2011

Sensory Discrimination Testing with Children

Karen Melissa Garcia
Louisiana State University and Agricultural and Mechanical College

SENSORY DISCRIMINATION TESTING WITH CHILDREN

A Dissertation

Presented to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillments of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Food Science

by
Karen Melissa Garcia
B.S., Chemical Engineering, Louisiana State University, 2004
M.S., Food Science, Louisiana State University, 2006
May 2011

DEDICATION

I dedicate this dissertation to my parents, Omer and Marta. You have provided me with everything I have, made me the person I am, and helped me achieve all I have today. I am truly blessed to have such supportive and loving parents. Thank you for everything you have done for me!

ACKNOWLEDGEMENTS

First of all I would like to thank God for giving me the strength to pursue my PhD degree. I thank my major professor, Dr. Witoon Prinyawiwatkul, for his invaluable guidance and support. It was a great pleasure having him as my advisor. I also thank Dr. John Caprio, Dr. Subramaniam Sathivel, Dr. Jing Wang, and Dr. Zhimin Xu for serving on my dissertation committee.

I want to express my gratitude to Dr. William Richardson, LSU AgCenter Chancellor, for the financial support. I thank Dr. David Boethel, Dr. David Morrison, and Dr. Michael Kennan for their support during the initial phase of my research. I am further indebted with Dr. Daniel Ennis for his time and advice. I would like to thank Carl Fink for his support and assistance with editing.

I am grateful to my family, especially my parents for their emotional and financial support. I appreciate Ashley Bond Gutierrez, Kennet Carabante, Phantipha Charoenthaikij, Luis Espinoza, Miguel Gutierrez, Wannita Jirangrat, Piyaporn Muangprasit, Chanchira Primparian, Sriwiang Tipkanon, Damir Torrico, Pamarin Waimaleongora-Ek, and Wisdom Wardy. I would like to thank my collaborators, Dr. Andres Herrara Corredor and Dr. Penkwan Chompreeda. I also thank Dr. Judith Williams, Prof. Marta Garcia, and Mrs. Onya Carter for their assistance at the different educational facilities.

Finally I thank all the children who participated in this research study; it would not have been possible without them.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF TABLES ix

LIST OF FIGURES xiv

ABSTRACT xvi

CHAPTER 1. INTRODUCTION xvii

CHAPTER 2. LITERATURE REVIEW 3
2.1 Introduction 3
2.2 Child Development and Its Effects on Sensory Test Performance 5
2.2.1 Physiological Development 5
2.2.2 Cognitive Development 5
2.2.2.1 Verbal Skills 6
2.2.2.2 Attention Span 7
2.2.2.3 Comprehension 7
2.2.3 Perception and Development of Food Preferences 8
2.2.4 Differences between Children and Adults 9
2.3 Sensory Testing with Children 10
2.3.1 Affective Methods 11
2.3.2 Discrimination Methods 11
2.4 Special Topics in Discrimination Testing with Children 14
2.4.1 Discrimination Testing 14
2.4.1.1 2-Alternative Forced Choice (2-AFC) Method 15
2.4.1.2 The Triangle Method 16
2.4.1.3 3-Alternative Forced Choice Method (3-AFC) 16
2.4.1.4 Method of Tetrads 17
2.4.2 The Paradox of Discriminatory Non-discriminators 17
2.4.2.1 Thurstonian Modeling 18
2.4.3 The Variants of the Methods of Tetrads 26
2.5 References 29

CHAPTER 3. THE PARADOX OF DISCRIMINATORY NON-DISCRIMINATORS WITH CHILDREN AS SUBJECTS USING EASILY DISCRIMINABLE STIMULI 40
3.1 Introduction 40
3.2 Materials and Methods 43

3.2.1 Subjects 43
3.2.2 Stimuli 44
3.2.2.1 Honduras, Mexico, USA 44
3.2.2.2 Thailand 44
3.2.3 Testing Conditions 45
3.2.4 Experimental Design 45
3.2.5 Testing Protocol 47
3.2.5.1 Experimental Session Overview 49
3.2.6 Data Analysis 50
3.3 Results 51
3.3.1 Honduras 51
3.3.1.1 Overall Performance 51
3.3.1.2 Individual Performance 55
3.3.2 Mexico 57
3.3.2.1 Overall Performance 58
3.3.2.2 Individual Performance 59
3.3.3 Thailand 60
3.3.3.1 Overall Performance 60
3.3.3.2 Individual Performance 63
3.3.4 USA 64
3.3.4.1 Overall Performance 64
3.3.4.2 Individual Performance 66
3.4 Discussion 67
3.4.1 Overall Performance 67
3.4.2 Individual Performance 69
3.5 Conclusion 72
3.6 Future Work 72
3.7 References 72

CHAPTER 4. THE PARADOX OF DISCRIMINATORY NON-DISCRIMINATORS WITH CHILDREN AS SUBJECTS USING HARDLY DISCRIMINABLE STIMULI 75
4.1 Introduction 75
4.2 Materials and Methods 77
4.2.1 Subjects 77
4.2.2 Stimuli 77
4.2.2.1 Honduras 78
4.2.2.1.1 Study I: Hardly Distinguishable Stimuli 78
4.2.2.1.2 Study II: Hardly Distinguishable Stimuli 78
4.2.2.2 Mexico 79
4.2.2.3 Thailand 79
4.2.3 Testing Conditions 80
4.2.4 Experimental Design 81
4.2.5 Testing Protocol 83
4.2.5.1 Experimental Session Overview 84

4.2.6 Data Analysis 86
4.3 Results 86
4.3.1 Honduras 86
4.3.1.1 Study I: Hardly Distinguishable Stimuli 86
4.3.1.1.1 Overall Performance 88
4.3.1.1.2 Individual Performance 91
4.3.1.2 Study II: Hardly Distinguishable Stimuli 93
4.3.2 Mexico 96
4.3.2.1 Overall Performance 96
4.3.2.2 Individual Performance 99
4.3.3 Thailand 100
4.3.3.1 Overall Performance 100
4.3.3.2 Individual Performance 103
4.3.4 USA 104
4.3.4.1 Overall Performance 104
4.3.4.2 Individual Performance 107
4.4 Discussion 108
4.4.1 Overall Performance 108
4.4.2 Individual Performance 109
4.5 Conclusion 112
4.6 Future Work 112
4.7 References 113

CHAPTER 5. CONFIRMING THE PARADOX OF DISCRIMINATORY NON-DISCRIMINATORS AND THURSTONIAN PREDICTIONS WITH CHILDREN AS SUBJECTS 115
5.1 Introduction 115
5.2 Materials and Methods 117
5.2.1 Subjects 117
5.2.2 Stimuli 118
5.2.3 Experimental Design 119
5.2.4 Testing Protocol 119
5.2.4.1 Experimental Session Overview 120
5.2.5 Data Analysis 122
5.3 Results and Discussion 122
5.3.1 Overall Performance 122
5.3.2 Individual Performance 123
5.4: Conclusion 127
5.5 Future Work 127
5.6 References 127

CHAPTER 6: CORROBORATING THURSTONIAN PREDICTIONS FOR THE VARIANTS OF THE TETRAD METHOD WITH CHILDREN AS SUBJECTS 130
6.1 Introduction 130

6.2 Materials and Methods ... 131
6.2.1 Subjects ... 131
6.2.2 Stimuli ... 132
6.2.3 Testing Procedure ... 132
6.2.4 Data Analysis ... 133
6.3 Results and Discussion ... 134
6.3.2 Overall Performance ... 134
6.3.2 Individual Performance ... 136
6.4 Conclusion ... 137
6.5 References ... 137

CHAPTER 7: DISCRIMINATION TESTING WITH CHILDREN: COMPARISON OF PERFORMANCE IN 2-AFC, TRIADIC AND TETRADS METHODS ... 139
7.1 Introduction ... 139
7.2 Materials and Methods ... 140
7.2.1 Subjects ... 140
7.2.2 Stimuli ... 141
7.2.3 Experimental Design ... 141
7.2.4 Testing Protocol ... 142
7.2.4.1 Experimental Session Overview ... 143
7.5.5 Data Analysis ... 146
7.3 Results and Discussion ... 146
7.3.1 Comparison of Unspecified Methods: Unspecified Method of Tetrads vs. Triangle. 146
7.3.2 Comparison of Specified Methods ... 149
7.3.2.1 Specified Method of Tetrads vs. 3-AFC ... 149
7.3.2.2 Specified Method of Tetrads vs. 2-AFC ... 152
7.3.2.3 3-AFC vs. 2-AFC ... 153
7.4 Conclusion ... 154
7.5 References ... 154

CHAPTER 8: SUMMARY AND CONCLUSIONS ... 157

APPENDIX A: PARENTAL CONSENT FORM ... 160

APPENDIX B: CHILD ASSENT FORM ... 165

APPENDIX C: LIST OF PRODUCT INGREDIENTS ... 166

APPENDIX D: SURVEYS ... 167
D.1 Triangle Test ... 167
D.2 3-AFC Test ... 168
D.3 2-AFC Test ... 169
D.4 Method of Tetrads ... 170

APPENDIX E: CHILDREN'S PERFORMANCE WHEN SWALLOWING VS. EXPECTORATING THE STIMULI 171

APPENDIX F: DETERMINING d' 174

APPENDIX G: VARIANCE OF d' 180

APPENDIX H: d' TEST STATISTIC 183

APPENDIX I: VARIANCE CORRECTION FOR REPLICATED TESTING 184

VITA 185

LIST OF TABLES

Table 3.1: Distribution of subjects per grade for each country 43

Table 3.2: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for Honduran children (first to sixth grade, N=339) 52

Table 3.3: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for Honduran children (first to sixth grade, N=339) 53

Table 3.4: Pooled proportions of correct responses, pooled *d′* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the Honduran population (N=678[a]) 53

Table 3.5: Proportions of correct responses, *d′* values and significance tests for the triangle and 3-AFC methods for the individual grades for the Honduran population 56

Table 3.6: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for Mexican children (first to sixth grade, N=172) 57

Table 3.7: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for Mexican children (first to sixth grade, N = 172) 58

Table 3.8: Pooled proportions of correct responses, pooled *d′* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the Mexican population (N=344[a]) 59

Table 3.9: Proportions of correct responses, *d′* values and significance tests for the triangle and 3-AFC methods for the individual grades for the Mexican population 59

Table 3.10: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for Thai children (first to sixth grade, N = 360) 61

Table 3.11: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for Thai children (first to sixth grade, N=360) 62

Table 3.12: Pooled proportion of correct responses, pooled *d′* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the Thai population (N= 720[a]) 62

Table 3.13: Proportions of correct responses, *d′* values, test statistics and p-values for the triangle and 3-AFC methods for the individual grades for the Thai population........................... 63

Table 3.14: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for US children (first to sixth grade, N = 86) 64

Table 3.15: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for US children (first to sixth grade, N = 86) 65

Table 3.16: Pooled proportions of correct responses, pooled *d′* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the USA population (N= 172[a]) 65

Table 3.17: Proportions of correct responses, *d′* values, test statistics and p-values for the triangle and 3-AFC methods for the individual grades for the USA population 66

Table 3.18: Summary of pooled proportions of correct responses, *d′* values and significance tests for the triangle and 3-AFC methods overall and for the individual countries 67

Table 3.19: Summary of P_c and *d′* values for the individual grades pooled from the Honduran, Mexican, Thai and US population. 70

Table 4.1: Distribution of subjects per grade for each country 77

Table 4.2: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for Honduran children (first to sixth grade) 87

Table 4.3: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for Honduran children (first to sixth grade) 88

Table 4.4: Pooled proportions of correct responses, pooled *d′* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the Honduran population (N=612[a]) 88

Table 4.5: Proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods for the individual grades for the Honduran population 92

Table 4.6: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for Honduran children (first to second grade) 94

Table 4.7: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC protocol for Honduran children (first to sixth grade) 95

Table 4.8: Proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods for the individual grades for the Honduran population 95

Table 4.9: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for Mexican children (first to sixth grade) ... 97

Table 4.10: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for Mexican children (first to sixth grade) ... 98

Table 4.12: Proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods for the individual grades for the Mexican population ... 99

Table 4.13: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle test for Thai children (first to sixth grade) ... 101

Table 4.14: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for Thai children (first to sixth grade) ... 102

Table 4.15: Pooled proportions of correct responses, pooled *d′* values, test statistic (N= 720[a]) and p-value for the triangle and 3-AFC methods for all six grades of the Thai population ... 102

Table 4.16: Proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods for the individual grades of the Thai population ... 103

Table 4.17: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for US children (first to sixth grade) ... 105

Table 4.18: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for US children (first to sixth grade) ... 106

Table 4.19: Pooled proportions of correct responses, pooled *d′* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the USA population (N= 186[a]) ... 106

Table 4.20: Proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods for the individual grades for the US population ... 107

Table 4.21: Summary of the proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods overall and for the different countries ... 108

Table 4.22: Summary of the proportions of correct responses, *d′* values and significance tests for the triangle and 3-AFC methods for the individual grades from the Honduran, Mexican, Thai, and US population ... 111

Table 5.1: Distribution of subjects per grade ... 117

Table 5.2: Pooled proportions of correct responses, pooled *d′* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades (N = 404) 122

Table 5.3: Proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods for the individual grades 123

Table 5.4: Summary of result for the individual populations and combined populations for all studies performed regarding the paradox of discriminatory non-discriminators and Thurstonian predictions for the triangle and 3-AFC methods.. 125

Table 5.5: Summary of result for the individual grades for all populations combined for all studies performed regarding the paradox of discriminatory non-discriminators and Thurstonian predictions for the triangle and 3-AFC methods.. 126

Table 6.1: Pooled proportion of correct responses, pooled *d’* values, test statistics (N=404) and p-values for the variants of the method of tetrads for all six grades.. 135

Table 6.2: Proportion of correct responses, *d′* values and significance tests for the variants of the tetrad method for the individual grades 136

Table 7.1: Distribution of subjects per grade 140

Table 7.2: Pooled proportion of correct responses, pooled *d′* values, test statistic and p-value for the unspecified method of tetrads and the triangle method for all six grades............................ 147

Table 7.3: Proportion of correct responses, *d′* values and significance tests for the unspecified method of tetrads and the triangle method for the individual grades 149

Table 7.4: Pooled proportion of correct responses, pooled *d′* values, test statistic and p-value for the specified method of tetrads and the 3-AFC method for all six grades (N=404).................. 150

Table 7.5: Proportion of correct responses, *d′* values and significance tests for the specified method of tetrads and the 3-AFC method for the individual grades.. 151

Table 7.6: Pooled proportion of correct responses, pooled *d′* values, test statistic and p-value for the specified method of tetrads and the 2-AFC method for all six grades (N=404).................. 152

Table 7.7: Proportion of correct responses, *d′* values and significance tests for the specified method of tetrads and the 2-AFC method for the individual grades.. 152

Table 7.8: Pooled proportion of correct responses, pooled *d′* values, test statistic and p-value for the specified method of tetrads and the 3-AFC method for all six grades (N=404) 153

Table 7.9: Proportion of correct responses, *d′* values and significance tests for the specified method of tetrads and the 3-AFC method for the individual grades 153

Table E.1: Proportions of correct responses, *d′* values and significance test for swallowing vs. expectoration of the samples in a triangle test 173

Table F.1: 2-AFC Method - Probability of a correct response ($x10^4$) as a function of δ (Ennis, 1993) 175

Table F.2: Triangle Method - Probability of a correct response ($x10^4$) as a function of δ (Ennis, 1993) 176

Table F.3: 3-AFC Method - Probability of a correct response ($x10^4$) as a function of δ (Ennis, 1993) 177

Table F.4: Unspecified Method of Tetrads - Probability of a correct response ($x10^4$) as a function of δ (Ennis and others, 1998) 178

Table F.5: Specified Method of Tetrads - Probability of a correct response ($x10^4$) as a function of δ (Ennis and others, 1998) 179

Table G.1: Triangle Method - B value for the estimation of the variance of *d′* (Bi and others, 1997) 181

Table G.2: 3-AFC Method - B value for the estimation of the variance of *d′* (Bi and others, 1997) 182

LIST OF FIGURES

Figure 2.1: Frequency distribution along flavor intensity axis representing variation in flavor of a stimulus on repeated tasting (adapted from O'Mahony and others 1994). 19

Figure 2.2: Frequency distributions representing distinguishable and confusable stimuli (adapted from O'Mahony and others 1994) 20

Figure 2.3: Thurstonian treatment of correct and incorrect triangle and 3-AFC tests when instructions are to identify the odd sample, indicating the paradox of discriminatory non-discriminators (adapted from O'Mahony, 1993) 22

Figure 2.4: Psychometric functions for the triangle and 3-AFC methods (source of data: Ennis, 1993) 24

Figure 2.5: Thurstonian treatment of correct and incorrect unspecified and specified method of tetrads tests (adapted from O'Mahony, 1993) 27

Figure 2.6: Psychometric functions for the specified and unspecified method of tetrads (source of data: Ennis and others 1998) 28

Figure 3.1: Presentation given to children at the beginning of the testing session 46

Figure 3.2: USA child performing the triangle test 48

Figure 3.3: Honduran children performing the 3-AFC test 49

Figure 3.4: Triangle test (a) observed proportions, (b) predicted probabilities by binomial distribution, and (c) predicted probabilities by beta-binomial distribution. 54

Figure 3.5: 3-AFC test (a) observed proportions, (b) predicted probabilities by binomial distribution, and (c) predicted probabilities by beta-binomial distribution. 55

Figure 4.1: Psychometric functions for the Triangle and 3-AFC methods (source of data: Ennis, 1993) 76

Figure 4.2: Presentation given to children at the beginning of the testing session 81

Figure 4.3: USA child performing a distractor test.. 82

Figure 4.4: Thai children performing the triangle test 84

Figure 4.4: Triangle test (a) observed proportions, (b) predicted probabilities by binomial distribution, and (c) predicted probabilities by beta-binomial distribution. 90

Figure 4.5: 3-AFC test (a) observed proportions, (b) predicted probabilities by binomial distribution, and (c) predicted probabilities by beta-binomial distribution. 90

Figure 5.1: Psychometric functions for the triangle and 3-AFC methods (source of data: Ennis, 1993) 116

Figure 5.2: Child filling out the child assent form 118

Figure 6.1: Psychometric functions for the specified and unspecified method of tetrads (source of data: Ennis and others, 1998)... 135

Figure 7.1: Psychometric functions for the unspecified method of tetrads and the triangle method (data from: Ennis 1993, Ennis and others 1998).. 147

Figure 7.2: Psychometric functions for the 2-AFC method, 3-AFC method, and specified method of tetrads (data from: Ennis 1993, Ennis and others 1998)... 151

ABSTRACT

Literature concerning children's performance in sensory discrimination methods is not as extensively published as performance with adults. Therefore, in this dissertation children's performance in discrimination methods was investigated. First, performance in the triangle and 3-AFC methods was explored with children 6 to 12 years old. The paradox of discriminatory non-discriminators states that the proportion of correct responses in the 3-AFC method will be higher than that of the triangle. However, Thurstonian theory predicts that despite the difference in proportions of correct responses the degree of difference between the stimuli will be similar. The paradox and the Thurstonian predictions were challenged for sweetness discrimination with three different sets of stimuli: easily discriminable (100% vs. 60% apple juice), confusable (100% vs. 75% apple juice) and hardly discriminable (carbonated beverages with different sweeteners). For easily and hardly discriminable stimuli the paradox and Thurstonian predictions were not fully confirmed. With confusable stimuli the paradox (3-AFC $P_c = 0.62$; triangle $P_c = 0.43$) and Thurstonian predictions (3-AFC $d' = 0.97$; triangle $d' = 1.09$, $p = 0.48$) were confirmed. It is not known if the same results would be observed when the number of samples increases. Secondly, Thurstonian predictions regarding the variants of the method of tetrads were challenged using confusable stimuli. Results show that the number of correct responses and degree of difference among the stimuli for the unspecified and specified method of tetrads were similar. Finally, children's performance in different discrimination methods was compared. Performance in the unspecified methods (tetrads and triangle) was as predicted by Thurstonian theory. For the specified methods the 2-AFC resulted in the highest proportion of correct responses, followed by the 3-AFC and the tetrads, respectively, as previously observed with adults as subjects. Results from this investigation give further support to the Thurstonian

predictions discussed above, which had been tested only with adult subjects. In conclusion, under the circumstances of this study, children between 6 to 12 years of age were capable of performing sensory discrimination methods for sweetness perception and they utilized the same decision rules as adults.

CHAPTER 1. INTRODUCTION

The market for children's food products, one of the largest markets in most parts of the world, is constantly expanding. Now more than ever children have more choices and control over their diets. Therefore, manufacturers within the food industry have to compete in order to attain and retain the interest of this market; they must cater to their likes in order to achieve repeated purchases (Craig-Petsinger 2005, Popper and Kroll 2007). When developing new food products specifically for children, it is important to find out what appeals to them, especially since children's food preferences are different from those of adults (Lavin and Lawless 1998). In order to achieve this, children must be involved directly and/or indirectly in the product development process, for either a new product or a reformulation of an existing product. Therefore, research regarding children's perception towards the product under development is essential.

Sensory testing with children is challenging; it poses some of the most difficult problems in sensory analysis and product testing. In a few words, procedural difficulties are encountered because children's cognitive, communicative, and social skills are still developing (Moskowitz 1985). Despite the difficulties, testing with children can provide important insights regarding key product attributes, marketing and exposure benefits, and defining popular products. Acceptance of food products among consumers is mainly determined by the sensory properties of foods and beverages; therefore, the need for sound methodology for sensory testing with children has increased (Moskowitz 1985, Guinard 2001). Little is known about the sensory perception of food by children (Popper and Kroll 2007) and few studies have been published concerning discrimination testing with children.

From a review of the literature it is evident that sensory testing with children is in its developing stage. Not much research has been conducted, especially concerning children's capabilities for performing discrimination methods. Therefore, the objectives of this research were to evaluate children's performance in discrimination methods.

This dissertation is divided into seven chapters. Chapter one provides a summarized introduction and discusses the justifications for this study. Chapter two presents the literature review with concepts and topics associated with this research. In Chapters three to five the paradox of discriminatory non-discriminators (triangle method vs. 3-AFC method) and the Thurstonian predictions associated with it are explored using three sets of stimuli varying in discrimination difficulty. These chapters also discuss age (first to sixth grade) and cultural (Honduras, Mexico, Thailand, and the USA) effects. Chapter three discusses the results using easily discriminable stimuli, chapter four using hardly distinguishable stimuli, and chapter five using confusable stimuli. Chapter six explores the variants of the method of tetrads and the Thurstonian predictions associated with these methods. Chapter seven compares performance in the 2-AFC method, triangle method, 3-AFC method, unspecified method of tetrads, and specified method of tetrads. All cited references are given at the end of each chapter and the appendices contain all supplementary information pertinent to the studies. To conclude, the VITA of the author is provided.

CHAPTER 2. LITERATURE REVIEW

2.1 Introduction

Since many foods and beverages are specifically developed for children, they must be tested by children. This is particularly important, since it was recognized that children and adults differ in their acceptance of various food products (Lavin and Lawless, 1998) due to their difference in sensitivities (Desor and Beauchamp, 1987). Acceptance for food products among consumers is mainly determined by the sensory properties of foods and beverages; therefore, the need for sound methodology for sensory testing with children has increased (Moskowitz, 1985; Guinard, 2001). Little is known about the sensory perception of food by children (Popper and Kroll 2007) and few studies were published concerning product testing with children. This can be partially explained by the limited availability of methodologies to measure food preference in children (Pagliarini and others, 2003).

From various studies it can be stated that the sensory evaluation methods to be used have to be simple in order to be understood by young children and simultaneously have to be robust enough to reliably measure food preferences (Kroll, 1990; Chen and Resurreccion, 1996; Leon and others, 1999; Guinard 2001). While some investigators may view children as unreliable sources of information regarding their own preferences, many others have observed that children do not hesitate to communicate their likes and dislikes about food (Birch 1979a). The use of children in consumer testing of children products is widely accepted, but controversy arises in the case of analytical tests (difference tests, scaling, time-intensity) or descriptive analysis. These methods are either complex and require vast concentration (time-intensity) or require sensory experience and a large sensation database in memory (descriptive analysis). Discrimination methods require more cognitive abilities than preference methods. Therefore, the researcher must consider the following questions: “What information do I want to obtain from children, and what information are they actually capable of providing?" (Propper and Kroll, 2005)

There are many important tools that the researcher needs to have in order to gather the information needed from children. The cognitive and physiological/linguistic development of a child needs to be understood as do the best methods to gather data from them at any age. It is unknown at which age children develop the cognitive ability to reliably and consistently perform analytical and/or consumer tests. Since the child is a "developing organism," different problems arise at different ages due to the evolving capabilities (Moskowitz, 2005). The researcher also needs to keep in mind the role of family habits and feeding practices, along with understanding the importance of familiarity, the role of repeated exposure and peer influence. In addition, the investigators also need to consider other factors that affect food choices, such as schools and teachers, media and advertising, and product cost and availability. The researcher must decide how to handle these interactions in order to avoid biased opinions.

During the sensory evaluation process, the importance of the testing conditions is also essential. Both an adequate testing area and waiting area must be provided. Putting children in foreign environments can lead to false or misleading results. Allowing the children to have fun with the "data collection task" and putting them at ease are critical to gathering their true insights and opinions. In order to meet children's needs for interactive testing, one-on-one interviews have been used, where better data have been obtained if the waiting and testing areas were fun (Moskowitz, 1985).

In conclusion, when using children as subjects in sensory evaluation methods, it is important to understand their limitations and select the appropriate methods for research and data analysis.

2.2 Child Development and Its Effects on Sensory Test Performance

2.2.1 Physiological Development

The sense of taste is directly related to nutrition, allowing humans to reject foods that are toxic (bitter taste) and seek those that are beneficial (sweet taste). There is evidence that the sense of taste is developed before birth and continues to develop during infancy and childhood. The preference of sweets and rejection of bitter during childhood is possibly the most outstanding taste difference between children and adults. This has been shown in several studies (Kroll, 1990; Tuorila-Ollikainen and others, 1984). After seven or eight weeks of gestation, the human fetus has specialized taste cells; more mature receptors are recognizable after 13 to 15 weeks, and information is transmitted to the central nervous system from the taste receptors by the last trimester of pregnancy (Mennella, 2009). It was observed that the fetus can detect saccharin present in the amniotic fluid (Windle, 1940). A newborn will make facial expression as a response to sweet, sour and bitter solutions two hours after birth (Rosenstein and Oster, 1988) and by the age of four months the initial aversion to saltiness is reduced and salty foods are preferred (Harris and others, 1990).

2.2.2 Cognitive Development

A better understanding of children's abilities to perform sensory methods could be accomplished through the comprehension of their general perceptions and cognitive abilities (Resurreccion, 1998). Individuals can be classified into Piaget's four age-related stages of intellectual or cognitive development: sensorimotor, preoperational, concrete operational, and formal operational (Piaget 1952, 1954).

During the sensorimotor stage (birth to 2 years) infants gain understanding of the world by coordinating sensory experiences with physical/motoric actions. At birth, infants have reflexive patterns which progress to complex sensorimotor patterns and use of primitive symbols by the age of 2 years. From the age of 2 years to the age of 7, children are in the preoperational stage. During this period, children represent the world with words, images, and drawings. Stable concepts are formed, beginnings of symbolic motor play are present, speech tends to be egocentric and transductive reasoning can be observed. Also, children still have a tendency to focus their attention only on one aspect of an object while ignoring the others (centration). During the concrete operational stage (7 to 11 years), there is indication of structured, rational thought. The child is able to perform multiple classification tasks and can order objects in logical sequence, and thinking becomes less transductive and less egocentric. At 11 years, the formal operational stage starts and from here to adulthood the reasoning is more abstract, idealistic, and logical (Santrock, 2007).

It is suggested from the sensory scientist's point of view, a critical stage for sensory and consumer testing with children is the preoperational stage, mainly due to centration (Guinard, 2001). However, there are other limitations encountered in sensory evaluation with children which include language (such as limited verbal skills), varying attention spans, and procedural difficulties in comprehending the methods (Guinard, 2001; Resurreccion, 1998).

2.2.2.1 Verbal Skills

Children's limited verbal skills can have an effect on the different parts of a sensory test. First, it can affect their understanding of the questions posed; therefore, the investigator needs to communicate in a language that the children can understand (Thomas, 1992). Children tend to have a problem understanding the meaning of sensory attributes. Attributes such as bitter and sour are often confused, and unconventional sweeteners are confused with off-taste (Moskowitz, 1985). Sensory tests have been adjusted to include suitable wording for children (Kroll 1990).

However, it is suggested that an alternative to modifying the test to fit the children's cognitive abilities is to increase the children's motivation to carry out the taste tests (Liem and Zandstra, 2010).

Because of the language limitations, it is essential to pay special attention during questionnaire design and the execution of sensory tests. Special attention must be given when conducting a test as children tend to answer a question in the way in which it was phrased; i.e., a question asked in a positive manner will be answered affirmatively (Mennella and Beauchamp, 1991). Also, one-on-one interviews may be a better option when conducting the test (Kimmel and others, 1994; Kroll, 1990).

2.2.2.2 Attention Span

In order to maximize the attention span the right balance between comfort and familiarity in the test environment needs to be achieved (Guinard, 2001). Moskowitz (1985) recommended a fun and colorful reception area to help the child relax before the test; however, an uncluttered or undecorated testing area was recommended by Kimmel and others (1994) in order to avoid distraction. Increasing the child's intrinsic motivation may increase their attention. Cooperative instructions resulted in higher intrinsic motivation when compared to individualistic instructions (Hom and others 1994). Liem and Zandstra (2010) showed that with 6 to 9 year old children results for sensory tests can be improved by slight differences in instruction (competitive, cooperative, competitive-cooperative and neutral). Competitive instructions resulted in better discrimination for liking, but resulted in the lowest discrimination for similarity tests.

2.2.2.3 Comprehension

The difficulties in comprehension can be lessened to some extent by using familiarization techniques. A warm-up (O'Mahony and others, 1998) or training exercise can be conducted before the actual study so that the children become familiar with the protocol. This procedure can also help evaluate the children's understanding of the procedure that will take place later and to separate children according to their understanding, not only to their age. Guinard (2001) suggested using visual stimuli before tasting the actual stimuli in order to diminish possible comprehension problems. Kimmel and others (1994) used pictures of food prior to the actual test, and Thomas and Murray (1980) demonstrated and practiced the 'same-different' procedure with colored wooden blocks. Group demonstrations, individual training sessions (Birch and Sullivan, 1991), and practice evaluations (Moskowitz, 1994) have also been recommended. Birch (1979a) divided the tests into smaller parts to help improve understanding.

2.2.3 Perception and Development of Food Preferences

Children's food intakes are predicted by their food preferences (Birch, 1979a, b; Dommel and others, 1996; Fisher and Birch, 1995; Resnicow, 1997). Various nutrition researchers believed that children were not able to provide reliable and valid information regarding their food preference and relied on maternal reports (Bryan and Lowenberg, 1958; Sanjur and Scoma, 1971). However, later research studies demonstrated that children actually can provide reliable and valid preference data (Birch, 1979a, b; Phillips and Kolasa, 1980).

In addition to the biological and psychological factors that affect children's food choices, choice is affected by sensory (taste, appearance, texture) and social factors. It is vital to keep in mind that individual preferences will change over time; i.e., young children's preferences will be different by the time they enter their teenage years. Perception along with learning and experience play a central role in the development of food preferences and aversions.

Taste and olfactory perceptions are present at birth, as children respond differently to sweet, sour, and bitter tastes, with sweet being avidly accepted but sour and bitter rejected. With the exception of foods with high sugar content, it is well known and documented that children do not easily accept new foods. Novel tastes generally receive a negative response. This fear and dislike of novel tastants is known as food neophobia (Mustonen and Tuorila, 2010). In an evolutionary sense, neophobia may have been an adaptive trait by which the ingestion of potentially harmful toxins was diminished; however, in the modern Western civilization, the risk associated with the consumption of new types of food was eliminated (Russell and Worsley, 2008). The development of food preferences is mainly influenced by the sensory properties of the food product, but sometimes these play a secondary role when compared to neophobia. Research has shown a link between neophobia and dislikes for food (Skinner and others, 2002); therefore, neophobia plays a vital role in children's day by day food preferences (Russell and Worsley, 2008).

Peer influence seems to play an important role on a child's likes, as in the case of a study by Birch (1980) that provides evidence concerning the role of peers in children's food preferences in which influences were not only momentary but also long lasting.

2.2.4 Differences between Children and Adults

Differences in food perception and preferences between children and adults exist, and children's preference for greater sweetness intensities than adults is the most documented difference (De Graaf and Zandstra, 1999; Desor and Beuchamp, 1986; Desor and others, 1975; Enns and others, 1979; James and others, 1999; Kimmel and others, 1994; Liem and others, 2004; Zandstra and De Graaf, 1998). Even though most studies have focused on the differences in sweetness perception between children and adults, sourness (Liem and Menella, 2003; Zandstra and De Graaf, 1998), saltiness (Desor and others, 1975; Beuchamp and Cowart, 1990), and bitterness (James and others, 1997) were the focal points of a few studies. In addition, differences in texture (Narain, 2005; Urbick, 2002) and olfactory (Moncrieff, 1966) preferences have been studied.

Several studies have reported that differences between adults and children have been observed when the complexity of the stimuli increased. Oram and others (2001) reported that children's performance declined as the complexity of the stimuli increased. Children performed better when they had to identify a single taste (sweet, sour, or salty) than when they had to identify one of these tastes in a mixture of two. Several investigators observed that when having to discriminate among mixtures of sucrose in water and sucrose in a non-carbonated orange drink, children perform more poorly than adults when using the latter mixture (James and others, 1999; James and others, 2003; Temple and others, 2002). However, results from other investigators resulted in children being less able to discriminate among aqueous solutions of sucrose than adults (De Graaf and Zandstra, 1999).

It is important to understand that there are differences between adult and children's perception of sensory attributes when developing or reformulating a product (Popper and Kroll, 2007). For example, an ingredient change that may be detected by adults might pass undetected by children. As stated by Kimmel and others (1994), "Does the company need to know whether an adult can perceive the difference between two products or whether a child can?"

2.3 Sensory Testing with Children

Children may perform discrimination and/or consumer tests as part of a sensory evaluation study. Several studies with children as subjects in which sensory evaluation were used as a tool were previously published. These tests have been conducted for preference or discrimination.

2.3.1 Affective Methods

Affective methods are used in sensory evaluation to assess preference and/or acceptance of an existing product or a product idea (Meilgaard and others, 2007). As consumers, children may be required to perform a preference or affective test; e.g., preference ranking, paired preference, and hedonic scaling. During the product development process, it is important to be able to effectively determine the level of liking of the product (Popper and Kroll 2007). Several studies used hedonic scales to determine the degree of liking of food products either assessing the methodology or using it as a tool for sensory evaluations (Bovell-Benjamin and others, 1999; Buhany and Bordi, 2004; Capaldi and Privetera, 2008; Chapman and Boor, 2001; Chen and Resurreccion, 1996; Cooper, 2002; Dansby and Bovell-Benjamin, 2003; Hough and others, 1997; Kimmel and others, 1994; Kroll, 1990; Leon and others, 1999; Liem and Zandstra, 2010; Monneuse and others, 1991; Natvaratat and others, 2007; Pagliarini and others, 2003, 2005; Palacios and others, 2010; Reverdy and others, 2010; Shaviklo and others, 2010; Swanson and others, 2002; Tuorilla-Ollikainen and others, 1984; Ward and others, 1999; Wardle and others, 2003). Likewise, paired comparison (Beauchamp and Cowart, 1990; Kimmel and others, 1994; Kroll, 1990; Reverdy and others, 2010) and ranking procedures (Anliker and others, 1991; Birch 1979a, b, 1980, 1990; De Graaf and Zandstra, 1999; Fisher and Birch, 1995; Kimmel and others, 1994; Leon and others, 1999; Liem and Zandstra, 2009; Liem and others, 2004; Popper and Kroll, 2002; Popper and others, 2002) were used to evaluate liking by children.

2.3.2 Discrimination Methods

Discrimination testing is used to determine whether panelists can detect the difference among confusable stimuli (Meilgaard and others, 2007; see Section 2.4.1 for further details). A search of the literature resulted in a limited number of studies that focused on methodologies for discrimination tests and the ability of children to perform such tests. Thomas and Murray (1980) used same-different tests with children 5 to 8 years old (N = 22, kindergarten through third grade) to evaluate their ability to discriminate among spices. Specifically under the conditions of this study, there was no significant difference in discrimination ability among the children (either among age groups or gender) or between children and adults. The investigators reported that the results indicated that the responses were consistent and replicable, except for third graders. This difference in performance by the third graders was unknown to the investigators where the only difference between this grade and the rest was the uneven distribution of gender. The study concluded that the procedure is a reliable method for assessing taste discrimination abilities in young children.

In addition to studying preference, Kimmel and others (1994) assessed discrimination with children (N= 111, 2 to 10 years). The three discrimination tests performed were paired-comparison, duo-trio, and intensity ranking. Stimuli consisted of Kool-aid® flavored drinks, ice cream, and processed cheese slices. All tests were performed using one-on-one interviews and children received training prior to the test to illustrate and strengthen the cognitive skill inherent in the sensory test. Results showed that children 6 years of age and older were able to reliably perform discrimination tests. Mixed results were found for the 4 to 5 age group and results from 2-3 year olds indicated that they should not be used for discrimination tests.

Paired comparison and rank-order tests were used to measure discriminatory ability of 4- and 5-year-old children (N=21, 47) by Liem and others (2004) using stimuli that only differed in sweetness. For the paired comparison tests, the subjects were to identify the sweeter sample, and for the rank-order test, subjects ranked by elimination the sweetest to least sweet samples. Results showed that 5-year-old children were able to discriminate between all solutions and showed high consistency between both tests. In contrast, 4-year-olds failed to distinguish sweetness intensities and also failed to carry out both tests in a consistent way.

From the above presented published studies, it is evident that sensory testing with children is in its developing stage. Not much research has been conducted, especially concerning children's capabilities for performing discrimination tests. Only a few studies tested the validity and repeatability of the methods used, along with a comparison among them (Kimmel and others, 1994; Kroll, 1990; Leon and others, 1999). While children can be a valuable tool for sensory testing, the appropriate protocols and environment must be provided, and children must be treated as a special population. Various studies simply utilize a scaling methodology as an aid in their research without questioning the validity of such. Many studies use published procedures because it was "shown" in other studies that they yielded reliable, valid data on children's food preference or were found to be appropriate for children (De Graaf and Zandstra, 1999). Could they be using the inadequate sensory evaluation method for testing with children? The following are excerpts of commentaries concerning the need for more scientific-based sensory research with children:

"Testing with children is in an embryonic stage. Over the years, a few sensory researchers have considered the problems involved in applying their science to this special population, but for the most part the field has been static. The need for serious investigation is pointed up by how little research has been done in this area." – Kroll 1990

"Some of Kroll's concerns remain true today… given the size of the market and the potential that reliable kid testing has for the food industry, there is still need for more research to help maximize the insights that research with children is able to provide." – Popper and Kroll 2005

"As a working practitioner in a company that has many kid targeted products, I hope that the future of testing with children continues to evolve as the sensory community focuses on developing methods and techniques that gather the richness of insights available from this vast, important segment of the consumer population." – Craig-Petsinger 2005

"I am pleased that Popper and Kroll (2005) specifically commented that more research is needed, because that certainly is true." – Chambers 2005

"…sensory analysis ought to concentrate on creating a repository of knowledge, and, along with that repository, ensure that the methods work…. data is not available for publication, because much of it was funded through contract research projects" – Moskowitz 2005

"Consumer tests with children are conducted routinely nowadays, but the results of such studies typically remain the property of the companies who order them" – Guinard 2001

2.4 Special Topics in Discrimination Testing with Children

2.4.1 Discrimination Testing

Discrimination testing is the basis for sensitivity measurements in psychophysics (comprehending how the human senses work) using the human senses as tools to evaluate food attributes) and investigating consumers' capability to discriminate between foods (O'Mahony, 1988; O'Mahony and Rousseau, 2002). In sensory science, discrimination or difference methods are designed to evaluate subjects' ability to discriminate slight sensory differences among food stimuli, i.e., measure perceptible differences. These differences tend to be small and as a result, the stimuli are highly similar or confusable and special tests protocols are required to establish whether the difference can be perceived or not. Only when the differences among stimuli are subtle are discrimination tests necessary.

Discrimination methods can be employed for product development, studying the effect of processing or ingredient changes, packaging change, storage and shelf life studies, quality assurance, and ingredient specification (Lawless and Heymann, 2010). Typically, discrimination tests are performed with only two different stimuli (for purposes of this dissertation these two stimuli will be referred to as **S** for Strong and **W** for Weak). For three or more stimuli, ranking or scaling techniques are more useful.

Earlier, Stone and Sidel (1992) stated that "one should be capable of handling any discrimination problem using the paired comparison, duo-trio, or triangle method." However, there are many situations in which the use of these tests is not appropriate. There are essentially two groups into which discrimination tests fall: (1) those that require the sensory attribute responsible for the difference among the two stimuli to be known and (2) those that do not have this constraint. Among the first group are the *n*-alternative forced choice methods and the specified method of tetrads. Within the second group we find the paired comparison, duo-trio, triangle, and the unspecified method of tetrads. Therefore, for every test that requires the difference to be identified, there is usually a counterpart that does not; e.g., the 2-AFC is the specified version of the paired-comparison test, the 3-AFC of the triangle test, etc. In this dissertation, the methods of interest are the 2-AFC method, the triangle method, the 3-AFC method, the unspecified method of tetrads, and the specified method of tetrads, which are described in more detail below.

