



University of Kentucky
UKnowledge

Theses and Dissertations--Rehabilitation
Sciences

Rehabilitation Sciences

2015

PROBLEM SOLVING OF TYPICALLY DEVELOPING CHILDREN ON AN ADAPTATION OF THE TWENTY QUESTIONS TASK

Janice Carter Smith
University of Kentucky, janice.smith@wku.edu

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Recommended Citation

Smith, Janice Carter, "PROBLEM SOLVING OF TYPICALLY DEVELOPING CHILDREN ON AN ADAPTATION OF THE TWENTY QUESTIONS TASK" (2015). *Theses and Dissertations--Rehabilitation Sciences*. 26.
https://uknowledge.uky.edu/rehabsci_etds/26

This Doctoral Dissertation is brought to you for free and open access by the Rehabilitation Sciences at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Rehabilitation Sciences by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Janice Carter Smith, Student

Dr. Robert Marshall, Major Professor

Dr. Richard Andreatta, Director of Graduate Studies

PROBLEM SOLVING OF TYPICALLY DEVELOPING CHILDREN ON AN
ADAPTATION OF THE TWENTY QUESTIONS TASK

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Health Sciences
at the University of Kentucky

By

Janice Carter Smith

Lexington, Kentucky

Director: Dr. Robert Marshall, Professor of Communication Sciences and Disorders

Lexington, KY

2015

Copyright © Janice Carter Smith 2015

ABSTRACT OF DISSERTATION

PROBLEM SOLVING OF TYPICALLY DEVELOPING CHILDREN ON AN ADAPTATION OF THE TWENTY QUESTIONS TASK.

This cross sectional study examined problem solving by typically developing children on the Rapid Assessment of Problem Solving test (RAPS). The RAPS, a modification of Mosher and Hornsby's 20Q task, requires the examinee to solve three problems. Each problem involves asking yes/no questions to identify a target picture from a 32-picture array with as few questions as possible. Participants were 73 young (ages 7-9), 79 early adolescent (ages 10-13) and 77 adolescent (ages 14-17) children residing in Kentucky. Children were seen in the summer months and administered the RAPS on a single occasion, with 22 of the children being testing twice. All children passed screening tasks and completed RAPS testing without difficulty. Test-retest stability for the RAPS was adequate for clinical purposes and no learning effects were seen on the test. Results were examined to identify group differences in components of executive functioning (planning, strategy selection, strategy execution, and strategy shifting) that impact problem solving efficiency. To determine how children went about solving problems, questions were classified by type and in terms of when they were asked in the sequence of questions leading to solving of a problem. Results revealed that the young group differed from the early adolescent and adolescent groups on several objective measures: number of questions to problem solve, use of constraint questions, problem solving efficiency, mean integration planning score, and overall RAPS efficiency. The young group also differed from the two older groups in terms of the types of questions asked and when certain types of questions were asked in solving a problem. Young children were more prone to guess on early questions whereas older children asked effective constraint questions. Many of the differences suggest young and older children and young and older adults differ in their ability to integrate information needed to solve RAPS problems effectively. Findings of this study suggest there are age-related differences in solving fixed-alternative 20Q problems and provide a normative data base for using the RAPS to assess problem solving of both normal and disabled children in the age range studied.

KEYWORDS: Problem Solving, Rapid Assessment of Problem Solving, Twenty Questions Task, Strategy, Typical Children

Janice Carter Smith
Student

April 20, 2015

PROBLEM SOLVING OF TYPICALLY DEVELOPING CHILDREN ON AN
ADAPTATION OF THE TWENTY QUESTIONS TASK

By

Janice Carter Smith

Robert C. Marshall
Director of Dissertation

Richard D. Andreatta
Director of Graduate Studies

April 20, 2015

DEDICATION

“We must begin thinking like a river if we are to leave a legacy
of beauty and life for future generations.”
— David Brower

This work is the product of those who came before me, those who pushed me onward and
those who have carried me. It is theirs more than mine.
To them, I dedicate my dissertation.

ACKNOWLEDGMENTS

The following dissertation is the result of faith, determination, perseverance and an incredibly supportive network of people in my life. First, my Dissertation Chair, Dr. Robert Marshall, exemplifies the high quality of scholarship to which I aspire. In addition, Dr. Marshall provided timely and instructive feedback and evaluation at every stage of the dissertation process, allowing me to learn from his modeling and expertise.

