avaliação assistida: uma abordagem promissora na avaliação cognitiva de crianças. *Temas em Psicologia*, 6 (3), 231-254.
- Malloy-Diniz, L. F., Sedo, M., Fuentes, D. & Leite, W. B. (2008). Neuropsicologia das funções executivas. Em D. Fuentes, L. F. Malloy-Diniz, C. H. P. Camargo, R. M. Consenza e colaboradores. *Neuropsicologia: Teoria e Prática* (pp. 187-206). Porto Alegre: Artes Médicas.
- McGrew, K. S. (1997). Analysis of the Major Intelligence Batteries According to a Proposed Comprehensive Gf-Gc Framework. Em: D. P. Flanagan, J. L. Gensheft, P. L. Harrison (Eds). *Contemporary Intellectual Assessment Theories, Tests, and Issues*. New York, London: The Guilford Press.
- McGrew, K.S. (2007) *Cattell-Horn-Carroll (CHC) Definition Project*. Disponível em 01/04/2007, em <http://www.iapsych.com/chcdef.htm>
- McGrew, K. S. & Flanagan, D. P. (1998). *The Intelligence Test Desk Reference (ITDR): Gf-Gc Cross-battery Assessment*. Needham Heights: Allyn & Bacon.
- Menezes, A. (2008). *Evidências de validade de instrumentos para avaliar funções executivas em alunos de 5ª a 8ª série*. Dissertação de Mestrado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba.

- Messick, S. (1989). Meaning and values in test validation: the science and ethics of assessment. *Educational Researcher*, 18, 2, 5-11.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H. & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cognitive Psychology*, 41, 49-100.
- Moore-Brown, B., Huerta, M., Uranga-Hernandez, Y. & Peña, E. D. (2006). Using dynamic assessment to evaluate children with suspected learning disabilities. *Intervention in School and Clinic*, 41(4), 209-217.
- Munhoz, A. M. H (2004). *Uma análise multidimensional da relação entre a inteligência e o desempenho acadêmico em universitários ingressantes*. Tese (doutorado). Universidade Estadual de Campinas – Faculdade de Educação.
- Muniz, J. (2004). La validación de los tests. *Metodología de las Ciencias del Comportamiento*, 5 (2), 121-141.
- Noronha, A. P. P. & Vendramini, C. M. M. (2003). Parâmetros psicométricos: estudo comparativo entre testes de inteligência e de personalidade. *Psicologia: Reflexão e Crítica*, 16 (1), 177-182.
- Posner, M. I. & DiGirolamo, G. J (2000). Cognitive neuroscience: Origins and promise. *Psychological Bulletin.*, 126(6), 873-889.

- Primi, R. (1995) *Inteligência, processamento de informação e teoria da gestalt: um estudo experimental*. Dissertação de Mestrado. Pontifícia Universidade Católica de Campinas.
- Primi, R. (2002a) Complexity of Geometric Inductive Reasoning Tasks: Contribution to the Understanding of the Fluid Intelligence. *Intelligence*, 30 (1), 41-70.
- Primi, R. (2002b). Relatório Parcial do projeto: *Avaliação Componencial Informatizada da Inteligência Fluida*. Laboratório de Avaliação Psicológica e Educacional – LabAPE. Universidade São Francisco.
- Primi, R. (2003). Relatório do projeto: *Avaliação Componencial Informatizada da Inteligência Fluida*. Laboratório de Avaliação Psicológica e Educacional – LabAPE. Universidade São Francisco.
- Primi,R. (2005). *Avaliação componencial informatizada da inteligência fluida*. Relatório final de bolsa Jovem Pesquisador. Fundação de Amparo à Pesquisa do Estado de São Paulo.
- Primi, R., & Almeida, L. S. (2000b). Estudo de Validação da Bateria de Provas de Raciocínio (BPR-5). *Psicologia: Teoria e Pesquisa*, 16 (2), 165-173.
- Primi, R., Santos, A. A. A. & Vendramini, C. M. M. (2002). Habilidades básicas e desempenho acadêmico em universitários ingressantes. *Estudos de Psicologia*, 7(1), 47-55.