2.4.1.1 2-Alternative Forced Choice (2-AFC) Method

Also known as the directional paired comparison method (Peryam, 1958), the 2-AFC method (Green and Swets, 1966) is used to determine whether two samples differ in a specific sensory attribute, such as sweetness, bitterness, crunchiness, etc. Since its objective is to determine how an attribute differs between samples, it can be referred to as an attribute difference test. There are two possible serving sequences for this method (SW, WS) and its chance probability is ½. The panelist is simultaneously presented with two coded samples and is instructed to select the sample that is stronger or weaker in the specified attribute, e.g., identify which sample tastes sweeter. The panelist has to choose one sample, even if a difference cannot be detected, hence it is a forced choice method.

2.4.1.2 The Triangle Method

An overall difference test (Meilgaard and others, 2007), the triangle method (Dawson and Harris, 1951; Peryam, 1958) is the most prominent discrimination method used. It was first developed in order to utilize sensory methods in the evaluation of beer at Carlsberg Breweries by Bengtsson and co-workers (Stone and Sidel, 1992). This three-sample test consists of determining which sample is different (chance probability = ⅓). The panelists are simultaneously served three samples, two of which are the same and one of which is different. The six possible serving sequences (SSW, SWS, WSS, WWS, WSW, SWW) for this method should be counterbalanced across the panelists. The panelist usually has to specify which is the odd sample, or, in some cases, which two are the same (Helm and Trolle, 1946). Contrary to the 2-AFC method, the triangle test does not indicate the direction of the difference (Lawless and Heymann, 2010).

2.4.1.3 3-Alternative Forced Choice Method (3-AFC)

The 3-AFC method (Green and Swets 1966) follows the same sample presentation as the triangle test except that the instructions specify the nature of the difference, just like the 2-AFC method. The panelists indicate which of the samples is strongest or weakest in a given sensory attribute, e.g., identify which sample tastes sweetest. A sample must be chosen even if a difference cannot be detected. The guessing probability for this test is ⅓.

2.4.1.4 Method of Tetrads

The method of tetrads (O'Mahony and others, 1994; Delwiche and O'Mahony, 1996) involves four stimuli, two of one kind and two of a different kind: SS-WW. There are two variants of the tetrad method: the unspecified and the specified tests. For the unspecified method, the subject is required to correctly separate the four stimuli into their two appropriate sets with a chance probability of $\frac{1}{3}$. For the specified method, with a chance probability = $^{1}/_{6}$, the subject has to correctly identify the two strongest stimuli from the group of four samples.

2.4.2 The Paradox of Discriminatory Non-discriminators

In a sensory study about discriminability between quinine sulfate solutions using the triad method, Byer and Abrams (1953) found their data to be internally inconsistent. An improvement in discrimination performance was observed when there was a change from the triangle method instructions to the 3-AFC instructions. Subjects were presented with tri-stimulus sets comprised of two different stimuli, A (0.005%) and B (0.006%), using all six permutations (AAB, ABA, ABB, BBA, BAB, BAA). In 24 out of 45 trials, when the subjects were instructed to select the odd stimulus, one stimulus of the identical pair was incorrectly chosen as the odd one; i.e., from set AAB stimulus A was chosen and from set ABB stimulus B was chosen. However, when the subjects were instructed to choose the weakest or strongest stimulus, only in 13 out of 45 trails was the incorrect stimulus chosen (A from AAB or B from set ABB). The most significant finding was that of the 24 that incorrectly identified the odd stimulus, 17 chose the strongest or weakest stimuli correctly. Even though a subject chose the incorrect odd stimulus, he/she was capable of choosing the correct weakest or strongest stimulus. Byer and Abrams concluded that the results were contradictory. This inconsistency became known as "the paradox of discriminatory non-discriminators" (Gridgeman, 1970).

In 1979, Fritjers resolved this paradox using Thurstonian (Thurstone, 1927a, b) arguments. The paradox arises in part because the triangle test requires the judge to estimate three sensory differences, whereas the forced-choice method just requires the accurate perception of the strongest or weakest of three stimuli. Using a psychometrical rationalization, Fritjers showed that the instruction given has drastic effects on the decision rule the subject used for stimulus selection. When instructed to select the odd stimulus, distances between momentary sensory values are compared, whereas, when instructed to select the weakest or strongest stimulus a comparison of the absolute momentary sensations must be made. Having established the above, Fritjers used Thurstonian models (Ura, 1960) and Signal Detection Theory (Green and Swets 1966) to reanalyze Byer and Abrams' data. Thurstonian modeling gives the relation between the probability of correct responses and the sensory distance (δ) between the physically different stimuli. Signal Detection Theory was used to obtain the same relation when the instruction is to select the weakest or strongest stimulus. Numerical solutions to these equations facilitated the estimation of the sensory distances between two types of triangle stimuli. Fritjer's results showed that Byer and Abrams' data was not conflicting as both proportions of correct responses (21/45 and 32/45) give virtually identical δ values (1.29 and 1.28). He concluded that the paradox of discriminatory non-discriminators is not real. Therefore, the inconsistency observed by Byer and Abrams is simply due to inappropriate data analysis.

2.4.2.1 Thurstonian Modeling

The differences in performance noted in discrimination methods can be explained by Thurstonian modeling and thus explain the paradox of discriminatory non-discriminators. This approach takes into consideration the brain's central processing (Ennis, 2003). At the same time, Thurstonian modeling provides a measure of degree of difference (d') which can be used to compare results obtained from different discrimination tests (O'Mahony and others, 1994).

Thurstonian modeling states that, everything else being equal, the measure of the degree of difference between products is independent of the methodology used. Thurstonian modeling is based on the ideas of variation in product perception and cognitive strategy or decision rule.

Each time a food product is tasted, the perception of its flavor will vary. This variation can be due to several sources. In regard to the panelist, there is noise in the nervous system, such as inconsistency in the quantity of receptors triggering a response at the peripheral level, physiological effects such as sensory adaptation particularly as a result of residuals from previously tasted stimuli, etc. (O'Mahony, 1995; O'Mahony and others, 1994, 1995; O'Mahony and Rousseau, 2002). In the product, there can be lack of ingredient homogeneity within and between the samples (Lee and O'Mahony, 2007; O'Mahony, 1995). Due to the above mentioned variations, the intensity of the product perception will not be constant, but will vary slightly according to a frequency distribution. For example, variations in sweetness intensity (weaker or stronger) can be represented by a continuous frequency distribution along a flavor intensity axis (Figure 1.1). The momentary intensity upon tasting will be some value along a univariate axis where the height of the distribution represents the frequency of the intensity at any given moment.

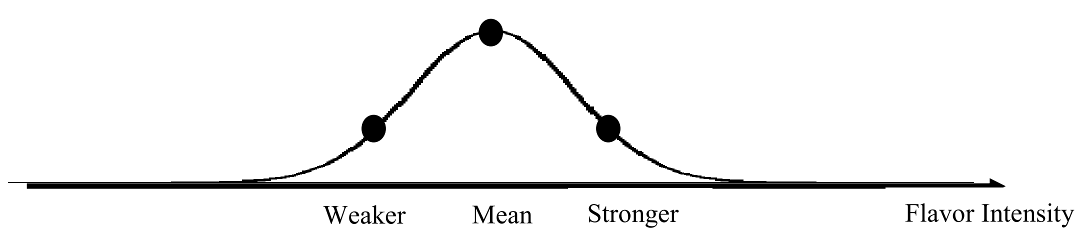


Figure 2.1: Frequency distribution along flavor intensity axis representing variation in flavor of a stimulus on repeated tasting (adapted from O'Mahony and others 1994).

Frequency distributions can either overlap (i.e. the stimuli are confusable) or not overlap (i.e. the stimuli are distinguishable) (Figure 2.1). The more two distributions overlap, the more confusable the stimuli. It is assumed that the two distributions are normally distributed and have equal variance. This is assumed in order to avoid confusion as to which of the distributions provides the standard deviation units. Also, when the stimuli are confusable, it is logical to assume that the two variances are the same. This assumption was confirmed experimentally (Hautus and Irwin, 1995; O'Mahony, 1972). The measure of the overlap and the degree of sensory difference between two products is δ (population parameter) or *d′*(experimental estimate), which is the distance between the two means of the intensity distributions measured in units of standard deviation (Ennis, 2003). The larger the *d′*, the more different the perception of the stimuli.

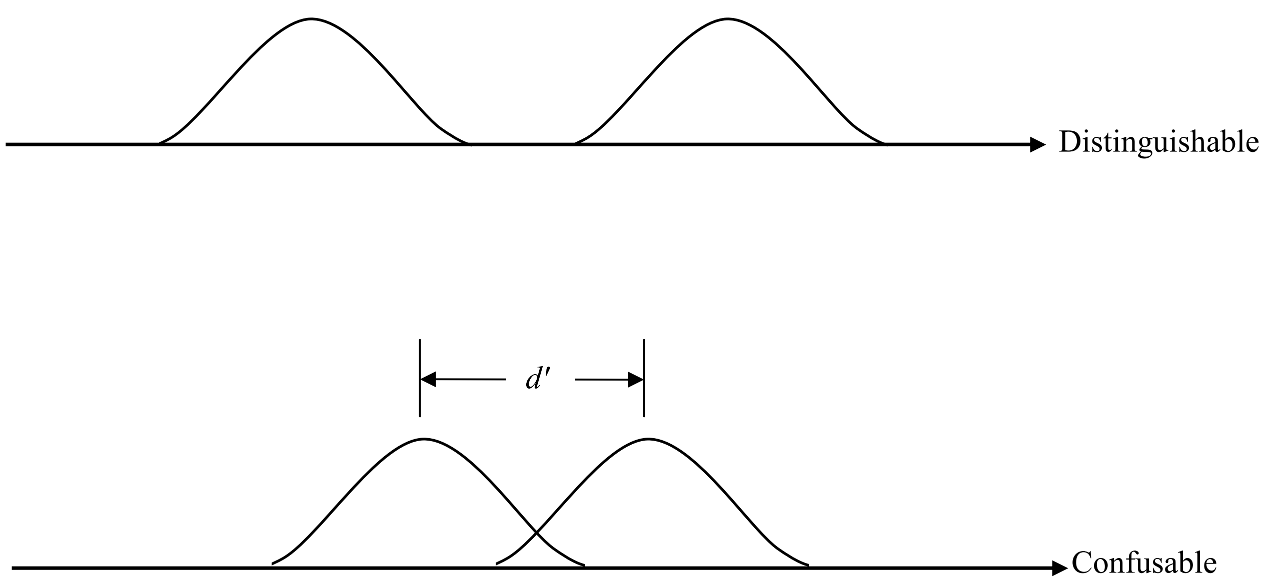
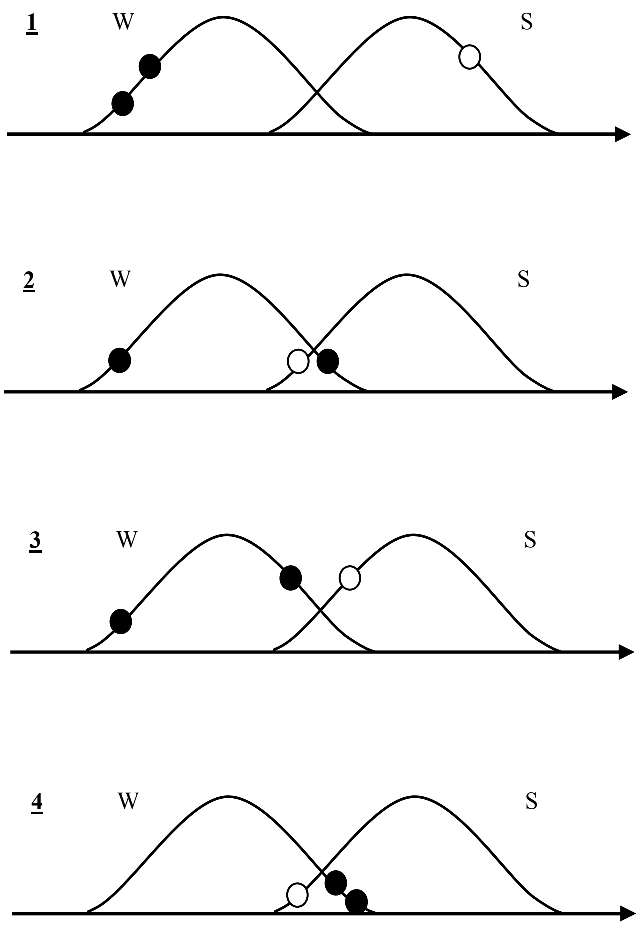


Figure 2.2: Frequency distributions representing distinguishable and confusable stimuli (adapted from O'Mahony and others 1994)

Estimates of δ from the proportion of correct responses will become more accurate as the sample size (N) increases; i.e., the variance of d' gets smaller as N gets larger. The greater the standard deviation of the distributions, the farther apart the means must be in order for the stimuli to be distinguishable. A procedure for calculating values for the variance of d' was explained by Bi and others (1997). With values of the variance, tests of significance can be made. Since d' is measured in units of standard deviation, an increase in the variance of the distributions will reduce d' for the same perceptual distance. For that reason, one can discuss experimental variables in terms of their effect upon variance. Neural noise, memory, pooling replicated data, and pooling data over judges are sources of variance. Thurstonian modeling takes into account variance and thus can account for experimental variables such as memory and product heterogeneity.

The second idea in Thurstonian modeling is the cognitive strategy adopted by the subjects when making decisions during the test. In order to better explain the differences in performance resulting from different instructions, consider the following illustration (Figure 2.3) in which the stimuli are confusable. O'Mahony and others (1993) identify the two stimuli as N and S, but S (strong) and W (weak) are used here for ease of understanding. If the subject is presented with a tri-stimulus set (WWS), the following could happen depending on whether the instruction is to select the odd stimulus (triangle test) or to select the stimulus with the less strong or stronger sensory characteristic (3-AFC).

In case 1, when the stimuli are tasted, both W stimuli are at their lower ebb and S at its higher ebb. In a triangle test, the subject will correctly identify S as the odd sample, and in a 3-AFC test, the subject will also choose S as the stronger stimulus. However, there are cases when the subject will perform differently depending on the sensory perception of the flavor intensity of the stimulus at the instant it is tasted. Three additional scenarios are illustrated in Figure 2.3. One can observe cases when both tests are performed incorrectly (case 2), when only the 3-AFC test is performed correctly (case 3), and when only the triangle test is performed correctly (case 4).



	Stimuli at Instant Tasted	**Triangle**	**3-AFC**
1	W - lower ebb S - higher ebb	Correctly identify S	Correctly identify S
2	W – one lower ebb and one higher ebb S - lower ebb (lower than S)	Incorrectly identify W (lower ebb)	Incorrectly identify S (higher ebb)
3	W – one lower ebb and one higher ebb S – lower ebb (higher than A)	Incorrectly identify W (lower ebb)	Correctly identify S
4	W – higher ebb S – lower ebb (lower than A)	Correctly identify S	Incorrectly identify W

Figure 2.3: Thurstonian treatment of correct and incorrect triangle and 3-AFC tests when instructions are to identify the odd sample, indicating the paradox of discriminatory non-discriminators (adapted from O'Mahony, 1993)

The case where more 3-AFCs will be performed correctly (case 3) will be more frequent than the case where more triangle tests will be performed correctly (case 4). Even though the judge's sensitivity (δ) will not have changed, the cognitive strategy used to find the target stimulus is different. These two strategies were introduced as the "comparison of distances strategy" and "skimming strategy". The comparison of distance strategy applies to the triangle test. The subject has to compare distances of the stimuli flavor intensities along the intensity axis and must choose the stimulus farthest away from the other two as the odd sample. For the 3-AFC test, the judge uses the skimming strategy to determine the strongest sample by moving up or down the axis until the strongest sample is encountered. The choice made is based on absolute sensory magnitude. The difference between these two strategies explains the difference in performance.

Because the cognitive strategies are different, the triangle and 3-AFC methods are not equivalent. The Thurstonian approach aids in understanding and modeling the decision involved in discrimination testing protocols, thereby predicting that the degree of difference between two samples is independent of the protocol used. One must keep in mind that being a "discriminator in the 3-AFC test is not the same as being a discriminator in the triangle test" (Lawless and Heymann, 2010). Being a discriminator in the triangle test requires a correct decision based on the comparison of differences across the stimuli. For the 3-AFC test, the discriminator simply selects the strongest or weakest sample; therefore, the relative differences among the stimuli need not be compared.

Using Thurstonian modeling, Ennis (1993) computed the percentage of correct responses for the triangle and 3-AFC test for different d' values. Psychometric functions relate the proportion of correct responses to δ or d'. Figure 2.4 shows a graphical representation of such functions. The general psychometric function for the triangle method is (Bi, 2006):

$$P_c = 2\int_0^{\infty}\{\Phi[-u\sqrt{3}+\delta\sqrt{(2/3)}]+\Phi[-u\sqrt{3}-\delta\sqrt{2/3}\,]\}\,e^{-(\frac{1}{2})u^2}/\sqrt{2\pi}du$$

The psychometric function for 3-alternative forced choice method is (Bi 2006):

$$P_c = \int_{-\infty}^{\infty}\Phi^2\,(u)\varphi(u-\delta)du$$

Where $\Phi(u)$ = standard normal cumulative distribution function
$\varphi(u)$ = standard normal density function

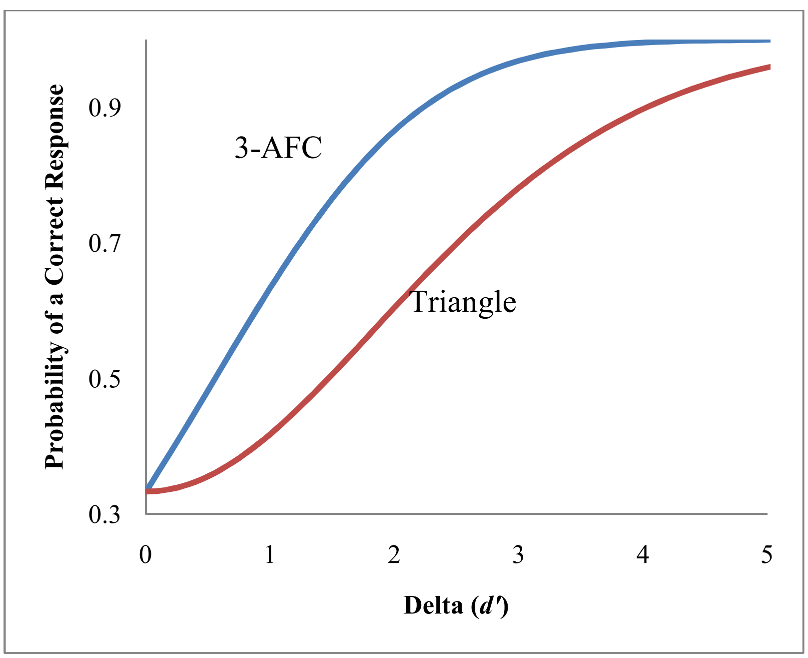


Figure 2.4: Psychometric functions for the triangle and 3-AFC methods (source of data: Ennis, 1993)

Thurstonian modeling creates a theoretical basis for understanding such differences in performance. Initially, Thurstonian modeling (Thurstone, 1927a, b) was applied to paired comparison, triangle, and duo-trio tests (Ura, 1960). Univariate and multivariate Thurstonian models were further developed for various discrimination tests (Bradely, 1963; Ennis, 1988a/b, 1990, 1992; Ennis and Mullen, 1985, 1986 a/b, 1992 a/b; Ennis and O'Mahony, 1995; Ennis and others, 1998; Fritjers, 1979a/b, 1980, 1982b; Kapenga and others, 1987; Mosteller, 1951a/b/c; Mullen and Ennis, 1987, 1991; Mullen and others, 1998; Vessereau, 1965). These models were used to produce tables relating δ and proportions of correct responses for many force-choice discrimination procedures (Elliot, 1964; Ennis, 1993; Ennis and Mullen, 1986; Ennis and others, 1998; Fritjers, 1980, 1982b; Fritjers and others, 1980; Hacker and Ratcliff, 1979; Ura, 1960). The application of Thurstonian modeling to discrimination testing has been reviewed (Ennis, 1990; Lee and O'Mahony, 2004; Lee and O'Mahony, 2007; O'Mahony, 1992, 1995a/b; O'Mahony and Rousseau, 2002; O'Mahony and others, 1994). Utilizing d' values, the comparison of subject performance on various discrimination tests has been performed (Braun and others, 2004; Delwiche and O'Mahony, 1996a/b; Dessirier and O'Mahony, 1999; Geelhoed and others, 1994; Hautus and Irwin,1995; Huang and Lawless, 1998; Ishii and others, 2007; Kim and others, 2006; Kuesten, 2001; Lau and others, 2004; MacRea and Geelhoed, 1992; Masuoka and others, 1995; Rousseau and O'Mahony, 1997, 2000; Rousseau and others, 1999, 2002; Stillman, 1993; Stillman and Irwin, 1995; Tedja and others, 1994).

In direct relation to the paradox of discriminatory non-discriminators, it was observed experimentally that the same subjects discriminating between the same stimuli perform a higher proportion of 3-AFC tests correctly than triangle tests, yet have the same d' value (Delwiche and O'Mahony, 1996; Fritjers 1981; MacRea and Geelhoed, 1992; Masouka and others, 1995;

Rousseau and O'Mahony, 1997; Stillman, 1993; Tedja and others, 1994). However, studies have been not been conducted with children as subjects.

2.4.3 The Variants of the Methods of Tetrads

Thurstonian theory predicts that the number of correct responses should increase when the instructions for triadic methods are altered from those of the triangle to those of the 3-AFC. O'Mahony and others (1994) reported that the paradox can be generalized to methods that have an uneven number of samples, but not to tests with an even number of samples that need to be arranged into two groups of equal size. Hence, an increase in performance is not predicted when the instructions for the tetrad method change from the unspecified to the specified method (O'Mahony and others, 1994). The strategy associated with the specified method of tetrads does not elicit a better performance. Figure 2.5 illustrates performance in the method of tetrads.

Consider the unspecified method, where the instructions are to separate the four samples (SSWW) into two groups of two (SS and WW). The judge would correctly sort the stimuli into the correct groups for case 1 and 2. For the specified method (select the two sweeter samples), cases 1 and 2 would also be scored correctly. However, in case 3, the S and W stimuli perceived weaker (left) would be grouped together as would the S and W perceived stronger (right), and both the specified and unspecified versions will be completed incorrectly. Case 4 illustrates the event where the two W stimuli would be perceived as stronger than the two S stimuli. Here the "comparison of distances" strategy (difference unspecified) would results in a correct response, but the "skimming" strategy (difference specified) would not. O'Mahony and others (1994) reported that since this last case would occur less often, the difference in performance expected due to the decision rules adopted will not be distinct.

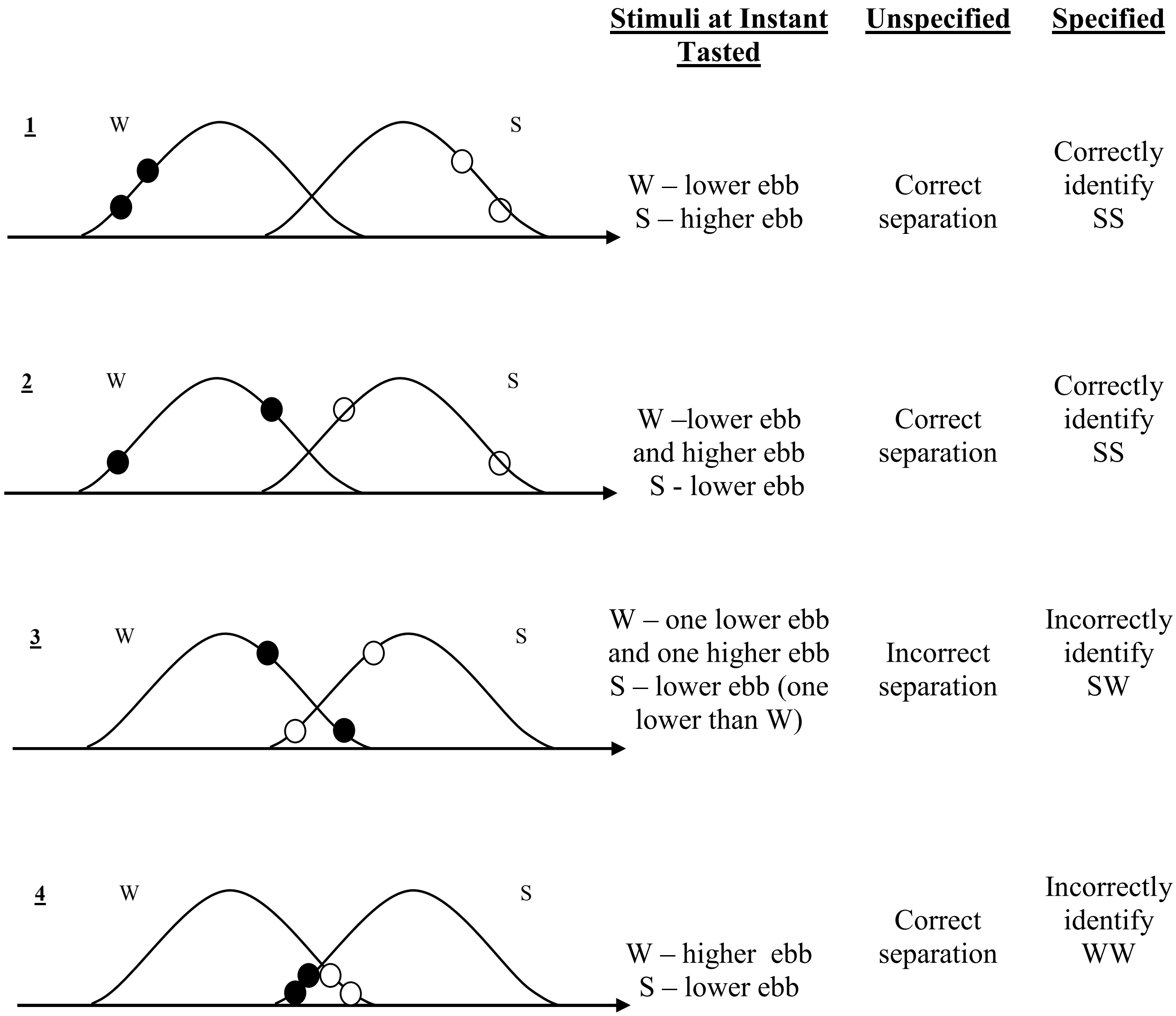


Figure 2.5: Thurstonian treatment of correct and incorrect unspecified and specified method of tetrads tests (adapted from O'Mahony, 1993)

The psychometric functions for the variants of the method of tetrads were derived by Ennis and others (1998) and are presented in Figure 2.6. These functions show how, for differences greater than 1 ($d' > 1$), the difference in the probability of correct responses is negligible.

The general psychometric function for the unspecified method is (Ennis and others, 1998):

$$P_c = 1 - 2\int_{-\infty}^{\infty} \varphi(x)[2\Phi(x)\Phi(x-\delta) - [\Phi(x-\delta)]^2\,]\,dx$$

The psychometric function for the specified method is (Ennis and others, 1998):

$$P_c = 1 - 2\int_{-\infty}^{\infty} \varphi(x)\Phi(x)[2\Phi(x-\delta) - [\Phi(x-\delta)]^2\,]\,dx$$

Where $\varphi(u)$ = standard normal density function
$\Phi(u)$ = standard normal cumulative distribution function

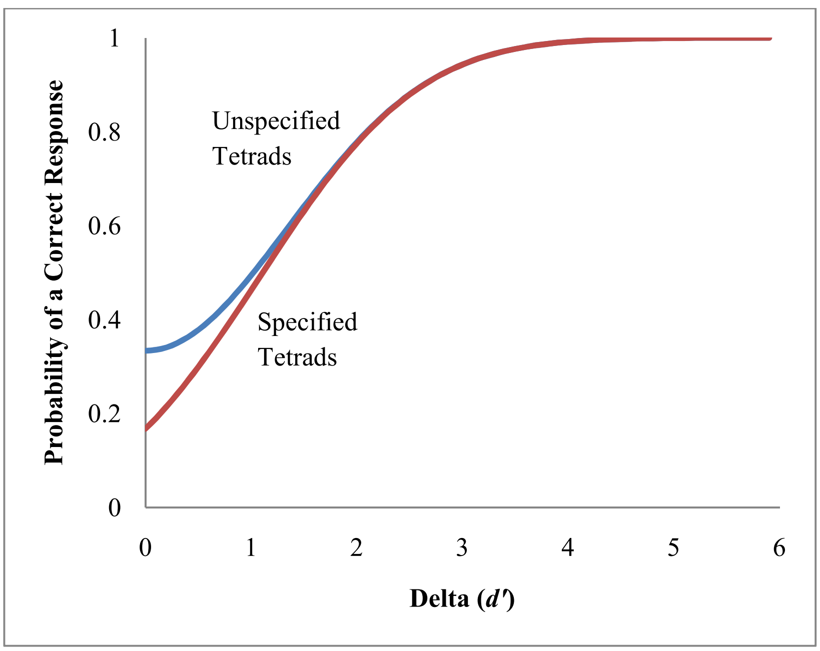


Figure 2.6: Psychometric functions for the specified and unspecified method of tetrads (source of data: Ennis and others 1998)

It was observed experimentally that the same adult subjects discriminating between the same stimuli did not perform a greater number of correct specified tests in relation to the unspecified tests. Delwiche and O'Mahony (1996) showed that specifying the attribute in a tetrads test did not elicit better discrimination using chocolate pudding varying in sweetness concentrations as the stimuli. Masouka and others (1995) also confirmed such predictions using beer with different bitterness concentrations as the stimuli. However, studies with children as subjects have not been performed to this date.

2.5 References

Anliker JA, Bartoshuk L, Ferris AM, Hooks LD. 1991. Children's food preferences and genetic sensitivity to the bitter taste of 6-n-propylthiouracil (PROP). American Journal of Clinical Nutrition. 54:316-320.

Beauchamp GK and Cowart BJ. 1990. Preference for high salt concentrations among children. Develop. Psychol. 26(4):539-545.

Bi J, Ennis DM and O'Mahony M. 1997. How to estimate the variance of d' from difference tests. Journal of Sensory Studies. 12:87-104.

Birch LL. 1979a. Dimensions of preschool children's food preferences and consumptions patterns. Journal of Nutritional Education. 11:77-80.

Birch LL. 1979b. Preschool children's food preferences and consumption patterns. Journal of Nutritional Education. 11: 189-192.

Birch, L.L. 1980. Effect of peer models' food choices and eating behaviors on preschoolers' food preferences. Child Dev. 51(2): 489-496.

Birch LL and Sullivan SA. 1991. Measuring children's food preferences. Journal of School Health. 61:212-214.

Bovell-Benjamin AC, Allen LH, Guinard JX. 1999. Toddlers' acceptance of whole maize meal porridge fortified with Ferrous Bisglycinate. Food Quality and Preference. 10, 123-128.

Buhaly MD and Bordi PL. 2004. Development and sensory evaluation of a high-protein, vitamin-fortified fruit roll-up for children with cystic fibrosis. Food Science Research International. 14:243-256.

Bradely RA. 1963. Some relationships among sensory difference tests. Biometrics. 19: 385-397.

Braun V, Rogeaux M, Schneid N, O'Mahony M, Rousseau B. 2004. Corroborating the 2-AFC and 2-AC Thurstonian models using both a model system and sparkling water. Food Quality and Preference. 15: 501-507.

Bryan MS, Lowenberg ME. 1958. The father's influence on young children's food preferences. Journal of the American Dietetic Association. 34(1):30-35.

Byer AJ and Abrams D. 1953. A comparison of the triangular and two-sample taste-test methods. Food Technology. 7:185-187.

Capaldi Ed, Privitera GJ. 2008. Decreasing dislike for sour and bitter in children and adults. Appetite. 50:139-145.

Chapman KW and Boor KJ. 2001. Acceptance of 2% ultra-pasteurized milk by consumers, 6 to 11 years old. Journal of Dairy Science. 84: 951-954.

Chambers E. 2005. Commentary: Conducting sensory research with children. Journal of Sensory Studies. 20, 90-92.

Chen AW and Resurreccion AVA. 1996. Age appropriate hedonic scales to measure food preference of young children. Journal of Sensory Studies. 11:141.

Cooper H. 2002. Designing successful diagnostic scales for children. Presented at Annual Meeting Institute of Food Technologists, Anaheim, CA June 15-19.

Craig-Petsinger D. 2005. Commentary: Sensory research with children. Journal of sensory Studies. 20, 88-89.

Dansby MY and Bovell-Benjamin AC. 2003. Physical properties and sixth graders' acceptance of an extruded ready-to-eat sweetpotato breakfast cereal.

Dawson EH and Harris BL. 1951. Sensory methods for measuring differences in food quality. Agriculture Information Bulletin 34, US Department of Agriculture, Washington, DC.

De Graaf C and Zandstra EH. 1999. Sweetness intensity and pleasantness in children, adolescents, and adults. Physiology and Behavior. 67 (4):513-520.

Delwiche J and O'Mahony M. 1996a. Flavour discrimination – an extension of Thurstonian paradoxes to the tetrad method. Food Quality and Preference. 7:1-5.

Delwiche J and O'Mahony M. 1996b. Changes in secreted salivary sodium are sufficient to alter salt taste sensitivity: use of signal detection measures with continuous monitoring of the oral environment. Physiology and Behavior. 59: 605-611.

Dessirier JM and O'Mahony M. 1999. Comparison of d' values for the 2-AFC (paired comparison) and 3-AFC discrimination methods: Thurstonian models, sequential sensitivity analysis and power. Food Quality and Preference. 10: 51-58.

Domel SB, Thompson WO, Davis HC, Baranowski T, Leonard SB, Baranowski J. 1996. Psychosocial predictors of fruit and vegetable consumption among elementary school children. Health Education Research. 11:299-308.

Elliot PB. 1964. Tables of d'. Signal Detection and Recognition by Human Observers. John Wiley and Sons, New York. pp 651-684.

Ennis DM. 1988a. Multivariate sensory analysis. Food Technology. 42(11):118-122.

Ennis DM. 1988b. Confusable and discriminable stimuli: Comment on Nofosky (1986) and Shepard (1986). Journal of Experimental Psychology. General. 117(4):408-411.

Ennis DM. 1990. Relative power of difference testing methods in sensory evaluation. Food Technology. 44(4):114-118.

Ennis DM. 1992. Modeling similarity and identification when there are momentary fluctuations in psychological magnitude. Multidimensional Models of Perception and Cognition. FG Ashby ed. Lawrence Erlbaum Assoc, Hillsdale, New Jersey. pp. 279-298.

Ennis DM. 1993. The power of sensory discrimination methods. Journal of Sensory Studies. 8:353-370.

Ennis DM. 2003. Foundations of sensory science. In: Viewpoints and controversies in sensory science and consumer product testing. Moskowitz HR, Munos AM, Gacula MC. Food and Nutrition Press, Inc. Connecticut. p. 477.

Ennis DM, Mullen K. 1985. The effect of dimensionality of results from the triangular method. Chemical Senses. 10: 605-608.

Ennis DM, Mullen K. 1986a. Theoretical aspects of sensory discrimination. Chemical Senses. 11:513-522.

Ennis DM, Mullen K. 1986b. A multivariate model for discrimination methods. Journal of Mathematical Psychology. 30(2):206-219.

Ennis DM, Mullen K. 1992a. Probabilistic psychophysics with noisy stimuli. Mathematical Social Sciences. 23(2): 221-234.

Ennis DM, Mullen K. 1992b. A general probabilistic model for triad discrimination, preferential choice, and two- alternative identification. Multidimensional Models of Perception and Cognition. FG Ashby ed. Lawrence Erlbaum Assoc, Hillsdale, New Jersey. pp. 115-112.

Ennis JM, Ennis DM, Yip D, O'Mahony M. 1998. Thurstonian models for variants for the method of tetrads. British Journal of Mathematical and Statistical Psychology. 51:205-215.

Enns MP, Van Itallie TB, Grinker, JA. 1979. Contributions of age, sex and degree of fatness on preferences and magnitude estimation of sucrose in humans. Physiology and Behavior. 22(5):999-1003.

Fisher JO and Birch LL. 1995. Fat preferences and fat consumption of 3- to 5-year-old children are related to parental adiposity. Journal of the American Dietetic Association. 95(7): 759-764.

Fritjers JER. 1979a. The paradox of discriminatory nondiscriminators resolved. Chemical Senses and Flavour. 4(4):355-358.

Fritjers JER. 1979b. Variations of the triangular method and the relationship of its unidimensional probabilistic models to three-alternative forced-choice signal detection theory models. British Journal of Mathematical and Statistical Psychology. 32:229-241.

Fritjers JER. 1980. Three-stimulus procedures in olfactory psychophysics: An experimental comparison of Thurstone-Ura and three-alternative forced choice models of signal detection theory. Perception & Psychophysics. 28: 390-397.

Fritjers JER, Kooistra A and Vereijken PFG. 1980. Tables of d' for the triangular method and the 3-AFC signal detection procedure. Perception and Psychophysics. 27:176-178.

Fritjers JER. 1982b. Expanded tables for conversions of a proportion correct responses (Pc) to the measure of sensory difference (d') for triangular method and the 3-alternative forced choice procedure. Journal of Food Science. 47:139-143.

Geelhoed EN, MacRea AW, and Ennis DM. 1994. Preference gives more consistent judgments than oddity only if the task can be modeled as forced choice. Perception and Psychophysics. 55: 473-477.

Green DM and Swets JA. 1966. Signal Detection Theory and Psychophysics. Wiley, New York.

Gridgeman NT. 1970. A re-examination of the two-stage triangle test for the perception of sensory differences. Journal of Food Science. 35:87-91.

Guinard JX. 2001. Sensory and consumer testing with children. Trends in Food Science and Technology 11, 273-283.

Hacker MJ, Ratcliff R. 1979. A revised table of d' for M-alternative forced choice. Perception and Psychophysics. 26:168-170.

Harris G, Thomas A, Booth DA. 1990. Development of salty taste in infancy. Developmental Psychology. 26; 534-538.

Hautus MJ, Irwin Rj. 1995. Two models for estimating the discriminability of foods and beverages. Journal of Sensory Studies. 26:168-170.

Helm E, Trolle B. 1946. Selection of a taste panel. Wallerstein Lab. Communiactions. 9:181-194.

Hough G, Sanchez R, Barbieri T, Martinez E. 1997. Sensory optimization of a powdered chocolate milk formula. Food Quality and Preference. 8(3):213-221.

Huang YT and Lawless HT. 1998. Sensitivity of the ABX discrimination test. Journal of Sensory Studies. 13: 229-239.

Ishii R, Kawaguchi H, O'Mahony M, and Rousseau B. 2007. Relating consumer and trained panels' discriminative sensitivities using vanilla flavored ice cream as a medium. Food Quality and Preference. 18:89-96.

James CE, Liang DG, and Oram N. 1997. A comparison of the ability of 8-9-year-old children and adults to detect taste stimuli. Physiol. Behavior. 62(1):193-197.

James CE, Laing DG, Oram N, Hutchinson I. 1999. Perception of sweetness in simple and complex taste stimuli by adults and children. Chemical Senses. 24(3):281-287.

Kapenga JA, de Doncker E, Mullen K, and Ennis DM. 1987. The integration of the multivariate normal density function for the triangular method. Numerical Integration. P. Keast and G. Fairweather, eds. D Reidel Publ. Co., Dordrecht, Holland. pp 321-328.

Kim H, Jeon SY, Kim K and O'Mahony M. 2006. Thurstonian models and variance I: Experimental confirmation of cognitive strategies for difference tests and effects of perceptual variance. Journal of Sensory Studies. 21:465-484.

Kim K, Ennis DM and O'Mahony M. 1998. A new approach to category scales of intensity II: Use of d' values. Journal of Sensory Studies. 13:251-267.

Kimmel S, Sigman-Grant MJ, Guinard JX. 1994. Sensory testing with young children. Food Technology. 48, 92-99

Kroll BJ. 1990. Evaluating rating scales for sensory testing with children. Food Technology. 44(11) 78-80, 82, 84, 86.

Kuesten CL. 2001. Sequential use of the triangle, 2-AC, 2-AFC, and same-different methods applied to a cost-reduction effort: Consumer learning acquired throughout testing and influence on preference judgments. Food Quality and Preference. 12:447-455.

Lavin JG and Lawless HT. 1998. Effects of color and odor on judgments of sweetness among children and adults. Food Quality and Preference. 9(4): 283-289.

Lawless HT, Heymann H. 2010. Sensory evaluation of food: principles and practices. Springer-Verlag, New York. P. 850.

Lau S, O'Mahony M, and Rousseau B. 2004. Are three-sample tasks less sensitive than two-sample tasks? Memory effects in the testing of taste discrimination. Perception and Psychophysics. 66: 464-474.

Lee H -S and O'Mahony M. 2004. Sensory difference testing: Thurstonian models. Food Science and Biotechnology. 13: 841-847.