Next, I wish to thank the complete Dissertation Committee, and outside reader, respectively: Dr. Judith Page, Dr. Peggy Wittman, Dr. Jane Kleinert, Dr. Lee Blonder, and Dr. David Berry. Each individual invested time and energy in the editing of my dissertation as well as the completion of my doctoral program. In addition to the assistance above, I am very grateful for the valuable support regarding statistical analyses that I received from Dr. Philip Westgate.

My colleagues at Western Kentucky University consistently afforded me support and encouragement without fail. I am incredibly grateful for their faithful kindness. Emily Smith and Beth Avis acted as Research Assistants who tirelessly joined in my doctoral pursuit. I also wish to thank the WKU CSD graduate classes of 2013 and 2014 who served as testers for my study as well as the participating parents and children (who remain anonymous for confidentiality purposes). Their comments and insights created an informative and interesting project with opportunities for future work.

Finally, without the support and guidance of my family and friends, I would not have finished this degree. I owe a debt of gratitude to my grandparents, Willis and Vera Mae Carter, who, at a very young age, taught me that through Christ Jesus all things are possible...even for me. This accomplishment is a direct result of their legacy. My mom, Patti Taylor Carter Baribeau, instilled in me the desire and skills to dream big and then work hard in order to make my dreams a reality. She was, is, and will always be my strongest role model, my exemplar of drive, determination and perseverance through adversity. My children, Joshua and Jayli, gave me a depth of purpose, inspiration and a healthy life balance between serious and silly that served as an anchor throughout the doctoral process. My hubster, Jim Smith, provided unwavering support and sacrificial devotion. He believed in me when I didn't. He drove me forward, carried my load and wiped my tears. This degree belongs unequivocally to both of us. Thank you, Jim, for giving me the strength to be a finisher.

TABLE OF CONTENTS

Acknowledgements.....	iii
List of Tables	v
List of Figures.....	vi
Chapter One: Introduction	
Background.....	1
Research Questions.....	10
Chapter Two: Review of Literature	
Mosher and Hornsby’s Task.....	12
Optimal Problem Solving Paradigm	15
Materials for the RAPS	16
Administration of the RAPS	18
Scores for the RAPS	20
Categorization of Questions on the RAPS	22
Strategy Group Assignment	23
Research with the RAPS	24
Chapter Three: Methods	
Participants.....	32
Screening Tasks	32
Procedures.....	35
Scoring	40
Strategy Group Assignment.....	50
Groupings of Participants by Age	51
Chapter Four: Results	
Reliability and Test Stability	52
Gender Differences	54
Within and Between Group Differences.....	55
Descriptive Analysis Results	63
Chapter Five: Discussion	
RAPS and the 20Q Task	72
Group Differences in Problem Solving.....	75
Differences in Adult and Child Performance on the RAPS.....	83
Limitations	87
Clinical Implications	89
Future Research	89
References.....	91
Vita.....	100

LIST OF TABLES

Table 2.1, Summary of differences between Twenty Questions Task and the RAPS.....	17
Table 3.1, Procedural review checklist for students	40
Table 3.2, Characteristics of the participants included in the data analysis	51
Table 4.1, Total number of questions needed to solve 3 RAPS problems and mean number of questions per problem solved.....	56
Table 4.2, Mean percentage of constraint questions for each age group.....	56
Table 4.3, Mean integration planning scores for each group.....	58
Table 4.4, Mean problem solving efficiency scores for young, early-adolescent, and adolescent groups for the three RAPS problems	59
Table 4.5, Mean overall efficiency scores	60
Table 4.6, Mean and standard deviations for question sequence efficiency scores for young, preadolescent, and adolescent participants on the RAPS	62
Table 4.7, Number and percentage of novel, category-limited, narrowing, inefficient-constraint questions and guesses for young, early-adolescent, and adolescent participants on the RAPS on the first four questions	64
Table 4.8, Percentages of types of questions asked by young, preadolescent and adolescent participants for the first six questions of the RAPS	66
Table 4.9, Problem solving strategies for each age group	69

LIST OF FIGURES

Figure 1.1, The Twenty Questions Task by Mosher and Hornsby (1966).	6
Figure 2.1, RAPS Recording Form and Problem Solving Board	19
Figure 3.1, RAPS Screening Protocol for Children	34
Figure 3.2, Sample Scoring for the RAPS	42
Figure 4.1, Mean QSE Scores for Questions 1 through 4 by Age Group	62
Figure 4.2, Percentage of times the first category-limited question targeted each category size available by age group	65
Figure 5.1, Overall RAPS efficiency scores for children in this study and adults	86