- Primi, R., Sternberg, R. J. & Grigorenko, E. L. (1997) *Experimental Software for Dynamic Assessment of Fluid Reasoning* [Computer Software]. New Haven: Authors.
- Raichle, M. E., Fiez, J. A., Vidden, T. O., MacLeod, A. K., Pardo, J. V., Fox, P. T. & Petersen, S. E. (1994). Practice-related changes in human brain function anatomy during nonmotor learning. *Cerebral Cortex* 4, 8-26.
- Raven, J. C. (1987) *Matrizes Progressivas Coloridas*. São Paulo: Casa do Psicólogo
- Richardson, K. (1991). *Understanding intelligence*. Milton Keynes: Open University Press.
- Salovey, P. & Mayer, J.D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, 9, 185-221.
- Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology*, 19(4), 532-545.
- Santos, M. A. dos & Primi, R. (2005). Desenvolvimento de um teste informatizado para avaliação do raciocínio, da memória e da velocidade do processamento. *Estudos de Psicologia*, 22 (3), 241-254.
- Sisto, F. F. (2006). *Teste de Raciocínio Inferencial: Manual Técnico*. São Paulo: Vetor Editora.
- Spearman, C. (1927). *Las Habilidades del Hombre Su Natureza y Medición*. Buenos Aires: Editorial Paidós.

- Stankov, L. (2000). Structural Extensions of a Hierarchical View on Human Cognitive Abilities. *Learning and Individual Differences*, 12, 35-51.
- Sternberg, J.R. (1992a). As capacidades intelectuais humanas: uma abordagem em processamento de informações. Porto Alegre: Artes Médicas.
- Sternberg, R. J. (1992b). A capacidade intelectual geral. Em R. J. Sternberg. *As capacidades intelectuais humanas uma abordagem processamento de informações* (pp. 17-42). Porto Alegre: Artes Médicas.
- Sternberg, R.J. (2000). *Psicologia cognitiva*. Trad. Maria Regina Borges Osório. Porto Alegre: Artes Médicas.
- Sternberg, R. J. (1977). A component process in analogical reasoning. *Psychological Review*, 84, 4, 353-378.
- Sternberg, R. J. (1979). The nature of mental abilities. *American Psychologist*, 34, 3, 214-230.
- Sternberg, R. J. (1980). Factor theories of intelligence are all right almost. *Educational Researcher*, 9, 6-13.
- Sternberg, R. J. (1981). The evolution of theories of intelligence. *Intelligence*, 5, 209-230.
- Sternberg, R. J. (1983). Components of human intelligence. *Cognition*, 15, 1-48.



- Sternberg, R. J. (1986). Toward a unified theory of human reasoning. *Intelligence, 10*, 281-314.
- Sternberg, J.R & Grigorenko, E.L. (2002). *Dynamic testing: the nature and measurement of learning potential*. Cambridge: University Press.
- Tavares, M. (2003). Validade clínica. *Pisco-USF, 8 (2)*, 125-136.
- Tortella, G., Dias, N. M., Capovilla, A. G. S., Trevisan, B. T., Hipólito, R. & Berberian, A. A. (2008). Teste de Geração Semântica: evidências de validade para um instrumento neuropsicológico para avaliação do controle inibitório. Em A. N. Noronha, C. Machado, L. S. Almeida, M. Gonçalves, S. Martins & V. Ramalho (Orgs.), *Actas da XIII Conferência Internacional de Avaliação Psicológica: Formas e Contextos*. Braga, Portugal: Psiquilíbrios.
- Thurstone, L. L. (1934). The Vectors of Mind. *Psychological Review, 41*, 1-32.
- Tzuriel, D. (2001). *Dynamic assessment of young children*. New York: Kluwer Academic/Plenum Publisher.
- Tzuriel, D. & Feuerstein, R. (1992). Dynamic assessment for prescriptive teaching. Em C. Haywood & D. Tzuriel (Eds.), *Interactive Assessment*. New York: Springer.
- Tzuriel, D. & Klein, P. S (1990). *The Children's Analogical Thinking Modifiability Test: Instrucion Manual*. Ramat-Gan: School of Education Bar Ilan University.