Lee H -S and O'Mahony M. 2007. The evolution of a model: A review of Thurstonian and conditional stimulus effects on difference testing. Food Quality and Preference. 18:369-383

Leon F, Couronne T, Marcuz MC, Koster EP. 1999. Measuring food liking in children: a comparison of non-verbal methods. Food Quality and Preference. 10, 93-100.

Liem DG, Mars M, De Graaf C. 2004. Consistency of sensory testing with 4- and 5-year old children. Food Quality and Preference. 15:541-548.

Liem DG, Mennella J. 2003. Heightened sour preferences during childhood. Chemical Senses. 28 (2): 173-180.

Liem DG, Zandstra LH. 2010. Children's liking and wanting of snack products: influence of shape and flavour. International Journal of Behavior Nutrition and Physical Activity. 6 (38)

MacRea AW and Geelhoed EN. 1992. Preference can be more powerful than detection of oddity as a test of discriminability. Perception and Psychophysiology. 51:79-81.

Masouka S, Hatjopoulos D and O'Mahony M. 1995. Beer bitterness detection: Testing Thurstonian and sequential sensitivity analysis models for triad and tetrad methods. Journal of Sensory Studies. 10:295-306.

Meilgaard MC, Civille GV, Carr BT. 2007. Sensory Evaluation Techniques. CRC Press: Boca Raton. p. 448.

Menella J. 2009. In: the child an encyclopedic companion. University of Chicago Press. Editor: Richard Shweder . 1144 p.

Mennella J, Beauchamp G. 1991. Maternal diet alters the sensory qualities of human milk and the nursling's behavior. Pediatrics. 88(4): 734-744.

Moncrieff RW. 1966. Odour preferences. Leonard Hill, London.

Monneuse MO, Bellisle F, Louis-Sylvestre J. 1991. Impact of sex and age on sensory evaluation of sugar and fat in dairy products. Psychology and Behavior. 50(6): 1111-1117.

Moskowitz HR. 1985. Product testing with children. In “New direction for product testing and sensory analysis of foods”. Food and Nutrition Press Inc. Westport, CT. 371p.

Moskowitz HR. 2005. Commentary: Whither now the grand sensory “project” in an age of improving methodology? Journal of Food Science. 20, 93-95.

Mosteller F. 1951a. Remarks on the method of paired-comparisons: I. The least squares solution assuming equal standard deviations and equal correlations. Psychometrika. 16:3-9

Mosteller F. 1951b. Remarks on the method of paired-comparisons: II. The effect of an aberrant standard deviation when equal standard deviations and equal correlations are assumed. Psychometrika. 16:203-206.

Mosteller F. 1951c. Remarks on the method of paired-comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed. Psychometrika. 16: 207-218.

Mullen K and Ennis Dm. 1987. Mathematical formulation of multivariate Euclidean models for discrimination methods. Psychometrika. 52: 235-249

Mullen K and Ennis DM. 1991. A simple multivariate probabilistic model for preferential and triadic choices. Psychometrika. 56:69-75.

Mullen K, Ennis DM, de Doncker E and Kapenga JA. 1988. Models for the duo-trio and triangular methods. Biometrics. 44:1169-1175.

Narain C. 2005. Texture preferences in children. Presented at the 6th Pangborn Sensory Science Symposium, Harrogate, UK, August 7-11.

Natvaratat M, Chompreeda P, Haruthaithanasan V, Rimkeeree H. 2007. Optimization of supplementary protein milk tablet formulation for rural school children under her Royal Highness Princess Maha Chakri Sirindhorn’s project. Kasetsart Journal Natural Science. 41:733-739.

O’Mahony M. 1972. Salt taste sensitivity: a signal detection approach. Perception. 1:459-464.

O’Mahony M. 1988. Sensory difference and preference testing: the use of signal detection measures. In Applied Sensory Analysis of Food. Ed. H Moskowitz. CRC Press, Boca Raton, Florida. Vol. 1. pp 145-175.

O’Mahony M. 1992. Understanding discrimination tests: a user-friendly treatment of response bias, rating and ranking R-index tests and their relationship to signal detection. Journal of Sensory Studies. 7:1-47.

O’Mahony M. 1995a. Who told you the triangle test was simple? Food Quality and Preference. 6:227-238.

O'Mahony M. 1995b. Sensory measurement in food science: fitting methods to goals. Food Technology. 49(4):72-82.

O'Mahony M, Masuoka S, Ishii R. 1994. A theoretical note on difference tests: models, paradoxes and cognitive strategies. Journal of Sensory Studies. 9: 247-272.

O'Mahony M and Rousseau B. 2002. Discrimination testing: A few ideas, old and new. Food Quality and Preference. 14:157-164.

O'Mahony M, Thieme U and Goldstein L. 1988. The warm-up effect as a means of increasing the discriminability of sensory difference tests. Journal of Food Science. 53: 1848-1850.

Oram N, Laing DG, Freeman MH and Hutchinson I. 2001. Analysis of taste mixtures by adults and children. Developmental Psychbiology. 38: 67-77.

Pagliarini E, Ratti S, Balzaretti C, Dragoni I. 2003. Evaluation of hedonic scaling method for measuring the acceptability of school lunches by children. Italian Journal of Food Science. 2(15), 215-224.

Pagliarini E, Gabbiadini N, Ratti S. 2005. Consumer testing with children on food combinations for school lunch. Food Quality and Preference. 16, 131-138.

Palacios OM, Bardan J, Spence J, Drake MA, Reisner M, Moskowitz HR. 2010. Measuring acceptance of milk and milk substitutes among younger and older children. Journal of Food Science. 75(9): S522-S526.

Peryam DR. 1958. Sensory difference tests. Food Technology. 12(5):231-236.

Phillips BK and Kolasa K. 1980. Vegetable preferences of preschoolers in day care. Journal of Nutrition Education. 12:192-195.

Piaget J. 1952. The origin of intelligence in children. International Universities Press, New York.

Piaget J. 1954. The construction of reality in the child. Basic, New York.

Popper R, Schraidt M, Kroll J. 2002. Testing with preschool children: The effect of the interviewer. Presented at Annual Meeting Institute of Food Technologists, Anaheim, CA, June 15-19.

Popper R, Kroll JJ. 2005. Issues and Viewpoints: Conducting Sensory Research with Children. Journal of Sensory Studies. 20, 75-87.

Popper R, Kroll JJ. 2007. Consumer testing of food products using children. In: consumer-led food product development. CRC Press Boca Raton.

Resnicow K, Davis-Hearn M, Smith M. 1997. Social cognitive predictors of fruit and vegetable intake in children. Health Psychology. 16: 272-276.

Resurreccion AVA. 1998. Affective Testing with Children. In: Consumer Sensory Testing for Product Development. Aspen Publishing.

ReverdyC, Schlich P, Koster EP, Ginon E, Lange C. 2010. Effect of sensory education on preferences in children. Food Quality and Preference. 21:795-804.

Rosenstein D and Oster H. 1988. Differential facial responses to four basic tastes in newborns. Child Development . 59: 1555-1568.

Rousseau B. 2001. The β-strategy: An alternative and powerful cognitive strategy when performing sensory discrimination tests. Journal of Sensory Science. 16:301-318.

Rousseau B, O'Mahony M. 1997. Sensory difference tests: Thurstonian and SSA predictions for vanilla flavored yogurts. Journal of Sensory Studies. 12:127-146.

Rousseau B, O'Mahony M. 2000. Investigation of the effect of within-trial retasting and comparison of the dual-pair, same-different and triangle paradigms. Food Quality and Preference. 11(6): 457-464.

Rousseau B, O'Mahony M. 2001. Investigation of the dual-pair method as a possible alternative to the triangle and same-different tests. Journal of Sensory Studies. 16:161-178.

Rousseau B, Stroh S, O'Mahony M. 2002. Investigating more powerful discrimination tests with consumers: effects of memory and response bias. Food Quality and Preference. 13(1):39-45.

Russell CG and Worsley A. 2008. A population-based study of preschoolers' food neophobia and its association with food preferences. J Nutr Educ Behav. 40:11-19

Sanjur D, Scoma AD. 1971. Food habits of low income children in Northern New York. Journal of Nutritional Education. 2: 85.

Santrock JW. 2007. Life Span Development. 3rd Ed. McGraw Hill, New York. P. 669.

Skinner JD, Carruth BR, Bounds W, Zeigler PJ. 2002. Children's food preferences: a longitudinal analysis. Journal of the American Dietetic Association. 102: 1638-1647.

Shavilko GR, Olafsdottir AO, Sveinsdottir K, Thorkelsson G, Rafipour F. 2010. Quality characteristics and consumer acceptance of a high fish protein puffed corn-fish snack. Journal of Food Science and Technology.

Stillman JA. 1993. Response selection, sensitivity, and taste test performance. Perception and Psychophysics. 54:190-194.

Stillman JA, Irwin RJ. 1995. Advantages of the same-different method over the triangle method for the measurement of taste discrimination. Journal of Sensory Studies. 10:261-272.

Stone H, Sidel JL. 1992. Sensory Evaluation Practices: Food and Science Technology Series. Elsevier Science and Technology Books. p. 338.

Swanson RB, Perry JM, Carden LA. 2002. Acceptability of reduced-fat brownies by school-aged children. Journal of the American Dietetic Association. 102(6): 856-859.

Tedja S, Ninaka R. Ennis DM, O'Mahony M. 1994. Triadic discrimination testing – refinement of Thurstonian and SSA approaches. Chemical Senses. 19:279-301.

Temple EC, Laing DG, Hutchinson I, Jinks AL. 2002. Temporal perception of sweetness by adults and children using computerized time-intensity measures. Chemical Senses. 27: 729-737.

Thomas RM. 1992. Comparing theories of Child Development. 3rd ed. Wadsworth, Belmont, CA.

Thomas MA, Murray FS. 1980. Taste Perception in Young Children. Food Technology. 34(3): 38-41.

Thurstone LL. 1927a. Psychophysical analysis. American Journal of Psychology. 38: 368-389.

Thurstone LL. 1927b. A law of comparative judgment. Psychology Review. 34:273-286.

Tuorila-Ollikainen H, Mahlamaki-Kultanen S, Kurela R. 1984. Relative importance of color, fruity flavor and sweetness in the overall liking of soft drinks. Journal of Food Science. 49: 1598-1600.

Ura S. 1960. Pair, triangle and duo-trio tests. Reports of Statistical Stat. Appl. Res., JUSE. 7:107-119.

Urbick B. 2002. Kids have great taste: An update to sensory work with children. Presented at Annual Meeting Institute of Food Technologists, Anaheim, CA June 15-19.

Vesserreau A. 1965. Les methodes statistiques appliquées au test des caracteres organoleptiques. Ann. Nutr. 19:103-140.

Ward CDW, Stampanoni Koeferli C, Piccinali Schwegler P, Schaeppi D, Plemmons LE. 1999. European strawberry yogurt market analysis with a case study on acceptance drivers for children in Spain using principal component analysis and partial least square regression. Food Quality and Preference. 10:387-400.

Wardle J, Cooke LJ, Gibson EL, Sapochnik M, Sheiham A, Lawson M. 2003. Increasing children's acceptance of vegetables; a randomized trial of parent-led exposure. Appetite. 40:155-162.

Wardle J, Herrera ML, Cooke LJ, Gibson EL. 2003. Modifying children's food preferences: the effects of exposure and reward on acceptance of an unfamiliar vegetable. European Journal of Clinical Nutrition. 57:341-348.

Windle WF. 1940. Physilogy of the human fetus. Philadelpia: W.B. Saunders

Zandstra EH and de Graaf C. 1998. Sensory perception and pleasantness of orange beverages from childhood to old age. Food Quality and Preference. 9, 5-12.

CHAPTER 3. THE PARADOX OF DISCRIMINATORY NON-DISCRIMINATORS WITH CHILDREN AS SUBJECTS USING EASILY DISCRIMINABLE STIMULI

3.1 Introduction

Discrimination tests are routinely used by the food industry for discrimination among confusable products. These tests are routinely used in product development, quality control and shelf life studies (O'Mahony and Rousseau, 2002). While performing forced-choice discrimination tests, subjects may perform better when detecting differences using a protocol than when using another. In food science, seemingly paradoxical results are observed in triangle and 3-alternative forced-choice (3-AFC) methods. These methods require that the subject distinguish a stimulus (S) from a pair of identical stimuli (N-N). The triangle method (Peryam and Swartz, 1950) instructions do not specify the nature of the difference between S & N and the subject is to identify the odd sample. The 3-AFC (Green and Swets, 1966) instructions specify the nature of the difference, and the subject is to select the strongest or weakest stimulus with respect to the specified attribute (e.g., sweetness).

Byers and Abrams (1953) first observed that a change in instructions from the triangle method to the 3-AFC resulted in an increased number of correct responses. This paradox became known as the paradox of discriminatory non-discriminators or Gridgeman's paradox (Gridgeman, 1970). However, Fritjers (1979) resolved this paradox using Thurstonian (Thurstone, 1927a, b) arguments. These arguments were reviewed in detail by O'Mahony (1995) and O'Mahony and others (1994). Thurstonian theory states that the perception of the intensity of a stimulus will vary each time the stimulus is tasted. This variation is represented by a normal frequency distribution along a univariate intensity axis. The degree of difference between the stimuli or difference between the means of the two distributions is called δ (δ is the population parameter and d' is the experimental estimate) and is measured in units of standard deviation. It is generally assumed that the two distributions have the same standard deviation. The precision of d' is conveyed by its variance, which depends on sample size (N) and the discrimination method used (B value) (Bi and others, 1997).

Thurstonian theory shows that there are different cognitive strategies associated with the triangle (comparison of distances strategy) and 3-AFC (skimming strategy) methods. In theory, d' is not affected by the cognitive strategy adopted. Thus, according to Thurstonian predictions, in spite of the larger proportion of correct responses in a 3-AFC method, its estimate of d' should not be significantly different from that of the triangle method.

Psychometric functions relate the proportion of correct responses to δ. The general psychometric function for the triangle method is (Bi, 2006):

$$P_c = 2\int_0^{\infty}\left\{\Phi\left[-u\sqrt{3}+\delta\sqrt{(2/3)}\right]+\Phi\left[-u\sqrt{3}-\delta\sqrt{2/3}\ \right]\right\}e^{-(\frac{1}{2})u^2}/\sqrt{2\pi}du$$

The psychometric function for the 3-AFC method is (Bi 2006):

$$P_c = \int_{-\infty}^{\infty}\Phi^2\,(u)\varphi(u-\delta)du$$

Where $\Phi(u)$ = standard normal cumulative distribution function

φ(u) = standard normal density function

It was observed experimentally that the same subjects discriminating between the same stimuli perform better (higher proportion of correct responses) in the 3-AFC than in triangle, yet have the same d' value (Delwiche and O'Mahony, 1996; Fritjers, 1981; Masouka and others, 1995; Rousseau and O'Mahony, 1997; Stillman 1993). However, all of these discrimination studies were conducted with adult subjects. Would the same predictions apply if the subjects were children?

The food industry would benefit greatly from taking into consideration the opinions of children when testing products that are intended specifically for them. Since many foods and beverages are specifically developed for children, they must be tested with children. This is particularly important, since it is recognized that children and adults differ in their acceptance of various food products (Lavin and Lawless, 2010). Acceptance for food products among consumers is mainly determined by the sensory properties of foods and beverages and the need for sound methodology for sensory testing with children has increased (Moskowitz, 1985; Guinard, 2001). However, there is no published data concerning 6 – 12 year old children's ability to perform discrimination tests.

Therefore, the research objectives were to (1) challenge the Thurstonian prediction for the triangle and 3-AFC protocols with children as subjects, (2) determine age effects and (3) determine cultural effects. With existing research collaboration children between 6 and 12 years old from Honduras (La Ceiba, Atlantida), Mexico (Tepatlaxco, Veracruz), Thailand (Nakhon-Sawan) and the United States of America (Baton Rouge, Louisiana) participated in the discrimination tests using apple juice with different sweetness concentrations as the stimuli. Throughout this dissertation the different locations will be referred to using only the name of the country.

3.2 Materials and Methods

3.2.1 Subjects

The panelists were first through sixth grade students from elementary schools in Honduras, Mexico, Thailand, and the United States of America (USA). Table 3.1 presents the distribution of subjects per country.

Table 3.1: Distribution of subjects per grade for each country

Age Group[a]	Honduras	Mexico	Thailand	USA
1	63	23	60	15
2	69	40	60	16
3	54	25	60	15
4	51	30	60	10
5	51	37	60	16
6	51	17	60	14
TOTAL	339	172	360	86

[a] Age group categories refer to grades first to sixth

Subjects were divided into six age groups based on the grade they were currently in. This was done in order to account for the variation encountered in cognitive skills among children of the same age (Gollick 2002). It was shown that with cognitive testing with children there can be up to a four year difference between the age at which 10% of the subjects can perform a determined task and the age at which 90% of the subjects can perform the same task (Gollick 2002).

Criteria for recruitment of participants were that they were between the first and sixth grade and were not allergic to any of the ingredients present in the juice products. Participants were required to have parental consent (see Appendix A) and to sign an assent form (see Appendix B) stating their willingness to participate. Both forms were approved by the Louisiana State University AgCenter Institutional Review Board prior to participating in the testing. No monetary incentive or rewards were given to subjects for participation.

3.2.2 Stimuli

All samples were presented in approximately 1.5 fl oz aliquots in 2 fl oz lidded plastic cups (ProPak$_{TM}$ Soufflé Cup Translucent Plastic 2oz, Independent Marketing Alliance, Houston, TX) and served at room temperature (approximately 25-27 ^{o}C, depending on the room temperature at the particular site, but constant in a given session).

3.2.2.1 Honduras, Mexico, USA

Stimuli consisted of regular apple juice (Mott's® Original "100% apple juice", Mott's LLP, Rye Brook, NY) and reduced sugar apple juice (Mott's for Tots®, "40% less sugar apple juice", Mott's LLP, Rye Brook, NY). From preliminary studies it was concluded that the regular product was sweeter than the reduced sugar product; therefore, sweetness was the attribute used to specify the difference among the stimuli. The former will be referred to as 'regular' apple juice or the strong (S) sample and the latter as the 'reduced sugar' apple juice or the weak (W) sample. See Appendix C for a complete list of ingredients.

The stimulus for the distractor tests and for palate cleansing was bottled water. For Honduras the water brand was Dasani (Producto Centroamericano elaborado y distribuido bajo licencia de The Coca Cola Company por Cerveceria Hondureña, S. A. de C. V., salida carretera a Puerto Cortes, San Pedro Sula, Honduras), for Mexico the water brand was Cielo® (Propinex S.A.de C.V. Guillermo Gonzalez Camarena N^{o} 600 7^{o} Piso Col. Centro de Ciudad Santa Fe Del Alvarado, Obregon, Mexico D.F. C.P. 01210 con la autorizacion de The Coca Cola Company) and for the USA the water brand was Dasani (The Coca Cola Company, Atlanta, GA, USA).

3.2.2.2 Thailand

Stimuli consisted of apple juice (Tipco, Ti&B Co.,Ltd. 90/1 Moo7, Phaholyothin Rd., Tambon Sanubtueb, Amphur Wangnoi, Phra Nakhon Si Ayutthaya 13170, Thailand). The discrimination was performed between the pure product and juice that was diluted with water to 60% by weight. From preliminary studies it was concluded that the regular product was sweeter than the diluted product; therefore, sweetness was the attribute used to specify the difference among the stimuli. The former will be referred to as 'regular' apple juice or the strong (S) sample and the latter as the 'reduced sugar' apple juice or the weak (W) sample. See Appendix C for complete list of ingredients.

The stimulus for the distractor tests and for palate cleansing was bottled water (Crystal ®, Serm Suk Co. Ltd. 700/369 Moo 6, Nongkhong, Muang, Chonburi, Thiailand).

3.2.3 Testing Conditions

In order to avoid having language as an additional source of variation in this study, all tests were conducted in the children's native language, i.e., Spanish in Honduras and Mexico, Thai in Thailand, and English in the USA. All tests were performed in the children's classroom settings at the schools. The triangle and 3-AFC tests were performed in the same session.

Before each experimental session, the children were given a presentation in order to ensure that they understood the basic logistics of the testing session (Figure 3.1). During the presentation, a review of the terms *same*, *different*, *sweet, sweeter, not sweet, sweeter, less sweet* and *less sugar* was given to the children. Also, the testing procedure was explained, along with an overview of the questionnaire (see Appendix D) and its proper fill-out.

3.2.4 Experimental Design

Each subject performed two triangle tests and two 3-AFC tests to discriminate regular apple juice from reduced sugar apple juice. Triangle and 3-AFC tests were presented in

Figure 3.1: Presentation given to children at the beginning of the testing session

succession during the same experimental session following an AB design (A = triangle and B = 3-AFC). This order was used to avoid the possible influence of the specified attribute on the unspecified method. In order to avoid fatigue and distraction, only two out of the six possible orders for triads were presented to each child. This means that the complete block of the six possible orders of tasting for the tests was divided among three different subjects. Child1: SSW – WWS, child 2: SWS – WSW, and child 3: SWW – WSS.

In order to prevent the judges from abandoning the triangle strategy, a distractor test was intermingled with the two target tests. Thus, each judge performed a total of four target tests to discriminate the regular sample from the reduced sugar sample (two triangle tests and two 3-AFC tests) and one distractor test. One particular experimental session was:

Triangle test: SSW, distractor, WWS

BREAK

3-AFC test: SSW, WWS

The hypotheses being tested for the d' values were:

H_o: Triangle $d' = $ 3-AFC d'

H_a: Triangle $d' \neq$ 3-AFC d'

3.2.5 Testing Protocol

The triadic tests were given under two possible sets of instructions. For the triangle test the nature of the difference was not specified and the instructions were: "here are three juice samples; two are the same and one is different: **circle the juice that tastes different.**" For the 3-AFC test a univariate question was posed. Sweetness was specified as the nature of the difference and the instructions were altered depending on which stimuli was the odd one. When the regular juice was the odd sample, the instructions were: "here are three juice samples; one is sweeter than the other two: **circle the juice that tastes sweeter.**" When the reduced-sugar sample was the odd one, the specific instructions were: "here are three juice samples; one is less sweet than the other two: **circle the juice that tastes less sweet.**"

Subjects began each session by cleansing their palate with water. Next, the three samples for each question were presented simultaneously to the subjects and they were instructed to taste from left to right (Figures 3.2 and 3.3). Judges cleansed their palate with water after tasting each sample. The same protocol was repeated for the other four questions. The subjects were allowed to retaste the sample with the condition that they always tasted all three samples in the order presented to them. The children were tested on the same day for both protocols and given a 10 minute break between sessions. Session lengths ranged from 30 to 60 minutes.

Research showed that sweet taste receptors, in addition to being present in the mouth cavity and esophagus, are also present in the gastrointestinal tract of humans (Jang and others, 2007; Margolskee and others, 2007). Our preliminary study showed that there was no significant difference in discrimination between swallowing and expectorating a sample (see Appendix E). In this study, the children were therefore instructed to swallow the samples as it was a more natural behavior that caused less distraction.

Figure 3.2: USA child performing the triangle test

Figure 3.3: Honduran children performing the 3-AFC test

3.2.5.1 Experimental Session Overview

The following is a detailed protocol of the testing session:

- Children filled out the child assent form
- Children were presented with the questionnaire and water
- Children filled out the top portion of the questionnaire (demographic information: name, age, gender, grade)
- Triangle Test
    - Overview presentation for the triangle test
    - The three samples for question 1 (Q1) were simultaneously presented
    - Children cleansed their palates
    - Children were instructed to taste the samples

        - Taste the first sample
        - Taste the second sample
        - Taste the third sample
    - Children were instructed to circle on the questionnaire the juice that tasted "different"
    - Children stopped
    - The process was repeated for Q2 and Q3
    - Children took a 10 minute mandatory break
- 3-AFC Test
    - Overview presentation for the 3-AFC test
    - The three samples for Q4 were simultaneously presented
    - Children cleansed their palates
    - Children were instructed to taste the samples
        - Taste the first sample
        - Taste the second sample
        - Taste the third sample
    - Child were instructed to circle the juice that was less sweet or had less sugar and stopped
    - The process was repeated again for Q5.
    - This time the child circled the juice that was sweeter or had more sugar
- End of session

3.2.6 Data Analysis

For each testing method, the number of correct responses was counted and recorded. The P_cs for both tests were used to determine the corresponding *d′* values (Ennis, 1993).

The experimental estimate *d′*, the variance of *d′*, and the test of significance were obtained using the IFPrograms software Version 8.5.0320 (The Institute for Perception, Richmond, VA, USA). The *d′* values were calculated under the assumptions that the intensity distributions for the two stimuli were unidimenasional normal distributions that had equal variance (Bi and others 1997). Alternatively, *d′* values for the triangle and 3-AFC tests could be obtained from published tables (e.g., Ennis 1993, see Appendix F). The variance of *d′* and the test statistic can be obtained by the approach described by Bi and others (1997) (see Appendix G and H). All analyses were evaluated at $\alpha = 0.05$.

3.3 Results

3.3.1 Honduras

For the Honduran population (N=339), P_cs, *d′* values, and the variance of *d′* for the triangle and 3-AFC protocols are presented in Tables 3.2 and 3.3. For Tables 3.2 and 3.3 the upper section refers to the questions where the odd sample was the reduced sugar apple juice/weak stimulus (WSS, SWS, or SSW) and the lower section refers to the questions where the odd sample was the regular apple juice/strong stimulus (SWW, WSW, or WWS).

3.3.1.1 Overall Performance

Table 3.4 presents the overall performance of the Honduran population for each method, the pooled P_cs, and the corresponding *d′* values. For this Honduran population, the paradox was confirmed since the P_c for the 3-AFC method (0.87) was higher than the P_c for the triangle method (0.58).

For the *d′* values of the entire population a more powerful approach of pooling the data of all subjects (N=339) was used (Braun, Rogeaux, Schnied & Rousseau, 2004). Due to two replications of each test per subject, the sample size used for each protocol was 678.

Table 3.2: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for Honduran children (first to sixth grade, N=339)

Age Group[a]	**Subjects**	**Correct Responses**	P_c[b]	*d′*	**variance *d′***
Identify the odd (WEAK[c]) sample					
1	63	28	0.44	1.16	0.134
2	69	40	0.58	1.87	0.090
3	54	39	0.72	2.63	0.126
4	51	35	0.69	2.43	0.127
5	51	29	0.57	1.82	0.122
6	51	41	0.80	3.16	0.162
Identify the odd (STRONG[d]) sample					
1	63	22	0.35	0.42	0.646
2	69	27	0.39	0.82	0.191
3	54	30	0.56	1.75	0.117
4	51	31	0.61	2.02	0.121
5	51	34	0.67	2.32	0.125
6	51	35	0.69	2.43	0.127

[a] Age group categories refer to grades first to sixth

[b] P_c = proportion of correct responses

[c] WEAK = Reduced Sugar Apple Juice

[d] STRONG = Regular Apple Juice

Table 3.3: Proportions of correct responses, *d'* values, and variance of *d'* for the 3-AFC method for Honduran children (first to sixth grade, N=339)

Age Group[a]	Subjects	Correct Responses	P_c[b]	*d'*	variance *d'*
Identify the WEAK[c] sample					
1	63	58	0.92	2.4	0.093
2	69	65	0.94	2.61	0.103
3	54	48	0.89	2.15	0.088
4	51	41	0.80	1.67	0.069
5	51	47	0.92	2.41	0.115
6	51	49	0.96	2.86	0.181
Identify the STRONG[d] sample					
1	63	42	0.67	1.12	0.045
2	69	58	0.84	1.86	0.057
3	54	50	0.93	2.45	0.113
4	51	41	0.80	1.67	0.069
5	51	45	0.88	2.11	0.091
6	51	46	0.90	2.25	0.101

[a] Age group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Reduced Sugar Apple Juice
[d] STRONG = Regular Apple Juice

Table 3.4: Pooled proportions of correct responses, pooled *d'* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the Honduran population (N=678[a])

Triangle Method			3-AFC Method			*d'* Sig Test[c]	
P_c[b]	*d'*	variance	P_c[b]	*d'*	variance	χ^2	p-value[d]
0.58	1.86	0.009	0.87	2.03	0.006	1.93	0.17

[a] N = 678 due to two replications of each of the 339 subjects
[b] P_c = proportion of correct responses
[c] Significance test for *d'* values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[d] All p-values > 0.05 are not significant

As replications from 339 different subjects were combined, it was necessary to take into account the possibility of overdispersion (Cox, 1983; Anderson, 1988) in the data. Overdispersion is added variance; therefore, it occurs when there is more than one source of variance in the data. The Binomial distribution is normally used for sensory difference or preference methods. However, the Binomial distribution cannot be applied to pooled data from multiple subjects when each subject performs more than one test. If the Binomial model is used for pooled data, the Type I error may be inflated; i.e., the probability of rejecting a true null hypothesis (H_o: Triangle d' = 3-AFC d') will increase (see Ennis and Bi, 1998 for further details).

Overdispersion can be accounted for by using the Beta-Binomial model (Bi and others, 1997), and a Thurstonian variant was developed by Bi and Ennis (1998) for replicated difference tests. The variance of d' is measured in terms of a gamma (γ) value (see Appendix I). A γ value of 0 indicates no overdispersion, and a γ value of 1 indicates full overdispersion. When comparing these two models the p-value indicates the statistical probability that the Beta-Binomial fits the data significantly better than the Binomial model. For the triangle (Figure 3.4) and 3-AFC (Figure 3.5) methods, it was observed that the Beta-Binomial model was not better (triangle: $p = 0.58$, $\gamma = 0$; 3-AFC: $p = 0.27$, $\gamma = 0.03$; IFPrograms). Therefore, overdispersion was not significant in these data and there was no need to adjust the variance of the d' values.

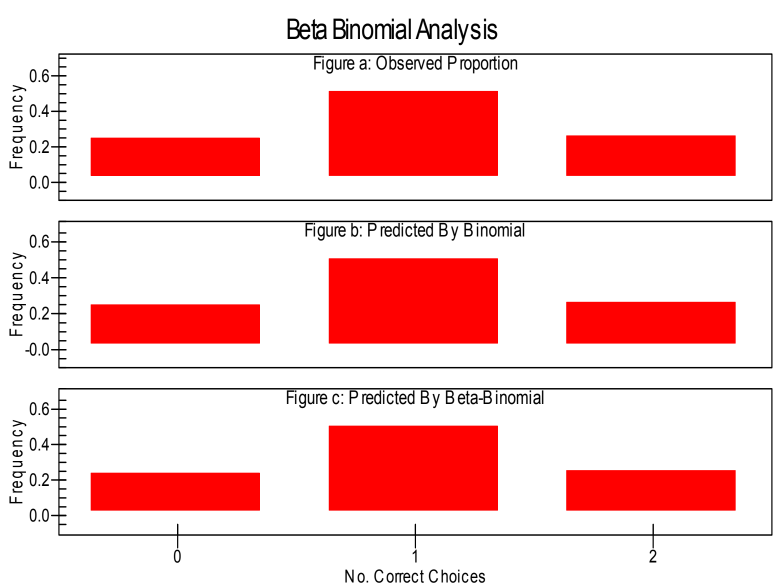


Figure 3.4: Triangle test (a) observed proportions, (b) predicted probabilities by binomial distribution, and (c) predicted probabilities by beta-binomial distribution.

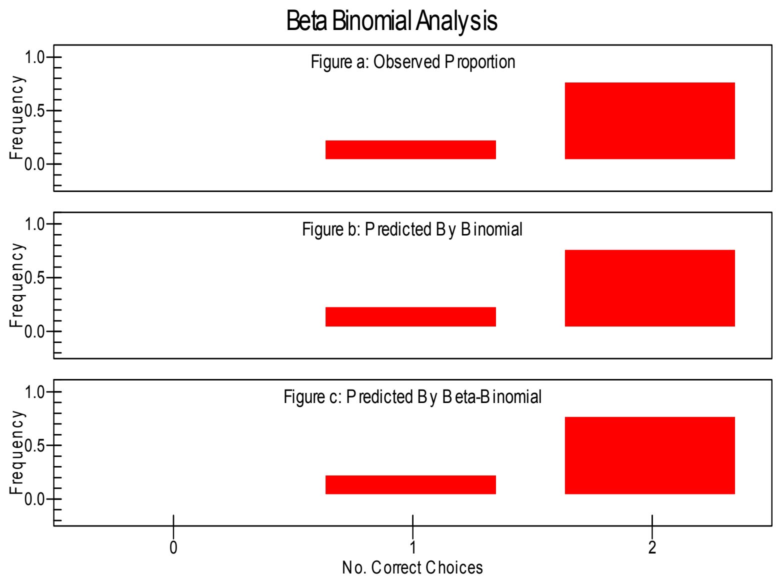


Figure 3.5: 3-AFC test (a) observed proportions, (b) predicted probabilities by binomial distribution, and (c) predicted probabilities by beta-binomial distribution.

For the triangle test, there were 58% correct responses, which corresponded to a *d′* of 1.86. In the 3-AFC task, 87% correct selections were made, which corresponded to a *d′* of 2.03. Comparing pooled *d′* values for each protocol showed no significant difference among them (p = 0.17) (Table 3.4). From these results, Thurstonian predictions were confirmed for Honduran children between the ages of 6 and 12, using apple juices with different sweetness concentrations (100% vs. 60%) as the stimuli. Furthermore, it confirms the assumptions regarding the decision rules that the Thurstonian approach associates with each method. The "skimming" strategy used for the 3-AFC method did elicit superior performance than the "comparison of distances" strategy associated with the triangle method.

3.3.1.2 Individual Performance

Performance of the different age groups was further explored for each protocol, with results presented in Table 3.5. For all age groups, the 3-AFC P_c was higher than the triangle P_c, thereby confirming the paradox. There was an overall increase in performance (P_c) with age for both protocols. For the triangle test, P_c increased from 40% for the first grade to 75% for sixth graders. The increase for the 3-AFC method was from 79% for first graders to 93% for sixth graders.

The *d′* values for the individual grades are also presented in Table 3.5. Overdispersion was assessed for each grade's pooled responses (from two replications) and was not significant. The *d′* values of the third to the sixth grade were not significantly different, confirming the Thurstonian predictions. However, the results for the first and second graders do not confirm these predictions ($p = 0.03$ and $p = 0.01$, respectively).

Table 3.5: Proportions of correct responses, *d′* values and significance tests for the triangle and 3-AFC methods for the individual grades for the Honduran population

		Triangle			3-AFC			*d′* Sig Test [c]	
Grade	**N[a]**	P_c [b]	***d′***	**variance**	P_c [b]	***d′***	**variance**	χ^2	**p-value** [d]
1	126	0.40	0.80	0.108	0.79	1.62	0.027	4.98	0.03
2	138	0.49	1.39	0.052	0.89	2.17	0.035	6.99	0.01
3	108	0.64	2.18	0.058	0.91	2.29	0.049	0.11	0.74
4	102	0.65	2.22	0.061	0.80	1.67	0.035	3.15	0.08
5	102	0.62	2.07	0.061	0.90	2.25	0.050	0.29	0.59
6	102	0.75	2.77	0.070	0.93	2.50	0.063	0.54	0.46

[a] N = twice the number of subjects per grade, due to two replications of each subject
[b] P_c = proportion of correct responses
[c] Significance test for *d′* values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[d] All p-values > 0.05 are not significant

The difference in performance could be due to the differences in cognitive abilities of children; in this study first and second grade children were mainly 6-8 years old and third to sixth grade children 9-12 years old. The degree to which children can perform triadic discrimination tests remains unknown because there is limited data in the literature. Further research in this area should be explored.

3.3.2 Mexico

For the Mexican population (N = 172), the proportions of correct responses, corresponding *d′* values, and the variance of *d′* for the triangle method are presented in Table 3.6. The same set of results is presented in Table 3.7 for the 3-AFC test. Again, the upper section of the table refers to the questions where the odd sample was the reduced sugar apple juice/weak stimulus (WSS, SWS, or SSW) and the lower section refers to the questions where the odd sample was the regular apple juice/strong stimulus (SWW, WSW, or WWS).

Table 3.6: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for Mexican children (first to sixth grade, N=172)

Age Group[a]	**Subjects**	**Correct Responses**	P_c[b]	***d′***	**variance *d′***
Identify the odd (WEAK[c]) sample					
1	23	14	0.61	2.02	0.269
2	40	26	0.65	2.23	0.157
3	25	18	0.72	2.62	0.272
4	30	26	0.87	3.67	0.351
5	37	30	0.81	3.21	0.227
6	17	17	1.00	5.33[c]	2.042
Identify the odd (STRONG[d]) sample					
1	23	8	0.35	0.4	1.929
2	40	17	0.43	1.04	0.237
3	25	12	0.48	1.36	0.291
4	30	23	0.77	2.9	0.248
5	37	22	0.59	1.95	0.167
6	17	13	0.76	2.89	0.436

[a]Age group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Reduced Sugar Apple Juice
[d] STRONG = Regular Apple Juice

Table 3.7: Proportions of correct responses, *d'* values, and variance of *d'* for the 3-AFC method for Mexican children (first to sixth grade, N = 172)

Age Group[a]	Subjects	Correct Responses	P_c[b]	*d'*	variance *d'*
Identify the WEAK[c] sample					
1	23	15	0.65	1.06	0.122
2	40	25	0.63	0.97	0.069
3	25	22	0.88	2.09	0.183
4	30	28	0.93	2.52	0.217
5	37	36	0.97	3.08	0.323
6	17	16	0.94	2.6	0.414
Identify the STRONG[d] sample					
1	23	20	0.87	2.02	0.19
2	40	28	0.70	1.24	0.074
3	25	24	0.96	2.85	0.364
4	30	30	1.00	3.35[c]	0.644
5	37	30	0.81	1.7	0.097
6	17	15	0.88	2.11	0.272

[a] Age Group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Reduced Sugar Apple Juice
[d] STRONG = Regular Apple Juice

3.3.2.1 Overall Performance

Overall performance data of the Mexican population is presented in Table 3.8. The P_c of the 3-AFC method (0.84) was higher than that of the triangle (0.66), confirming the paradox of discriminatory non-discriminators. Overdispersion was not present in the triangle data ($p = 0.86$, $\gamma = 0$), but was present in the 3-AFC data ($p = 0.0007$, $\gamma = 0.24$). Therefore, the 3-AFC variance of *d'* increased from 0.011 to 0.014. The variance of 0.014 is the one shown in Table 3.8. For the triangle test there were 66% correct responses, which corresponded to a *d'* of 2.27. For the 3-AFC task, 84% correct selections were made, which corresponded to a *d'* of 1.85. Thurstonian predictions were not confirmed (p = 0.02) for this Mexican population. This population detected a greater degree of difference among the stimuli with the triangle test. These results are different than those obtained for the Honduran population.

Table 3.8: Pooled proportions of correct responses, pooled d' values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the Mexican population (N=344[a])

Triangle Method			3-AFC Method			d' Sig Test [c]	
P_c [b]	d'	variance	P_c [b]	d'	variance	χ^2	p-value [d]
0.66	2.27	0.018	0.84	1.85	0.014	5.5	0.02

[a] N = 344 due to two replications of each of the 172 subjects
[b] P_c = proportion of correct responses
[c] Significance test for d' values, Critical χ^2 value = 3.84, α = 0.05
[d] All p-values > 0.05 are not significant

3.3.2.2 Individual Performance

For the individual age groups, results showed that the 3-AFC P_c was higher than the triangle P_c, except for the second and sixth grades (Table 3.9). The higher P_cs observed for the 3-AFC confirm the superior performance elicited by the "skimming" strategy. As with the Honduran population, it was observed that there was an increase in performance (P_c) with age for both protocols. For the triangle method the P_cs increased from 48% for the first grade to 88% for sixth graders. The increase for the 3-AFC method was from 76% to 91% for first to sixth graders.

Table 3.9: Proportions of correct responses, d' values and significance tests for the triangle and 3-AFC methods for the individual grades for the Mexican population

		Triangle			3-AFC			d' Sig Test [c]	
Grade	N[a]	P_c [b]	d'	variance	P_c [b]	d'	variance	χ^2	p-value [d]
1	46	0.48	1.35	0.159	0.76	1.48	0.070	0.07	0.79
2	80	0.54	1.66	0.08	0.66	1.10	0.053	2.36	0.12
3	50	0.60	1.98	0.124	0.92	2.39	0.116	0.7	0.40
4	60	0.82	3.25	0.143	0.97	2.96	0.172	0.27	0.61
5	74	0.70	2.52	0.09	0.89	2.17	0.065	0.79	0.37
6	34	0.88	3.83	0.336	0.91	2.32	0.262	3.81	0.05

[a] N twice the number of subjects per grade; due to two replications of each subject
[b] P_c = proportion of correct responses
[c] Significance test for d' values, Critical χ^2 value = 3.84, α = 0.05
[d] All p-values > 0.05 are not significant

The d' values are also presented in Table 3.9. Overdispersion was assessed for each grade's pooled responses for the triangle test and was found not to be significant (all $p > 0.05$). For the 3-AFC method, overdispersion was only present in second ($p = 0.0008$, $\gamma = 0.497$) and sixth ($p = 0.005$, $\gamma = 0.63$) grade. The variance for the second grade increased from 0.035 to 0.053 and the variance for the sixth grade increased from 0.161to 0.262. The adjusted variances are the ones show in Table 3.9. Results show that d' values were not significantly different for the first to the sixth grade (all $p > 0.05$), as predicted by Thurstonian modeling. However, as noted above, the pooled d' values for the entire Mexican population were significantly different from each other.