Chapter 1: Introduction

Background

Throughout the life span humans are confronted with problems to solve. M. Scott Peck, author of the best-selling book *The Road Less Traveled*, suggests solving problems brings out both our courage and our wisdom and is crucial to distinguishing success and failure (1978). From an early age, children are given problems to solve by their parents. Whether it be something as simple as deciding what shirt to wear in the morning or as significant as dealing with an unreasonable employer most individuals experience satisfaction in solving a problem. Some neuropsychologists contend that problem solving comes into play when routine or automatic behaviors are inadequate for attaining a desired goal (Kiel & Kaszniak, 2002; Lezak, Howieson, Loring, Hannay, & Fischer 2004; Shallice, 1982). It is also thought that problem solving entails the use of overlapping metacognitive skills such as problem identification, goal setting, planning, strategic thinking, and generating alternative solutions (Allen, Chinsky, Larcen, Lochman, & Selinger, 1976).

The effectiveness, with which problems are solved, apart from complexity of the problems themselves, varies from person to person. D’Zurilla and Goldfried (1971, p. 107) cite Socrates’ observation that competent individuals are “those who manage well the circumstances which they encounter daily, and who possess a judgment which is accurate in meeting occasions as they arise and rarely miss the expedient course of action.” Conversely, these authors point out that much of what is considered as “abnormal behavior” reflects ineffective problem solving that may sometimes result in

undesirable effects such as anxiety, depression, or creation of additional problems. This distinction in problem solving abilities of normal individuals lends support to Lezak and colleagues' (2004) assertion that problem solving is a component of executive functioning and something that "can occur at any place on the complexity and abstraction continua" (p. 30).

Problem solving has been found to be compromised, to greater or lesser degrees, in persons with neurologic damage and disease. This has been well-documented in individuals with traumatic brain injuries (Ben-Yishay & Diller, 1983; Oddy, 1984; Glosser & Goodglass, 1990; Goldstein & Levin, 1991), strokes (Prescott, Gruber, Olson, & Fuller, 1987; Prescott, Loverso, & Selinger, 1992; Purdy, 2002; Wade, Legh-Smith, & Hewer, 1986), progressive neurological disease (Benson et al., 1983; Kuhn et al., 1998; Lubinski, 1995), and severe mental illness (Gold & Harvey, 1993; Saykin, Shtasel, Gur, Kester, Mozley, Stafiniak, & Gur, 1994). While more is known about the problem solving limitations of neurologically compromised adults, children with mental retardation (Borys, 1979), learning disabilities (Agran, Blanchard, Wehmeyer, & Hughes, 2002; Barton, 1988; Glago, Mastropieri, & Scruggs, 2008; Simmonds, 1990), and autism spectrum disorders (Alderson-Day et al., 2011; Alderson-Day & McGonigle-Chalmers, 2011; Minshew, Siegel, Goldstein, & Weldy, 1994) have also been found to have difficulties solving problems.

Lezak and colleagues (2004) have suggested the simplest issues of daily life call for problem solving and that problem solving involves executive functions. They define executive functions as "capacities that enable a person to engage successfully in

independent, purposive, self-serving behavior” (p. 35). These authors further point out that one’s ability to solve problems may have greater consequences for living independently than physical or cognitive limitations. For these and other reasons, cognitive psychologists, neuropsychologists, educational psychologists, and others have developed assessment tools, both formal and informal, to obtain information about one’s ability to solve problems. Rehabilitation specialists, educators, and others use this information to pinpoint difficulties in planning, strategy selection, use, shifting, and other components of the problem solving process to plan interventions and maximize the patient’s social, vocational, and educational integration.

Instruments that assess problem solving use both conventional neuropsychological and socially based approaches. Conventional approaches typically present the examinee with problems that are complex, highly structured, well-defined, impersonal, and emotionally neutral (Channon & Crawford, 1999; Lezak et al., 2004). These could involve standardized tests such as the Wisconsin Card Sorting Test (WCST; Grant & Berg, 1948), Porteus Maze Test (PMT; Porteus, 1965), and Colored Progressive Matrices (CPM; Raven, Raven, & Court, 1984) or problem solving tasks as contained on measures such as the Tinkertoy Test (Lezak, 1995), Tower of London (Shallice, 1982), Tower of Hanoi (Prescott et al. 1987), and Wheelbarrow Assembly Test (Butler, Rorsman, Hill, & Tuma, 1989).