- Tzuriel, D. & Shamir, A. (2002). The effect of mediation in computer assisted assessment. *Journal of Computer Assisted Learning*, 18, 21-32.
- Urbina, S. (2006). *Fundamentos da Testagem Psicológica*. Porto Alegre: Artes Médicas.
- Vernon, P. E. (1961) *The structure of human abilities*. (2<sup>nd</sup> Ed.). London: Methuen.
- Vygotsky, L. S. (1991). *A formação social da mente: o desenvolvimento dos processos psicológicos superiores* (4<sup>a</sup> Ed., J. C. Neto, L. S. M. Barreto & S. C. Afeche, Trans.). São Paulo, Martins Fontes. Trabalho original publicado em 1970.
- Wilhelm, O. (2005). Measuring reasoning ability. In: O. Wilhelm & R. W. Engle (Ed.), *Handbook of Understanding and Measuring Intelligence* (pp.373-407). London, Sage Publications.
- Winfred, A. Jr. & Woehr, D. J. (1993). A confirmatory factor analytic study examining the dimensionality of the Raven's advanced progressive matrices. *Educational and Psychological Measurement*, 53, 471-478.
- Wright, B. D. & Stone, M. H. (2004). *Making measure*. Chicago: The Phaneron Press.

## **5- ANEXOS**

## TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

### Projeto de pesquisa: Desenvolvimento e validação de instrumentos de avaliação neuropsicológica

Eu, \_\_\_\_\_, RG \_\_\_\_\_, abaixo assinado, responsável pelo menor \_\_\_\_\_, dou meu consentimento livre e esclarecido para autorizar a participação do referido menor no projeto de pesquisa supracitado, sob a responsabilidade da pesquisadora Dra. Alessandra G. S. Capovilla, docente do Programa de Pós-Graduação *Stricto Sensu* em Psicologia da Universidade São Francisco. Esta pesquisa está sendo financiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), processo número 03/11767-9.

O objetivo da pesquisa é desenvolver instrumentos para a avaliação neuropsicológica, verificando qual é o desempenho médio de alunos do ensino fundamental sem distúrbios neurológicos. Serão aplicados ao aluno instrumentos psicológicos para avaliar diferentes habilidades, tais como linguagem, atenção, memória, processamento visual e espacial, funções motoras, raciocínio, funções executivas, formação de conceitos e habilidades intelectuais. Todos os instrumentos serão aplicados na própria escola, durante o período escolar regular, com a autorização prévia da diretora da escola. Esta aplicação não oferece qualquer risco conhecido para os alunos, apesar de poder causar fadiga. Somente participarão da pesquisa os alunos que voluntariamente concordarem em responder aos instrumentos.

Os dados pessoais do menor serão mantidos em sigilo e os resultados gerais obtidos através da pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, expostos acima, incluída sua publicação na literatura científica especializada.

Obtive as informações necessárias para poder decidir conscientemente sobre a participação do menor acima citado na referida pesquisa, sabendo que tanto eu quanto o menor estamos livres para interrompê-la a qualquer momento.

Poderei entrar em contato com a pesquisadora responsável pelo estudo, Dra. Alessandra G. S. Capovilla, sempre que julgar necessário pelo telefone (11) 4534-8046, ou com o Comitê de Ética em Pesquisa da Universidade São Francisco pelo telefone (11) 4451-9518, com o Sr. Edson Azevedo.

Este Termo de Consentimento é feito em duas vias, sendo que uma permanecerá em meu poder e a outra, com o pesquisador responsável.

**Prof<sup>a</sup> Dra. Alessandra G. S. Capovilla**

Universidade São Francisco, Itatiba-SP

Telefone: (11) 4534-8046

São Paulo, \_\_\_\_ de \_\_\_\_\_ de 2007.

Nome do menor: \_\_\_\_\_

Data de nascimento do menor: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

\_\_\_\_\_  
Assinatura do responsável

1ª VIA

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO  
**Projeto de pesquisa: Desenvolvimento e validação de instrumentos de avaliação  
neuropsicológica**

Eu, \_\_\_\_\_,  
RG \_\_\_\_\_, abaixo assinado, responsável pelo  
menor \_\_\_\_\_, dou meu  
consentimento livre e esclarecido para autorizar a participação do referido menor no projeto de  
pesquisa supracitado, sob a responsabilidade da pesquisadora Dra. Alessandra G. S. Capovilla,  
docente do Programa de Pós-Graduação *Stricto Sensu* em Psicologia da Universidade São  
Francisco. Esta pesquisa está sendo financiada pela Fundação de Amparo à Pesquisa do Estado de  
São Paulo (FAPESP), processo número 03/11767-9.