3.3.3 Thailand

For the Thai population (N = 360), the proportions of correct responses, d' values, and the variance of d' for the triangle method are presented in Table 3.10. The 3-AFC results are presented in Table 3.11.

3.3.3.1 Overall Performance

For the entire Thai population, the 3-AFC P_c (0.96) was higher than the triangle P_c (0.86) (Table 3.12). This confirms the existence of the paradox of discriminatory non-discriminators. For the pooled d' values of the entire population the sample size used was 720 due to two replications of each protocol per subject. Overdispersion was present in the triangle test data ($p = 0.0005$, $\gamma = 0.1732$) and the 3-AFC data ($p = 0.02$, $\gamma = 0.1082$). Therefore, the triangle method variance of d' increased from 0.014 to 0.043. For the 3-AFC method the variance increased from 0.013 to 0.014. These adjusted variances are the ones shown in Table 3.12.

Table 3.10: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for Thai children (first to sixth grade, N = 360)

Age Group[a]	**Subjects**	**Correct Responses**	P_c[b]	***d′***	**variance *d′***
Identify the odd (WEAK[c]) sample					
1	60	40	0.67	2.32	0.106
2	60	45	0.75	2.8	0.12
3	60	51	0.85	3.52	0.162
4	60	59	0.98	5.86	0.794
5	60	53	0.88	3.84	0.192
6	60	56	0.93	4.49	0.283
Identify the odd (STRONG[d]) sample					
1	60	51	0.85	3.52	0.162
2	60	46	0.77	2.9	0.124
3	60	55	0.92	4.24	0.242
4	60	57	0.95	4.8	0.348
5	60	59	0.98	5.86	0.794
6	60	50	0.83	3.38	0.152

[a] Age group categories refer to grades first to sixth

[b] P_c = proportion of correct responses

[c] WEAK = Reduced Sugar Apple Juice

[d] STRONG = Regular Apple Juice

Table 3.11: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for Thai children (first to sixth grade, N=360)

Age Group[a]	**Subjects**	**Correct Responses**	P_c[b]	***d′***	**variance *d′***
Identify the WEAK[c] sample					
1	60	56	0.93	2.52	0.108
2	60	55	0.92	2.36	0.094
3	60	59	0.98	3.35	0.283
4	60	60	1.00	3.71	0.537
5	60	56	0.93	2.52	0.108
6	60	58	0.97	2.96	0.172
Identify the STRONG[d] sample					
1	60	58	0.97	2.96	0.172
2	60	53	0.88	2.11	0.077
3	60	60	1.00	3.71	0.537
4	60	59	0.98	3.35	0.283
5	60	60	1.00	3.71	0.537
6	60	58	0.97	2.96	0.172

[a] Age group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Reduced Sugar Apple Juice
[d] STRONG = Regular Apple Juice

The *d′* values of the triangle (3.65) and 3-AFC (2.87) methods were significantly different ($p < 0.001$). From these results, Thurstonian predictions were not confirmed for the triadic tests among Thai children. A greater degree of discrimination was detected among the stimuli with the triangle test, which was the same result observed for the Mexican population.

Table 3.12: Pooled proportion of correct responses, pooled *d′* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the Thai population (N= 720[a])

Triangle Method			**3-AFC Method**			***d′* Sig Test** [c]	
P_c [b]	***d′***	**variance**	P_c [b]	***d′***	**variance**	χ^2	**p-value** [d]
0.86	3.65	0.042	0.96	2.87	0.014	22.53	< 0.001

[a] N = 722 due to two replications of each of the 360 subjects
[b] P_c = proportion of correct responses
[c] Significance test for *d′* values, Critical χ^2 value = 3.84, α = 0.05
[d] All p-values > 0.05 are not significant

3.3.3.2 Individual Performance

Performance in each individual grade was further explored. The paradox of discriminatory non-discriminators was confirmed for all grades because the P_cs for the 3-AFC method were higher than those for the triangle. Results show that P_c increased with age for both protocols. For the triangle test the P_c increased from 76% for the first grade to 88% for sixth graders. The increase for the 3-AFC method was from 95% for first graders to 97% for sixth graders.

Table 3.13: Proportions of correct responses, *d′* values, test statistics and p-values for the triangle and 3-AFC methods for the individual grades for the Thai population

		Triangle			3-AFC			*d′* Sig Test [c]	
Grade	**N[a]**	P_c [b]	*d′*	**variance**	P_c [b]	*d′*	**variance**	χ^2	**p-value** [d]
1	**120**	0.76	2.85	0.061	0.95	2.71	0.065	0.16	0.69
2	**120**	0.76	2.85	0.061	0.90	2.23	0.053	3.73	0.05
3	**120**	0.88	3.84	0.130	0.99	3.71	0.239	0.05	0.82
4	**120**	0.97	5.21	0.234	0.99	3.71	0.239	4.76	0.03
5	**120**	0.93	4.49	0.141	0.97	2.96	0.086	10.31	< 0.001
6	**120**	0.88	3.84	0.096	0.97	2.96	0.086	4.25	0.04

[a] N twice the number of subjects per grade; due to two replications of each subject
[b] P_c = proportion of correct responses
[c] Significance test for *d′* values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[d] All p-values > 0.05 are not significant

The *d′* values of the individual grades are presented in Table 3.13. Overdispersion was found to be present only in the second grade 3-AFC data ($\gamma = 0.35$, $p = 0.003$) and the third grade triangle data ($\gamma = 0.26$, $p = 0.022$). The variance increased from 0.096 to 0.13 for the second grade 3-AFC data and from 0.042 to 0.053 for the third grade triangle data. The adjusted variances are the ones shown in Table 3.13. Results show that *d′* values for the first to third grade are not significantly different from each other. However, *d′* values of the fourth to sixth grade were found to be significantly different. Therefore, Thurstonian predictions were only confirmed for first to third graders, but not for the higher grades.

3.3.4 USA

For the United States population (N = 86), the P_cs, d' values, and the variance of d' for the triangle and 3-AFC methods are presented in Tables 3.14 and 3.15, respectively.

Table 3.14: Proportions of correct responses, d' values, and variance of d' for the triangle method for US children (first to sixth grade, N = 86)

Age Group[a]	Subjects	Correct Responses	P_c[b]	d'	variance d'
Identify the odd (WEAK[c]) sample					
1	15	7	0.47	1.29	0.508
2	16	10	0.63	2.10	0.387
3	15	13	0.87	3.67	0.701
4	10	8	0.80	3.13	0.814
5	16	15	0.94	4.56	1.111
6	14	14	1.00	5.14	2.147
Identify the odd (STRONG[d]) sample					
1	15	10	0.67	2.32	0.424
2	16	6	0.38	0.69	1.075
3	15	10	0.67	2.32	0.424
4	10	10	1.00	4.8	2.347
5	16	13	0.81	3.22	0.529
6	14	12	0.86	3.59	0.718

[a] Age group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Reduced Sugar Apple Juice
[d] STRONG = Regular Apple Juice

3.3.4.1 Overall Performance

For the overall USA population the 3-AFC P_c (0.86) was larger than the triangle P_c (0.74), confirming the paradox. This outcome was the same one observed for the Honduran, Mexican and Thai population. Therefore, regardless of the children's culture the paradox was present. This confirms the argument for triadic methods that knowing the sensory attribute responsible for the difference among the stimuli elicits a decision rule that results in better performance in the 3-AFC method.

Table 3.15: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for US children (first to sixth grade, N = 86)

Age Group[a]	**Subjects**	**Correct Responses**	P_c [b]	***d′***	**variance *d′***
		Identify the WEAK[c] sample			
1	15	10	0.67	1.12	0.19
2	16	14	0.88	2.06	0.279
3	15	11	0.73	1.37	0.206
4	10	10	1.00	2.71	0.914
5	16	14	0.88	2.06	0.279
6	14	14	1.00	2.92	0.814
		Identify the STRONG[d] sample			
1	15	12	0.80	1.65	0.233
2	16	12	0.75	1.43	0.198
3	15	14	0.93	2.52	0.433
4	10	10	1.00	2.71	0.914
5	16	14	0.88	2.06	0.279
6	14	13	0.93	2.47	0.445

[a] Age group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Reduced Sugar Apple Juice
[d] STRONG = Regular Apple Juice

Table 3.16: Pooled proportions of correct responses, pooled *d′* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the USA population (N= 172[a])

Triangle Method			**3-AFC Method**			***d′* Sig Test** [c]	
P_c [b]	***d′***	**variance**	P_c [b]	***d′***	**variance**	χ^2	**p-value** [d]
0.74	2.76	0.041	0.86	1.97	0.029	8.9	0.003

[a] N = 172 due to two replications of each of the 86 subjects
[b] P_c = proportion of correct responses
[c] Significance test for *d′* values, Critical χ^2 value = 3.84, α = 0.05
[d] All p-values > 0.05 are not significant

For the pooled *d′* values of the entire population the sample size used was 172 due to the two replications per subject. Overdispersion was not significant in the triangle data ($p = 0.42$, $\gamma = 0.03$) but was in the 3-AFC data ($p = 0.02$, $\gamma = 0.2252$). Due to overdispersion, the variance of *d′* for the 3-AFC method increased from 0.024 to 0.029. The adjusted variance is the one shown in Table 3.16. For the triangle test the 74% correct selections made of the odd stimulus was equivalent to a d' of 2.76. In the 3-AFC task, 86% correct selections were made, which corresponded to a d' of 1.97.

There was a significant difference among these d' values (p = 0.002) (Table 3.16). From these results, Thurstonian predictions were not confirmed for the triadic tests among USA children using apple juice with different sweetness concentrations as the stimulus. As with the Mexican and Thai populations, the USA population detected a greater degree of difference among the stimuli with the triangle method.

3.3.4.2 Individual Performance

Results from the assessment of performance in each individual grade show that the P_cs were higher for the 3-AFC method when compared to those of the triangle test and the paradox was confirmed. There was also an increase in performance (P_c) with age for both protocols. For the triangle test the increase was from 57% for the first grade to 93% for sixth graders. The increase for the 3-AFC method was from 73% to 96% from first to sixth grade.

Table 3.17: Proportions of correct responses, d' values, test statistics and p-values for the triangle and 3-AFC methods for the individual grades for the USA population

		Triangle			**3-AFC**			**d' Sig Test** [c]	
Grade	**N**[a]	**P_c** [b]	**d'**	**variance**	**P_c** [b]	**d'**	**variance**	**χ^2**	**p-value** [d]
1	**30**	0.57	1.81	0.208	0.73	1.37	0.103	0.62	0.43
2	**32**	0.50	1.47	0.215	0.81	1.71	0.113	0.18	0.68
3	**30**	0.77	2.9	0.357	0.83	1.82	0.128	3.10	0.08
4	**20**	0.90	4.03	0.639	0.98	3.13	0.726	0.59	0.44
5	**32**	0.88	3.76	0.343	0.88	2.06	0.139	5.99	0.01
6	**28**	0.93	4.42	0.577	0.96	2.92	0.351	2.42	0.12

[a] N twice the number of subjects per grade; due to two replications of each subject
[b] P_c = proportion of correct responses
[c] Significance test for d' values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[d] All p-values > 0.05 are not significant

The d' values of the individual grades are presented in Table 3.17. Overdispersion was assessed for each grade's pooled responses and was found to be significant only in the third grade triangle data ($\gamma = 0.441$, $p = 0.04$). The variance increased from 0.248 to 0.357, and is the one presented in Table 3.17. Results show that d' values are not significantly different for all age groups, except for fifth grade. Fifth graders had equal P_cs (0.88) for both tests; hence, the d' values were significantly different. Therefore, Thurstonian predictions were confirmed for the Thai children from the first to fourth and sixth grades.

3.4 Discussion

3.4.1 Overall Performance

This study explored the paradox of discriminatory non-discriminators by extending it to children as subjects using apple juices (100% vs. 60%) as the stimuli. The difference in P_cs observed for the triangle and 3-AFC tests is attributed to the cognitive strategy associated with each method. However, Thurstonian theory states that the degree of difference (d') between the two stimuli should be independent of the methodology used.

Table 3.18: Summary of pooled proportions of correct responses, d' values and significance tests for the triangle and 3-AFC methods overall and for the individual countries

		Triangle			**3-AFC**			**d' Sig Test** [b]	
Country	**N**	**P_c** [a]	**d'**	**variance**	**P_c** [a]	**d'**	**variance**	**χ^2**	**p-value** [c]
Honduras	**678**	0.58	1.86	0.009	0.87	2.03	0.006	1.93	0.17
Mexico	**344**	0.66	2.27	0.018	0.84	1.85	0.014	5.50	0.02
Thailand	**720**	0.86	3.65	0.042	0.96	2.87	0.014	22.53	< 0.001
USA	**172**	0.74	2.76	0.041	0.86	1.97	0.029	8.90	0.003
OVERALL	**1914**	0.71	2.58	0.004	0.90	2.22	0.003	18.50	< 0.001

[a] P_c = proportion of correct responses
[b] Significance test for d' values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[c] All p-values > 0.05 are not significant

For these stimuli, the 3-AFC P_c was significantly higher than the triangle P_c for each individual population and for all populations combined, thus confirming the paradox (Table 3.18) and indicating that cultural effects were not present under the conditions of this study. These results for the P_cs are in accord with studies performed with adults by Byer and Abrams (1953) for bitterness perception, Fritjers (1981) for the linalylacetate discrimination (N = 24), Stillman (1993) and Tedja and others (1994) for saltiness discrimination (N =144, N = 3, respectively), Masouka and others (1995) for beer bitterness detection (N=9), Delwiche and O'Mahony (1996) for sweetness discrimination using chocolate pudding (N =13), and Rousseau and O'Mahony (1997) for sweetness discrimination using vanilla flavored yogurts (N = 15).

For all countries combined, the d' values of the triangle (2.58) and 3-AFC (2.22) methods were significantly different from each other; consequently, the Thurstonian predictions were not confirmed. The same result was observed for the Mexican, Thai and USA individual populations. Only for the Honduran population were Thurstonian predictions (triangle d' = 1.86, 3-AFC d' = 2.03) confirmed. For the Mexican, Thai and USA populations the d' values for the triangle method were higher, indicating that a greater degree of difference was detected among the regular (strong) and reduced sugar (weak) apple juices with the triangle method. Pooled results from the Honduran population confirmed those observed in other studies utilizing adult subjects (Delwiche and O'Mahony 1995, Masouka and others 1995, Rousseau and O'Mahony 1997, Stillman 1993, Tedja and others 1994). However, these results were not observed for the other three populations or for the overall population. Contrary to the absence of cultural effects regarding the paradox, it was observed that cultural affects were present regarding Thurstonian predictions. More research is needed to determine of this is an occasional result or if a cultural effect is in fact present.

Confirmation of the paradox, but not the similarity of the d' values for 100% vs. 60% apple juice can be explained by the nature of the stimuli. The difference between the two stimuli (100% vs. 60%) was probably obvious to the panelists and therefore resulted in higher triangle d' values than can be expected. Even if the difference between the samples was not specified it was "easy" to detect the odd sample. For example, for a 3AFC P_c of 0.84 for the Mexican population the expected triangle P_c is around 0.57. For the Thai and the USA populations the expected triangle P_cs are 0.76 (3-AFC $P_c = 0.96$) and 0.60 (3-AFC $P_c = 0.86$), respectively. For the Honduran population, where the actual triangle P_c was 0.58, the expected triangle P_c for a 3-AFC P_c of 0.87 is approximately 0.60. This study should be performed with more confusable stimuli (Chapter 4).

It could also be advantageous to perform a more in depth examination of the children's performance on triads with the weak stimulus as the odds sample versus triads with the strong stimulus as the odd sample. A lack of agreement among the methods may be due to sequence effects (Ennis, 2003). It would be interesting to see if Sequential Sensitivity Analysis predictions for triads are confirmed for children as subjects.

3.4.2 Individual Performance

For each population, the 3-AFC P_cs for each individual grade were higher than those of the triangle test, confirming the paradox. It was observed as a general trend that performance (P_c) increased with age for the triangle and 3-AFC methods in all countries. For the Honduran population the P_cs for the triangle test increased from first to sixth grade from 0.40 to 0.75 and for the 3-AFC test from 0.79 to 0.93. For the Mexican population the increase from first to sixth grade was from 0.48 to 0.88 for the triangle method and 0.76 to 0.91 for the 3-AFC method. Similar trends were also observed for the Thai and the USA populations.

Thurstonian theory states that regardless of the higher number of correct responses on the 3-AFC test the *d′* values for this test and the triangle test should not be significantly different. A significant difference among *d′* values was not present in the third to sixth grade for the Honduran population, first to sixth grade for the Mexican population, first to third grade for the Thai population, and first to fourth and sixth grade for the USA population. Therefore, individual grade performance in the different countries showed an inconsistent trend.

Table 3.19 shows the pooled results from all four populations for all the individual grades. These pooled results show that there was an increase of performance with age. The P_c increased for the first to the sixth grade for the triangle (56% - 84%) and 3-AFC (84% - 95%) tests. All the 3-AFC P_cs were significantly higher than the triangle test P_cs, confirming the paradox. Thurstonian predictions were confirmed for the first to third and fifth grades. However, the triangle test's *d′* values were significantly higher for the fourth and sixth grades. In conclusion, the paradox and Thurstonian predictions were confirmed for the first, second, third, and fifth graders.

Table 3.19: Summary of P_c and *d′* values for the individual grades pooled from the Honduran, Mexican, Thai and US population.

		Triangle			3-AFC			*d′* Sig Test [b]	
Grade	**N**	P_c [a]	***d′***	**variance**	P_c [a]	***d′***	**variance**	χ^2	**p-value** [c]
1	**322**	0.56	1.77	0.020	0.84	1.86	0.012	0.25	0.61
2	**370**	0.59	1.91	0.017	0.84	1.84	0.010	0.18	0.67
3	**308**	0.74	2.74	0.023	0.94	2.54	0.021	0.91	0.34
4	**302**	0.79	3.07	0.026	0.92	2.43	0.020	8.9	0.003
5	**328**	0.75	2.78	0.022	0.92	2.40	0.018	3.61	0.06
6	**284**	0.84	3.42	0.033	0.95	2.67	0.027	9.38	0.002

[a] P_c = proportion of correct responses
[b] Significance test for *d′* values, Critical χ^2 value = 3.84, α = 0.05
[c] All p-values > 0.05 are not significant

Effects of the testing procedure were observed. Children did not want to terminate the taste test. Sometimes children were influenced by their peers. Even though they were strictly instructed to keep to themselves, in certain occasions they were easily distracted and difficult to control them. It was observed that first and second grade children were inclined to second guess themselves, resulting in their selecting the incorrect sample as the odd one after first having selected the correct one. Testing might have been more effective if the children were tested individually. However, this was not possible due the number of children that participated and the restrictions of the school officials. The desk arrangement may have distracted the younger children. Younger children may have chosen the sample they preferred or not preferred instead of the odd sample.

While the ability of individual children varied, the results of this study strongly suggest that children of different ages can perform triangle and 3-AFC tests. Children can be a valuable measuring tool for sensory testing, but one must keep in mind that they are a special population and the appropriate testing environment and protocols must be used. Kimmel and others (1994) made a statement with which the investigator of this study agrees with: "it is apparent when working with children that one is faced with a wide range of observable differences in attention span, intelligence, problem solving abilities, parental upbringing, and cognitive abilities that must be addressed". This study has given rise to additional questions. Would one expect the same results if the complexity of the stimuli increased? How would children perform in different discrimination tests? This study will be performed with more confusable stimuli in order to corroborate the findings of this study (Chapter 4).

3.5 Conclusion

The paradox of discriminatory non-discriminators was present for the total population and the individual countries using apple juice with different sweetness concentrations as the stimuli. However, Thurstonian predictions that the degree of difference among the two stimuli should be independent of the discrimination test utilized were only confirmed for the Honduran population.

3.6 Future Work

The paradox of discriminatory non-discriminators and Thurstonian predictions regarding the degree of difference among the stimuli detected with the triangle and 3-AFC methods will be further explored using more confusable stimuli (i.e., carbonated beverages sweetened with different sweeteners).

3.7 References

Anderson DA. 1988. Some models for overdispersed binomial data. Australian Journal of Statistics. 30: 125-148.

ASTM Committee E-18.2003.Standard guide for sensory evaluation of products by children E 2299-03. ASTM, Philadelphia, PA.

Bi J. 2006. Sensory discrimination tests and measurements: statistical principles, procedures and tables. Blackwell Publishing. Iowa. 298 p.

Bi J, Ennis DM, O’Mahony M. 1997. How to estimate the variance of d’ from difference tests. Journal of Sensory Studies. 12:87-104.

Bi J, Ennis DM. 1998. A Thurstonian variant of the beta-binomial model for replicated difference tests. Journal of Sensory Studies. 13: 461-466.

Braun V, Rogeaux M, Schneid N, O’Mahony M, Rousseau B. 2004. Corroborating the 2-AFC and 2-AC Thurstonian models using both a model system and sparkling water. Food Quality and Preference. 15: 501-507.

Byer AJ, Abrams D. 1953. A comparison of the triangular and two-sample taste-test methods. Food Technology. 7:185-187.

Cox DR.1983. Some remarks on overdispersion. Biometrika. 70: 269-274.

Delwiche J, O'Mahony M. 1996. Flavour discrimination – an extension of Thurstonian paradoxes to the tetrad method. Food Quality and Preference. 7:1-5.

Ennis DM. 1990. Relative power of difference testing methods in sensory evaluation. Food Technology. 44(4):114-118.

Ennis DM. 1993. The power of sensory discrimination methods. Journal of Sensory Studies. 8:353-370.

Ennis DM. 2003. Foundations of sensory science. In: Viewpoints and controversies in sensory science and consumer product testing. Moskowitz HR, Munos AM, Gacula MC. Food and Nutrition Press, Inc. Connecticut. p. 477.

Ennis DM, Bi J. 1998. The beta-binomial model: accounting for intertribal variation in replicated difference and preference tests. Journal of Sensory Studies. 13: 389-412.

Fritjers JER. 1979a. The paradox of discriminatory nondiscriminators resolved. Chemical Senses and Flavour. 4(4):355-358.

Fritjers JER. 1981. An olfactory investigation of the compatibility of oddity instructions with the design of a 3-AFC signal detection task. Acta Psychologica. 49: 1-16.

Green DM, Swets JA. 1966. Signal Detection Theory and Psychophysics. Wiley, New York.

Gridgeman NT. 1970. A re-examination of the two-stage triangle test for the perception of sensory differences. Journal of Food Science. 35:87-91.

Guinard JX. 2001. Sensory and consumer testing with children. Trends in Food Science and Technology 11, 273-283.

Jang HJ, Kokrashvili Z, Theodorakis MJ, Carlson O, Kim BJ, Zhou J, Kim HH, Xu X, Chan SL, Juhaszova M, Bernier M, Mosinger B, Robert F. Margolskee RF, Egan JM. 2007. Gut-expressed gustducin and taste receptors regulate secretion of glucagon-like peptide-1. Proceedings National Academy of Science U.S.A. 104 (38): 15069-15074.

Lavin JG, Lawless HT. 1998. Effects of color and odor on judgments of sweetness among children and adults. Food Quality and Preference. 9(4): 283-289.

Margolskee RF, Dyer J, Kokrashvili Z, Salmon KSH, Ilegems E, Daly K, Maillet EL, Ninomiya Y, Mosinger B, Shirazi-Beechey SP. 2007. T1R3 and gustducin in gut sense sugars to regulate expression of Na+-glucose cotransporter 1. Proceedings of the National Academy of Science U.S.A. 104 (38): 15075-15080.

Masouka S, Hatjopoulos D, O'Mahony M. 1995. Beer bitterness detection: Testing Thurstonian and sequential sensitivity analysis models for triad and tetrad methods. Journal of Sensory Studies. 10:295-306.

Moskowitz HR. 1985. Product testing with children. In "New direction for product testing and sensory analysis of foods". Food and Nutrition Press Inc. Westport, CT. 371p.

O'Mahony M. 1995a. Who told you the triangle test was simple? Food Quality and Preference. 6:227-238.

O'Mahony M, Masuoka S, Ishii R. 1994. A theoretical note on difference tests: models, paradoxes and cognitive strategies. Journal of Sensory Studies. 9: 247-272.

O'Mahony M, Rousseau B. 2002. Discrimination testing: A few ideas, old and new. Food Quality and Preference. 14:157-164.

O'Mahony M, Thieme U, Goldstein L. 1988. The warm-up effect as a means of increasing the discriminability of sensory difference tests. Journal of Food Science. 53: 1848-1850.

Peryam DR, Swartz VW. 1950. Measurement of sensory differences. Food Technology. 4: 390-395.

Rousseau B, O'Mahony M. 1997. Sensory difference tests: Thurstonian and SSA predictions for vanilla flavored yogurts. Journal of Sensory Studies. 12:127-146.

Stillman JA. 1993. Response selection, sensitivity, and taste test performance. Perception and Psychophysics. 54:190-194.

Tedja S, Ninaka R. Ennis DM, O'Mahony M. 1994. Triadic discrimination testing – refinement of Thurstonian and SSA approaches. Chemical Senses. 19:279-301.

Thurstone LL. 1927a. Psychophysical analysis. American Journal of Psychology. 38: 368-389.

Thurstone LL. 1927b. A law of comparative judgment. Psychology Review. 34:273-286.

CHAPTER 4. THE PARADOX OF DISCRIMINATORY NON-DISCRIMINATORS WITH CHILDREN AS SUBJECTS USING HARDLY DISCRIMINABLE STIMULI

4.1 Introduction

Seemingly paradoxical results are observed when discriminating confusable stimuli using the triangle (Peryam and Swartz, 1950) and 3-alternative forced choice (3-AFC) (Green and Swets 1966) methods. Subjects perform better with the 3-AFC method than with the triangle method despite having the same guessing probability (1/3). Both procedures are forced-choice methods that require the subjects to select one stimulus (S) from a set of three (SNN). For the triangle test, the nature of the difference among the stimuli is not specified, and for the 3-AFC method it is. The increase in the number of correct selections made when the instruction switched from those of the triangle method to those of the 3-AFC method was first observed by Byers and Abrams (1953). This became known as the paradox of discriminatory non-discriminators (Gridgeman, 1970) and was later resolved by Fritjers (1979) using Thurstonian (Thurstone, 1927a, b) arguments.

The perceptual intensity variations of the two stimuli are represented by normal frequency distributions (with equal variance) along a univariate intensity axis. The difference between the means of the two distributions is called δ (δ is the population parameter and d' is the experimental estimate). Thurstonian arguments state that the difference in performance is due to the decision rules associated with each test; the "comparison of distances" strategy is associated with the triangle test and the "skimming" strategy with the 3-AFC (O'Mahony and others, 1994). Thus, Thurstonian modeling explains the superior performance elicited by the "skimming" strategy and for a given d' value the proportion of correct responses for the triangle test will be less than for the 3-AFC (see Figure 4.1).

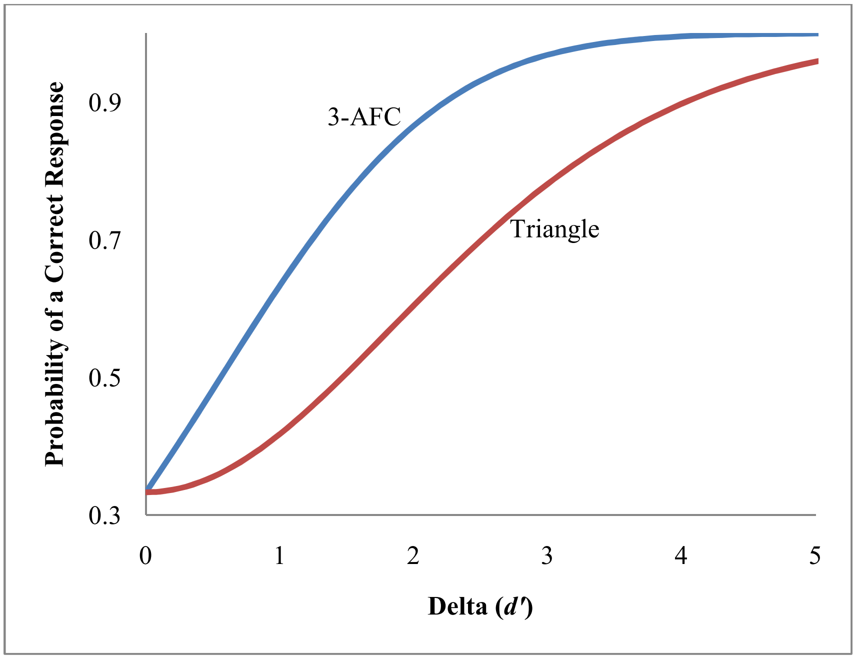


Figure 4.1: Psychometric functions for the Triangle and 3-AFC methods (source of data: Ennis, 1993)

The paradox and the Thurstonian predictions regarding this paradox were observed experimentally (Delwiche and O'Mahony, 1996; Fritjers, 1981, Masouka and others, 1995, Rousseau and O'Mahony, 1997; Stillman, 1993). However, all of these discrimination studies were conducted with adult subjects. Our study involving children discriminating between easily discriminable stimuli (apple juices with different sweetness concentrations, Chapter 3) revealed overall that children did perform a higher number of 3-AFC tests correctly than triangle tests; however, the *d′* values were significantly different.

Therefore, the research objectives were to (1) challenge the Thurstonian prediction for the triangle and 3-AFC protocols with children as subjects using stimuli whose discrimination presents higher difficulty, (2) determine age effects and (3) determine cultural effects. For this purpose, children between 6 and 12 years old in Honduras, Mexico, Thailand and The United

States of America (USA) participated in the discrimination tests using carbonated beverages with different sweeteners as the stimuli.

4.2 Materials and Methods

4.2.1 Subjects

The panelists were first through sixth grade students from elementary schools in Honduras, Mexico, Thailand, and the USA (Table 4.1). Criteria for recruitment of participants were that they were between the first and sixth grade and were not allergic to any of the ingredients present in the products. Participants were required to have parental consent and to sign an assent form stating their willingness to participate, both approved by Louisiana State University AgCenter Institutional Review Board prior to participating in the testing. No monetary incentive or rewards were given to subjects for participation.

Table 4.1: Distribution of subjects per grade for each country

Age Group[a]	Honduras	Mexico	Thailand	USA
1	51	21	60	7
2	51	38	60	22
3	51	25	60	16
4	51	35	60	12
5	51	37	60	18
6	51	23	60	18
TOTAL	306	176	360	93

[a] Age group categories refer to grades first to sixth

4.2.2 Stimuli

All samples were presented in approximately 1.5 fl oz aliquots in 2 fl oz lidded plastic cups (ProPak$_{TM}$ Soufflé Cup Translucent Plastic 2oz, Independent Marketing Alliance, Houston, TX) and served at room temperature (between 25 and 27^{o}C, depending on the room temperature at the particular site, but constant in a given session). See Appendix C for the complete list of ingredients for each product.

4.2.2.1 Honduras

4.2.2.1.1 Study I: Hardly Distinguishable Stimuli

Stimuli for the target tests consisted of a carbonated beverage sweetened with sucrose (Coca Cola® Classic, Producto Centroamericano elaborado y distribuido bajo licencia de the Coca Cola Company por: Cerveceria Hondureña, S. A. de C. V., salida carretera a Puerto Cortes, San Pedro Sula, Honduras) and a carbonated beverage sweetened with aspartame and acesulfame-K (Coca Cola® Light, Producto Centroamericano elaborado y distribuido bajo licencia de the Coca Cola Company por: Cerveceria Hondureña, S. A. de C. V., salida carretera a Puerto Cortes, San Pedro Sula, Honduras). From preliminary studies it was concluded that the Coca Cola® Classic product was sweeter than the Coca Cola® Light product; therefore, sweetness was the attribute used to specify the difference among the stimuli. The former will be referred to as the “strong” (S) sample and the latter as the “weak” (W) sample.

The stimulus for the distractor tests and for palate cleansing was bottled water (Dasani, Producto Centroamericano elaborado y distribuido bajo licencia de the Coca Cola Company por: Cerveceria Hondureña, S. A. de C. V., salida carretera a Puerto Cortes, San Pedro Sula, Honduras).

4.2.2.1.2 Study II: Hardly Distinguishable Stimuli

Stimuli for the target tests consisted of carbonated beverages sweetened with aspartame and acesulfame-K (Coca Cola Zero ® and Coca Cola® Light, Producto Centroamericano elaborado y distribuido bajo licencia de the Coca Cola Company por: Cerveceria Hondureña, S. A. de C. V., salida carretera a Puerto Cortes, San Pedro Sula, Honduras). From preliminary studies, it was concluded that the Coca Cola® Zero product was slightly sweeter that the Coca Cola® Light product; therefore, sweetness was the attribute used to specify the difference among the stimuli. The former will be referred to as the “strong” sample and the latter as the “weak” sample.

The stimulUS for the distractor tests and for palate cleansing was bottled water (Dasani, Producto Centroamericano elaborado y distribuido bajo licencia de the Coca Cola Company por: Cerveceria Hondureña, S. A. de C. V., salida carretera a Puerto Cortes, San Pedro Sula, Honduras).

4.2.2.2 Mexico

Stimuli for the target tests consisted of a carbonated beverage sweetened with sucrose (Coca Cola®, Mexico, The Coca Cola Company) and a carbonated beverage sweetened with aspartame and acesulfame-K (Coca Cola® Light, Mexico, The Coca Cola Company). From preliminary studies, it was concluded that the Coca Cola® product was sweeter than the Coca Cola® Light product; therefore, sweetness was the attribute used to specify the difference among the stimuli. The former will be referred to as the “strong” sample and the latter as the “weak” sample.

The stimulus for the distractor tests and for palate cleansing was bottled water (Cielo®, Propinex S.A.de C.V. Guillermo Gonzalez Camarena N^{o} 600 7^{o} Piso Col. Centro de Ciudad Santa Fe Del Alvarado, Obregon, Mexico D.F. C.P. 01210 con la autorizacion de The Coca Cola Company).

4.2.2.3 Thailand

Stimuli for the target tests consisted of a carbonated beverage sweetened with sucrose (Coca Cola® Classic, Thai Nam Thip Company Ltd. 214 Moo 5, Viphawadeerungsit Rd., Laksri, Bangkok, Thailand) and a carbonated beverage sweetened with aspartame, acesulfame-K, and sucralose (Coca Cola® Light, Thai Nam Thip Company Ltd.). From preliminary studies it was concluded that the Coca Cola ® Classic product was sweeter that the Coca Cola® Light product; therefore, sweetness was the attribute used to specify the difference among the stimuli. The former will be referred to as the “strong” sample and the latter as the “weak” sample.

The stimulus for the distractor tests and for palate cleansing was bottled water (Crystal ®, Serm Suk Co. Ltd. 700/369 Moo 6, Nongkhong, Muang, Chonburi, Thiailand).

4.2.2.4 USA

Stimuli consisted of a carbonated beverage sweetened with high fructose corn syrup (Coca Cola ® Classic, The Coca Cola Company, Atlanta, GA, USA) and a carbonated beverage sweetened with aspartame (Diet Coke®, The Coca Cola Company, Atlanta, GA, USA). From preliminary studies, it was concluded that the Coca Cola ® Classic product was sweeter that the Diet Coke® product; therefore, sweetness was the attribute used to specify the difference among the stimuli. The former will be referred to as the “strong” sample and the latter as the “weak” sample.

The stimulus for the distractor tests and for palate cleansing was bottled water (Dasani, The Coca Cola Company, Atlanta, GA, USA).

4.2.3 Testing Conditions

All tests were conducted in the children’s native language in order to avoid having language as an additional source of variation in this study; i.e., Spanish in Honduras and Mexico, Thai in Thailand, and English in the USA. All discrimination tests were performed in the children’s classroom settings at the schools.

Before each experimental session the children were given a presentation in order to ensure that they understood the basic logistics of the testing session (see Figure 4.2). During the presentation a review of the terms *same*, *different*, *sweet*, *not sweet*, *sweeter*, *less sweet,* and *less sugar* was given to the children. Also, the testing procedure was explained along with an overview of the questionnaire (see Appendix D) and its proper fill-out.

Figure 4.2: Presentation given to children at the beginning of the testing session

4.2.4 Experimental Design

Each judge performed two triangle tests and two 3-AFC tests to discriminate among the carbonated beverages. Triangle and 3-AFC tests were presented in succession during the same experimental session following an AB design (A = triangle and B = 3-AFC). This order was used to avoid the possible influence of the specified attribute on the unspecified method.

In order to avoid fatigue and distraction with the children, only two out of the six possible orders of tasting were presented to each child. This means that the complete block of the six possible orders of tasting for the target tests was divided among three different subjects. Child1: SSW – WWS, child 2: SWS – WSW, and child 3: SWW – WSS.

In order to prevent the judges from abandoning the triangle strategy, a distractor test (Figure 4.3) was intermingled with the two target tests. Thus, each judge performed a total of four (4) target tests to discriminate the strong sample from the weak (two triangle tests and two 3-AFC tests) and one (1) distractor test. One particular experimental session was:

- Triangle test: SSW, distractor, WWS
- BREAK
- 3-AFC test: SSW, WWS

The hypotheses being tested for the d' values were:

H_o: Triangle d' = 3-AFC d'

H_a: Triangle $d' \neq$ 3-AFC d'

Figure 4.3: USA child performing a distractor test

4.2.5 Testing Protocol

The triadic tests were given under two possible sets of instructions. For the triangle test the nature of the difference was not specified and the specific instructions were: "here are three samples, two are the same and one is different: **circle the beverage that tastes different**". For the 3-AFC test the nature of the difference was specified and the instructions were altered depending on which stimuli was the odd one. When the strong stimulus was the odd sample the instructions were: "here are three samples, one is sweeter than the other two: **circle the one that tastes sweeter**". When the weak stimulus was the odd sample the instructions were: "here are three samples, one is less sweet than the other two: **circle the one that tastes less sweet**".

Subjects began each session by drinking bottled water and repeated it before each triadic test. Next, the three samples for each question were presented simultaneously to the subjects, and the judges were instructed to taste from left to right (Figure 4.4). Judges rinsed with water after tasting each sample. The same protocol was repeated for the other four questions. The subjects were allowed to retaste the sample with the condition that they always tasted all three samples in the order presented to them. However, they had to be consistent over all the experimental sessions. The children were tested on the same day for both protocols given a 10 minute break between sessions. Session lengths ranged from 30 to 60 minutes.

Research showed that sweet taste receptors, in addition to being present in the mouth cavity and esophagus, are also present in the gastrointestinal tract of humans (Jang and others 2007, Margolskee and others 2007). Our preliminary study showed that there was no significant difference in discrimination between swallowing and expectorating a sample (see Appendix E). In this study, the children were therefore instructed to swallow the samples as it is a more natural behavior and causes less distraction.

Figure 4.4: Thai children performing the triangle test

4.2.5.1 Experimental Session Overview

The following is a detailed protocol of the testing session:

- Children filled out the child assent form
- Children were presented with the questionnaire and water
- Children filled out the top portion of the questionnaire (demographic information: name, age, gender, grade)
- Triangle Test
  - Overview presentation for the triangle test
  - Three samples for question 1 (Q1) were simultaneously presented
  - Children cleansed their palates

    - Children were instructed to taste the samples
        - Taste the first sample
        - Taste the second sample
        - Taste the third sample
    - Children were instructed to circle on the questionnaire the beverage that tasted "different"
    - Children stopped
    - The process was repeated for Q2 and Q3
    - Children took a 10 minute mandatory break
- 3-AFC Test
    - Overview presentation for the 3-AFC test
    - Three samples for Q4 were simultaneously presented
    - Children cleansed their palates
    - Children were instructed to taste the samples
        - Taste the first sample
        - Taste the second sample
        - Taste the third sample
    - Children circled the beverage that was less sweet or had less sugar and stopped
    - The process was repeated again for Q5.
    - This time the children circled the beverage that is sweeter or had more sugar
- End of session

4.2.6 Data Analysis

For each testing method, the number of correct responses provided by the subjects was counted and recorded. The P_cs for both tests were used to determine the corresponding *d′* values (Ennis, 1993).

The experimental estimate *d′*, the variance of *d′*, and the test of significance were obtained using the IFPrograms software Course Version 8.5.0320 (The Institute for Perception, Richmond, VA, USA). The *d′* values were calculated under the assumptions that the intensity distributions for the two stimuli were unidimenasional normal distributions that had equal variance (Bi and others 1997). Alternatively, *d′* values for the triangle and 3-AFC tests could be obtained from published tables (e.g., Ennis, 1993; see Appendix F). The variance of *d′* and the test statistic can be obtained by the approach described by Bi and others (1997) (see Appendix G and H). All analyses were evaluated at $\alpha = 0.05$.

4.3 Results

4.3.1 Honduras

4.3.1.1 Study I: Hardly Distinguishable Stimuli

For the Honduran population (N = 306), P_cs, *d′* values, and the variance of *d′* for the triangle and 3-AFC methods are presented in Tables 4.2 and 4.3. For Tables 4.2 and 4.3, the upper section refers to the questions where the odd sample was the weak stimulus (WSS, SWS, or SSW) and the lower section refers to the questions where the odd sample was the strong stimulus (SWW, WSW, or WWS). For a P_c less than the guessing probability (⅓) the corresponding *d′* value is zero and variance of *d′* cannot be calculated.