Socially motivated problem solving assessment tools present the examinee with problems that might occur in living one’s life (e.g., intrapersonal/nonsocial problems, impersonal problems, interpersonal problems, and broader community and societal

problems (D’Zurilla, Maydeu-Olivaris, & Kant, 1998). Representative examples include the Functional Assessment of Verbal Reasoning and Executive Strategies test (FAVRES; McDonald & Johnson, 2005), Rusk Problem Solving Role Play Test (RPSRPT; Sherr, Rath, Langenbahn, Simon, Biderman, & Diller, 1998), Multiple Errand Test (MET; Aitken, Chase, McCue, & Ratcliff, 1993), Six Element Test (Shallice & Burgess, 1991), Route Finding Task (Boyd & Sautter, 1993), Test for Functional Abilities (TOFA; Bamdad, Ryan, & Warden, 2003), Assessment of Interpersonal Problem-Solving Skills (AIPSS; Donahoe, Carter, Bloem, Hirsch, Laasi, & Wallace, 1990), and the Hospital Version of the Multiple Errand Test (MET-HV; Knight, Alderman, & Burgess, 2002). Finally, questionnaire/inventories such as the Heppner Problem Solving Inventory (Heppner, 1988), Everyday Problem Solving Inventory (Cornelius & Caspi, 1987), and Behavioral Assessment of Executive Function (Wilson, Alderman, Burgess et al., 1996) have been used to assess problem solving.

Regardless of whether a neuropsychological or socially motivated approach is used to assess problem solving, many of the tests and tasks used to assess adults are not suitable for children. This is particularly true when it comes to assessing problem solving capabilities of young children with learning disabilities, developmental delays, and autism spectrum disorders. Some of the reasons children have difficulties with problem solving measures designed for adults relate to (1) difficulties understanding the test instructions, (2) receptive and expressive language demands of the task, (3) motor skills needed to perform the task, and (4) the fact that tasks used to assess problem solving in adults are not always interesting or engaging to children. In many situations, rather than

objectively assessing children's problem solving with a test, clinicians rely on observations and/or parental or teacher reports.

While assessment of problem solving abilities in children, particularly young children, poses a formidable challenge, there is a task that has stood the test of time in this regard. Over the years, this task has become known as the twenty questions task (20Q). The 20Q task, based on the parlor game of the same name, was introduced to the scientific community in 1966 by two educational researchers, Frederic A. Moshier and Joan Rigney Hornsby (1966). It consisted of a single 6 x 7 array of watercolor drawings of common objects (see Figure 1.1). Moshier and Hornsby used the 20Q task to study the development of question-asking strategies of six, eight, and 11-year-old children. Child participants were instructed that they were going to play a "question asking game" and that the object of the game was to try to identify a picture the examiner was thinking of by asking questions that could be answered "yes" or "no" with the goal of doing this with as few questions as possible (Denney, 1985). Moshier and Hornsby (1966) found that children asked three types of questions: constraint-seeking (CS), hypothesis-scanning (HS), and pseudoconstraint (PC). CS questions (e.g., Is it an animal?) eliminated more than one picture from consideration regardless of whether they were answered "yes" or "no." HS and PC questions were guesses that targeted one picture resulting in a high payoff for a "yes" answer and having little effect on target picture identification if answered "no." HS questions, however, targeted a picture by name (e.g., Is it the dog?), whereas PC questions were phrased as CS questions (e.g., Is it something that barks?). Moshier and Hornsby found that children systematically reduced their use of HS and PC questions and increased their use of CS questions over the age range studied.

Although Mosher and Hornsby (1966) developed their 20Q task to study children's ability to ask informative questions to seek information about their world, subsequent investigators have conceptualized the 20Q task as a problem solving endeavor (Carroll, 1993; Denney, 1985; Horn & Cattell, 1967; Siegler, 1977). In the last half century, the Mosher and Hornsby 20Q task has been used to (a) examine problem solving in normal children and adults across the life span (Denney, 1985), (b) compare problem solving in normal and clinical populations, and (c) determine the effects of training on problem solving. Some researchers have created their own versions of the 20Q task. For

Figure 1.1 The Twenty Questions Task by Mosher and Hornsby (1966)



example, Delis, Kaplan, & Kramer (2001) used a 30-picture array for a 20Q task included in a test battery for assessing executive function. McKinney's (1973) twenty question task, the Flowers test, is made up of 16 pictures of flowers differing along four

dimensions: size, color of petals, color of stem, and number of petals. Drumm, Jackson, and Magley's (1995) task involves a set of 24 pictures of turtles, walruses, lizards, and dogs that differ by habitat (natural/biological) and maturity (young/old). Zelniker and colleagues (Zelniker, Renan, Sorer, & Shavit, 1977) used a set of 36 pictures of common objects belonging to six conceptual categories (fruit, animals, toys, musical instruments, clothes, and dinnerware) to study problem solving of reflexive and impulsive children. Rather than use a closed set of pictures, Klouda and Cooper (1990) used an open-ended task in which brain injured and normal subjects were instructed to ask yes/no questions to identify an animal of which the examiner was thinking. Thus it appears safe to say that several twenty question *tasks* have been used to assess problem solving of children and adults, but there is no specific twenty questions *test*.