O objetivo da pesquisa é desenvolver instrumentos para a avaliação neuropsicológica,  
verificando qual é o desempenho médio de alunos do ensino fundamental sem distúrbios  
neurológicos. Serão aplicados ao aluno instrumentos psicológicos para avaliar diferentes  
habilidades, tais como linguagem, atenção, memória, processamento visual e espacial, funções  
motoras, raciocínio, funções executivas, formação de conceitos e habilidades intelectuais. Todos os  
instrumentos serão aplicados na própria escola, durante o período escolar regular, com a autorização  
prévia da diretora da escola. Esta aplicação não oferece qualquer risco conhecido para os alunos,  
apesar de poder causar fadiga. Somente participarão da pesquisa os alunos que voluntariamente  
concordarem em responder aos instrumentos.

Os dados pessoais do menor serão mantidos em sigilo e os resultados gerais obtidos através  
da pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, expostos acima, incluída  
sua publicação na literatura científica especializada.

Obtive as informações necessárias para poder decidir conscientemente sobre a participação  
do menor acima citado na referida pesquisa, sabendo que tanto eu quanto o menor estamos livres  
para interrompê-la a qualquer momento.

Poderei entrar em contato com a pesquisadora responsável pelo estudo, Dra. Alessandra G. S.  
Capovilla, sempre que julgar necessário pelo telefone (11) 4534-8046, ou com o Comitê de Ética  
em Pesquisa da Universidade São Francisco pelo telefone (11) 4451-9518, com o Sr. Edson  
Azevedo.

Este Termo de Consentimento é feito em duas vias, sendo que uma permanecerá em meu  
poder e a outra, com o pesquisador responsável.

**Profª Dra. Alessandra G. S. Capovilla**  
Universidade São Francisco, Itatiba-SP  
Telefone: (11) 4534-8046

São Paulo, \_\_\_\_ de \_\_\_\_\_ de 2007.

Nome do menor: \_\_\_\_\_

Data de nascimento do menor: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

\_\_\_\_\_  
Assinatura do responsável

2ª VIA

**TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO AREA DE CIÊNCIAS HUMANAS E EXATAS CEP – CHE/USF (primeira via)**

**PROJETO DE PESQUISA: Construção de um teste dinâmico e informatizado de raciocínio indutivo para crianças de 7 a 11 anos.**

Eu, \_\_\_\_\_, (nome, idade, RG),

abaixo assinado, responsável pelo menor \_\_\_\_\_ (nome da criança) dou meu consentimento livre e esclarecido para autorizar sua participação como voluntário do projeto de pesquisa supra-citado, sob a responsabilidade da pesquisadora Monalisa Muniz Nascimento, aluna do doutorado do Programa de Pós-Graduação Stricto Sensu em Psicologia da Universidade São Francisco, e da Prof<sup>ª</sup>. Dr<sup>ª</sup>. Alessandra Gotuzo Seabra Capovilla. Esta pesquisa está sendo financiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP.

Assinando este Termo de Consentimento estou ciente de que:

**1** – O objetivo da pesquisa é avaliar o potencial de aprendizagem da criança em relação ao raciocínio indutivo, ou seja, mensurar o quanto a criança consegue aprender a partir de instruções dadas em um treinamento.

**2** – Serão aplicados aos participantes instrumentos psicológicos para avaliar habilidade do raciocínio indutivo, bem como o potencial de aprendizagem deste raciocínio, também serão aplicados dois testes de funções executivas para avaliar controle inibitório e atenção seletiva. Todos os testes serão aplicados na própria escola, durante o período regular das aulas, em horário previamente agendado com a escola. Apenas um dos testes é de forma coletiva, os outros serão individuais.

**3** – Esta aplicação não oferece qualquer risco conhecido para os participantes, apesar de poder causar fadiga. Somente participarão da pesquisa os participantes que voluntariamente concordarem em responder aos instrumentos;

**4** – Obtive todas as informações necessárias para poder decidir conscientemente sobre a participação do menor acima citado na referida pesquisa, sabendo que tanto eu quanto o menor estamos livres para interrompê-la a qualquer momento.