Table 4.2: Proportions of correct responses, *d'* values, and variance of *d'* for the triangle method for Honduran children (first to sixth grade)

Age Group[a]	**Subjects**	**Correct Responses**	P_c[b]	*d'*	**variance** *d'*
Identify the odd (WEAK[c]) sample					
1	51	19	0.37	0.67	0.355
2	51	15	0.29	0	---[e]
3	51	10	0.20	0	---[e]
4	51	14	0.27	0	---[e]
5	51	39	0.76	2.89	0.145
6	51	19	0.37	0.67	0.355
Identify the odd (STRONG [d]) sample					
1	51	37	0.73	2.65	0.135
2	51	31	0.61	2.02	0.121
3	51	29	0.57	1.82	0.122
4	51	22	0.43	1.08	0.178
5	51	45	0.88	3.83	0.224
6	51	37	0.73	2.65	0.135

[a] Age Group categories refer to grades first to sixth

[b] P_c = proportion of correct responses

[c] WEAK = Coca Cola® Light

[d] STRONG = Coca Cola ® Classic

[e] $P_c < ⅓$; *d'* values = 0 and their variance cannot be calculated

Table 4.3: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for Honduran children (first to sixth grade)

Age Group[a]	Subjects	Correct Responses	P_c[b]	*d′*	variance *d′*
Identify the WEAK[c] sample					
1	51	32	0.63	0.98	0.054
2	51	25	0.49	0.52	0.052
3	51	21	0.41	0.27	0.052
4	51	27	0.53	0.65	0.052
5	51	39	0.76	1.50	0.064
6	51	21	0.41	0.27	0.052
Identify the STRONG[d] sample					
1	51	15	0.29	0	---[e]
2	51	23	0.45	0.4	0.052
3	51	40	0.78	1.58	0.066
4	51	24	0.47	0.46	0.052
5	51	38	0.75	1.41	0.062
6	51	36	0.71	1.26	0.058

[a] Age Group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Coca Cola® Light
[d] STRONG = Coca Cola ® Classic
[e] $P_c < ⅓$; *d′* values = and their variance cannot be calculated

4.3.1.1.1 Overall Performance

Table 4.4 shows the overall performance of the Honduran population for each method; the pooled P_cs and the corresponding *d′* values are presented. The P_c for the 3-AFC method (0.56) was higher than that for the triangle method (0.52), confirming the paradox was present in this population.

Table 4.4: Pooled proportions of correct responses, pooled *d′* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the Honduran population (N=612[a])

Triangle Method			3-AFC Method			*d'* Sig Test[c]	
P_c[b]	*d′*	variance	P_c[b]	*d′*	variance	χ^2	p-value[d]
0.52	1.56	0.011	0.56	0.74	0.004	44.83	< 0.001

[a] N = 612 due to two replications of each of the 306 subjects
[b] P_c = proportion of correct responses
[c] Significance test for *d′* values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[d] All p-values > 0.05 are not significant

The d' values of the entire population were pooled. Therefore, the sample sized used for each protocol was 612, due to two replications of each test per subject. As replications from 306 different subjects were combined, it was necessary to take into account the possibility of overdispersion (Cox, 1983; Anderson, 1988) in the data. Overdispersion is added variance. The Binomial distribution, normally used for sensory difference or preference tests, cannot be applied to pooled data from multiple subjects that have performed replicate tests. If the Binomial model is used for pooled data, the Type I error may be inflated; i.e., the probability of rejecting a true null hypothesis (H_o: Triangle d' = 3-AFC d') will be greater (see Ennis and Bi, 1998 for further details). Overdispersion can be accounted for by using the Beta-Binomial model and a Thurstonian variant has been developed by Bi and Ennis (1998) for replicated difference tests. The variance of d' is then measured in terms of a gamma (γ) value (see Appendix I). Gamma ranges from zero (no overdispersion) to one (full overdispersion). When comparing these two models, the p-value indicates the statistical probability that the Beta-Binomial model fits the data significantly better than the Binomial model.

After assessing the presence of overdispersion in the data of both protocols, the Beta Binomial model was found not significantly better than the Binomial model for these data; therefore, overdispersion was not significant. For the triangle (Figure 4.4) and 3-AFC (Figure 4.5) tests γ was equal to zero (triangle: $p = 0.95$; 3-AFC: $p = 0.82$; IFPrograms).

Comparing pooled d' values of the triangle (1.56) and 3-AFC (0.74) protocols showed a significant difference between them ($p < 0.001$) (Table 4.4). These results from the Honduran population do not confirm those observed in the previous study (Chapter 3) using easily discriminable stimuli (apple juice with different sweetness concentrations). However, it is in

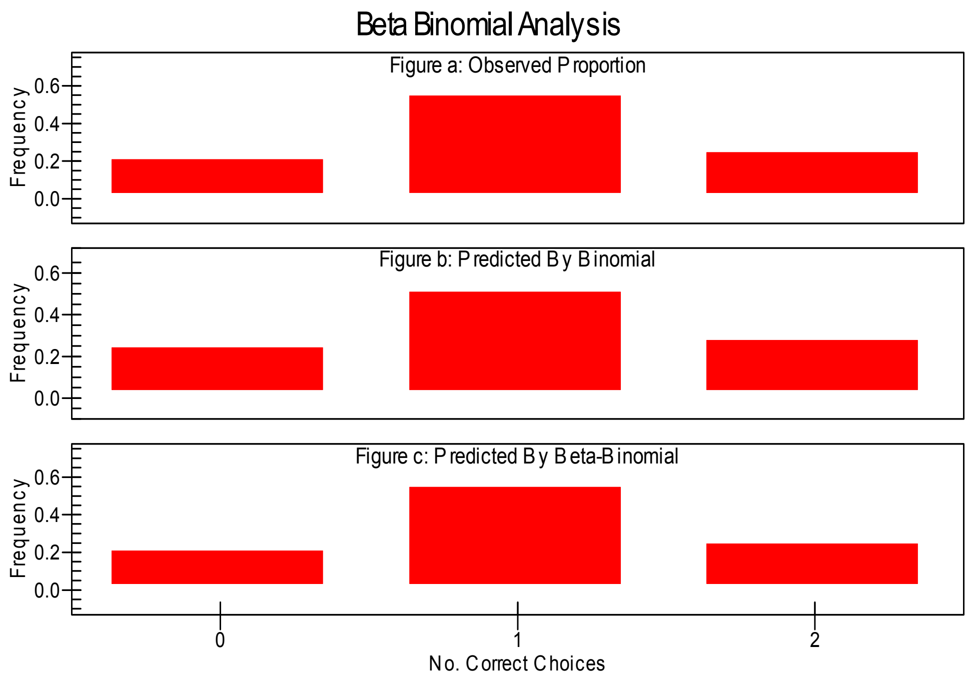


Figure 4.4: Triangle test (a) observed proportions, (b) predicted probabilities by binomial distribution, and (c) predicted probabilities by beta-binomial distribution.

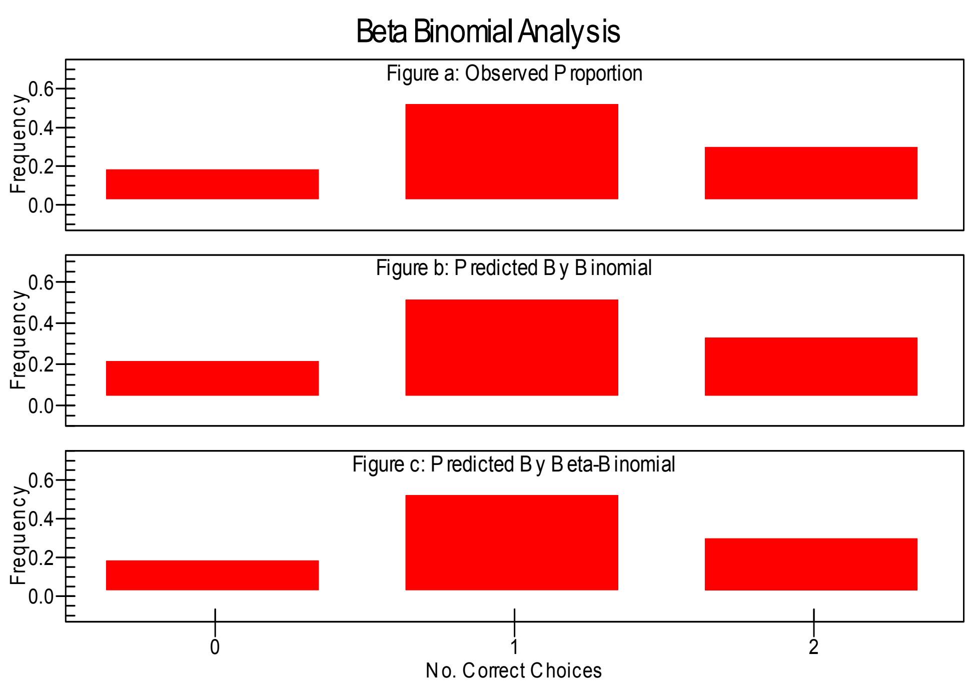


Figure 4.5: 3-AFC test (a) observed proportions, (b) predicted probabilities by binomial distribution, and (c) predicted probabilities by beta-binomial distribution.

accord with the results of the easily discriminable study for the Mexican, Thai, and USA populations where the triangle d' was significant higher than that for the 3-AFC method. Thurstonian predictions that the two triadic protocols should not yield significantly different d' values regardless of the difference in P_c were not confirmed for Honduran children between the ages of 6 and12, using carbonated beverages sweetened with different sweeteners.

From these results, it was observed that a larger degree of difference among the stimuli was detected with the triangle test (sweetness not specified as the difference) than with the 3-AFC method (sweetness specified as the difference). Could it be that the difference detected would not be labeled "sweetness" by the children? As the stimuli used were hard to distinguish, children may not be able to differentiate the two test stimuli. They may have utilized other easier attributes such as "preference" to complete their task.

4.3.1.1.2 Individual Performance

Table 4.5 presents the results of the different age groups. The paradox was only confirmed for the second, third, fourth and sixth grades (3-AFC P_c > Triangle P_c). For first and fifth graders the P_c for the triangle method was higher than that of the 3-AFC method. Therefore, for the first and fifth graders the paradox was not confirmed.

In the previous study regarding the paradox using apple juice with different sweetness concentrations as the stimuli, an increase in performance with age was observed for all populations. However, for the Honduran population in this study, there was no clear trend regarding the relationship between correct responses and age. For instance, in the triangle test the first and sixth graders performed equally (P_c = 0.55), with the highest P_c observed in the fifth grade (0.82) and the lowest P_c observed for the fourth grade (0.35). For the 3-AFC test, the highest P_c was also observed for the fifth graders (0.75). It can be argued that for the hardly discriminable stimuli used in this study, performance does not appear to increase with age.

Table 4.5: Proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods for the individual grades for the Honduran population

		Triangle			**3-AFC**			***d′* Sig Test** [c]	
Grade	**N** [a]	P_c [b]	***d′***	**variance**	P_c [b]	***d′***	**variance**	χ^2	**p-value** [d]
1	102	0.55	1.72	0.062	0.46	0.43	0.026	18.90	< 0.001
2	102	0.45	1.20	0.080	0.47	0.46	0.026	4.98	0.03
3	102	0.38	0.75	0.148	0.60	0.88	0.027	0.09	0.76
4	102	0.35	0.47	0.328	0.50	0.56	0.026	0.02	0.88
5	102	0.82	3.30	0.086	0.75	1.45	0.031	29.50	< 0.001
6	102	0.55	1.72	0.062	0.56	0.75	0.026	10.23	< 0.001

[a] N = 102 due to two replications of each of the 51 subjects per grade
[b] P_c = proportion of correct responses
[c] Significance test for *d′* values, Critical χ^2 value = 3.84, α = 0.05
[d] All p-values > 0.05 are not significant

The *d′* values for the individual grades are also presented in Table 4.5. Overdispersion was assessed for each grade's pooled responses (from two replications) and was found not to be significant (all $p > 0.05$). Since the 3-AFC P_cs were not higher in the first and fifth grades the *d′* value's significance tests will not be discussed. For those grades (second, third, fourth, sixth) for which the 3-AFC test P_c was higher than the triangle test P_c, the *d′* values were not significantly different for the third and fourth grade, thereby confirming the Thurstonian predictions that the *d′* values should not be significantly different despite the different P_cs. However, the results for the second and sixth grades do not confirm these predictions (all $p < 0.05$). As observed for the overall Honduran population, a larger degree of difference among the stimuli was detected by the second and sixth graders with the triangle test (sweetness not specified as the difference) than with the 3-AFC method (sweetness specified as the difference).

When a difference in performance between the age groups is observed in the data it can be suggested that it could be due to differences in cognitive abilities or sensory acuity. However, for the data presented above, the difference in performance cannot be attributed to these factors because a clear trend is not observed. For example, for the triangle test, sixth graders performed equally to first graders, and for the 3-AFC method, we have third graders performing better than sixth graders. The P_cs for both protocols as well as the *d′* values present an uncertain trend. It is not known to what degree children can perform triadic discrimination tests. Further research in this area should be explored.

4.3.1.2 Study II: Hardly Distinguishable Stimuli

For the Honduran population (N = 303), P_cs, *d′* values, and the variance of *d′* for the triangle and 3-AFC methods are presented in Tables 4.6 and 4.7. The upper section of the table refers to the questions where the odd sample was the weak stimulus (WSS, SWS, or SSW) and the lower section refers to the questions where the odd sample was the strong stimulus (SWW, WSW, or WWS). Table 4.8 presents the overall data and the data for the individual grades. For a P_c less than the guessing probability (⅓), the corresponding *d′* value is zero and variance of *d′* cannot be calculated.

As can be observed from Tables 4.6, 4.7 and 4.8, there were a significant number of age groups for which *d′* values could not be calculated. For the triangle test, there were 5/12 occasions for which the *d′* value was zero and the variance could not be calculated (Table 4.6); for the 3-AFC method there were 6/12 occasions (Table 4.7). Due to the instances when the P_c was less than ⅓ for pooled data for the individual grades (Table 4.8), comparisons (significance tests) among the protocols were only possible for the second and fifth grades.

Table 4.6: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for Honduran children (first to second grade)

Age Group[a]	**Subjects**	**Correct Responses**	P_c[b]	***d′***	**variance *d′***
Identify the odd (WEAK[c]) sample					
1	51	32	0.63	2.12	0.122
2	54	25	0.46	1.27	0.143
3	51	25	0.49	1.41	0.138
4	51	9	0.18	0	---[e]
5	48	16	0.33	0	---[e]
6	48	15	0.31	0	---[e]
Identify the odd (STRONG[d]) sample					
1	51	19	0.37	0.67	0.355
2	54	28	0.52	1.56	0.123
3	51	10	0.20	0	---[e]
4	51	11	0.22	0	---[e]
5	48	29	0.60	2.00	0.129
6	48	12	0.25	0	---[e]

[a] Age Group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Coca Cola® Light
[d] STRONG = Coca Cola ® Classic
[e] $P_c < ⅓$; *d′* values and their variance cannot be calculated

Table 4.7: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC protocol for Honduran children (first to sixth grade)

Age Group[a]	Subjects	Correct Responses	P_c[b]	*d′*	variance *d′*
Identify the WEAK [c] sample					
1	51	20	0.39	0.2	0.053
2	54	28	0.52	0.62	0.049
3	51	11	0.22	0	---[e]
4	51	16	0.31	0	---[e]
5	48	24	0.50	0.56	0.055
6	48	23	0.48	0.49	0.055
Identify the STRONG [d] sample					
1	51	10	0.20	0	---[e]
2	54	9	0.17	0	---[e]
3	51	10	0.20	0	---[e]
4	51	11	0.22	0	---[e]
5	48	22	0.46	0.42	0.055
6	48	20	0.42	0.28	0.056

[a] Age Group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Coca Cola® Light
[d] STRONG = Coca Cola ® Classic
[e] P_c < ⅓; *d′* values = 0 and their variance cannot be calculated

Table 4.8: Proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods for the individual grades for the Honduran population

		Triangle			3-AFC			*d′* Sig Test [c]	
Grade	**N[a]**	P_c [b]	*d′*	**variance**	P_c [b]	*d′*	**variance**	χ^2	**p-value** [d]
1	102	0.50	1.47	0.067	0.29	0	---[e]	--- [e]	--- [e]
2	102	0.49	1.42	0.065	0.34	0.03	0.026	21.23	< 0.001
3	102	0.34	0.33	0.63	0.21	0	---[e]	--- [e]	--- [e]
4	102	0.20	0	---[e]	0.26	0	---[e]	--- [e]	--- [e]
5	96	0.47	1.3	0.079	0.48	0.49	0.027	6.19	0.01
6	96	0.28	0	---[e]	0.45	0.39	0.028	--- [f]	--- [e]
overall	606	0.38	0.74	0.025	0.34	0.01	0.005	17.76	< 0.001

[a] N = 102 due to two replications of each of the 51 subjects per grade
[b] P_c = proportion of correct responses
[c] Significance test for *d′* values, Critical χ^2 value = 3.84, α = 0.05
[d] All p-values > 0.05 are not significant
[e] P_c < ⅓; *d′* values = 0, the variance of *d′* cannot be calculated, *d′* significance cannot be tested

For the overall data of this population, the comparison among the triangle and 3-AFC methods was possible. However, the 3-AFC method P_c (0.34) is slightly above the guessing probability (⅓); subsequently, the d' value (0.01) was close to zero. The evident reason for this outcome is that the stimuli used in this study (Coca Cola® Light, Coca Cola® Cero; aspartame and acesulfame-K were the sweeteners for both samples) were too confusing and therefore not appropriate for discrimination testing with children. As a result, this study was not continued for the Mexican, Thai, and USA populations.

4.3.2 Mexico

For the Mexican population (N = 176), the proportions of correct responses, corresponding d' values, and the variance of d' for the triangle and 3-AFC tests are presented in Tables 4.9 and 4.10, respectively. Again, the upper section of the table refers to the questions were the odd sample was the weak stimulus and the lower section refers to the questions where the odd sample was the strong stimulus. Note that when the P_c for a test is less than the guessing probability (1/3), the corresponding d' value is zero and the variance of d' cannot be calculated.

4.3.2.1 Overall Performance

Overall performance data for the Mexican population are presented in Table 4.9. The P_c of the 3-AFC method (0.49) was higher than that of the triangle method (0.48), confirming the paradox.

Due to the two replications per subject, the sample size used for calculations of pooled d' values was 352. Overdispersion was assessed and was not significant in the triangle data (p = 0.77, $\gamma = 0$) or the 3-AFC data (p = 0.09, $\gamma = 0.10$). For the triangle test, the $P_c = 0.48$ corresponded to a d' of 1.35.

Table 4.9: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for Mexican children (first to sixth grade)

Age Group[a]	**Subjects**	**Correct Responses**	P_c[b]	***d′***	**variance *d′***
		Identify the odd (WEAK [c]) sample			
1	21	8	0.38	0.74	0.734
2	38	9	0.24	0	---[e]
3	25	4	0.16	0	---[e]
4	32	17	0.53	1.63	0.203
5	37	17	0.46	1.25	0.212
6	23	8	0.35	0.40	1.929
		Identify the odd (STRONG [d]) sample			
1	21	9	0.43	1.07	0.441
2	38	21	0.55	1.74	0.166
3	25	16	0.64	2.18	0.25
4	32	18	0.56	1.79	0.196
5	37	30	0.81	3.21	0.227
6	23	11	0.48	1.35	0.318

[a]Age Group categories refer to grades first to sixth

[b] P_c = proportion of correct responses

[c] WEAK = Coca Cola® Light

[d] STRONG = Coca Cola ® Classic

[e] $P_c < ⅓$; *d′* values and their variance cannot be calculated

Table 4.10: Proportions of correct responses, *d'* values, and variance of *d'* for the 3-AFC method for Mexican children (first to sixth grade)

Age Group[a]	Subjects	Correct Responses	P_c[b]	*d'*	variance *d'*
Identify the WEAK[c] sample					
1	21	7	0.33	0	0.133
2	38	29	0.76	1.49	0.085
3	25	11	0.44	0.36	0.106
4	32	14	0.44	0.35	0.083
5	37	14	0.38	0.16	0.073
6	23	14	0.61	0.91	0.119
Identify the STRONG[d] sample					
1	21	4	0.19	0	---[e]
2	38	25	0.66	1.08	0.074
3	25	2	0.08	0	---[e]
4	32	19	0.59	0.86	0.085
5	37	24	0.65	1.05	0.076
6	23	10	0.43	0.34	0.115

[a] Age Group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Coca Cola® Light
[d] STRONG = Coca Cola ® Classic
[e] $P_c < ⅓$; *d'* values = 0 and their variance cannot be calculated

Table 4.11: Pooled proportions of correct responses, pooled *d'* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the Mexican population (N=352[a])

Triangle Method			3-AFC Method			*d'* Sig Test[c]	
P_c[b]	*d'*	variance	P_c[b]	*d'*	variance	χ^2	p-value[d]
0.48	1.35	0.021	0.49	0.53	0.007	24.01	< 0.001

[a] N = 352 due to two replications of each of the 176 subjects
[b] P_c = proportion of correct responses
[c] Significance test for *d'* values, Critical χ^2 value = 3.84, α = 0.05
[e] All p-values > 0.05 are not significant

For the 3-AFC task, the P_c = 0.49% corresponded to a *d'* of 0.53. As with the Honduran population, Thurstonian predictions were not confirmed ($p < 0.001$) for this Mexican population. A greater degree of difference between the stimuli was detected by this population with the triangle test. The Honduran population also presented this same result. Also, a higher *d'* for the triangle tests was observed for the Mexican population for the discrimination of apple juice samples with different sweetness concentrations. The reasons for this outcome are not clear, but more than likely are related to the nature of the stimuli.

4.3.2.2 Individual Performance

For the individual age groups, results showed that the P_c for the 3-AFC method was higher than that of the triangle test in the second and sixth grades (Table 4.12), confirming the paradox. For all the other grades, the triangle method P_c was higher.

Table 4.12: Proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods for the individual grades for the Mexican population

		Triangle			3-AFC			*d′* Sig Test [c]	
Grade	**N[a]**	P_c [b]	***d′***	**variance**	P_c [b]	***d′***	**variance**	χ^2	**p-value [d]**
1	42	0.40	0.91	0.268	0.26	0	---[e]	---[e]	---[e]
2	76	0.39	0.84	0.166	0.71	1.28	0.039	0.84	0.33
3	50	0.40	0.88	0.237	0.26	0	---[e]	---[e]	---[e]
4	64	0.55	1.71	0.099	0.52	0.61	0.041	8.64	< 0.001
5	74	0.64	2.16	0.084	0.51	0.60	0.036	33.67	< 0.001
6	46	0.41	0.97	0.227	0.52	0.63	0.057	0.41	0.52

[a] N twice the number of subjects per grade; due to two replications of each subject
[b] P_c = proportion of correct responses
[c] Significance test for *d′* values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[d] All p-values > 0.05 are not significant
[e] $P_c < ⅓$; *d′* values, the variance of *d′* cannot be calculated, and *d′* significance cannot be tested

As with the Honduran population, a clear relationship was not observed between performance (P_c) and age. In the triangle test, sixth graders performed similarly to first, second and third graders. For the 3-AFC test, first and third grader's P_cs were lower than the guessing probability of this test (1/3).

The *d′* values for the individual age groups are also presented in Table 4.12. The *d′* values for the 3-AFC method could not be calculated for the first and third grade since the P_cs (0.26) are less than the guessing probability of this test. Overdispersion was assessed for each grade's pooled responses for the triangle and 3-AFC methods and was found not to significant (all p >

0.05). Overdispersion could not be assessed for 3-AFC data of the first and third grade, also due to a $P_c < ⅓$. Due to the lack of a 3-AFC test *d′* value for the first and third grades, a significance test among *d′* values could only be performed for second, fourth, fifth and sixth grade. For these grades, results show that *d′* values were not significantly different for the second and sixth grade (all $p > 0.05$), confirming the Thurstonian predictions. However, these Thurstonian predictions could not be confirmed for fourth and fifth grades (all $p < 0.001$). As observed for the overall data (Table 4.8), the *d′* value for the triangle test was higher than that for the 3-AFC tests for the fourth to sixth grade. A greater degree of difference between the stimuli was detected by these age groups with the triangle test.

4.3.3 Thailand

For the Thai population (N = 360), the proportions of correct responses, *d′* values, and the variance of *d′* for the triangle test are presented in Table 4.13. The 3-AFC results are presented in Table 4.14.

4.3.3.1 Overall Performance

For the entire Thai population (N = 360), the P_c for the triangle (0.49) was higher than that for the 3-AFC (0.42) methods (Table 4.15). This is an unexpected result; since Thurstonian theory predicts that the 3-AFC will results in more correct responses than the triangle method (see Figure 4.1). Therefore, the paradox of discriminatory non-discriminators was not confirmed for this population.

For the pooled *d′* values of the entire population, the sample size used was 720 due to two replications of each protocol per subject (overdispersion was not significant; triangle and 3-AFC $\gamma = 0$, all $p > 0.05$).

Table 4.13: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle test for Thai children (first to sixth grade)

Age Group[a]	Subjects	Correct Responses	P_c [b]	*d′*	variance *d′*
Identify the odd (WEAK[c]) sample					
1	60	28	0.47	1.29	0.127
2	60	25	0.42	0.99	0.169
3	60	31	0.52	1.55	0.111
4	60	39	0.65	2.23	0.105
5	60	40	0.67	2.32	0.106
6	60	32	0.53	1.64	0.108
Identify the odd (STRONG[d]) sample					
1	60	29	0.48	0.50	0.044
2	60	32	0.53	0.66	0.044
3	60	23	0.38	0.17	0.045
4	60	23	0.38	0.17	0.045
5	60	22	0.37	0.12	0.045
6	60	31	0.52	0.61	0.044

[a] Age group categories refer to grades first to sixth

[b] P_c = proportion of correct responses

[c] WEAK = Coca Cola® Light

[d] STRONG = Coca Cola ® Classic

Table 4.14: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for Thai children (first to sixth grade)

Age Group[a]	Subjects	Correct Responses	P_c[b]	*d′*	variance *d′*
		Identify the WEAK[c] sample			
1	60	26	0.43	0.34	0.044
2	60	22	0.37	0.12	0.045
3	60	31	0.52	0.61	0.044
4	60	8	0.13[e]	0	---[e]
5	60	39	0.65	1.06	0.047
6	60	12	0.20[e]	0	---[e]
		Identify the STRONG[d] sample			
1	60	32	0.53	0.66	0.044
2	60	39	0.65	1.06	0.047
3	60	15	0.25[e]	0	---[e]
4	60	25	0.42	0.28	0.044
5	60	24	0.40	0.23	0.045
6	60	28	0.47	0.45	0.044

[a]Age Group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Coca Cola® Light
[d] STRONG = Coca Cola ® Classic
[e] P_c < ⅓; *d′* values = 0 and their variance cannot be calculated

Table 4.15: Pooled proportions of correct responses, pooled *d′* values, test statistic (N= 720[a]) and p-value for the triangle and 3-AFC methods for all six grades of the Thai population

Triangle Method			3-AFC Method			*d′* Sig Test [c]	
P_c [b]	*d′*	variance	P_c [b]	*d′*	variance	χ^2	p-value [d]
0.49	1.43	0.01	0.42	0.29	0.004	92.83	< 0.001

[a] N = 720 due to two replications of each of the 360 subjects
[b] P_c = proportion of correct responses
[c] Significance test for *d′* values, Critical χ^2 value = 3.84, α = 0.05
[d] All p-values > 0.05 are not significant

The *d′* values of the triangle (1.43) and 3-AFC (0.43) methods were significantly different (p < 0.001). From these results, Thurstonian predictions were not confirmed for the triadic tests among Thai children. Since the *d′* value was higher for the triangle test, this population detected a significantly greater degree of difference among the beverages when the instruction was to choose the odd sample instead of specifying sweetness as the attribute responsible for the difference. This result is in accord with the results for the overall populations of the Honduras and Mexico.

4.3.3.2 Individual Performance

Results of performance in the individual grades show that only for the second and fifth grades were the 3-AFC test P_cs higher than the triangle test P_cs (Table 4.16). For the third, fourth and sixth grades the triangle test P_cs were higher than those of the 3-AFC method. Even though this is not the outcome predicted by the paradox, this is the same result observed for the overall data of this population (triangle P_c 0.49 > 3-AFC P_c 0.42; Table 4.15).

A clear trend was not observed for the relationship between the proportions of correct responses and age. For the triangle test, results show that P_c increased with age (48% first grade - 53% sixth grade). However, an increase was not observed for the 3-AFC data. For instance, the first graders performed better (0.48) than the sixth graders (0.33) and the fourth graders had a P_c (0.28) lower than the guessing probability (1/3) of this test.

Table 4.16: Proportions of correct responses, *d'* values and significance test for the triangle and 3-AFC methods for the individual grades of the Thai population

		Triangle			**3-AFC**			***d'* Sig Test** [c]	
Grade	**N** [a]	P_c [b]	***d'***	**variance**	P_c [b]	***d'***	**variance**	χ^2	**p-value** [d]
1	**120**	0.48	1.33	0.062	0.48	0.50	0.022	8.20	< 0.001
2	**120**	0.48	1.33	0.062	0.51	0.58	0.022	6.70	0.01
3	**120**	0.45	1.19	0.068	0.38	0.17	0.023	11.43	< 0.001
4	**120**	0.52	1.55	0.055	0.28	0	--- [e]	--- [e]	--- [e]
5	**120**	0.52	1.55	0.055	0.53	0.64	0.022	10.75	< 0.001
6	**120**	0.53	1.60	0.055	0.33	0	0.023	32.82	< 0.001

[a] N twice as the number of subjects per grade (60); due to two replications of each subject
[b] P_c = proportion of correct responses
[c] Significance test for *d'* values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[d] All p-values > 0.05 are not significant
[e] $P_c < ⅓$; *d'* values = 0, the variance of *d'* cannot be calculated, and *d'* significance cannot be tested

The d' values of the individual grades are presented in Table 4.16. The d' values for the 3-AFC method could not be calculated for fourth grade since the P_c (0.28) was less than the guessing probability of this test. Overdispersion was assessed for each grade's pooled responses and was not found to be present in the triangle or 3-AFC data (all $p > 0.05$). Also, due to the low P_c ($0.28 < 0.33$) of the fourth grade 3-AFC data, overdispersion could not be assessed for this age group. Results for the individual grades show that all triangle d' values were significantly greater than those for the 3-AFC. Therefore, Thurstonian predictions were not confirmed for the individual age groups of the Thai population. Higher triangle d' values were also observed for the Honduran and Mexican populations.

4.3.4 USA

For the United States population (N = 93) the P_cs, d' values, and the variance of d' for the triangle and 3-AFC tests are presented in Tables 4.17 and 4.18. The upper section of the table refers to the questions were the odd sample was the weak stimulus and the lower section refers to the questions where the odd sample was the strong stimulus.

4.3.4.1 Overall Performance

For the overall performance of the USA population, the P_c for the triangle method (0.52) was greater than that for the 3-AFC (0.51) method (Table 4.19). This confirmed that the paradox was not observed for this population. This result is supported by previous findings in the overall Thai population, where the triangle test P_c (0.49) was also higher than the P_c for the 3-AFC method (0.42). For these two populations, the children performed better when the difference among the stimuli was not specified.

Table 4.17: Proportions of correct responses, *d′* values, and variance of *d′* for the triangle method for US children (first to sixth grade)

Age Group[a]	**Subjects**	**Correct Responses**	P_c[b]	***d′***	**variance *d′***
Identify the odd (WEAK[c]) sample					
1	7	4	0.57	1.83	0.891
2	22	10	0.45	1.22	0.364
3	16	7	0.44	1.12	0.547
4	12	5	0.42	0.99	0.842
5	18	15	0.83	3.38	0.506
6	18	9	0.50	1.47	0.382
Identify the odd (STRONG[d]) sample					
1	7	2	0.29	0	---[e]
2	22	9	0.41	0.94	0.491
3	16	9	0.56	1.79	0.392
4	12	9	0.75	2.80	0.599
5	18	9	0.50	1.47	0.382
6	18	9	0.50	1.47	0.382

[a] Age Group categories refer to grades first to sixth

[b] P_c = proportion of correct responses

[c] WEAK = Coca Cola® Light

[d] STRONG = Coca Cola ® Classic

[e] $P_c < ⅓$; *d′* values = 0 and their variance cannot be calculated

Table 4.18: Proportions of correct responses, *d′* values, and variance of *d′* for the 3-AFC method for US children (first to sixth grade)

Age Group[a]	Subjects	Correct Responses	P_c [b]	*d′*	variance *d′*
Identify the WEAK[c] sample					
1	7	2	0.29	0	---[e]
2	22	8	0.36	0.11	0.124
3	16	6	0.38	0.14	0.170
4	12	6	0.50	0.56	0.220
5	18	10	0.56	0.74	0.148
6	18	11	0.61	0.92	0.152
Identify the STRONG[d] sample					
1	7	4	0.57	0.79	0.383
2	22	8	0.36	0.11	0.124
3	16	7	0.44	0.35	0.166
4	12	11	0.92	2.36	0.471
5	18	13	0.72	1.32	0.169
6	18	9	0.50	0.56	0.146

[a] Age Group categories refer to grades first to sixth
[b] P_c = proportion of correct responses
[c] WEAK = Coca Cola® Light
[d] STRONG = Coca Cola ® Classic
[e] $P_c < ⅓$; *d′* values = 0 and their variance cannot be calculated

Table 4.19: Pooled proportions of correct responses, pooled *d′* values, test statistic and p-value for the triangle and 3-AFC methods for all six grades of the USA population (N= 186[a])

Triangle Method			3-AFC Method			*d′* Sig Test [c]	
P_c [b]	*d′*	variance	P_c [b]	*d′*	variance	χ^2	p-value [d]
0.52	1.58	0.035	0.51	0.59	0.014	20.00	< 0.001

[a] N = 186 due to two replications of each of the 93 subjects
[b] P_c = proportion of correct responses
[c] Significance test for *d′* values, Critical χ^2 value = 3.84, α = 0.05
[d] All p-values > 0.05 are not significant

For the pooled *d′* values of the entire population the sample size used was 186 (two replications per subject). Overdispersion was not present in the triangle test data or the 3-AFC data (all $p > 0.05$). For the triangle test there were 52% correct selections of the odd stimulus, which was equivalent to a *d′* of 1.58. In the 3-AFC task, 59% correct selections were made, which corresponded to a *d′* of 0.59. There was a significant difference between these two *d′*

values ($p < 0.001$) (Table 4.19). From these results, Thurstonian predictions were not confirmed for the triadic tests among the USA population, using carbonated beverages with different sweeteners as the stimuli. The higher triangle test *d′* implies that a greater degree of difference was detected among the carbonated beverages with the triangle test than with the 3-AFC test. This means that with the triangle test the samples were detected as different, but with the 3-AFC method the same samples were detected as more confusable.

4.3.4.2 Individual Performance

Results from the assessment of performance in each individual grade (Table 4.20) show that the P_cs were higher for the 3-AFC method for the fourth and sixth grade; therefore the paradox was present for these age groups. However, the triangle P_cs were higher for the second, third and fifth grade. For the first grade, the P_cs for both methods of triads were the same.

Table 4.20: Proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods for the individual grades for the US population

		Triangle			3-AFC			*d′* Sig Test [c]	
Grade	**N[a]**	P_c [b]	***d′***	**variance**	P_c [b]	***d′***	**variance**	χ^2	**p-value** [d]
1	**14**	0.43	1.07	0.661	0.43	0.32	0.190	0.66	0.42
2	**44**	0.43	1.09	0.206	0.36	0.11	0.062	3.58	0.06
3	**32**	0.50	1.47	0.215	0.41	0.25	0.084	4.98	0.03
4	**24**	0.58	1.89	0.258	0.71	1.27	0.124	1.01	0.32
5	**36**	0.67	2.32	0.177	0.64	1.02	0.077	6.65	0.01
6	**36**	0.50	1.47	0.191	0.56	0.74	0.074	2.01	0.16

[a] N twice as the number of subjects per grade; due to two replications of each subject
[b] P_c = proportion of correct responses
[c] Significance test for *d′* values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[d] All p-values > 0.05 are not significant

The unexpected result of a higher triangle test P_c in some grades was also observed for other grades in the Honduran, Mexican and Thai populations. In addition, these results show that in this population there was also not a clear trend regarding the relationship between performance and age. The *d′* values of the individual grades are also presented in Table 4.20

(overdispersion was not significant; all $p > 0.05$). Results show that d' values are not significantly different for all age groups, except for the third and fifth grades. As with the other populations, the d' values for the triangle test were greater than those for the 3-AFC method.

4.4 Discussion

4.4.1 Overall Performance

This study explored the paradox of discriminatory non-discriminators with children as subjects using moderately discriminable stimuli, i.e., carbonated beverages with different sweeteners. Thurstonian theory states that regardless of differences in P_cs among the triangle and 3-AFC methods, the degree of difference (d') between a pair of stimuli should be independent of the methodology used. For the Honduran and Mexican populations, the 3-AFC P_c was higher than the triangle P_c (Table 4.21). On the other hand, for Thailand and the USA the P_c for the triangle method was higher than that for the 3-AFC. Therefore, the paradox was only confirmed for the Honduran and Mexican populations. For all countries combined, the triangle P_c (0.50) was higher than the 3-AFC P_c (0.49). It can be concluded that the children performed equally in the triangle and 3-AFC tests indicating that overall the paradox was not confirmed for the hardly discriminable stimuli used in this study.

Table 4.21: Summary of the proportions of correct responses, d' values and significance test for the triangle and 3-AFC methods overall and for the different countries

		Triangle			3-AFC			d' Sig Test [c]	
Country	**N[a]**	P_c [b]	d'	**variance**	P_c [b]	d'	**variance**	χ^2	**p-value [d]**
Honduras	**612**	0.52	1.56	0.011	0.56	0.74	0.004	44.83	< 0.001
Mexico	**352**	0.48	1.35	0.021	0.49	0.53	0.007	24.01	< 0.001
Thailand	**720**	0.49	1.43	0.01	0.42	0.29	0.004	92.83	< 0.001
USA	**186**	0.52	1.58	0.035	0.51	0.59	0.014	20.00	< 0.001
OVERALL	**1870**	0.50	1.47	0.004	0.49	0.51	0.001	184.32	< 0.001

[a] N twice as the number of subjects per country; due to two replications of each subject

[b] P_c = proportion of correct responses

[c] Significance test for d' values, Critical χ^2 value = 3.84, $\alpha = 0.05$

[d] All p-values > 0.05 are not significant

The overall *d′* values for the triangle and 3-AFC protocols were 1.47 and 0.49, respectively. Because these values were significantly different from each other, the Thurstonian predictions were not confirmed. The same trend was observed in each of the individual countries, the triangle *d′* value was significantly higher than the 3-AFC *d′* value. Therefore, Thurstonian predictions were not confirmed under the circumstances of this study. Overall and for the individual countries, a significantly larger degree of difference was detected among the beverages with the triangle test (the nature of the difference among the stimuli not specified) than with the 3-AFC test (nature of the difference specified). This is the same overall result observed in our previous study involving apple juice with different sweetness concentrations (100% vs. 60%) as the stimuli (Chapter 3). These results are not supported by those of other studies comparing subject performance and *d′* values in the triangle and 3-AFC methods (Delwiche and O'Mahony 1995, Masouka and others 1995, Rousseau and O'Mahony 1997, Stillman 1993, Tedja and others 1994).

4.4.2 Individual Performance

For the discrimination tests with apple juice with different sweetness concentrations (100% vs. 60%) as the stimulus, it was observed that the P_cs increased with age for the triangle and 3-AFC tests for all countries. However, for this study involving carbonated beverages with different sweeteners as the stimuli, this trend of increasing P_cs with age was not observed. Rather, an uncertain trend was present. This trend could be due to the children's cognitive abilities or sensory acuity. From the outcomes of the previous study (Chapter 3) and this study, I am comfortable concluding that children are able to perform triangle tests, since the results show that a difference among the stimuli was detected when it was present. As far as the children's sensory acuity, it is clear that the children were not guessing and in fact could perform the tests, as all P_cs were greater than the guessing probability (⅓). It could then be argued that the lack of a clear trend is most certainly due to the nature of the stimuli.

The P_cs of the individual grades of all populations were not always higher for the 3-AFC protocol. In about half of the cases the triangle test's P_cs were higher. For the Honduran population first and fifth grade had higher triangle tests P_cs. For the Mexican population first, third, fourth and fifth grades had higher triangle P_cs. For the Thai population third, fourth and fifth grades had higher triangle P_cs. For the US population second, third and fifth grades had higher proportions for the triangle test. Therefore the paradox could not be confirmed for all the individual grades in all the countries.

Thurstonian theory states that regardless of the higher number of correct responses on the 3-AFC test, the d' values among this test and the triangle test should not be significantly different. The 3-AFC P_c was not higher than the triangle P_c for all cases in this study. For these cases the results for the significance tests for the d' values will not be discussed. For the grades for which the 3-AFC P_c was higher than the triangle P_c, a significant difference was not present only in the second and sixth grades for the Mexican population. The d' values were significantly different for all the grades in the other three populations, where the d' value for the triangle test was higher than that for the 3-AFC test.

Table 4.22 shows the pooled results from all four populations for all the individual grades (N = 1870). These pooled results show again an inconsistent trend in regards to the relationship of P_cs and age. The only grades for which the 3-AFC P_c was higher were the second and third grades. Therefore, the paradox was only confirmed for these two grades. For the rest of the age groups, the triangle P_c was higher than that for the 3-AFC method. For the second and third grades the Thurstonian predictions were not confirmed; i.e., the *d′* values were significantly different from each other with higher *d′* values for the triangle method.