In 2003, Marshall and colleagues introduced the Rapid Assessment of Problem Solving test (RAPS) as an alternative to the Mosher and Hornsby's 20Q task (Marshall, Karow, Morelli, Iden, & Dixon, 2003). The RAPS was intended to be used to assess problem solving of brain injured adults in clinical settings where these individuals are often hard to test and/or need to be assessed quickly to inform clinical decision-making. While the RAPS is similar to the Mosher and Hornsby (1966) 20Q task, it was designed as a *test* of problem solving and its materials, administration, and scoring differ markedly from those of the Mosher and Hornsby (1966) task and its variants. Research with the RAPS, to be presented in greater detail in Chapter 2, has found the test to be useful for clinicians in need of information about an adult client's problem solving abilities. Two studies have shown that neurologically intact subjects reflect a range of normally distributed scores on the RAPS, good test-retest reliability, an absence of learning effects,

and that performance on the test is minimally affected by age, gender, and years of education (Marshall et al., 2003a; Marshall & Karow, 2008). Research with neurologically compromised subjects has shown the RAPS to be sensitive to brain damage (Marshall & Karow, 2013). Most studies have found neurologically compromised subjects ask more questions, ask different types of questions, and use different question-asking strategies when solving problems on the RAPS than healthy controls (Marshall, Karow, Morelli, Iden, & Dixon, 2003b; Marshall, McGurk, Karow, Kairy, & Flashman, 2006; Marshall, McGurk, Karow, & Kairy, 2007). Finally, clinical studies have demonstrated that the RAPS can be used to document changes in problem solving after interventions (Marshall, Capilouto, & McBride, 2007; Marshall, Dixon, Iden, Karow, & Morelli, 1999; Marshall, Karow, Morelli, Iden, Dixon, & Cranfill, 2004; Marshall, et al., 2009).

This normative study had two primary goals. One was to examine the performance of typically developing children on the RAPS. The second was to establish a normative database for the RAPS with children similar to that available for adults so as to eventually inform use of the RAPS to assess problem solving of children with disabilities. Since the RAPS is designed similarly to the 20Q task, and the 20Q task has been used successfully with children with disabilities, there is every reason to assume this clinical test would be a useful tool to assess problem solving of children with disabilities. This assumption, however, does not preclude the need to carry out a normative study to determine how typically developing children perform on the RAPS. Kafer and Hunter (1997) have cautioned against accepting results of clinical tests at face value. These researchers stressed the importance of providing normative data that demonstrate the

test's ability to reliably measure executive functions such as problem solving in normal populations before using the test with clinical populations.

Aside from the fact that there are no published data delineating how typically developing children perform on the RAPS, a gap which this study hopes to fill, there are additional reasons to support carrying out the study. First, validating the RAPS as a tool for assessing problem solving of children could potentially provide clinicians with information about children's problem solving that is quantitatively and qualitatively superior to that provided by 20Q task and its variants. Information obtained from the RAPS, as will be shown in Chapter 2, permits the clinician/examiner to identify component processes of problem solving such as planning, strategy selection, and strategy shifting (Scholnick & Friedman, 1993). Secondly, the RAPS is not simply a single problem solving task, but a *test* of problem solving. Because it was designed to be a test, it contains multiple similarly constructed problems, which allows the clinician to assess the same individual repeatedly. Clinicians are able to measure changes in problem solving abilities over time and/or after an intervention designed to improve problem solving. Finally, and on a more personal note, the investigator has spent the better part of a lifetime assessing and treating children with autism in the public schools, university clinics, and through home health care agencies. During this time, two clinical needs were apparent: assessments that were quick, motivating, and accurate and assessments that translate to real-life skills. The RAPS has potential to provide clinicians with rich information regarding cognitive planning and mental flexibility used in problem solving. These skills have a direct impact on many activities of daily living. With approximately 50,000 children with autism turning 18 each year, clinician-friendly measurement tools

are needed to assess results of interventions focused on helping children with autism transition into adulthood (Shattuck et al., 2012). The RAPS can provide meaningful data regarding question-asking efficiency, prioritization of information, and strategy use that can inform intervention plans targeting independence.