**5** – Os dados pessoais do menor serão mantidos em sigilo e os resultados gerais obtidos através da pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, expostos acima, incluída sua publicação na literatura científica especializada;

**6** – Poderei contatar o Comitê de Ética em Pesquisa da Universidade São Francisco para apresentar recursos ou reclamações em relação à pesquisa através do telefone: 11 - 4534-8023;

**7** – Poderei entrar em contato com a responsável pelo estudo, Monalisa Muniz Nascimento, sempre que julgar necessário pelo telefone (11) 4534-8040, ou com o Comitê de Ética em Universidade São Francisco pelo telefone (11) 4542-8040;

**8** – Este Termo de Consentimento é feito em duas vias, sendo que uma permanecerá em meu poder e outra com o pesquisador responsável.

Monalisa Muniz Nascimento

Universidade São Francisco, Itatiba-SP Telefone: (11) 4534-8040 Itatiba, \_\_\_\_\_ de \_\_\_\_\_ de 2007.

Nome do menor: \_\_\_\_\_

Data de nascimento do menor \_\_\_/\_\_\_/\_\_\_ Assinatura do Voluntário: \_\_\_\_\_

**TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO AREA DE CIÊNCIAS HUMANAS E EXATAS CEP – CHE/USF (segunda via)**

**PROJETO DE PESQUISA: Construção de um teste dinâmico e informatizado de raciocínio indutivo para crianças de 7 a 11 anos.**

Eu, \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_ (nome, idade, RG),

abaixo assinado, responsável pelo menor \_\_\_\_\_ (nome da criança) dou meu consentimento livre e esclarecido para autorizar sua participação como voluntário do projeto de pesquisa supra-citado, sob a responsabilidade da pesquisadora Monalisa Muniz Nascimento, aluna do doutorado do Programa de Pós-Graduação Stricto Sensu em Psicologia da Universidade São Francisco, e da Prof<sup>a</sup>. Dr<sup>a</sup>. Alessandra Gotuzo Seabra Capovilla. Esta pesquisa está sendo financiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP.

Assinando este Termo de Consentimento estou ciente de que:

**1** – O objetivo da pesquisa é avaliar o potencial de aprendizagem da criança em relação ao raciocínio indutivo, ou seja, mensurar o quanto a criança consegue aprender a partir de instruções dadas em um treinamento.

**2** – Serão aplicados aos participantes instrumentos psicológicos para avaliar habilidade do raciocínio indutivo, bem como o potencial de aprendizagem deste raciocínio, também serão aplicados dois testes de funções executivas para avaliar controle inibitório e atenção seletiva. Todos os testes serão aplicados na própria escola, durante o período regular das aulas, em horário previamente agendado com a escola. Apenas um dos testes é de forma coletiva, os outros serão individuais.

**3** – Esta aplicação não oferece qualquer risco conhecido para os participantes, apesar de poder causar fadiga. Somente participarão da pesquisa os participantes que voluntariamente concordarem em responder aos instrumentos;

**4** – Obtive todas as informações necessárias para poder decidir conscientemente sobre a participação do menor acima citado na referida pesquisa, sabendo que tanto eu quanto o menor estamos livres para interrompê-la a qualquer momento.

**5** – Os dados pessoais do menor serão mantidos em sigilo e os resultados gerais obtidos através da pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, expostos acima, incluída sua publicação na literatura científica especializada;

**6** – Poderei contatar o Comitê de Ética em Pesquisa da Universidade São Francisco para apresentar recursos ou reclamações em relação à pesquisa através do telefone: 11 - 4534-8023;

**7** – Poderei entrar em contato com a responsável pelo estudo, Monalisa Muniz Nascimento, sempre que julgar necessário pelo telefone (11) 4534-8040, ou com o Comitê de Ética em Universidade São Francisco pelo telefone (11) 4542-8040;

**8** – Este Termo de Consentimento é feito em duas vias, sendo que uma permanecerá em meu poder e outra com o pesquisador responsável.

Monalisa Muniz Nascimento

Universidade São Francisco, Itatiba-SP Telefone: (11) 4534-8040 Itatiba, \_\_\_\_\_ de \_\_\_\_\_ de 2007.

Nome do menor: \_\_\_\_\_

Data de nascimento do menor \_\_\_/\_\_\_/\_\_\_ Assinatura do Voluntário: \_\_\_\_\_