Table 4.22: Summary of the proportions of correct responses, *d′* values and significance tests for the triangle and 3-AFC methods for the individual grades from the Honduran, Mexican, Thai, and US population

		Triangle			3-AFC			*d′* Sig Test [c]	
Grade	**N[a]**	P_c [b]	***d′***	**variance**	P_c [b]	***d′***	**variance**	χ^2	**p-value** [d]
1	**278**	0.49	1.41	0.025	0.44	0.36	0.010	31.50	< 0.001
2	**342**	0.44	1.16	0.025	0.55	0.73	0.008	5.60	0.02
3	**304**	0.42	1.04	0.031	0.44	0.35	0.009	11.90	< 0.001
4	**310**	0.47	1.33	0.024	0.35	0.06	0.009	48.87	< 0.001
5	**332**	0.65	2.25	0.019	0.61	1.00	0.008	57.87	< 0.001
6	**304**	0.51	1.53	0.022	0.46	0.44	0.009	38.32	< 0.001

[a] N twice as the number of subjects per grade per country; due to two replications of each subject

[b] P_c = proportion of correct responses

[c] Significance test for *d′* values, Critical χ^2 value = 3.84, α = 0.05

[d] All p-values > 0.05 are not significant

The reason for the outcome of higher triangle P_cs is not known and requires further investigation. Several factors may have contributed to these results. The children's attention span might not have been enough for proper performance in the 3-AFC after having performed three triangle tests. As performance in the triangle test did yielded higher degrees of difference (*d′*) between the set of different stimuli, cognitive ability cannot be pinpointed with certainty as the reason for poor performance. It is possible that the children were not familiar with the sweeteners involved in this study, particularly those of the "weak" samples. Therefore, according to the children, sweetness might not have been the attribute responsible for the difference between the stimuli and caused them not to perform as predicted. The children may have used other easier attributes such as "preference" or "no preference" to complete their task.

Several studies reported that as the complexity of the stimulus increased, children's performance in sensory tests was different than that of adults. Several studies regarding sweetness scaling reported that that perception of sweetness increased slower when sucrose was diluted in orange drinks than when sucrose was diluted in water (De Graaf and Zandstra 1999, James and others 1999, James and others 2003, Temple and others 2002). Oram and others (2001) also reported that children performed more poorly than adults when stimuli changed from a single taste (e.g., sweet) to a binary taste (e.g., sweet and salty). It has been suggested that children probably performed differently due to either their differences in perception, cognitive ability, or their discrimination tactic (Popper and Kroll 2007). It would be advantageous to perform a paradox study using more confusable non-carbonated beverages (like the apple juice used in the first study, Chapter 3) as the stimuli in order to corroborate the findings of this study and those of Chapter 3.

4.5 Conclusion

The paradox of discriminatory non-discriminators was not present for the total population in this study using carbonated beverages sweetened with different sweeteners with children as subjects. Thurstonian predictions that the degree of difference among the two stimuli should be independent of the discrimination test utilized were not confirmed.

4.6 Future Work

The paradox of discriminatory non-discriminators and Thurstonian predictions regarding the degree of difference detected with the triangle and 3-AFC methods will be further explored using apple juice with different sweetness concentrations (100% vs. 75%) as the stimuli.

4.7 References

Anderson DA. 1988. Some models for overdispersed binomial data. Australian Journal of Statistics. 30: 125-148.

Bi J, Ennis DM, O'Mahony M. 1997. How to estimate the variance of d' from difference tests. Journal of Sensory Studies. 12:87-104.

Bi J, Ennis DM. 1998. A Thurstonian variant of the beta-binomial model for replicated difference tests. Journal of Sensory Studies. 13: 461-466.

Braun V, Rogeaux M, Schneid N, O'Mahony M, Rousseau B. 2004. Corroborating the 2-AFC and 2-AC Thurstonian models using both a model system and sparkling water. Food Quality and Preference. 15: 501-507.

Byer AJ, Abrams D. 1953. A comparison of the triangular and two-sample taste-test methods. Food Technology. 7:185-187.

Cox DR.1983. Some remarks on overdispersion. Biometrika. 70: 269-274.

Delwiche J, O'Mahony M. 1996a. Flavour discrimination – an extension of Thurstonian paradoxes to the tetrad method. Food Quality and Preference. 7:1-5.

Ennis DM, Bi J. 1998. The beta-binomial model: accounting for intertribal variation in replicated difference and preference tests. Journal of Sensory Studies. 13: 389-412.

Fritjers JER. 1979a. The paradox of discriminatory nondiscriminators resolved. Chemical Senses and Flavour. 4(4):355-358.

Green DM, Swets JA. 1966. Signal Detection Theory and Psychophysics. Wiley, New York.

Gridgeman NT. 1970. A re-examination of the two-stage triangle test for the perception of sensory differences. Journal of Food Science. 35:87-91.

Jang HJ, Kokrashvili Z, Theodorakis MJ, Carlson O, Kim BJ, Zhou J, Kim HH, Xu X, Chan SL, Juhaszova M, Bernier M, Mosinger B, Robert F. Margolskee RF, Egan JM. 2007. Gut-expressed gustducin and taste receptors regulate secretion of glucagon-like peptide-1. Proceedings National Academy of Science U.S.A. 104 (38): 15069-15074.

Margolskee RF, Dyer J, Kokrashvili Z, Salmon KSH, Ilegems E, Daly K, Maillet EL, Ninomiya Y, Mosinger B, Shirazi-Beechey SP. 2007. T1R3 and gustducin in gut sense sugars to regulate expression of Na+-glucose cotransporter 1. Proceedings of the National Academy of Science U.S.A. 104 (38): 15075-15080.

Masouka S, Hatjopoulos D, O'Mahony M. 1995. Beer bitterness detection: Testing Thurstonian and sequential sensitivity analysis models for triad and tetrad methods. Journal of Sensory Studies. 10:295-306.

O'Mahony M. 1995. Who told you the triangle test was simple? Food Quality and Preference. 6:227-238.

O'Mahony M, Masuoka S, Ishii R. 1994. A theoretical note on difference tests: models, paradoxes and cognitive strategies. Journal of Sensory Studies. 9: 247-272.

O'Mahony M, Rousseau B. 2002. Discrimination testing: A few ideas, old and new. Food Quality and Preference. 14:157-164.

O'Mahony M, Thieme U, Goldstein L. 1988. The warm-up effect as a means of increasing the discriminability of sensory difference tests. Journal of Food Science. 53: 1848-1850.

Peryam DR, Swartz VW. 1950. Measurement of sensory differences. Food Technology. 4: 390-395.

Rousseau B, O'Mahony M. 1997. Sensory difference tests: Thurstonian and SSA predictions for vanilla flavored yogurts. Journal of Sensory Studies. 12:127-146.

Stillman JA. 1993. Response selection, sensitivity, and taste test performance. Perception and Psychophysics. 54:190-194.

Tedja S, Ninaka R. Ennis DM, O'Mahony M. 1994. Triadic discrimination testing – refinement of Thurstonian and SSA approaches. Chemical Senses. 19:279-301.

Thurstone LL. 1927a. Psychophysical analysis. American Journal of Psychology. 38: 368-389.

Thurstone LL. 1927b. A law of comparative judgment. Psychology Review. 34:273-286.

CHAPTER 5. CONFIRMING THE PARADOX OF DISCRIMINATORY NON-DISCRIMINATORS AND THURSTONIAN PREDICTIONS WITH CHILDREN AS SUBJECTS

5.1 Introduction

When subjects perform forced-choice discrimination tests, a greater number of correct responses are observed with one test protocol than with another. This seemingly paradoxical result is observed in the case of the triangle and 3-alternative forced-choice (3-AFC) tests. Both tests are triadic methods that require correct discrimination of one stimulus (S) from a pair of identical stimuli (N-N). The difference between S and N is not specified under the instructions of the triangle test (Peryam and Swartz 1950); therefore, the subjects are required to select the odd stimulus. The 3-AFC (Green and Swets 1966) instructions do specify the nature of the difference and the subject is to select the strongest or weakest stimulus with respect to the specified attribute.

An increased number of correct responses as a result of change in instruction from the triangle method to the 3-AFC method was first observed by Byers and Abrams (1953), a paradox that became known as "the paradox of discriminatory non-discriminators" (Gridgeman 1970). Using Thurstonian (Thurstone 1927a, b) arguments, Fritjers (1979) resolved this paradox (for a detailed review see O'Mahony (1995) and O'Mahony and others (1994)). Thurstonian arguments predict that the measured degree of difference (δ) between two stimuli is independent of the methodology used. Thus, according to Thurstonian predictions, in spite of the larger proportion of correct responses in a 3-AFC test, the experimental estimate of δ (d'), when compared to that of the triangle test, should not be significantly different (Figure 5.1).

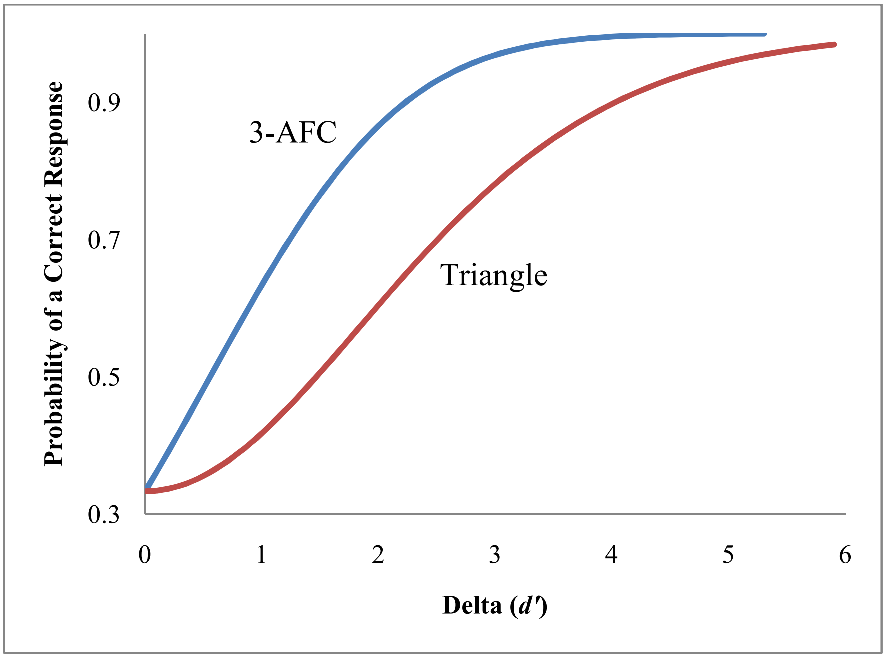


Figure 5.1: Psychometric functions for the triangle and 3-AFC methods (source of data: Ennis, 1993)

It was observed experimentally that the same subjects discriminating between the same stimuli perform a higher proportion of 3-AFC tests correctly than triangle tests, yet have the same *d′* value (Delwiche and O'Mahony 1996, Fritjers 1981, Masouka and others 1995, Rousseau and O'Mahony 1997, Stillman 1993). However, all of these discrimination studies were conducted with adult subjects. Our study involving children discriminating between easily discriminable stimuli (apple juices with different sweetness concentrations, 100% vs. 60%, Chapter 3) revealed that more often than not, children performed a higher number of 3-AFC tests correctly than triangle tests; however, the *d′* values were significantly different. On the other hand, our study involving hardly discriminable stimuli (carbonated beverages with different sweeteners, Chapter 4) resulted in children performing a greater number of triangle tests correctly than 3-AFC tests, and their *d′* values were significantly different. For the former study it was concluded that performing a study with more confusable stimuli would confirm or disprove the paradox and Thurstonian predictions. Results from the latter study revealed that the carbonated beverages used as hardly discriminable stimuli may have been too confusable for the children. We predict that perhaps a third study involving non carbonated beverages (e.g., apple juice) with more similar sweetness concentrations, yet still confusable, will corroborate or contradict our previous findings.

Therefore, the research objectives were to (1) challenge the Thurstonian prediction for the triangle and 3-AFC protocols with children as subjects using apple juice with different sweetness concentrations (100% vs. 75%) as the stimuli, and (2) determine age effects. For this purpose, children (N =404) between 6-12 years old from Baton Rouge, Louisiana, USA participated in the discrimination tests.

5.2 Materials and Methods

5.2.1 Subjects

The panelists (N = 404) were first through sixth grade students from elementary schools in Baton Rouge, Louisiana, USA. Table 5.1 presents the distribution of subjects per grade.

Table 5.1: Distribution of subjects per grade

Age Group[a]	1	2	3	4	5	6
N	96	55	48	80	51	74

[a] Age group categories refer to grades 1st -6th

Criteria for recruitment of participants were that they were between the first and sixth grade and were not allergic to any of the ingredients present in the juice products. Participants were required to have parental consent (see Appendix A) and to sign an assent form (Figure 5.2 , also see Appendix B) stating their willingness to participate, both approved by Louisiana State University AgCenter Institutional Review Board prior to participating in the testing. No monetary incentive or rewards were given to subjects for participation.

Figure 5.2: Child filling out the child assent form

5.2.2 Stimuli

All samples were presented in approximately 1.5 fl oz aliquots in 2 fl oz lidded plastic cups (Dart ® Conex ® Complements 2oz black portion containers, Dart Container Corporation, Mason, Mi 48854 USA) and served at room temperature (approximately 25-27^{o}C, depending on the room temperature at the particular site, but constant in a given session). Samples were labeled with three-digit codes.

Stimuli consisted of apple juice (Mott's® Original 100% apple juice, Mott's LLP, Rye Brook, NY). The discrimination was performed between the actual product and juice that had been diluted with water to 75% by weight. The former will be referred to as 'regular' apple juice and the latter as the 'reduced sugar' apple juice. See Appendix C for a complete list of ingredients.

5.2.3 Experimental Design

Each subject performed one triangle test and one 3-AFC test to discriminate regular apple juice from the reduced sugar apple juice. Triangle and 3-AFC tests were performed on two different days; the triangle test was performed on the first day and the 3-AFC test on the second day using the following combinations: WWS, WSW, SWW.

The hypotheses being tested for the d' values were:

H_o: Triangle $d' =$ 3-AFC d'

H_a: Triangle $d' \neq$ 3-AFC d'

5.2.4 Testing Protocol

The discrimination tests were conducted in English and performed in the children's classroom settings at the schools. Before each experimental session the children were given a presentation in order to ensure that they understood the basic logistics of the testing session. During the presentation a review of the terms *same*, *different*, *sweet*, *not sweet*, *sweeter*, *less sweet,* and *less sugar* was given to the children. Also, the testing procedure was explained along with an overview of the questionnaire (see Appendix D) and its proper fill-out.

The triadic tests were given under two possible sets of instructions. For the triangle test, the nature of the difference was not specified and the instructions were: "here are three juice samples; two are the same and one is different: **circle the juice that tastes different**". For the 3-AFC test the nature of the difference was specified and the instructions were: "here are three juice samples; one is sweeter than the other two: **circle the juice that tastes sweeter**".

Subjects began each session by filling out the demographic information section of the questionnaire and cleansing their palate with water (Ozarka ® natural Spring Water, Ozarka Spring Water Company, Division of Nestle Waters North America Inc., Greenwch, CT 06830, USA). Next, the three samples were presented simultaneously to the subjects and they were instructed to taste them from left to right. The subjects were allowed to retaste the sample with the condition that they always tasted all three samples in the order presented to them. Session lengths ranged from 20 to 30 minutes.

Research showed that sweet taste receptors, in addition to being present in the mouth cavity and esophagus, are also present in the gastrointestinal tract of humans (Jang and others 2007; Margolskee and others 2007). Our preliminary study showed that there was no significant difference in discrimination between swallowing and expectorating a sample (see Appendix E). In this study, the children were therefore instructed to swallow the samples as it was a more natural behavior that caused less distraction.

5.2.4.1 Experimental Session Overview

The following is a detailed protocol of the testing session:

- Day 1: Triangle Test
    - Children filled out the child assent form
    - Children were presented with the questionnaire and water
    - Children filled out the top portion of the questionnaire (demographic information: name, age, gender, grade)
    - Overview presentation for the triangle test
    - The three samples for the test were simultaneously presented

- Children cleansed their palates
- Children were instructed to taste the samples
    - Taste the first sample
    - Taste the second sample
    - Taste the third sample
- Children were instructed to circle on the questionnaire the juice that tasted "different"
- Children stopped
- End of session

- Day 2: 3-AFC Test
    - Children filled out the child assent form
    - Children were presented with the questionnaire and water
    - Children filled out top portion of questionnaire (Demographic information: name, age, gender, grade)
    - Overview presentation for the 3-AFC test
    - The three samples for the test were simultaneously presented
    - Children cleansed their palates
    - Children were instructed to taste the samples
        - Taste the first sample
        - Taste the second sample
        - Taste the third sample
    - Children circled the juice that was sweeter or had more sugar
    - Children stopped
    - End of session

5.2.5 Data Analysis

For each testing method, the number of correct responses provided by the subjects was counted and recorded. The P_cs for both tests were used to determine the corresponding d' values (Ennis 1993). The experimental estimate d', the variance of d', and the test of significance were obtained using the IFPrograms software Course Version 8.5.0320 (The Institute for Perception, Richmond, VA, USA). The d' values were calculated under the assumptions that the intensity distributions for the two stimuli were unidimenasional normal distributions that had equal variance (Bi and others 1997). Alternatively, d' values for the triangle and 3-AFC tests could be obtained from published tables (e.g., Ennis 1993, see Appendix F). The variance of d' and the test statistic can be obtained by the approach described by Bi and others (1997) (see Appendix G and H). All analyses were evaluated at $\alpha = 0.05$.

5.3 Results and Discussion

5.3.1 Overall Performance

Table 5.2 shows the overall performance of this USA population for each method; the pooled P_cs and corresponding d' values are presented. The P_c for the 3-AFC method (0.62) was higher than the P_c for the triangle method (0.43). Therefore, the paradox was confirmed under the circumstance of this study using apple juice with different sweetness concentrations (100% vs. 75%) as the stimuli.

Table 5.2: Pooled proportions of correct responses, pooled d' values, test statistic and p-value for the triangle and 3-AFC methods for all six grades (N = 404)

Triangle Method			**3-AFC Method**			**d' Sig Test** [b]	
P_c [a]	**d'**	**variance**	**P_c** [a]	**d'**	**variance**	**χ^2**	**p-value** [c]
0.43	1.09	0.022	0.62	0.97	0.007	0.5	0.48

[a] P_c = proportion of correct responses

[b] Significance test for d' values, Critical χ^2 value = 3.84, $\alpha = 0.05$

[c] All p-values > 0.05 are not significant

For the triangle method, there were 43% correct responses, which corresponded to a *d′* of 1.09. In the 3-AFC task, 62% correct selections were made, which corresponded to a *d′* of 0.97. Comparing pooled *d′* values for each protocol showed no significant difference among them (p = 0.48) (Table 5.2). Therefore, Thurstonian predictions were confirmed for this population of USA children between the ages of 6-12, using apple juice with different sweetness concentrations (100% vs. 75%) as the stimulus.

5.3.2 Individual Performance

Table 5.3 presents the results of the different age groups. For all grades, the 3-AFC P_c was greater than the triangle method P_c; therefore, the paradox was confirmed for all age groups.

Table 5.3: Proportions of correct responses, *d′* values and significance test for the triangle and 3-AFC methods for the individual grades

		Triangle			3-AFC			*d′* Sig Test [c]	
Grade	**N**	P_c [a]	***d′***	**variance**	P_c [a]	***d′***	**variance**	χ^2	**p-value** [d]
1	96	0.35	0.48	0.330	0.45	0.39	0.028	0.02	0.88
2	55	0.34	0.37	0.954	0.56	0.76	0.049	0.15	0.70
3	48	0.44	1.12	0.182	0.63	0.97	0.057	0.09	0.76
4	80	0.53	1.60	0.082	0.80	1.65	0.044	0.02	0.89
5	51	0.35	0.47	0.656	0.67	1.12	0.056	0.59	0.44
6	74	0.55	1.74	0.085	0.68	1.15	0.039	2.81	0.09

[a] P_c = proportion of correct responses
[b] Significance test for *d′* values, Critical χ^2 value = 3.84, α = 0.05
[c] All p-values > 0.05 are not significant

In our previous studies regarding the paradox of discriminatory non-discriminators with apple juice with different sweetness concentrations as the stimuli (100% vs. 60%, Chapter 3), an increase in performance with age was observed for all populations. Even though a consistently increasing trend of P_c with age was not present for the population in this study, an overall increase was observed. For the triangle test, there was an increase from 35% correct selections made by the first graders to 55% in the sixth grade. For the 3-AFC method the increase was from

45% correct selections made by first graders to 68% for the sixth graders. For the triangle test data, a decrease in P_c (0.35) occurred for the fifth graders with respect to the trend observed. Similarly, for the 3-AFC data, an increase in P_c (0.80) was observed for the fourth grade. The reason for this outcome is not obvious and may require further investigation. These results could consist of a onetime incident and probably do not indicate a repetitive occurrence. Overall, there was an increase in performance (P_c) with an increase in age, which lead to higher d' values as age increased. For that reason, children in the higher grades detected a greater degree of difference among the stimuli than the lower grade children.

The d' values for the individual grades are also presented in Table 5.3. The d' values of all age groups were not significantly different from each other, thereby confirming Thurstonian predictions. Results from this investigation confirm those observed in other studies utilizing adult subjects (Delwiche and O'Mahony 1995, Masouka and others 1995, Rousseau and O'Mahony 1997, Stillman 1993, Tedja and others 1994).

An overview of our findings for all studies is presented in Table 5.4. For easily discriminable stimuli (apple juice 100% vs. 60%, Chapter 3), the paradox was present in each country and for the combined populations; yet, the Thurstonian predictions were not confirmed. We hypothesized that utilizing more confusable stimuli would yield results that agreed with the theory. When the difficulty of the stimuli increased to hardly discriminable (Chapter 4) the paradox was only present for the Honduran and Mexican populations and the Thurstonian predictions were not confirmed for any of the countries. Due to this outcome it was concluded that the carbonated beverages used as the stimuli were too confusable for the children; therefore, a noncarbonated beverage was chosen for a final study. Apple juice with different sweetness concentrations (100% vs. 75%) categorized as confusable stimuli was used. For this last study both the paradox and the Thurstonian predictions were confirmed.

Table 5.4: Summary of result for the individual populations and combined populations for all studies performed regarding the paradox of discriminatory non-discriminators and Thurstonian predictions for the triangle and 3-AFC methods

	Chapter 3	**Chapter 4**	**Chapter5**
	Easily[a] Discriminable Stimuli	**Hardly[b] Discriminable Stimuli**	**Confusable[c] Stimuli**
		HONDURAS	
Paradox[d]	Y[f]	Y	--- [g]
Th Pred[e]	Y	N	--- [g]
		MEXICO	
Paradox	Y	Y	--- [g]
Th Pred	N	N	--- [g]
		THAILAND	
Paradox	Y	N	--- [g]
Th Pred	N	N	--- [g]
		USA	
Paradox	Y	N	Y
Th Pred	N	N	Y
		OVERALL	
Paradox	Y	N	Y
Th Pred	N	N	Y

[a] Apple Juice 100% vs. 60%
[b] Carbonated Beverages
[c] Apple Juice 100% vs. 75%
[d] Paradox = Paradox of discriminatory non-discriminators
[e] Th Pred = Thurstonian predictions that despite the difference in proportions of correct responses among the triangle and 3-AFC methods the degrees of difference detected among the set of stimuli by each method should not be significantly different from each other
[f] Y = Yes, N = No
[g] This study was not performed for this population

Table 5.5 summarizes the results for the individual age groups for all populations combined. The paradox was present in all grades in both studies performed with apple juice as the stimuli. However, this paradox was only observed for the second and third grades when the carbonated beverages were used as the stimuli. The confirmation of Thurstonian predictions was not consistent across all age groups with the easily discriminable stimuli but was present in all grades for the confusable stimuli. On the contrary, such predictions were not confirmed for any of the grades for hardly discriminable stimuli.

Table 5.5: Summary of result for the individual grades for all populations combined for all studies performed regarding the paradox of discriminatory non-discriminators and Thurstonian predictions for the triangle and 3-AFC methods

	Chapter 3		**Chapter 4**		**Chapter5**	
	Easily[a] Discriminable Stimuli		**Hardly[b] Discriminable Stimuli**		**Confusable[c] Stimuli**	
Age Group	**Paradox[d]**	**TP[e]**	**Paradox**	**TP**	**Paradox**	**TP**
1	Y[f]	Y	N	N	Y	Y
2	Y	Y	Y	N	Y	Y
3	Y	Y	Y	N	Y	Y
4	Y	N	N	N	Y	Y
5	Y	Y	N	N	Y	Y
6	Y	N	N	N	Y	Y

[a] Apple Juice 100% vs. 60%
[b] Carbonated Beverages
[c] Apple Juice 100% vs. 75%
[d] Paradox = Paradox of discriminatory non-discriminators
[e] Th Pred = Thurstonian predictions that despite the difference in proportions of correct responses among the triangle and 3-AFC methods the degrees of difference detected among the set of stimuli by each method should not be significantly different from each other
[f] Y = Yes, N = No

The paradox and Thurstonian predictions regarding this paradox were corroborated, both overall and in the individual grades, using confusable stimuli (100% vs. 75% apple juice) as the stimuli (Table 5.5). From our studies (Chapters 3, 4, 5) we may conclude that the paradox of discriminatory non-discriminators and Thurstonian predictions may or may not be confirmed depending on the stimuli and degree of confusability as observed in Table 5.5. Would one expect the same results if the complexity of the discrimination task (i.e., more samples served per set) increased? How would children perform in different discrimination tests?

5.4: Conclusion

The paradox of discriminatory non-discriminators was present, overall and in the individual grades, in this study using apple with different sweetness concentrations (100% vs. 75%) with children as subjects. Thurstonian predictions that the degree of difference among the two stimuli should be independent of the discrimination test utilized were also confirmed.

5.5 Future Work

Using apple juice with different sweetness concentrations (100% vs. 75%) as the stimuli, the paradox of discriminatory non-discriminators and Thurstonian predictions regarding the variants of the method of tetrads will be explored with children as subjects. There is no published research regarding children's performance in the method of tetrads and it will be of great scientific interest to know whether Thurstonian predictions for this method hold with children as subjects.

5.6 References

Anderson DA. 1988. Some models for overdispersed binomial data. Australian Journal of Statistics. 30: 125-148.

Bi J, Ennis DM, O'Mahony M. 1997. How to estimate the variance of d' from difference tests. Journal of Sensory Studies. 12:87-104.

Bi J, Ennis DM. 1998. A Thurstonian variant of the beta-binomial model for replicated difference tests. Journal of Sensory Studies. 13: 461-466.

Braun V, Rogeaux M, Schneid N, O'Mahony M, Rousseau B. 2004. Corroborating the 2-AFC and 2-AC Thurstonian models using both a model system and sparkling water. Food Quality and Preference. 15: 501-507.

Byer AJ, Abrams D. 1953. A comparison of the triangular and two-sample taste-test methods. Food Technology. 7:185-187.

Cox DR.1983. Some remarks on overdispersion. Biometrika. 70: 269-274.

Delwiche J, O'Mahony M. 1996a. Flavour discrimination – an extension of Thurstonian paradoxes to the tetrad method. Food Quality and Preference. 7:1-5.

Ennis DM. 1993. The power of sensory discrimination methods. Journal of Sensory Studies. 8:353-370.

Ennis DM, Bi J. 1998. The beta-binomial model: accounting for intertribal variation in replicated difference and preference tests. Journal of Sensory Studies. 13: 389-412.

Ennis DM. 1990. Relative power of difference testing methods in sensory evaluation. Food Technology. 44(4):114-118.

Fritjers JER. 1979a. The paradox of discriminatory nondiscriminators resolved. Chemical Senses and Flavour. 4(4):355-358.

Green DM, Swets JA. 1966. Signal Detection Theory and Psychophysics. Wiley, New York.

Gridgeman NT. 1970. A re-examination of the two-stage triangle test for the perception of sensory differences. Journal of Food Science. 35:87-91.

Jang HJ, Kokrashvili Z, Theodorakis MJ, Carlson O, Kim BJ, Zhou J, Kim HH, Xu X, Chan SL, Juhaszova M, Bernier M, Mosinger B, Robert F. Margolskee RF, Egan JM. 2007. Gut-expressed gustducin and taste receptors regulate secretion of glucagon-like peptide-1. Proceedings National Academy of Science U.S.A. 104 (38): 15069-15074.

Margolskee RF, Dyer J, Kokrashvili Z, Salmon KSH, Ilegems E, Daly K, Maillet EL, Ninomiya Y, Mosinger B, Shirazi-Beechey SP. 2007. T1R3 and gustducin in gut sense sugars to regulate expression of Na+-glucose cotransporter 1. Proceedings of the National Academy of Science U.S.A. 104 (38): 15075-15080.

Masouka S, Hatjopoulos D, O'Mahony M. 1995. Beer bitterness detection: Testing Thurstonian and sequential sensitivity analysis models for triad and tetrad methods. Journal of Sensory Studies. 10:295-306.

O'Mahony M. 1995. Who told you the triangle test was simple? Food Quality and Preference. 6:227-238.

O'Mahony M, Masuoka S, Ishii R. 1994. A theoretical note on difference tests: models, paradoxes and cognitive strategies. Journal of Sensory Studies. 9: 247-272.

O'Mahony M, Rousseau B. 2002. Discrimination testing: A few ideas, old and new. Food Quality and Preference. 14:157-164.

O'Mahony M, Thieme U, Goldstein L. 1988. The warm-up effect as a means of increasing the discriminability of sensory difference tests. Journal of Food Science. 53: 1848-1850.

Peryam DR, Swartz VW. 1950. Measurement of sensory differences. Food Technology. 4: 390-395.

Rousseau B, O'Mahony M. 1997. Sensory difference tests: Thurstonian and SSA predictions for vanilla flavored yogurts. Journal of Sensory Studies. 12:127-146.

Stillman JA. 1993. Response selection, sensitivity, and taste test performance. Perception and Psychophysics. 54:190-194.

Tedja S, Ninaka R. Ennis DM, O'Mahony M. 1994. Triadic discrimination testing – refinement of Thurstonian and SSA approaches. Chemical Senses. 19:279-301.

Thurstone LL. 1927a. Psychophysical analysis. American Journal of Psychology. 38: 368-389.

Thurstone LL. 1927b. A law of comparative judgment. Psychology Review. 34:273-286.

CHAPTER 6: CORROBORATING THURSTONIAN PREDICTIONS FOR THE VARIANTS OF THE TETRAD METHOD WITH CHILDREN AS SUBJECTS

6.1 Introduction

Within the forced-choice discrimination methods, there are certain protocols under which subjects demonstrate superior performance. In the case of triadic methods, the triangle method results in lower proportion of correct responses than the 3-AFC method. For the triangle method the subject has to identify the odd sample without the nature of the difference among the stimuli being specified, and for the 3-AFC method the subject has to select the sample that is stronger in a given sensory dimension. Fritjers (1979) showed that the decision rule the subject uses for selecting the appropriate stimulus is significantly affected by the instruction given. When instructed to select the odd stimulus, distances between momentary sensory perceptions are compared, thus bringing forth the "comparison of distances" strategy. On the other hand, when instructed to select the weakest or strongest stimulus a comparison of the absolute momentary sensations must be made.

Thurstonian arguments are based on the assumptions of normal perceptual distributions for the stimuli and cognitive strategies or decision rules (Thurstone, 1927a,b). They provide an integral measure of the difference between two stimuli. Thus, according to Thurstonian predictions, in spite of the larger probability of a correct response in the 3-AFC method when compared to that of the triangle test, δ values should be similar (O'Mahony and others 2002). The paradox can be generalized to tests with more than one kind of stimulus (two-out-of-five, two-out-of-six, three-out-of seven, etc.) and tests with an unequal number of stimuli in each class (triangle, 3-AFC, 4-AFC, etc) (O'Mahony and others 1994). However, this generalization does not apply when each class has an equal number of stimuli, as is observed for the tetrad method (Delwiche and O'Mahony 1996).

The tetrad method involves four stimuli, two of one kind and two from a different kind (AA-BB). There are two variants of the tetrad method: the unspecified and the specified tests. For the unspecified method, with a chance probability = ⅓, the subject is required to correctly separate the four stimuli into their two appropriate sets. For the specified method, with a chance probability = $^1/_6$, the subject has to correctly identify the two strongest stimuli from a group of four samples. O'Mahony and others (1994) thoroughly discussed with Thurstonian arguments why an increase in performance is not observed over a certain range of δ when the instructions for the tetrad method change from the unspecified to the specified method. The psychometric functions for the variants of the method of tetrads have been derived by Ennis and others (1998).

It was observed experimentally that there was no significant difference in performance between the variants of the tetrad method for the same subjects discriminating between the same stimuli. These Thurstonian predictions for the tetrad method were tested and confirmed for bitterness discrimination in beer (Masouka and others 1995) and flavor discrimination in puddings (Delwiche and O'Mahony 1996). However, all of these tests were conducted with adult subjects. Would the same predictions apply if the subjects were children?

The research objective was to challenge the Thurstonian predictions for the specified and unspecified method of tetrads. For this purpose children between 6-12 years old were tested using apple juice with different sweetness concentrations as the stimuli.

6.2 Materials and Methods

6.2.1 Subjects

Four hundred and four (404) subjects (age range 6-11 years) participated in this study. They were first through sixth grade students from elementary schools in Baton Rouge, Louisiana, USA. The breakdown by grade was: 96 first graders, 55 second graders, 48 third graders, 80 fourth graders, 51 fifth graders, and 74 sixth graders.

Criteria for recruitment of participants were that they were between the first and sixth grade and were not allergic to any of the ingredients of the product. Participants were required to have parental consent and to sign an assent form, both approved by Louisiana State University AgCenter Institutional Review Board prior to participating in the testing. No monetary incentive or rewards were given to subjects for participation.

6.2.2 Stimuli

Stimuli consisted of apple juice (Mott's® Original "100% apple juice", Mott's LLP, Rye Brook, New York, USA). The discrimination was performed between the pure product and juice that had been diluted with water to 75% by weight. All samples were presented in approximately 1.5 fl oz aliquots in lidded black plastic cups (Dart® Conex Complements® Black Portion 2oz, Dart Container Corporation, Mason, Michigan, USA) and served at room temperature (approximately 25-27^{o}C, depending on the room temperature at the particular site, but constant in a given session).

6.2.3 Testing Procedure

The tetrad methods were performed in two sessions in different days; i.e., session 1/day 1: unspecified method and session 2/day 2: specified method. This order was used to avoid the possible influence of the specified attribute on the unspecified method. Session length ranged between 15 and 25 minutes, varying according to grade. Before each experimental session, the children were given a presentation in order to explain the basic logistics of the testing session. During the presentation a review of the terms *same* and *different* (session 1) and *sweet* and *not sweet/less sweet* (session 2) was given to the children. For session 1, the children were asked to define *same* and *different*, to give examples, and to determine if the foods shown were same or different. For session 2, they were asked to define *sweet* and *not sweet/less sweet*, give examples of such foods, and to identify which of the foods shown were sweeter. Also, a demonstration of the tasting procedure was given along with an overview of the questionnaire (see Appendix D) and its proper fill-out.

The tetrad tests were given under two possible sets of instructions. For the unspecified method, the nature of the difference was not specified and the instructions were: "here are four juice samples; two belong to one group and the other two belong to a different group: **separate them according to their taste into two groups of two**." For the specified method the nature of the difference was specified and the instructions were: "here are four juices; two are sweeter than the rest: **identify the two juices that are sweeter**".

During each session the subjects performed one tetrad test, either unspecified or specified. Each subject performed one of the 24 possible permutations, with each child evaluating the same permutation during both sessions. Subjects began each session by rinsing their mouths with room temperature water (Ozarka ® natural Spring Water, Ozarka Spring Water Company, Division of Nestle Waters North America Inc., Greenwch, CT 06830, USA). The children were then simultaneously presented with the four samples and instructed to taste them in the order given from left to right. Retasting was allowed given that child tasted all the samples in the given order.

6.2.4 Data Analysis

For each testing method, the number of correct responses provided by the subjects was counted and recorded. Significance among the proportion of correct responses (P_c) for both methods was assessed using a Z test. The P_cs for both tests were used to determine the corresponding *d′* values (Ennis and others, 1998).

The experimental estimate *d′*, the variance of *d′*, and the test of significance were obtained using the IFPrograms Version 8.5.0320 (The Institute for Perception, Richmond, VA, USA). The *d′* values were calculated under the assumptions that the two intensity distributions for the stimuli are unidimenasional normal distributions that have equal variance (Bi and others 1997). Alternatively, *d′* values could be obtained from published tables (e.g., Ennis and others, 1998; see Appendix F). The variance of *d′* is obtained from IFPrograms which inverts the second derivative of the likelihood function at the estimate value. The test statistic can be obtained by the approach described by Bi and others (1997) (see Appendix H). All analyses were evaluated at $\alpha = 0.05$.

6.3 Results and Discussion

6.3.2 Overall Performance

Performance data and *d′* values for the unspecified and specified methods of tetrads are presented in Table 6.1 for the overall population. Thurstonian modeling indicates that for the method of tetrads specifying the difference between the stimuli will not result in better performance at $\delta > 1$ and therefore the *d′* values should be similar (O'Mahony and others, 1994). For the entire population (N=404), a significant difference did not exist between the P_cs for the unspecified and specified tests ($p = 0.62$, Table 6.1). For the *d′* values of the 404 subjects a more powerful approach by pooling the data and calculating pooled *d′* values was performed (Braun and others, 2004). The pooled *d′* values for the unspecified and specified tests were also not significantly different from each other ($p = 0.32$, Table 1). These findings are supported by Thurstonian modeling, which indicates that for *d′* values greater than 1 for the tetrad methods differences in performance should be negligible (see Figure 6.1). In the case of this study the pooled d' values for both tests are greater than 1.

Table 6.1: Pooled proportion of correct responses, pooled *d'* values, test statistics (N=404) and p-values for the variants of the method of tetrads for all six grades

Unspecified Method			**Specified Method**			**P_c Sig Test [b]**		**d' Sig Test [c]**	
P_c [a]	d'	**variance**	P_c [a]	d'	**variance**	**Z**	**p-value [d]**	χ^2	**p-value [d]**
0.54	1.18	0.007	0.56	1.29	0.005	0.50	0.62	1.01	0.32

[a] P_c = proportion of correct responses
[b] Significance test for P_c, Critical Z value = 1.96, $\alpha = 0.05$
[c] Significance test for d' values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[d] All p-values > 0.05 are not significant

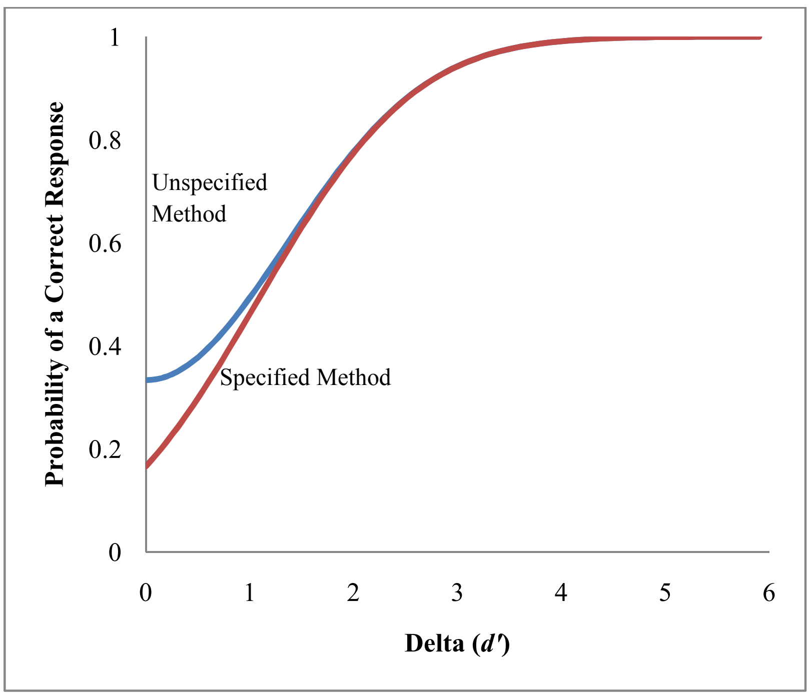


Figure 6.1: Psychometric functions for the specified and unspecified method of tetrads (source of data: Ennis and others, 1998)

When comparing the variants of the method of tetrads, Masouka and others (1995) reported that for bitterness discrimination in beer (N=9, 22-37 years of age), specifying the nature of the difference among the stimuli did not result in significantly better performance between the specified and unspecified methods. Delwiche and O'Mahony (1996) also reported that for discrimination among chocolate puddings with different added flavors (N=13, 20-45 years of age) there was no significant difference in performance. Our results are in accord with those presented above. Based on a larger sample size, these results support the argument that the 3-AFC method results in a greater P_c because of the decision rule used not because the nature of the difference among the stimuli was specified.

6.3.2 Individual Performance

Performance in each of the individual grades was further explored (Table 6.2). Results clearly show that there was an increase in performance with an increase in age for subjects in the first through fourth grade. For the unspecified method *d'* values increased from 0.70 for first grade to 1.40 for fourth grade and for the specified method the values increased from 0.61 to 1.64.