Research Questions

Accordingly, this cross sectional study examined the performance of typically developing young, early-adolescent, and adolescent children on the Rapid Assessment of Problem Solving test (RAPS) and sought to answer the following general research questions:

1. Do young, early-adolescent, and adolescent children differ in the efficiency with which they solve problems on the RAPS?
2. Do young, early-adolescent, and adolescent children exhibit differences in planning, strategy selection, strategy execution, and strategy shifting when solving problems on the RAPS?
3. Do young, early-adolescent, and adolescent children differ in the types of questions asked to solve problems on the RAPS?

Chapter 2: Review of the Literature

Problem solving is a complex composite ability. It taps several cognitive skill sets, some of which include planning, thinking ahead, understanding the consequences of one's actions, and making choices; all are components of executive functioning (Brookshire, 2007). As pointed out in Chapter 1, there are many ways to assess problem solving, but most of the assessment tools available are better suited for use with adults rather than children. The present study does not provide detailed information on the construct of problem solving per se, nor does it address how executive functioning is assessed. Its goals were to examine the performance of typically developing children on a specific clinical measure of problem solving, the Rapid Assessment of Problem Solving test (RAPS; Marshall et al., 2003a), establish a normative data base for the RAPS for children equivalent to that available for adults (Marshall & Karow, 2008), and provide clinicians and educators a simple tool to assess problem solving of both typically developing children and children with disabilities.

This chapter supplies (1) background information on the 20Q task of Mosher and Hornsby (1966), the paradigm on which the RAPS is based; (2) description of materials, administration, scoring, and other aspects of the RAPS; (3) and a summary of findings from research carried out with the RAPS with adult participants.

Mosher and Hornsby's Task

Mosher and Hornsby's 20Q task (1966) included a single page of 42 watercolor drawings of common objects (see Figure 1.1.) When the task was administered, the picture was placed in front of the examinee with these instructions:

Now we're going to play a question-asking game. I'm thinking of one of these pictures and your job is to find out which one it is that I have in mind. To do this you can ask any questions at all that I can answer by saying "yes" or "no," but I can't give any other answer but "yes" or "no." You can have as many questions as you need, but try to find out with as few questions as possible (Denney, 1985).

Questions from the examinee were classified as constraint-seeking (CS), hypothesis-scanning (HS), or pseudoconstraint questions (PC). CS questions were those that eliminated more than one object at a time regardless of whether they were answered "yes" or "no," (e.g., Is it living?). HS and PC were forms of guessing that eliminated only one picture with a "no" answer and solved the problem with a "yes" answer. HS questions, however, targeted the picture by name (e.g., Is it the clock?). PC questions (e.g., Is it something that tells time?) were phrased like CS questions, but did not mention the picture by name.

Life span studies with the 20Q task. Mosher and Hornsby (1966) used the 20Q task to examine information seeking strategies of six, eight, and 11 year-old boys. Children were found to decrease their use of HS questions from nearly 100% to approximately 10%, increase their use of CS questions, and slightly increase their use of PC questions over this age range. Earliest studies with adults using the 20Q task found that elderly adults (mean age, 82.5 years) asked more of the less-efficient HT questions

and fewer CS questions than middle-aged adults (Mean, 38.2 years), (Denney & Denney, 1973; Kesler, Denney, & Whitley, 1976). Subsequent investigations with adults spanning a wider age range found there was a linear decrease in the use of CS questions on the 20Q task from young adulthood through old age (Denney, 1982; Denney & Palmer, 1981).

Clinical studies with the 20Q task. The 20Q task has also been used to examine and compare problem solving of children and adults with and without cognitive-communicative disabilities. A study by Barton (1988) found boys with learning disabilities performed less efficiently on the task than matched normal controls. Borys (1979) reported the performance of young adults with mental retardation was below or comparable to that of first grade normal children on the task. Investigations with young and adolescent children with traumatic brain injuries found brain injured children did significantly more guessing and asked fewer CS questions than neurologically intact controls (Levin, et al., 1997; Levin, et al., 1993). Marschark and Everhart (1999) examined the performance of deaf and hearing students across multiple administrations of the 20Q task. Deaf students asked significantly more questions and often failed to solve the problems. They also asked significantly fewer CS questions than hearing students and significantly less efficient questions as evidenced by the number of pictures targeted by their questions.