Table 6.2: Proportion of correct responses, *d'* values and significance tests for the variants of the tetrad method for the individual grades

		Unspecified Method			Specified Method			P_c Sig Test [b]		*d'* Sig Test [c]	
Grade	**Subjects**	P_c [a]	*d'*	**variance**	P_c [a]	*d'*	**variance**	**Z**	**p-value** [d]	χ^2	**p-value** [d]
1	**96**	0.42	0.70	0.051	0.33	0.61	0.023	1.19	0.23	0.109	0.74
2	**55**	0.47	0.92	0.063	0.45	0.98	0.039	0.19	0.85	0.035	0.85
3	**48**	0.56	1.24	0.058	0.67	1.61	0.049	1.05	0.29	1.280	0.26
4	**80**	0.61	1.40	0.033	0.68	1.64	0.030	0.83	0.41	0.910	0.34
5	**51**	0.73	1.80	0.053	0.53	1.20	0.042	2.05	0.04	3.790	0.05
6	**74**	0.55	1.21	0.038	0.77	1.98	0.037	2.78	0.01	7.900	0.01

[a] P_c = proportion of correct responses
[b] Significance test for P_c, Critical Z value = 1.96, $\alpha = 0.05$
[c] Significance test for *d'* values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[d] All p-values > 0.05 are not significant

For the individual grades the P_cs and the *d'* values between the methods were not significantly different from each other. A different trend is observed for the fifth and sixth grades. For fifth graders the *d'* values among the methods are not significantly different from each other; however, the P_c were significantly different. For sixth graders there was a lower choice probability than observed for the fourth and fifth grades for the unspecified method, for this method the sixth graders performed more like third graders.

6.4 Conclusion

Thurstonian predictions regarding the method of tetrads were confirmed for sweetness discrimination in apple juice with children as subjects under the test conditions in this study. These results once again confirm that specifying the nature of the difference among the stimuli, thus eliciting the 'skimming strategy', does not result in a greater proportion of tests scored correctly. Further investigation with children would increase our knowledge regarding the established and corroborated theoretical basis of discrimination testing with adults. Our future goal is to compare performance in the methods of tetrads with the 2-AFC method, triangle method, and 3-AFC method using the same children and stimuli.

6.5 References

Bi J, Ennis DM, O'Mahony M. 1997. How to estimate and use the variance of d' from difference tests. Journal of Sensory Studies. 12: 87-104.

Birch LL .1979a. Dimensions of preschool children's food preferences and consumptions patterns. Journal of Nutritional Education. 11: 77-80.

Birch LL. 1979b. Preschool children's food preferences and consumption patterns. Journal of Nutritional Education. 11: 189-192.

Braun V, Rogeaux M, Schneid N, O'Mahony M, Rousseau B. 2004. Corroborating the 2-AFC and 2-AC Thurstonian models using both a model system and sparkling water. Food Quality and Preference. 15, 501-507.

Delwiche J, O'Mahony M. 1996. Flavour Discrimination: an extension of Thurstonian 'paradoxes' to the tetrad method. Food Quality and Preference. 7: 1-5.

Ennis J, Ennis D M, Yip D, O'Mahony M. 1998. Thurstonian models for variants of the method of tetrads. British Journal of Mathematical and Statistical Psychology. 51: 205-215.

Fritjers JER. 1979. The paradox of discriminatory nondiscriminators resolved. Chemical Senses and Flavour. 4: 355-358.

Gridgeman NT. 1956. Group size in taste sorting trials. Food Research. 21: 534-539.

Lockhart EE. 1951. Binomial systems and organoleptic analysis. Food Technology. 5: 428-431.

Masuoka S, Hatjopoulos D, O'Mahony M. 1995. Beer bitterness detection: testing Thurstonian and sequential sensitivity analysis models for triad and tetrad methods. Journal of Sensory Studies. 10: 295-306.

O'Mahony M, Masouka S, Ishii R. 1994. A theoretical note on difference tests: models, paradoxes and cognitive strategies. Journal of Sensory Studies. 12: 127-146.

Stillman JA. 1993. Response selection, sensitivity, and taste test performance. Perception and Psychophysics. 54: 190-194.

Thurstone LL. (1927a). Psychophysical analysis. American Journal of Psychology. 38: 368-389.

Thurstone LL. (1927b). A law of comparative judgment. Psychological Review. 34: 273-286.

CHAPTER 7: DISCRIMINATION TESTING WITH CHILDREN: COMPARISON OF PERFORMANCE IN 2-AFC, TRIADIC AND TETRADS METHODS

7.1 Introduction

Forced-choice discrimination tests are intended for discrimination among confusable stimuli. Comparison of performance, in terms of P_c and d', on these discrimination methods was the subject of several studies (Braun and others, 2004; Buchanan and others, 1987; Byer and Abrams, 1953; Delwiche and O'Mahony, 1996; Dessirier and O'Mahony, 1999; Francois and Sauvageot, 1988; Geelhoed and others, 1994; Hautus and Irwin, 1995; Huang and Lawless, 1998; Kim and others, 2006; Kuesten, 2001; Lau and others, 2004; MacRea and Geelhoed, 1992; Masouka and others, 1995; Rousseau and O'Mahony, 1997, 2000, 2001; Rousseau and O'Mahony, 1997; Rousseau and others, 2002; Stillman and Irwin, 1995; Stillman, 1993; Tedja and others, 1994).

Thurstonian modeling (Thurstone, 1927a, b) facilitates the comparison of performance in discrimination tests. The degree of difference between the means of the perceptual distributions of the two stimuli (δ) is independent of the methodology used and, therefore, can be used to compare performance among different methods even if their guessing probabilities are different (Ennis, 1990; O'Mahony, 1995; O'Mahony and others, 1994).

Comparison of performance in the unspecified method of tetrads (O'Mahony and others, 1994) vs. the triangle test (Peryam and Swartz, 1950) has not been studied. However, some studies were conducted separately with these methods, and it is possible to establish a comparison from the reported results (Delwiche and O'Mahony, 1996; Masouka and others 1995). There is also lack of studies regarding the comparison of the specified method of tetrads (O'Mahony and others 1994) with the 3-AFC (Green and Swets, 1966) and 2-AFC methods (Meilgaard and others, 2007). However, performance was compared for the 2-AFC and 3-AFC methods (Dessirier and O'Mahony, 1999; Rousseau and O'Mahony, 1997). All the studies regarding the comparison among discrimination methods were performed with adult subjects. Would the same results be obtained if performance in discrimination methods was compared with children as subjects?

Therefore, the research objective was to compare children's performance in unspecified discrimination methods (unspecified method of tetrads and triangle test) and the specified methods (specified tetrads, 3-AFC and 2-AFC). For this purpose children (N = 404) between 6 and 12 years old from Baton Rouge, Louisiana, USA participated in the discrimination tests using apple juice (confusable stimuli) with different sweetness concentrations as the stimuli.

7.2 Materials and Methods

7.2.1 Subjects

The panelists (N = 404) were first through sixth grade students from elementary schools in Baton Rouge, Louisiana, USA. Table 7.1 presents the distribution of subjects per grade.

Table 7.1: Distribution of subjects per grade

Age Group[a]	**1**	**2**	**3**	**4**	**5**	**6**
N	96	55	48	80	51	74

[a] Age group categories refer to grades 1^{st} -6^{th}

Criteria for recruitment of participants were that they were between the first and sixth grade and were not allergic to any of the ingredients present in the juice products. Participants were required to have parental consent (see Appendix A) and to sign an assent form (Appendix B) stating their willingness to participate, both approved by Louisiana State University AgCenter Institutional Review Board prior to participating in the testing. No monetary incentive or rewards were given to subjects for participation.

7.2.2 Stimuli

Stimuli consisted of apple juice (Mott's® Original "100% apple juice", Mott's LLP, Rye Brook, New York, USA). The discrimination was performed between the pure product and juice that had been diluted with water to75 % by weight. The former will be referred to as 'regular' apple juice and the latter as the 'reduced sugar' apple juice. See Appendix C for a complete list of ingredients.

All samples were presented in approximately 1.5 fl oz aliquots in lidded black plastic cups (Dart® Conex Complements® Black Portion 2oz, Dart Container Corporation, Mason, Michigan, USA) and served at room temperature (approximately 25-27^{o}C, depending on the room temperature at the particular site, but constant in a given session).

7.2.3 Experimental Design

The unspecified and specified tests were performed in two sessions in different days; i.e., session 1/day 1: unspecified tetrads and triangle test and session 2/day 2: specified tetrads, 3-AFC and 2-AFC. Each subject performed one test of each method to discriminate regular apple juice from the reduced-sugar apple juice. Subjects performed one of the 24 possible permutations of the unspecified and specified methods of tetrads, one of the six possible combinations (SSW, SWS, WSS, WWS, WSW, SWW) of the triangle test, one of the three possible orders of presentation (WWS, WSW, SWW) for the 3-AFC method, and one of the two possible combinations (AB, BA) for the 2-AFC method.

The hypotheses being tested for the d' values were:

H_o: Unspecified Tetrads d' = Triangle d'

H_a: Unspecified Tetrads $d' \neq$ Triangle d'

H_o: Specified Tetrads $d' = $ 3-AFC d'

H_a: Specified Tetrads $d' \neq $ 3-AFC d'

H_o: Specified Tetrads $d' = $ 2-AFC d'

H_a: Specified Tetrads $d' \neq $ 2-AFC d'

H_o: 3-AFC $d' = $ 2-AFC d'

H_a: 3-AFC $d' \neq $ 2-AFC d'

7.2.4 Testing Protocol

The discrimination tests were conducted in English and performed in the children's classroom settings at the schools. Before each experimental session, the children were given a presentation in order to ensure that they understood the basic logistics of the testing session. During the presentation a review of the terms *same*, *different*, *sweet*, *not sweet*, *sweeter*, *less sweet*, *and less sugar* was given to the children. Also, the testing procedure was explained along with an overview of the questionnaire (see Appendix D) and its proper fill-out.

The tetrad tests were given under two possible sets of instructions. For the unspecified method, sweetness was not specified as the difference and the instructions were: "here are four juice samples, two belong to one group and the other two belong to another group: separate them into the group they belong to according to how they taste". For the specified method, sweetness was specified as the difference and the instructions were: "here are four samples; two are sweeter than the other two samples: identify the two juices that taste sweeter". For the triangle test, the nature of the difference was not specified and the instructions were: "here are three juice samples; two are the same and one is different: **circle the juice that tastes different**". For the 3-AFC test, the nature of the difference was specified and the instructions were: "here are three juice samples; one is sweeter than the other two: **circle the juice that tastes sweeter**". For the 2-AFC test, the instructions were: "here are two juices, one is sweeter that the other, identify the one that is sweeter".

Subjects began each session by filling out the demographic information section of the questionnaire and cleansing their palate with water (Ozarka ® natural Spring Water, Ozarka Spring Water Company, Division of Nestle Waters North America Inc., Greenwch, CT 06830, USA). Next, the samples for a given test were presented simultaneously to the subjects and they were instructed to taste them from left to right. The subjects were allowed to retaste the sample with the condition that they always tasted all three samples in the order presented to them. Session lengths ranged from 30 to 60 minutes.

Research showed that sweet taste receptors, in addition to being present in the mouth cavity and esophagus, are also present in the gastrointestinal tract of humans (Jang and others, 2007; Margolskee and others, 2007). Our preliminary study showed that there was no significant difference in discrimination between swallowing and expectorating a sample (see Appendix E). In this study, the children were therefore instructed to swallow the samples as it was a more natural behavior that caused less distraction.

7.2.4.1 Experimental Session Overview

The following is a detailed protocol of the testing session:

- Day 1: Unspecified Tetrads and Triangle Test
    - Children filled out the child assent form
    - Children were presented with the questionnaire and water
    - Children filled out top portion of questionnaire (demographic information: name, age, gender, grade)

    - Overview presentation for the unspecified tetrads and triangle tests
    - Depending on the first question for the individual subject the four samples for the unspecified tetrads test or the three samples for the triangle test were simultaneously presented
    - Children cleansed their palates
    - For the unspecified tetrads test
        - Children were instructed to taste the samples
            - Taste the first sample
            - Taste the second sample
            - Taste the third sample
            - Taste the fourth sample
        - Children were instructed to separate the sample into two equal groups
        - 10 minute mandatory break
    - For the triangle test
        - Children were instructed to taste the samples
            - Taste the first sample
            - Taste the second sample
            - Taste the third sample
        - Children were instructed to identify the juice that tasted different
    - End of session
- Day 2: Specified tetrads, 3-AFC and 2-AFC tests
    - Children filled out the CHILD ASSENT FORM
    - Children were presented with the questionnaire and water

- Children filled out top portion of questionnaire (Demographic information: name, age, gender, grade)
- Overview presentation for the specified tetrads, 3-AFC and 2-AFC tests
- Depending on the first question for the individual subjects the four samples for the unspecified tetrads test, the three samples for the triangle test or the two samples for the 2-AFC test were simultaneously presented
- Children cleansed their palates
- For the specified tetrads test
  - Children were instructed to taste the samples
    - Taste the first sample
    - Taste the second sample
    - Taste the third sample
    - Taste the fourth sample
  - Children were instructed to select the two sweeter samples
  - 10 minute mandatory break
- For the triangle test
  - Children were instructed to taste the samples
    - Taste the first sample
    - Taste the second sample
    - Taste the third sample
  - Children were instructed to select the juice that tasted sweeter
  - 10 minute mandatory break
- For the 2-AFC test
  - Children were instructed to taste the samples

- Taste the first sample
- Taste the second sample

▪ Children were instructed to select the juice that tasted sweeter

o End of session

7.5.5 Data Analysis

For each testing method, the number of correct responses provided by the subjects was counted and recorded. The P_cs for both tests were used to determine the corresponding d' values (Ennis 1993). The experimental estimate d', the variance of d', and the test of significance were obtained using the IFPrograms Version 8.5.0320 (The Institute for Perception, Richmond, VA, USA). The d' values were calculated with the assumptions that the two intensity distributions for the stimuli are one-dimensional normal distributions that have equal variance (Bi and others 1997). Alternatively, d' values for the triangle and 3-AFC tests could be obtained from published tables (e.g., Ennis, 1993; Ennis and others, 1998; see Appendix F). For the 2-AFC, 3-AFC and triangle methods the variance of d' can be obtained by the approach described by Bi and others (1997) (see Appendix G). For the methods of tetrads the variance of d' is obtained from IFPrograms which inverts the second derivative of the likelihood function at the estimate value. For all methods the test statistic can be obtained by the approach described by Bi and others (1997) (see Appendix H). All analyses were evaluated at $\alpha = 0.05$.

7.3 Results and Discussion

7.3.1 Comparison of Unspecified Methods: Unspecified Method of Tetrads vs. Triangle

Table 7.2 presents the pooled results for the unspecified method of tetrads and the triangle test. The P_c for the unspecified method of tetrads (0.54) was higher than that of the triangle test (0.49). The d' values for the unspecified method of tetrads and the triangle test were not significantly different from each other. The psychometric functions show that for a given d' value the unspecified method of tetrads will result in a higher P_c than the triangle test (see Figure 7.1). This was confirmed in our data.

Table 7.2: Pooled proportion of correct responses, pooled d' values, test statistic and p-value for the unspecified method of tetrads and the triangle method for all six grades

Unspecified Tetrads				**Triangle**				**d' Sig Test** [b]	
N	**P_c** [a]	**d'**	**variance**	**N**	**P_c** [a]	**d'**	**variance**	**χ^2**	**p-value** [c]
404	0.54	1.18	0.007	808	0.49	1.41	0.009	3.31	0.07

[a] P_c = proportion of correct responses
[b] Significance test for d' values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[c] All p-values > 0.05 are not significant

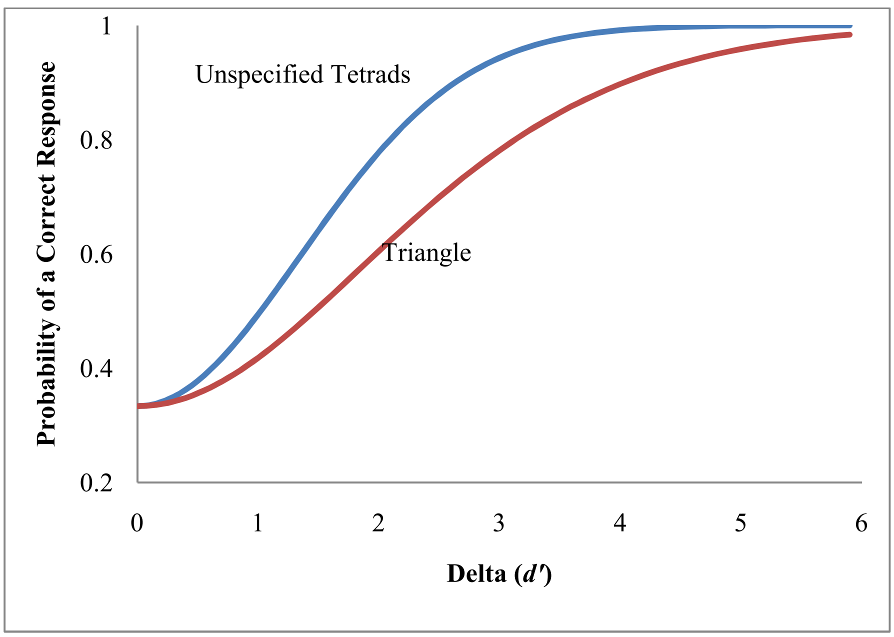


Figure 7.1: Psychometric functions for the unspecified method of tetrads and the triangle method (data from: Ennis 1993, Ennis and others 1998)

Overdispersion (Cox, 1983; Anderson, 1988) was assessed for the triangle data due to two replications of each test per subject. This possible added variance was accounted for by using the Thurstonian variant of the Beta Binomial model for replicated difference tests (Bi and

Ennis, 1998). The variance of d' prime is then measured in terms of a gamma (γ) value (see Appendix I). Gamma ranges from zero to one; no overdispersion to full overdispersion. When comparing these two models, the p-value indicates the statistical probability that the Beta-Binomial model fits the data significantly better than the Binomial model. After assessing the presence of overdispersion in triangle data, the Beta-Binomial model was found not significantly better ($p = 0.07$, $\gamma = 0.07$) than the Binomial model for this data; therefore, overdispersion was not significant.

Delwiche and O'Mahony (1996) conducted a study with these two methods (N =12, 20-45 years) using chocolate pudding as the stimulus. Even though the investigators did not conduct a direct comparison among the unspecified method of tetrads and the triangle tests, the reported data were used to make such comparisons. Here, the reported mean number of correct responses for the unspecified method of tetrads (9.8/12 = 0.68) was lower than that for the triangle test (8.1/12 = 0.82). Since we do not have access to the raw data of this study, the exact d' values corresponding to the both methods cannot be calculated. However, utilizing an approximate number of correct responses of 10/12 for the tetrads method and 8/12 for the triangle test the approximate d' values are 2.25 and 2.32, respectively (the authors reported that the predicted d' values for these tests were around 2). These d' values are not significantly different from each other ($p = 0.94$); therefore supporting our findings.

Our findings are also supported by those of Masouka and others (1995) for beer bitterness detection (N = 9, 21 – 37 years). The reported mean number of triangle tests performed correctly (5.6/12) was lower than that for the unspecified method of tetrads (7.4/12). The reported d' value for the triangle test (1.27, approximated by us to 0.99 using 5/12 for the P_c) is not significantly different (p = 0.76) from the approximated d' value for the unspecified method of tetrads (1.31, using 7/12 as the approximate P_c).

Table 7.3: Proportion of correct responses, d' values and significance tests for the unspecified method of tetrads and the triangle method for the individual grades

	Unspecified Tetrads				Triangle				d' Sig Test [b]	
Grade	**N**	P_c [a]	d'	**variance**	**N**	P_c [a]	d'	**variance**	χ^2	**p-value** [c]
1	96	0.42	0.70	0.051	192	0.41	0.92	0.058	0.45	0.51
2	55	0.47	0.92	0.063	110	0.44	1.11	0.080	0.25	0.62
3	48	0.56	1.24	0.058	96	0.50	1.47	0.072	0.41	0.52
4	80	0.61	1.40	0.033	160	0.56	1.79	0.039	2.11	0.15
5	51	0.73	1.80	0.053	102	0.48	1.36	0.071	1.56	0.21
6	74	0.55	1.21	0.038	148	0.55	1.74	0.043	3.38	0.07

[a] P_c = proportion of correct responses
[b] Significance test for d' values, Critical χ^2 value = 3.84, α = 0.05
[c] All p-values > 0.05 are not significant

Performance in the individual grades was further explored (Table 7.3). Overdispersion was assessed for the triangle data of all the age groups and was found not to be present (all $p > 0.05$). For all age groups the P_cs and d' values for the unspecified tetrads and triangle tests were not significantly different from each other (all $p > 0.05$). This is the same result observed for the overall population (Table 7.2) and the studies abovementioned. The theoretical predictions that the number of correct responses for the unspecified method of tetrads will be should be higher than those for the triangle method are observed in all grades.

7.3.2 Comparison of Specified Methods

7.3.2.1 Specified Method of Tetrads vs. 3-AFC

Pooled results for all six grades for the specified method of tetrads and the 3-AFC test are presented in Table 7.4. The P_c for the 3-AFC method (0.62) was higher than that for the unspecified method of tetrads (0.56). The d' values were significantly different ($p = 0.003$). According to the psychometric functions (Figure 7.2), a higher P_c is expected for the 3-AFC method than the specified method of tetrads. This theoretical prediction is observed in our data. However, these findings are only partially supported by those of Delwiche and O'Mahony (1996) and Masouka and others (1995). Again, these investigators did not conduct a direct comparison between the specified method of tetrads and the 3-AFC test, but their reported data was used to make such comparisons. For the data reported by Delwiche and O'Mahony (1996), the specified tetrads P_c (9.6/12) was lower than the 3-AFC P_c (11.1/12). The corresponding d' values are 2.25 (using an approximate $P_c = 10/12$) for the specified tetrads and 2.36 (using an approximate $P_c = 11/12$) for the 3-AFC. These d' values are not significantly different from each other ($p = 0.90$).

Masouka and others reported a specified tetrads P_c of 8.2/12 and a 3-AFC P_c of 8.3/12; P_cs which are not significantly different from each other ($p = 0.97$, Z-test). The reported d' value of the 3-AFC test (1.22, approximated us to 1.12 using a $P_c = 8/12$) is not significantly different ($p = 0.46$) from the approximated d' value for the specified method of tetrads (1.61, using 8/12 as the approximate P_c).

Table 7.4: Pooled proportion of correct responses, pooled d' values, test statistic and p-value for the specified method of tetrads and the 3-AFC method for all six grades (N=404)

Specified Tetrads			**3-AFC**			**d' Sig Test** [b]	
P_c [a]	**d'**	**variance**	**P_c** [a]	**d'**	**variance**	**χ^2**	**p-value** [c]
0.56	1.29	0.005	0.62	0.97	0.007	8.53	0.003

[a] P_c = proportion of correct responses
[b] Significance test for d' values, Critical χ^2 value = 3.84, $\alpha = 0.05$
[c] All p-values > 0.05 are not significant

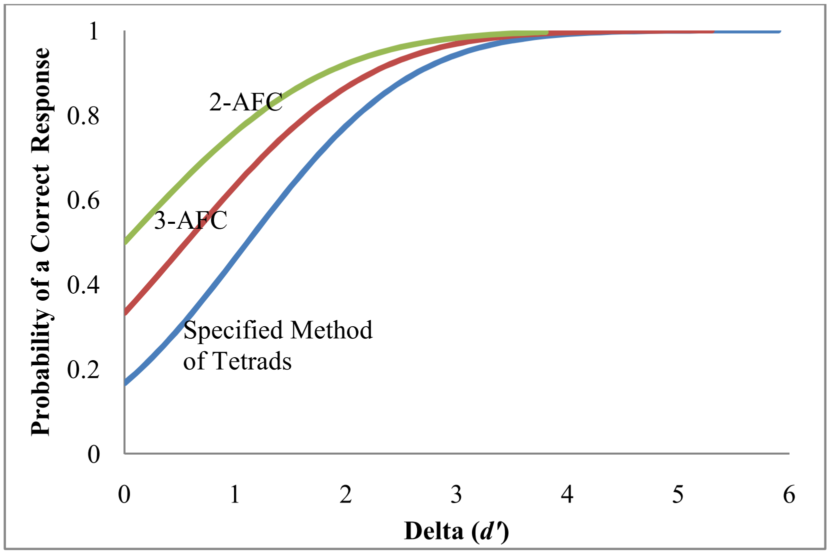


Figure 7.2: Psychometric functions for the 2-AFC method, 3-AFC method, and specified method of tetrads (data from: Ennis 1993, Ennis and others 1998)

Table 7.5 presents the results for the individual age groups. All P_cs for 3-AFC method were higher than the Pcs for the specified tetrads, except for the sixth grade. The d' values were also not significantly different for all grades, except sixth grade (p = 0.003). These individual results (except for the sixth grade) are supported by the overall findings of Delwiche and O'Mahony (1996) and Masouka and others (1995) that the P_cs and d' values of these two methods were not significantly different for adult populations.

Table 7.5: Proportion of correct responses, d' values and significance tests for the specified method of tetrads and the 3-AFC method for the individual grades

		Specified Tetrads			**3-AFC**			**d' Sig Test** [b]	
Grade	**N**	**P_c** [a]	**d'**	**variance**	**P_c** [a]	**d'**	**variance**	**χ^2**	**p-value** [c]
1	96	0.33	0.61	0.023	0.45	0.39	0.028	0.95	0.33
2	55	0.45	0.98	0.039	0.56	0.76	0.049	0.55	0.46
3	48	0.67	1.61	0.049	0.63	0.97	0.057	3.86	0.05
4	80	0.68	1.64	0.030	0.80	1.65	0.044	0.001	0.97
5	51	0.53	1.20	0.042	0.67	1.12	0.056	0.07	0.80
6	74	0.77	1.98	0.037	0.68	1.15	0.039	9.06	0.003

[a] P_c = proportion of correct responses
[b] Significance test for d' values, Critical χ^2 value = 3.84, α = 0.05
[c] All p-values > 0.05 are not significant

7.3.2.2 Specified Method of Tetrads vs. 2-AFC

Pooled results for all six grades for the specified method of tetrads and the 2-AFC test are presented in Table 7.6. The P_c for the 2-AFC method (0.77) was higher than that of the method of specified tetrads (0.56). The *d′* values were significantly different ($p = 0.04$). According to the psychometric functions (Figure 7.2), for a given *d′* value, a higher P_c is expected for the 2-AFC method than the specified tetrads. This theoretical prediction is observed in our data.

Table 7.6: Pooled proportion of correct responses, pooled *d′* values, test statistic and p-value for the specified method of tetrads and the 2-AFC method for all six grades (N=404)

Specified Tetrads			2-AFC			*d′* Sig Test [b]	
P_c [a]	*d′*	variance	P_c [a]	*d′*	variance	χ^2	p-value [c]
0.56	1.29	0.005	0.77	1.04	0.010	4.17	0.04

[a] P_c = proportion of correct responses
[b] Significance test for *d′* values, Critical χ^2 value = 3.84, α = 0.05
[c] All p-values > 0.05 are not significant

Results for the individual grades are presented in Table 7.7. The P_cs for the specified tetrads and 2-AFC methods are significantly different except for the fourth and sixth grades (both $p = 0.01$). The *d′* values were not significantly different between these two methods for all grades, expect sixth (p =0.01).

Table 7.7: Proportion of correct responses, *d′* values and significance tests for the specified method of tetrads and the 2-AFC method for the individual grades

		Specified Tetrads			2-AFC			*d′* Sig Test [b]	
Grade	N	P_c [a]	*d′*	variance	P_c [a]	*d′*	variance	χ^2	p-value [c]
1	96	0.33	0.61	0.023	0.68	0.65	0.035	0.03	0.87
2	55	0.45	0.98	0.039	0.73	0.86	0.065	0.14	0.71
3	48	0.67	1.61	0.049	0.90	1.78	0.119	0.17	0.68
4	80	0.68	1.64	0.030	0.80	1.19	0.051	2.50	0.11
5	51	0.53	1.20	0.042	0.78	1.11	0.077	0.07	0.79
6	74	0.77	1.98	0.037	0.80	1.18	0.055	6.96	0.01

[a] P_c = proportion of correct responses
[b] Significance test for *d’* values, Critical χ^2 value = 3.84, α = 0.05
[c] All p-values > 0.05 are not significant

7.3.2.3 3-AFC vs. 2-AFC

Table 7.8 presents the results for the 3-AFC and 2-AFC methods for the entire population (N= 404). The P_c for the 2-AFC method (0.77) was higher than that for the 3-AFC method (0.62), as expected from theory (see Figure 7.2). However, there was not a significant difference between the d' values. Dessirier and O'Mahony (1999) compared the d' values of the 2-AFC and 3-AFC methods reporting significant differences among them (2-AFC d' > 3-AFC d'). It was demonstrated that the higher d' values observed for 2-AFC test can be attributed to its advantageous sequence effects (Dessirier and O'Mahony, 1999; Rousseau and O'Mahony, 1997). For the individual grades, all 2-AFC P_cs were higher than 3-AFC P_cs (Table 7.9). However, d' values were only significantly higher for the first to third grade and sixth grade.

Table 7.8: Pooled proportion of correct responses, pooled d' values, test statistic and p-value for the specified method of tetrads and the 3-AFC method for all six grades (N=404)

3-AFC			2-AFC			d' Sig Test [b]	
P_c [a]	d'	variance	P_c [a]	d'	variance	χ^2	p-value [c]
0.62	0.97	0.007	0.77	1.04	0.010	0.29	0.59

[a] P_c = proportion of correct responses
[b] Significance test for d' values, Critical χ^2 value = 3.84, α = 0.05
[c] All p-values > 0.05 are not significant

Table 7.9: Proportion of correct responses, d' values and significance tests for the specified method of tetrads and the 3-AFC method for the individual grades

		3-AFC			2-AFC			d' Sig Test [b]	
Grade	N	P_c[a]	d'	variance	P_c[a]	d'	variance	χ^2	p-value [c]
1	96	0.45	0.39	0.028	0.68	0.65	0.035	1.07	0.30
2	55	0.56	0.76	0.049	0.73	0.86	0.065	0.09	0.77
3	48	0.63	0.97	0.057	0.90	1.78	0.119	3.73	0.05
4	80	0.80	1.65	0.044	0.80	1.19	0.051	2.23	0.13
5	51	0.67	1.12	0.056	0.78	1.11	0.077	0.001	0.98
6	74	0.68	1.15	0.039	0.80	1.18	0.055	0.01	0.92

[a] P_c = proportion of correct responses
[b] Significance test for d' values, Critical χ^2 value = 3.84, α = 0.05
[c] All p-values > 0.05 are not significant

7.4 Conclusion

The findings of this study are important when deciding on discrimination protocols to be used with children, whether the difference among the products is known or not. Significant differences in performance were not observed between unspecified methods. Among specified methods, as the number of samples to be compared decreased performance increased, however, the specified method of tetrads resulted in a greater degree of difference among the stimuli. It is important to note that these conclusions are only valid under the circumstances of this experiment.

7.5 References

Anderson DA. 1988. Some models for overdispersed binomial data. Australian Journal of Statistics. 30: 125-148.

Bi J, Ennis DM, O'Mahony M. 1997. How to estimate the variance of d' from difference tests. Journal of Sensory Studies. 12:87-104.

Bi J, Ennis DM. 1998. A Thurstonian variant of the beta-binomial model for replicated difference tests. Journal of Sensory Studies. 13: 461-466.

Braun V, Rogeaux M, Schneid N, O'Mahony M, Rousseau, B. (2004). Corroborating the 2-AFC and 2-AC Thurstonian models using both a model system and sparkling water. *Food Quality and Preference, 15*, 501-507.

Buchanan B, Givon M, Goldman A. (1987). Measurement of discrimination ability in taste tests: An empirical investigation. *Journal of Marketing Research 24*, 154-163.

Byer AJ, Abrams D. 1953. A comparison of the triangular and two-sample taste-test methods. Food Technology. 7:185-187.

Cox DR.1983. Some remarks on overdispersion. Biometrika. 70: 269-274.

Delwiche J, O'Mahony M. 1996. Flavour discrimination – an extension of Thurstonian paradoxes to the tetrad method. Food Quality and Preference. 7:1-5.

Dessirier JM, O'Mahony M. (1999). Comparison of d' values for the 2-AFC (paired comparison) and 3-AFC discrimination methods: Thurstonian models, sequential sensitivity analysis and power. *Food Quality and Preference, 10*, 51-58.

Ennis DM. 1993. The power of sensory discrimination methods. Journal of Sensory Studies. 8:353-370.

Ennis DM. 1990. Relative power of difference testing methods in sensory evaluation. Food Technology. 44(4):114-118.

Ennis J, Ennis D M, Yip D, O'Mahony M. (1998). Thurstonian models for variants of the method of tetrads. *British Journal of Mathematical and Statistical Psychology, 51,* 205-215.

Francois P, Sauvageot F. (1988). Comparison of the efficiency of pair, duo-trio, and triangle tests. *Journal of Sensory Studies*, *3*, 81-94.

Geelhoed EN, MacRea AW, Ennis DM. (1994). Preference gives more consistent judgments than oddity only if the task can be modeled as forced choice. *Perception and Psychophysics, 55*, 473-477.

Green DM, Swets JA. 1966. Signal Detection Theory and Psychophysics. Wiley, New York.

Hautus MJ, Irwin RJ. (1995). Two models for estimating the discriminability of foods and beverages. *Journal of Sensory Studies, 26*, 168-170.

Huang YT, Lawless HT. (1998). Sensitivity of the ABX discrimination test. *Journal of Sensory Studies, 13,* 229-239.

Jang HJ, Kokrashvili Z, Theodorakis MJ, Carlson O, Kim BJ, Zhou J, Kim HH, Xu X, Chan SL, Juhaszova M, Bernier M, Mosinger B, Robert F. Margolskee RF, Egan JM. 2007. Gut-expressed gustducin and taste receptors regulate secretion of glucagon-like peptide-1. Proceedings National Academy of Science U.S.A. 104 (38): 15069-15074.

Kim H, Jeon SY, Kim K, O'Mahony M. (2006). Thurstonian models and variance I: Experimental confirmation of cognitive strategies for difference tests and effects of perceptual variance. *Journal of Sensory Studies, 21,* 465-484.

Kuesten, C. L. (2001). Sequential use of the triangle, 2-AC, 2-AFC, and same-different methods applied to a cost-reduction effort: Consumer learning acquired throughout testing and influence on preference judgments. *Food Quality and Preference, 12*, 447-455.

Lavin J G, Lawless HT. (1998). Effects of color and odor on judgments of sweetness among children and adults. *Food Quality and Preference*, *9*, 283-289.

MacRea AW, Geelhoed EN. (1992). Preference can be more powerful than detection of oddity as a test of discriminability. *Perception and Psychophysiology, 51,* 79-81.

Margolskee RF, Dyer J, Kokrashvili Z, Salmon KSH, Ilegems E, Daly K, Maillet EL, Ninomiya Y, Mosinger B, Shirazi-Beechey SP. 2007. T1R3 and gustducin in gut sense sugars to regulate expression of Na+-glucose cotransporter 1. Proceedings of the National Academy of Science U.S.A. 104 (38): 15075-15080.

Masuoka S, Hatjopoulos D, O'Mahony M. (1995). Beer bitterness detection: testing Thurstonian and sequential sensitivity analysis models for triad and tetrad methods. *Journal of Sensory Studies, 10*, 295-306.

O'Mahony M. 1995. Who told you the triangle test was simple? Food Quality and Preference. 6:227-238.

O'Mahony M, Masuoka S, Ishii R. 1994. A theoretical note on difference tests: models, paradoxes and cognitive strategies. Journal of Sensory Studies. 9: 247-272.

Peryam DR, Swartz VW. 1950. Measurement of sensory differences. Food Technology. 4: 390-395.

Rousseau B, O'Mahony M. 1997. Sensory difference tests: Thurstonian and SSA predictions for vanilla flavored yogurts. Journal of Sensory Studies. 12:127-146.

Rousseau B, O'Mahony M. (2000). Investigation of the effect of within-trial retasting and comparison of the dual-pair, same-different and triangle paradigms. *Food Quality and Preference, 11*, 457-464.

Rousseau B, O'Mahony M. (2001). Investigation of the dual-pair method as a possible alternative to the triangle and same-different tests. *Journal of Sensory Studies, 16*, 161-178.

Rousseau B, Stroh S, O'Mahony M. (2002). Investigating more powerful discrimination tests with consumers: effects of memory and response bias. *Food Quality and Preference, 13*, 39-45.

Stillman JA. 1993. Response selection, sensitivity, and taste test performance. Perception and Psychophysics. 54:190-194.

Stillman JA, Irwin RJ. (1995). Advantages of the same-different method over the triangle method for the measurement of taste discrimination. *Journal of Sensory Studies, 10*, 261-272.

Tedja S, Ninaka R. Ennis DM, O'Mahony M. 1994. Triadic discrimination testing – refinement of Thurstonian and SSA approaches. Chemical Senses. 19:279-301.

Thurstone LL. 1927a. Psychophysical analysis. American Journal of Psychology. 38: 368-389.

Thurstone LL. 1927b. A law of comparative judgment. Psychology Review. 34:273-286.

CHAPTER 8: SUMMARY AND CONCLUSIONS

Sensory discrimination methods are used to determine if a subtle but perceptible difference exists among confusable stimuli. There are numerous published studies investigating discrimination methods using adults as subjects. However, children's performance in such methods has not been fully explored. Therefore, in this dissertation several sensory evaluation studies were conducted to investigate children's performance in different discrimination methods.

First, the paradox of discriminatory non-discriminators was challenged. This paradox theoretically states that the change in instructions from the triangle method (not specifying the nature of the difference among the stimuli) to the 3-AFC method (specifying the nature of the difference) will results in a higher proportion of correct responses (P_c) regardless of the same guessing probability (1/3). This paradox has been resolved using Thurstonian theory, which in the case of these triadic methods predicts that despite the differences in P_c the degree of difference between the two stimuli (δ) should not be significantly different between the methods. This paradox and the Thurstonian predictions were investigated for sweetness discrimination with three different sets of stimuli: easily discriminable, confusable, and hardly discriminable. Age (first to sixth grade children) and cultural (Honduras, Mexico, Thailand, USA) effects were also explored. The paradox was first tested (N = 1914) using easily discriminable stimuli (100% vs. 60% apple juice). The paradox was confirmed for all four populations together and individually, but Thurstonian predictions were only confirmed for the Honduran population. Results indicated a high degree of discrimination by the subjects among the stimuli and it was concluded that the stimuli were too easy to discriminate, and, thereby, not appropriate for discrimination testing. The degree of discriminability was then increased. Using carbonated beverages (regular cokes vs. diet cokes and diet coke vs. coke zero) with different sweeteners (N

= 1870) as the stimuli (hardly discriminable) the paradox was only confirmed for the Honduran and Mexican populations and Thurstonian predictions were not confirmed for any of the populations. It was concluded that it was too difficult for the children to discriminate among the carbonated beverages and a set of stimuli that was still confusable yet not as multidimensional in nature was chosen and tested with a USA children population.

When confusable stimuli (100% and 75% apple juice) were used the paradox (3-AFC P_c = 0.62; triangle P_c = 0.43, $p < 0.001$) and Thurstonian predictions (3-AFC d' = 0.97; triangle d' = 1.09, p = 0.48) were confirmed for the USA population (N = 404). These studies demonstrated the Thurstonian prediction that judges will get a higher proportion of correct responses in a 3-AFC than in a triangle method, yet with the statistically same corresponding d' values was dependent on the nature of the stimuli and the degree of discriminability among them. For the studies involving the easily and hardly discriminable stimuli a consistent cultural trend was not observed; however, for all three studies performance generally improved with age. It was then questioned if children's performance would be the same if the number of samples to be evaluated increased.

Secondly, Thurstonian predictions regarding the variants of the method of tetrads were challenged. Unidimenasional Thurstonian theory predicts similar choice probabilities for the unspecified and specified method of tetrads provided that $\delta > 1$. Using the confusable stimuli (100% vs. 75% apple juice) these predictions were tested with a USA population (N = 404). These results confirmed the predictions that specifying the difference among the stimuli will not result in superior performance (unspecified P_c = 0.54; specified P_c = 0.56, p = 0.62) in the case of the methods of tetrads. The degree of difference among the stimuli was also not significantly different among the stimuli (unspecified d' = 1.18; specified d' = 1.29, p = 0.32).

Finally, children's performance within unspecified methods or specified methods was compared. For the unspecified methods, performance in the unspecified method of tetrads and the triangle method was compared; results were as predicted by Thurstonian theory (tetrads P_c = 0.54; triangle P_c = 0.49; p = 0.10) (tetrads d' = 1.18; triangle d' = 1.41; p = 0.07). For the specified methods the 2-AFC, 3-AFC and specified methods of tetrads were compared. The 2-AFC resulted in the highest proportion of correct responses (0.77) followed by the 3-AFC (0.62) and the tetrads (0.56), respectively. Performance was not significantly different between the specified tetrads and the 3-AFC but a difference was present between the specified tetrads and the 2-AFC method. The degree of difference (d') among the stimuli was significantly different between the specified tetrads and the 3-AFC and between the specified tetrads and the 2-AFC. Results for the comparison between the 3-AFC and 2-AFC are in line with what has been observed with adult subjects.

Results from this investigation give further support to the Thurstonian predictions discussed above, which had been tested only with adult subjects. In conclusion, under the circumstances of this study, for sweetness discrimination given that the appropriate stimuli were used children between 6 to 12 years of age were capable of performing sensory discrimination methods, presented the same decision rules as adults and their performance increased with age. Overall, the findings from this study would provide insights to the food industries and sensory scientists when performing discrimination testing with children (first to sixth grade).

APPENDIX A: PARENTAL CONSENT FORM

Dear Parent or Legal Guardian:

We are asking your consent for your child to participate in a taste test of food products. With this study we hope to expand our understanding of appropriate sensory methodologies applicable to food testing with children. This research will be carried on a time frame of approximately two weeks. During this period the children may evaluate different samples during different sessions. However, a consent form will be sent to you for each different session specifying the protocol and the food products to be evaluated. Please read the description of the study below and then indicate if you would allow your child to participate.

Project Title: Sensory Evaluation with Children

Performance Site:

Investigators: This research will be conducted by Ms. Karen Garcia, a Ph.D. student in the Food Science Department at the Louisiana State University, under the guidance of Dr. Witoon Prinyawiwatkul.
The investigators are available to answer any questions anytime through e-mail or via a telephone call Monday - Friday 8:30AM – 4:30PM to 225-578-5188.
Karen Garcia (kgarci2@tigers.lsu.edu)
Dr. Witoon Prinyawiwatkul (wprinya@lsu.edu)

Purpose of Study: The purpose of the study is to validate the use of existing sensory evaluation methodologies with children. Results from this investigation will provide sensory scientists with a better understanding of the techniques appropriate for usage with children. Results from usage of the appropriate techniques will in turn provide the product developer with a more accurate direction for increasing consumer appeal specifically for children's market. In addition, knowledge about children's food sensory perceptions and food preferences provide important cues for the design and implementation of interventions aimed at promoting healthy eating.

Test Samples: Only commercially available food products will be used for this research. Please see attachment for complete ingredient and nutritional information.
Mott's ® 100% Apple Juice
Mott's® for Tots Apple Juice – 40% Less Sugar
Dasani ® Water

Description of Study: All procedures are the standard methods as published by the American Society for Testing and Materials and the Sensory Evaluation

Division of the Institute of Food Technologists.

The following is a detailed description of the process specific to this session:
-Orientation Session: Explain purpose of study, demonstrate the testing process and explain rating card fill-out procedure.

-Warm-up Session: Make children comfortable with evaluation procedure. A visual session will be conducted, where children are shown how pictures and or colored blocks differ from each other.

-Sensory Evaluation Session: This is the actual evaluation session and will have an approximate duration of 45 minutes.
Child rinses mouth with water
Three samples are simultaneously presented (randomly coded)
Child is instructed to taste the samples
Child is to record on the rating card which is the "odd" sample
The process is repeated three (3) times
Child takes a 10 minute mandatory break
Three more samples are simultaneously presented (randomly coded)
Child is instructed to taste the samples
Child is to record on the survey the "sweeter" or "least sweet" sample
This process is repeated twice (2)
The child tastes five (5) sets of triads (3) for a total of 15 samples

Risks: The only risk that can be foreseen is an allergic reaction to apple juice, coca cola® classic, coca cola diet ®, and/or coca cola zero®. However, because you know beforehand what your child will be testing this risk can be eliminated.

PLEASE MAKE SURE YOUR CHILD IS NOT ALLERGIC TO APPLE JUICE. IF YOU ARE NOT SURE PLEASE DO NOT SIGN THE CONSENT FORM.

Privacy: Pictures may be taken during the sessions and will be used for illustration purposes only; e.g., as part of a power point presentation during Ms.

Garcia's dissertation defense. Results of the study will be published but no names or identifying information will be included in the publication. Subject identity will remain confidential unless disclosure is required by law.

Financial Costs: There is no cost for participation in this study, nor will there be any compensation for participation.

Right to Refuse: Participation in this study is voluntary, and a child will become part of this study only if both child and parent agree to the child's participation. At any time, either the subject may withdraw from the study or the subject's parent may withdraw the subject from the study without penalty or loss of any benefit to which they might otherwise be entitled. Whether or not your child participates in this study will have no bearing on you or your child's status with the school or Louisiana State University.

We appreciate your cooperation and we look forward to having your child participate.
If you allow their participation, please complete the attached form and return it to your child's teacher.

Thank you.
Sincerely,

____

Dr. Witoon Prinyawiwatkul
Professor
Department of Food Science
Louisiana State University
Agricultural Center

____

Karen Garcia
Ph.D. Student
Department of Food Science
Louisiana State University
Agricultural Center

Mott's® Original 100% apple juice Nutrition Facts

100% apple juice

Serving Size: 8fl oz (240mL)
Servings per container: 8

Amount per serving: 120 Calories

Contents	Amount	% Daily Value
Total fat	0 mg	0%
Sodium	10 mg	0%
Potassium	240 mg	6%
Total Carb	29 g	10%
Sugars	0 g	
Protein	0 g	

- Vitamin A 0%
- Vitamin C 20%
- Calcium 2%
- Iron 6%

Ingredients:
Water
Apple Juice Concentrate
Ascorbic Acid (Vitamin C)

Mott's For Tots® Juice Nutrition Facts

Apple - Contains 54% Juice

Serving Size: 8 fl oz (240mL)
Servings per container: 8

Amount per serving: 60 Calories

Contents	Amount	% Daily Value
Total fat	0 mg	0%
Sodium	10 mg	0%
Potassium	160 mg	5%
Total Carb	15 g	5%
Sugars	15 g	
Protein	0 g	

- Vitamin A 0%
- Vitamin C 100%
- Calcium 2%
- Iron 2%

Ingredients:
Purified Water
Apple Juice Concentrate
Vitamin C
Natural Flavors

Dasani ® Water Nutrition Facts

Serving Size: 8 fl oz (240mL)
Servings per container: 2.5

Amount per serving: 0 Calories

Contents	Amount	% Daily Value
Total fat	0 mg	0%
Sodium	0 mg	0%
Total Carb	0 g	0%
Protein	0 g	0%

Ingredients:
Purified Water
Magnesium Sulfate
Potassium Chloride
Salt*†

*Adds negligible amount of sodium
†Minerals added for taste

<u>**PARENTAL** CONSENT FORM – SENSORY EVALUATION WITH CHILDREN</u>

I have read the study discussed above and all of my questions have been answered. I may direct any additional questions regarding study specifics to the investigators. I will allow my child to participate in this taste test.

The products that my child will be tasting are Mott's ® 100% apple juice, Mott's ® for Tots Apple Juice-40% less sugar and Dasani ® water. I verify that my child is not allergic to any of these products.

Child's name: ____________________________ Age:____________ Gender:___________

Parent's/Legal Guardian's Name: ____

***Parent's/Legal Guardian's Signature*:** _____________________ Date:______________

I verify that my child is not allergic to apple juice.

PARENT SIGNS THIS FORM

The parent/guardian has indicated to me that he/she is unable to read. I certify that I have read this consent from to the parent/guardian and explained that by completing the signature line above he/she has given permission for the child to participate in the study.

Signature of Reader:________________________________ Date:____________________

If you would like a copy of your signed consent form please indicate so by providing us with your mailing address or e-mail address:

Mailing Address: ___________________________________

____

____

E-Mail Address: ___________________________________

APPENDIX B: CHILD ASSENT FORM

I, _________________________, agree to participate in a study in which I have to taste apple juice and drink water. I will have to answer a question after I taste the samples and record it on paper. I can decide to stop at any time without getting in trouble.

Child's Signature: ______________________ Age: _________ Date:_______

Witness*: _____________________________ Date: _________

*Witness must be present for the assent process, the signature of the minor is not sufficient. The witness may be a parent or guardian, school teacher or a member of the investigation group. This signed assent form is invalid without the signed parental consent form

APPENDIX C: LIST OF PRODUCT INGREDIENTS

Apple Juice - Honduras, Mexico, USA

Mott's® Original 100% apple juice: water, apple juice concentrate

Mott's For Tots® Juice: purified water, apple juice concentrate, vitamin C, natural flavors

Apple Juice - Thailand

Tipco Apple Juice: water, apple juice concentrate

Carbonated Beverages - Honduras

Coca Cola ®: Agua carbonatada, azucar, color de caramel, acido fosforico como acidulante y saborizantes (incluyendo cafeina)

Coca Cola Light ®: Agua carbonatada, color de caramel, acido fosforico y citric como acidulantes, aspartame y acesulfame k como educolorantes. Saborizantes (incluyendo cafeina), benzoate de sodio como preservante

Coca Cola Cero®: agua carbonatada, color caramel, acido fosforico como acidulante, aspartame y acesulfame k como educolorantes, saborizantes (incluyendo cafeina), benzoate de sodio como preservante y citrate de sodio como regulardor de acidez.

Carbonated Beverages - Mexico

Coca Cola ®: Agua carbonatada, azucares y concentrados coca cola

Coca Cola Light ®: Agua carbonatada, concentrados coca cola light, mezcla de asparartame y acesulfame k (40mg/100g)

Carbonated Beverages – Thailand

Coca Cola®: carbonated water, sugar, coca cola syrup

Coca Cola Light ®: carbonated water, coca cola light syrup, aspartame, acesulfame K, sucralose

Carbonated Beverages - USA

Coca Cola® Classic: carbonated water, high fructose corn syrup, caramel color, phosphoric acid, natural flavors, caffeine

Diet Coke ®: carbonated water, caramel color, aspartame, phosphoric acid, potassium benzoate, natural flavors, citric acid, caffeine

APPENDIX D: SURVEYS

D.1 Triangle Test

Name ____ **Age**________

 **Girl**__________  **Boy** __________ **Grade** ______

1. Circle the juice that tastes different.

A **B** **C**

2. Circle the juice that tastes different.

A **B** **C**

3. Circle the juice that tastes different.

A **B** **C**

D.2 3-AFC Test

Name ____ **Age**________

 **Girl**__________  **Boy** __________ **Grade** ______

1. Circle the juice that is LESS sweet.

A **B** **C**

2. Circle the juice that is SWEETEST.

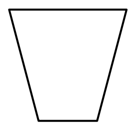
A **B** **C**

D.3 2-AFC Test

Name ____ **Age________**

 **Girl__________**  **Boy __________** **Grade ______**

Which juice is sweeter?



D.4 Method of Tetrads

Name ____ **Age________**

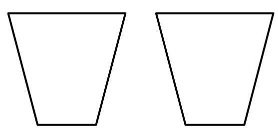
 **Girl__________**  **Boy __________** **Grade ______**

Unspecified Method of Tetrads

Separate the juices into two groups.

Specified Method of Tetrads

Which two juices taste sweeter?



APPENDIX E: CHILDREN'S PERFORMANCE WHEN SWALLOWING VS. EXPECTORATING THE STIMULI

Research showed that sweet taste receptors, in addition to being present in the mouth cavity and esophagus, are also present in the gastrointestinal tract of humans (Jang and others 2007, Margolskee and others 2007). Therefore, a preliminary study was performed in order to assess if there was a significant difference in discrimination between swallowing and expectorating the stimuli.

The subjects were third (N= 20) and fourth grade (N =22) Honduran children. Criteria for recruitment of participants were that they were not allergic to any of the ingredients present in the juice products. Participants were required to have parental consent (see Appendix A) and to sign an assent form (see Appendix B) stating their willingness to participate, both approved by Louisiana State University AgCenter Institutional Review Board prior to participating in the testing. No monetary incentive or rewards were given to subjects for participation.

Stimuli consisted of regular apple juice (Mott's® Original, "100% apple juice", Mott's LLP, Rye Brook, NY) and reduced sugar apple juice (Mott's for Tots®, "40% less sugar apple juice", Mott's LLP, Rye Brook, NY). Bottled water was used for palate cleansing (Dasani, producto Centroamericano elaborado y distribuido bajo licencia de The Coca Cola Company por Cerveceria Hondureña, S. A. de C. V., salida carretera a Puerto Cortes, San Pedro Sula, Honduras).

The discrimination tests were conducted in the children's native language; i.e., Spanish. All tests were performed in the children's classroom settings at the schools. The swallowing and expectoration sessions were performed on the same day. Before the experimental session the children were given a presentation in order to ensure that they understood the basic logistics of the testing session. During the presentation a review of the terms *same* and *different* was given to the children. Also, the testing procedure was explained along with an overview of the questionnaire (see Appendix C and D) and its proper fill-out. The children performed two triangle tests (swallowed and expectorated) using the SWS sequence. Both triangle tests were presented in succession during the same experimental session.

The hypotheses being tested for the proportions of correct responses (P_c) were:

H_o: Triangle P_c = 3-AFC P_c

H_a: Triangle $P_c \neq$ 3-AFC P_c

The hypotheses being tested for the d' values were:

H_o: Triangle d' = 3-AFC d'

H_a: Triangle $d' \neq$ 3-AFC d'

The triangle method instructions for the "swallowing" test were: "here are three juice samples, two are the same and one is different: **circle the juice that tasted different**". For the "expectoration" test the instructions were: "here are three juice samples, two are the same and one is different. Introduce the sample into your mouth but do not swallow the sample. Aided by your tongue swirl the sample in your mouth for a few seconds and expectorate it into the empty cup. Then **circle the juice that tasted different**".

Subjects began each session by cleansing their palate with water. Next, the three samples for the first test were presented simultaneously to the subjects and instructed to taste from left to right. The same protocol was repeated for the second test. The children were tested on the same day for both protocols given a 10 minute mandatory break between sessions. Session lengths lasted approximately 30 minutes.

The results are presented in Table E.1. The P_cs for the "swallowing" protocol were not significantly different than those for the "expectoration" protocol, for the individual grades and overall. Likewise, a difference was not present among the *d′* values (all $p > 0.05$). Therefore, the children did not perform significantly different whether they swallowed or expectorated the samples.

Table E.1: Proportions of correct responses, *d′* values and significance test for swallowing vs. expectoration of the samples in a triangle test

		Triangle - Swallow			**Triangle – Not Swallow**			**P_c Sig Test [b]**		**d' Sig Test [c]**	
Grade	**N**	**P_c [a]**	**d'**	**variance**	**P_c [a]**	**d'**	**variance**	**Z**	**p-value [d]**	**χ^2**	**p-value [d]**
3	**20**	0.65	2.23	0.314	0.60	1.98	0.309	0.33	0.74	0.1	0.75
4	**22**	0.77	2.94	0.344	0.55	1.70	0.289	1.54	0.12	2.43	0.12
overall	**42**	0.71	2.58	0.16	0.57	1.83	0.148	1.34	0.18	1.83	0.18

[a] P_c = proportion of correct responses

[b] Significance test for P_c, Critical Z value = 1.96, $\alpha = 0.05$

[c] Significance test for *d′* values, Critical χ^2 value = 3.84, $\alpha = 0.05$

[d] All p-values > 0.05 are not significant

APPENDIX F: DETERMINING *d′*

Utilizing the 2-AFC psychometric function as an example

$P_c = \Phi\ (\delta/\sqrt{2})$

where:

P_c = proportion of correct responses
Φ = cumulative distribution function of the standard normal distribution, area under the normal curve from $-\infty$ to some value (in this case $\delta/\sqrt{2}$)
δ = index of discrimination, the estimate is *d′*

For example, if $P_c = 0.76$

The Z value for this area under the normal curve is approximately 0.71

$\delta/\sqrt{2} = 0.71$

$\delta = 0.71(\sqrt{2})$

$\delta = 1.0$

With the P_c and then the calculated δ value one can construct tables of P_c as a function of δ. Tables F.1 and F.2 present the exact tables from Ennis (1993) for the triangle and 3-AFC methods. Tables F.3 and F.4 present the exact tables from Ennis and others (1998) for the unspecified and specified method of tetrads.

Table F.1: 2-AFC Method - Probability of a correct response (x10^4) as a function of δ (Ennis, 1993)

δ	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	5000	5028	5056	5085	5113	5141	5169	5197	5226	5254
0.1	5282	5310	5338	5366	5394	5422	5450	5478	5506	5534
0.2	5562	5590	5618	5646	5674	5702	5729	5757	5785	5812
0.3	5840	5868	5895	5923	5950	5977	6005	6032	6059	6086
0.4	6114	6141	6168	6195	6221	6248	6275	6302	6329	6355
0.5	6382	6408	6434	6461	6487	6513	6539	6565	6591	6617
0.6	6643	6669	6695	6720	6746	6771	6796	6822	6847	6872
0.7	6897	6922	6947	6971	6996	7021	7045	7069	7094	7118
0.8	7142	7166	7190	7214	7237	7261	7284	7308	7331	7354
0.9	7377	7400	7423	7446	7469	7491	7514	7536	7558	7580
1.0	7602	7624	7646	7668	7689	7711	7732	7754	7775	7796
1.1	7817	7837	7858	7879	7899	7919	7940	7960	7980	8000
1.2	8019	8039	8058	8078	8097	8116	8135	8154	8173	8192
1.3	8210	8229	8247	8265	8283	8301	8319	8337	8354	8372
1.4	8389	8406	8423	8440	8457	8474	8491	8507	8523	8540
1.5	8556	8572	8588	8603	8619	8635	8650	8665	8681	8696
1.6	8711	8725	8740	8755	8769	8783	8798	8812	8826	8840
1.7	8853	8867	8881	8894	8907	8920	8933	8946	8959	8972
1.8	8985	8997	9009	9022	9034	9046	9058	9070	9081	9093
1.9	9104	9116	9127	9138	9149	9160	9171	9182	9193	9203
2.0	9214	9224	9234	9244	9254	9264	9274	9284	9293	9303
2.1	9312	9321	9331	9340	9349	9358	9367	9375	9384	9393
2.2	9401	9409	9418	9426	9434	9442	9450	9458	9465	9473
2.3	9481	9488	9495	9503	9510	9517	9524	9531	9538	9545
2.4	9552	9558	9565	9571	9578	9584	9590	9596	9603	9609
2.5	9615	9620	9626	9632	9638	9643	9649	9654	9659	9665
2.6	9670	9675	9680	9685	9690	9695	9700	9705	9710	9714
2.7	9719	9723	9728	9732	9737	9741	9745	9749	9753	9757
2.8	9761	9765	9769	9773	9777	9781	9784	9788	9791	9795
2.9	9798	9802	9805	9809	9812	9815	9818	9821	9824	9828
3.0	9831	9833	9836	9839	9842	9845	9848	9850	9853	9856
3.1	9858	9861	9863	9866	9868	9870	9873	9875	9877	9880
3.2	9882	9884	9886	9888	9890	9892	9894	9896	9898	9900

Table F.2: Triangle Method - Probability of a correct response (x10^4) as a function of δ (Ennis, 1993)

δ	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	3333	3333	3334	3334	3335	3336	3337	3338	3339	3341
0.1	3343	3344	3347	3349	3351	3354	3357	3360	3363	3366
0.2	3370	3374	3378	3382	3386	3390	3395	3400	3405	3410
0.3	3415	3421	3427	3432	3439	3445	3451	3458	3464	3471
0.4	3478	3486	3493	3501	3508	3516	3524	3533	3541	3550
0.5	3558	3567	3576	3586	3595	3604	3614	3624	3634	3644
0.6	3654	3665	3676	3686	3697	3708	3719	3731	3742	3754
0.7	3766	3778	3790	3802	3814	3827	3839	3852	3865	3878
0.8	3891	3905	3918	3932	3945	3959	3973	3987	4001	4016
0.9	4030	4045	4059	4074	4089	4104	4119	4134	4149	4165
1.0	4180	4196	4212	4228	4244	4260	4276	4292	4309	4325
1.1	4342	4358	4375	4392	4409	4426	4443	4460	4477	4494
1.2	4512	4529	4547	4564	4582	4600	4618	4636	4654	4672
1.3	4690	4708	4726	4745	4763	4782	4800	4819	4837	4856
1.4	4875	4893	4912	4931	4950	4969	4988	5007	5026	5045
1.5	5065	5084	5103	5122	5142	5161	5180	5200	5219	5239
1.6	5258	5278	5297	5317	5337	5356	5376	5396	5415	5435
1.7	5455	5474	5494	5514	5534	5554	5573	5593	5613	5633
1.8	5653	5672	5692	5712	5732	5752	5771	5791	5811	5831
1.9	5851	5870	5890	5910	5930	5950	5969	5989	6009	6028
2.0	6048	6068	6087	6107	6127	6146	6166	6185	6205	6224
2.1	6244	6263	6283	6302	6321	6341	6360	6379	6398	6418
2.2	6437	6456	6475	6494	6513	6532	6551	6570	6589	6608
2.3	6627	6645	6664	6683	6701	6720	6739	6757	6776	6794
2.4	6812	6831	6849	6867	6885	6903	6922	6940	6958	6976
2.5	6993	7011	7029	7047	7064	7082	7100	7117	7135	7152
2.6	7169	7187	7204	7221	7238	7255	7272	7289	7306	7323
2.7	7340	7356	7373	7390	7406	7423	7439	7455	7472	7488
2.8	7504	7520	7536	7552	7568	7584	7600	7616	7631	7647
2.9	7662	7678	7693	7709	7724	7739	7754	7769	7784	7799
3.0	7814	7829	7844	7859	7873	7888	7902	7917	7931	7945
3.1	7960	7974	7988	8002	8016	8030	8044	8057	8071	8085
3.2	8098	8112	8125	8139	8152	8165	8179	8192	8205	8218
3.3	8231	8243	8256	8269	8282	8294	8307	8319	8332	8344
3.4	8356	8368	8381	8393	8405	8417	8428	8440	8452	8464
3.5	8475	8487	8498	8510	8521	8532	8544	8555	8566	8577

Table F.3: 3-AFC Method - Probability of a correct response (x10^4) as a function of δ (Ennis, 1993)

δ	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	3333	3362	3390	3418	3447	3475	3504	3533	3562	3591
0.1	3620	3649	3678	3707	3737	3766	3795	3825	3855	3884
0.2	3914	3944	3974	4003	4033	4063	4093	4124	4154	4184
0.3	4214	4244	4275	4305	4336	4366	4396	4427	4458	4488
0.4	4519	4549	4580	4611	4641	4672	4703	4734	4764	4795
0.5	4826	4857	4888	4918	4949	4980	5011	5042	5072	5103
0.6	5134	5165	5195	5226	5257	5288	5318	5349	5380	5410
0.7	5441	5471	5502	5532	5563	5593	5624	5654	5684	5714
0.8	5745	5775	5805	5835	5865	5895	5925	5955	5985	6014
0.9	6044	6074	6103	6133	6162	6191	6221	6250	6279	6308
1.0	6337	6366	6395	6423	6452	6481	6509	6538	6566	6594
1.1	6622	6650	6678	6706	6734	6761	6789	6816	6844	6871
1.2	6898	6925	6952	6979	7005	7032	7059	7085	7111	7137
1.3	7163	7189	7215	7241	7266	7292	7317	7342	7367	7392
1.4	7417	7442	7466	7491	7515	7539	7563	7587	7611	7635
1.5	7658	7682	7705	7728	7751	7774	7796	7819	7842	7864
1.6	7886	7908	7930	7952	7973	7995	8016	8037	8058	8079
1.7	8100	8121	8141	8162	8182	8202	8222	8242	8261	8281
1.8	8300	8319	8339	8357	8376	8395	8413	8432	8450	8468
1.9	8486	8504	8522	8539	8556	8574	8591	8608	8624	8641
2.0	8658	8674	8690	8706	8722	8738	8754	8769	8785	8800
2.1	8815	8830	8845	8860	8874	8889	8903	8917	8931	8945
2.2	8959	8973	8986	9000	9013	9026	9039	9052	9065	9077
2.3	9090	9102	9114	9127	9138	9150	9162	9174	9185	9197
2.4	9208	9219	9230	9241	9252	9262	9273	9283	9293	9304
2.5	9314	9324	9333	9343	9353	9362	9372	9381	9390	9399
2.6	9408	9417	9426	9434	9443	9451	9460	9468	9476	9484
2.7	9492	9500	9508	9515	9523	9530	9538	9545	9552	9559
2.8	9566	9573	9580	9587	9593	9600	9606	9613	9619	9625
2.9	9631	9637	9643	9649	9655	9661	9666	9672	9677	9683
3.0	9688	9693	9698	9703	9709	9713	9718	9723	9728	9733
3.1	9737	9742	9746	9751	9755	9759	9764	9768	9772	9776
3.2	9780	9784	9788	9791	9795	9799	9802	9806	9809	9813

Table F.4: Unspecified Method of Tetrads - Probability of a correct response (x10[4]) as a function of δ (Ennis and others, 1998)

δ	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	3333	3334	3334	3335	3336	3338	3340	3342	3345	3348
0.1	3352	3356	3360	3364	3369	3375	3380	3386	3393	3399
0.2	3406	3414	3422	3430	3438	3447	3456	3466	3476	3486
0.3	3497	3508	3519	3530	3542	3555	3567	3580	3593	3607
0.4	3621	3635	3650	3665	3680	3695	3711	3727	3743	3760
0.5	3777	3794	3812	3830	3848	3867	3885	3904	3924	3943
0.6	3963	3983	4003	4024	4045	4066	4087	4109	4131	4153
0.7	4175	4198	4221	4244	4267	4290	4314	4338	4362	4386
0.8	4411	4436	4461	4486	4511	4536	4562	4588	4614	4640
0.9	4666	4693	4720	4746	4773	4800	4828	4855	4883	4910
1.0	4938	4966	4994	5022	5050	5079	5107	5136	5165	5193
1.1	5222	5251	5280	5309	5338	5368	5397	5426	5456	5485
1.2	5515	5545	5574	5604	5634	5663	5693	5723	5753	5783
1.3	5813	5843	5872	5902	5932	5962	5992	6022	6052	6082
1.4	6112	6142	6172	6201	6231	6261	6291	6320	6350	6380
1.5	6409	6439	6468	6498	6527	6556	6586	6615	6644	6673
1.6	6702	6731	6760	6789	6817	6846	6874	6903	6931	6959
1.7	6987	7015	7043	7071	7099	7126	7154	7181	7209	7236
1.8	7263	7290	7316	7343	7370	7396	7422	7449	7475	7501
1.9	7526	7552	7578	7603	7628	7653	7678	7703	7728	7752
2.0	7777	7801	7825	7849	7873	7896	7920	7943	7966	7989
2.1	8012	8035	8058	8080	8102	8124	8146	8168	8190	8211
2.2	8233	8254	8275	8296	8316	8337	8357	8377	8397	8417
2.3	8437	8456	8476	8495	8514	8533	8552	8570	8588	8607
2.4	8625	8643	8660	8678	8695	8713	8730	8747	8764	8780
2.5	8797	8813	8829	8845	8861	8877	8892	8907	8923	8938
2.6	8953	8967	8982	8996	9011	9025	9039	9053	9066	9080
2.7	9093	9106	9119	9132	9145	9158	9170	9183	9195	9207
2.8	9219	9231	9242	9254	9265	9277	9288	9299	9310	9320
2.9	9331	9341	9352	9362	9372	9382	9392	9401	9411	9420
3.0	9430	9439	9448	9457	9466	9475	9483	9492	9500	9508
3.1	9517	9525	9533	9540	9548	9556	9563	9571	9578	9585
3.2	9592	9599	9606	9613	9620	9626	9633	9639	9646	9652

Table F.5: Specified Method of Tetrads - Probability of a correct response (x10^4) as a function of δ (Ennis and others, 1998)

δ	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	1667	1689	1711	1734	1757	1779	1803	1826	1849	1873
0.1	1897	1921	1945	1969	1994	2019	2044	2069	2094	2119
0.2	2145	2171	2197	2223	2249	2276	2302	2329	2356	2383
0.3	2410	2438	2465	2493	2521	2549	2577	2606	2634	2663
0.4	2692	2721	2750	2779	2809	2838	2868	2898	2928	2958
0.5	2988	3019	3049	3080	3110	3141	3172	3203	3235	3266
0.6	3298	3329	3361	3393	3425	3457	3489	3521	3553	3586
0.7	3618	3651	3684	3716	3749	3782	3815	3848	3881	3915
0.8	3948	3981	4015	4048	4082	4116	4149	4183	4217	4251
0.9	4285	4319	4352	4386	4421	4455	4489	4523	4557	4591
1.0	4625	4660	4694	4728	4762	4797	4831	4865	4900	4934
1.1	4968	5002	5037	5071	5105	5139	5174	5208	5242	5276
1.2	5310	5344	5378	5412	5446	5480	5514	5548	5581	5615
1.3	5649	5682	5716	5749	5783	5816	5849	5883	5916	5949
1.4	5982	6015	6048	6080	6113	6145	6178	6210	6243	6275
1.5	6307	6339	6371	6403	6434	6466	6497	6529	6560	6591
1.6	6622	6653	6684	6714	6745	6775	6805	6836	6866	6896
1.7	6925	6955	6984	7014	7043	7072	7101	7130	7158	7187
1.8	7215	7243	7271	7299	7327	7354	7382	7409	7436	7463
1.9	7490	7517	7543	7569	7596	7622	7647	7673	7699	7724
2.0	7749	7774	7799	7824	7848	7873	7897	7921	7945	7968
2.1	7992	8015	8038	8061	8084	8107	8129	8151	8174	8195
2.2	8217	8239	8260	8282	8303	8324	8344	8365	8385	8405
2.3	8425	8445	8465	8485	8504	8523	8542	8561	8580	8598
2.4	8617	8635	8653	8670	8688	8706	8723	8740	8757	8774
2.5	8791	8807	8823	8840	8856	8871	8887	8903	8918	8933
2.6	8948	8963	8978	8992	9007	9021	9035	9049	9063	9077
2.7	9090	9103	9117	9130	9142	9155	9168	9180	9193	9205
2.8	9217	9229	9240	9252	9263	9275	9286	9297	9308	9319
2.9	9329	9340	9350	9360	9371	9381	9390	9400	9410	9419
3.0	9429	9438	9447	9456	9465	9474	9482	9491	9499	9508
3.1	9516	9524	9532	9540	9547	9555	9563	9570	9577	9585

APPENDIX G: VARIANCE OF d'

Var (d') = B/N

where:

B = B-value which specific to each method. It can be obtained from tables (A.3.1, A.3.2) of B values as a function of P_c (Bi and others, 1997)
N = sample size

<u>2-AFC Method</u>

$$B = \frac{2Pc(1-Pc)}{\Phi^2(\frac{d'}{\sqrt{2}})}$$

where:
P_c = proportion of correct responses
$\Phi(d'/\sqrt{2})$ = density function of standard normal distribution evaluated at $d'/\sqrt{2}$

<u>3-AFC Method</u>

$$B = \frac{Pc(1-Pc)}{Pc'^2(d')}$$

where:
P_c = proportion of correct responses
$P_c' = \int_{-\infty}^{\infty} \Phi^2(u)\varphi(u-d'\_(u-d')du$

<u>Triangular Method</u>

$$B = \frac{Pc(1-Pc)}{P'^2(d')}$$

where:
Pc = proportion of correct responses
$P' = \sqrt{\frac{2}{3}}\Phi\left(\frac{d'}{\sqrt{6}}\right)\left[\Phi\left(\frac{d'}{\sqrt{2}}\right) - \Phi\left(\frac{-d'}{\sqrt{2}}\right)\right]$

Table G.1: Triangle Method - B value for the estimation of the variance of *d′* (Bi and others, 1997)

d′	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	*******	65800.	16452.	7314.	4115.	2635.	1831.	1346.	1031.	815.
0.1	661.08	546.88	460.03	392.44	338.81	295.54	260.13	230.78	206.19	185.38
0.2	167.61	152.31	139.06	127.50	117.35	108.40	100.45	93.38	87.05	81.36
0.3	76.236	71.597	67.387	63.554	60.054	56.850	53.910	51.205	48.711	46.406
0.4	44.273	42.294	40.455	38.742	37.146	35.655	34.261	32.954	31.729	30.578
0.5	29.496	28.477	27.517	26.611	25.755	24.945	24.179	23.452	22.764	22.110
0.6	21.489	20.898	20.336	19.801	19.291	18.805	18.341	17.897	17.474	17.069
0.7	16.681	16.310	15.954	15.614	15.287	14.973	14.672	14.383	14.106	13.839
0.8	13.582	13.335	13.097	12.868	12.647	12.435	12.230	12.032	11.841	11.657
0.9	11.479	11.308	11.142	10.982	10.827	10.677	10.532	10.392	10.256	10.125
1.0	9.998	9.875	9.756	9.641	9.529	9.421	9.316	9.214	9.115	9.019
1.1	8.926	8.836	8.749	8.664	8.582	8.502	8.424	8.349	8.276	8.205
1.2	8.136	8.069	8.004	7.941	7.880	7.820	7.762	7.706	7.651	7.598
1.3	7.547	7.497	7.448	7.401	7.355	7.310	7.267	7.225	7.184	7.144
1.4	7.106	7.068	7.032	6.997	6.962	6.929	6.897	6.866	6.835	6.806
1.5	6.778	6.750	6.723	6.697	6.672	6.648	6.624	6.601	6.579	6.558
1.6	6.538	6.518	6.499	6.480	6.462	6.445	6.429	6.413	6.398	6.383
1.7	6.369	6.355	6.342	6.330	6.318	6.307	6.296	6.286	6.276	6.267
1.8	6.258	6.250	6.242	6.235	6.228	6.222	6.216	6.211	6.206	6.201
1.9	6.197	6.193	6.190	6.187	6.185	6.183	6.181	6.180	6.179	6.179
2.0	6.178	6.179	6.179	6.180	6.182	6.184	6.186	6.188	6.191	6.194
2.1	6.198	6.201	6.206	6.210	6.215	6.220	6.226	6.231	6.238	6.244
2.2	6.251	6.258	6.265	6.273	6.281	6.289	6.298	6.307	6.316	6.326
2.3	6.336	6.346	6.356	6.367	6.378	6.389	6.401	6.413	6.425	6.438
2.4	6.450	6.463	6.477	6.490	6.504	6.518	6.533	6.548	6.563	6.578
2.5	6.594	6.609	6.626	6.642	6.659	6.676	6.693	6.710	6.728	6.746
2.6	6.765	6.783	6.802	6.821	6.841	6.861	6.881	6.901	6.921	6.942
2.7	6.963	6.985	7.007	7.029	7.051	7.073	7.096	7.119	7.143	7.166
2.8	7.190	7.214	7.239	7.264	7.289	7.314	7.340	7.366	7.392	7.418
2.9	7.445	7.472	7.500	7.527	7.555	7.584	7.612	7.641	7.670	7.700
3.0	7.729	7.760	7.790	7.821	7.852	7.883	7.914	7.946	7.979	8.011
3.1	8.044	8.077	8.111	8.144	8.179	8.213	8.248	8.283	8.318	8.354
3.2	8.390	8.427	8.463	8.500	8.538	8.576	8.614	8.652	8.691	8.730
3.3	8.770	8.810	8.850	8.890	8.931	8.973	9.014	9.056	9.099	9.142
3.4	9.185	9.228	9.272	9.317	9.361	9.406	9.452	9.498	9.544	9.591
3.5	9.638	9.685	9.733	9.781	9.830	9.879	9.929	9.979	10.029	10.080

Table G.2: 3-AFC Method - B value for the estimation of the variance of d' (Bi and others, 1997)

d′	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	2.7925	2.7862	2.7801	2.7740	2.7681	2.7624	2.7568	2.7513	2.7460	2.7408
0.1	2.7357	2.7307	2.7259	2.7212	2.7167	2.7123	2.7080	2.7038	2.6998	2.6959
0.2	2.6921	2.6884	2.6849	2.6814	2.6782	2.6750	2.6719	2.6690	2.6662	2.6635
0.3	2.6610	2.6585	2.6562	2.6540	2.6519	2.6499	2.6481	2.6464	2.6448	2.6433
0.4	2.6419	2.6406	2.6395	2.6384	2.6375	2.6367	2.6360	2.6355	2.6350	2.6347
0.5	2.6344	2.6343	2.6343	2.6345	2.6347	2.6351	2.6355	2.6361	2.6368	2.6376
0.6	2.6385	2.6396	2.6407	2.6420	2.6434	2.6449	2.6465	2.6483	2.6501	2.6521
0.7	2.6542	2.6564	2.6587	2.6611	2.6637	2.6664	2.6692	2.6721	2.6751	2.6783
0.8	2.6815	2.6849	2.6884	2.6921	2.6958	2.6997	2.7037	2.7079	2.7121	2.7165
0.9	2.7210	2.7256	2.7304	2.7353	2.7403	2.7454	2.7507	2.7561	2.7616	2.7673
1.0	2.7731	2.7790	2.7851	2.7913	2.7976	2.8041	2.8107	2.8175	2.8244	2.8314
1.1	2.8386	2.8459	2.8534	2.8610	2.8688	2.8767	2.8847	2.8930	2.9013	2.9098
1.2	2.9185	2.9273	2.9363	2.9454	2.9547	2.9642	2.9738	2.9836	2.9936	3.0037
1.3	3.0140	3.0244	3.0351	3.0459	3.0569	3.0680	3.0794	3.0909	3.1026	3.1145
1.4	3.1265	3.1388	3.1512	3.1639	3.1767	3.1898	3.2030	3.2164	3.2301	3.2439
1.5	3.2580	3.2722	3.2867	3.3014	3.3163	3.3314	3.3468	3.3623	3.3781	3.3942
1.6	3.4104	3.4269	3.4437	3.4607	3.4779	3.4954	3.5131	3.5311	3.5493	3.5678
1.7	3.5866	3.6056	3.6249	3.6445	3.6643	3.6845	3.7049	3.7256	3.7466	3.7679
1.8	3.7895	3.8114	3.8336	3.8561	3.8790	3.9021	3.9256	3.9494	3.9735	3.9980
1.9	4.0229	4.0480	4.0736	4.0995	4.1257	4.1523	4.1793	4.2067	4.2344	4.2626
2.0	4.2911	4.3201	4.3494	4.3792	4.4093	4.4399	4.4710	4.5025	4.5344	4.5667
2.1	4.5996	4.6328	4.6666	4.7008	4.7356	4.7708	4.8065	4.8427	4.8794	4.9167
2.2	4.9545	4.9928	5.0317	5.0711	5.1111	5.1517	5.1929	5.2346	5.2769	5.3199
2.3	5.3635	5.4077	5.4525	5.4981	5.5442	5.5910	5.6386	5.6868	5.7357	5.7853
2.4	5.8357	5.8868	5.9386	5.9912	6.0446	6.0988	6.1538	6.2096	6.2662	6.3236
2.5	6.3819	6.4411	6.5012	6.5622	6.6241	6.6869	6.7507	6.8154	6.8811	6.9479
2.6	7.0156	7.0843	7.1542	7.2250	7.2970	7.3701	7.4443	7.5196	7.5961	7.6737
2.7	7.7526	7.8327	7.9141	7.9967	8.0806	8.1658	8.2524	8.3403	8.4296	8.5204
2.8	8.6125	8.7061	8.8012	8.8979	8.9960	9.0957	9.1971	9.3000	9.4047	9.5110
2.9	9.619	9.729	9.840	9.954	10.069	10.186	10.305	10.426	10.549	10.674
3.0	10.801	10.930	11.062	11.195	11.331	11.469	11.609	11.752	11.897	12.045
3.1	12.195	12.347	12.503	12.660	12.821	12.984	13.150	13.319	13.491	13.666
3.2	13.844	14.025	14.209	14.396	14.587	14.781	14.978	15.179	15.383	15.591
3.3	15.802	16.018	16.237	16.460	16.687	16.918	17.154	17.393	17.637	17.886
3.4	18.138	18.396	18.658	18.925	19.196	19.473	19.755	20.042	20.334	20.632
3.5	20.936	21.245	21.559	21.880	22.206	22.539	22.878	23.224	23.575	23.934

APPENDIX H: *d′* TEST STATISTIC

Chi square test

$$\chi^2 = \sum \frac{(d'_{est} - d'_{exp})^2}{\sigma^2}$$

where:

d'_{est} = estimated d' value
d'_{exp} = expected d' value
σ^2 = variance of d' value

$$d'_{exp} = \frac{\frac{d'_a}{\sigma_a^2} + \frac{d'_b}{\sigma_b^2}}{\frac{1}{\sigma_a^2} + \frac{1}{\sigma_b^2}}$$

APPENDIX I: VARIANCE CORRECTION FOR REPLICATED TESTING

Estimate of the variance of d'

$$\sigma^2 (d') = \frac{B[1+\gamma(n-1)]}{N}$$

where:

B = B-value which specific to each method. It can be obtained from tables of B values as a function of P_c (Bi and others, 1997)
γ = gamma value, variation among trials. Ranges from zero to one
n = number of replications
N = number of panelists * n

Reference: Bi and Ennis, 1998

VITA

Karen Melissa Garcia was born in Tegucigalpa, Honduras. She attended Brassavola Bilingual School in La Ceiba, Honduras. Upon graduating high school, she attended Louisiana State University, receiving a Bachelor of Science in Chemical Engineering degree in May 2004. In fall 2004 she enrolled as a Master of Science degree candidate in the Department of Food Science at Louisiana State University. The focus of her research was the quality characterization of cholesterol-free mayonnaise-type spreads containing rice bran oil. She was awarded the Master of Science degree August 2006. She enrolled in the doctoral program at Louisiana State University in the Department of Food Science in January 2008. The primary focus of her research was sensory discrimination testing with children. She is currently a candidate for the Doctor of Philosophy degree to be awarded May 20, 2011.