Three studies have compared the performance of adults with histories of alcohol abuse with neurologically intact subjects on the 20Q task (Laine & Butters, 1982; Goldman & Goldman, 1988; Saarnio, 1993). All of the studies found that abstinent alcoholic subjects needed significantly more questions and asked significantly fewer CS

questions in solving 20Q problems than matched normal controls. A study by Marshall and colleagues (Marshall, Harvey, Freed, & Phillips, 1996) compared the performances of stroke survivors with mild aphasia and non-brain-damaged (NBD) participants on the 20Q task. Results of this study indicated that the individuals with aphasia asked significantly fewer CS questions and obtained less information per question than the NBD subjects.

Training effects. The effects of different training strategies on problem solving have also been examined using the 20Q task as a dependent variable. Some studies found that performance of young children could be improved using strategy modeling techniques. When employing strategy modeling, the clinician demonstrates the use of effective CS question asking strategies while playing a twenty questions game (Denney, Denney, & Ziobrowsky, 1973; Denney, 1972; Denney & Turner, 1979). Denny (1975) compared the effects of three modeling techniques (exemplary modeling, cognitive modeling, and a combination of the two) on six, eight, and 10 year-old children's use of CS questions on the 20Q task. Cognitive modeling alone was the most effective procedure and differences between cognitive and exemplary modeling were greater for the younger children. Denney and Connors (1974) compared the effects of exemplary strategy modeling, non-exemplary strategy modeling, and a control condition on CS usage by pre-school children who asked no CS in a pre-test condition. Participants in the strategy-modeling condition asked significantly more CS questions on the post-test than the control participants. Strategy-modeling training was found to be effective with children as young as four and five years old.

Studies examining the effects of strategy modeling on adults' performance on the 20Q task have yielded results similar to those with children. Denney and Denney (1974) trained adults between the ages of 70 and 90 to ask CS questions in two conditions, after a model and after a model accompanied by an explanation of the strategy behind the question. Elderly adults responded positively in both conditions. Moreover, elderly individuals that asked no CS questions on a pre-test asked CS questions on a post-test. Denny, Jones, and Krigel (1979) tested a hypothesis put forth by Denney and Wright (1976) that children might need to be taught to ask CS questions whereas older adults already possessed the knowledge to ask these types of questions, but did not always do so. These researchers compared the performance of six-year-old children and elderly adults on the 20Q task following a training procedure in which strategy modeling was broken down into three components: classification training, exemplification of CS questions based on classification, and use of information obtained in response to questions asked. Results revealed that both young children and older adults possessed the knowledge to ask CS questions and did not support the Denney and Wright's (1976) hypothesis that older adults had an advantage over children in asking CS questions based on life experience.

Optimal Problem Solving Paradigm

According to Carroll (1993) and others, the types of problems presented on the 20Q task and the RAPS test tap into general sequential reasoning, an important aspect of fluent intelligence and executive function tests (Horn & Cattell, 1967). The paradigm for solving problems on the RAPS stipulates that the most efficient or optimal way to solve

problems on the test is to ask constraint questions that systematically reduce the number of pictures from consideration in near 50% increments (Hartley & Anderson, 1983; Marshall and Karow, 2008). Using this paradigm, question 1 would reduce the number of pictures from 32 to 16; question 2 would reduce the number of pictures from 16 to eight and so forth. Question efficiency scores are designed to reflect the examinee's ability to use this optimal problem solving strategy.

Differences between the 20Q task and the RAPS. While the RAPS is similar to the 20Q task, the test contains a number of modifications that distinguish it from the older procedure. These modifications, differences, and the rationales for incorporating them into the RAPS were provided in previous publications on the RAPS (Marshall et al., 2003a; Marshall & Karow, 2008) and will not be provided in detail here for the sake of brevity. Table 2.1, however, highlights the differences in the 20Q task and the RAPS. This table has been included to aid the reader's understanding of the materials, administration, and scoring of the RAPS, which follows.

Materials for the RAPS

The RAPS has nine problem solving boards similar to the example shown in Figure 2.1. Each board is made up of 32 pictures of common objects derived from 18 known semantic categories (animals, birds, body parts, clothing, desserts, food, furniture, gardening equipment, insects, kitchen items, medical equipment, musical instruments, plants, sea creatures, sports balls, tools, toys, and transportation). The 32 pictures are arranged in a 4 x 8 grid. Half of the pictures are colored and half are black and white. Each board contains pictures from 6 of the 18 semantic categories with the number of

pictures from given categories controlled. Every board is composed of one category of 8, two categories of 6, and three categories of 4 pictures from the same category.

TABLE 2.1 Summary of differences between Twenty Questions Task (Mosher & Hornsby, 1966) and the RAPS (Marshall et al., 2003).

Feature	Twenty Questions	RAPS
Picture Stimuli	42 pictures in a 7 x 6 matrix	32 pictures in a 4 x 8 matrix
Picture features	Black and white line drawings	16 colored and 16 black and white pictures; slightly larger pictures
Picture categories	No control of number of pictures specified	Number of pictures in categories is controlled
Instructions	No modifications for disabled populations	Contains modifications for disabled populations; directs attention to problem solving board; stress on the word <i>few</i>
Screening	No screening for oral naming or picture recognition deficits; no practice on task	Screening for oral naming and picture recognition deficits; practice in yes/no question asking
Procedures	Does not cover pictures eliminated by questions	Covers pictures eliminated by questions
Repeat administration	Uses same 42-item picture repeatedly	Nine unique problem solving boards
Scoring	Number of questions needed to solve problem; % of constraint seeking questions	Adds question-asking efficiency scores; adds integration planning score

Categories of 8, 6, and 4 pictures are varied across the nine boards and pictures in the same category do not appear adjacently. Each board has a designated recording form to record the examinee's questions and other information needed to score the test. Figure 2.1 shows one problem solving board from the RAPS and its accompanying recording form to provide a representative example of how a problem might be solved.

Administration of the RAPS

To complete a RAPS test, the examinee solves three problems. To solve each problem, the examinee asks yes/no questions to identify a target picture of which the examiner is thinking. For each problem, the examiner selects a different problem solving board and preselects a different target picture. To start the test, the examiner places the first problem solving board in front of the examinee and gives these instructions:

“We are going to play a question-asking game. I am thinking of one of these pictures (examiner gestures to the pictures) and your job is to figure out which one it is. The way to do this is to ask me questions that I can answer “yes” or “no.” You can ask me any question you want so long as I can answer it “yes” or “no.” Try to ask as *few* questions as possible. When you are ready, go ahead and ask your first question.”

After each question from the examinee, the examiner answers “yes” or “no” and covers the pictures eliminated by the question with blank cards before allowing the examinee to ask the next question. This process continues until the examinee’s questions have reduced the number of pictures to two or three. The problem is considered solved at this point because guessing is the only option left. Second and third problems are then presented for solving in a similar manner. No time limit is set for solving a problem. However, in cases where the examinee’s questions are only guesses, a 10-question limit is imposed and the examiner defaults to a “yes” answer after the tenth question to terminate the task.

Administration guidelines for the RAPS (Marshall et al., 2003a) also specify the actions to be taken by the examiner when the examinee does not ask yes/no questions, or asks ambiguous questions, or if the examiner cannot determine the pictures targeted by the examinee’s question. Because these difficulties occur rarely when administering the RAPS and, for the sake of brevity, details on “examiner actions” will not be included

here. For this information, the reader is instead referred to papers on the RAPS by Marshall et al. (2003a) and Marshall and Karow (2008).

Figure 2.1 RAPS Recording Form and Problem Solving Board

RAPS
Robert C. Marshall & Colleen M. Karow

Board #8 Recording Form









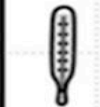























Patient Initials: **018** Gender: **F** Date of Testing: **[redacted]**
 D.O.B.: **[redacted]** (Age) **7.10** Target: **tennis ball** Problem: 1 2 3

Question Asked (write out)	Y/N	(C)onstraint / (G)uess	# Pictures Considered	# Pictures Targeted	# Pictures Eliminated	Question Efficiency	
1. Is it a dessert?	N	CL	32	8	8	0.50	
2. Is it a bird?	N	CL	24	6	6	0.50	
3. Do you use it to help people?	N	NOV	18	10	10	1.11	
4. Do you wear it on your feet?	N	IC	8	1	1	0.25	
5. Do you wear it on your body?	N	CL	7				
6. Can you play with it?	Y	IC	4				
7. Is it the tennis ball?	Y	FG	4				
8.							
9.							
10.				0.50	0.50	1.11	0.25

Comments: **#3: medical + tools**

Tally Types of Questions

Constraints				Guesses	
Category-Limited	Narrowing	Novel	Inefficient Constraint	Frank Guess	Classic constraint
3		1	2	1	

Scores for the RAPS

Four types of scores have been used to quantify performance on the RAPS. Scores are computed for each problem and averaged for the three problems. These scores are: (1) number of questions needed to solve the problem, (2) percentage of constraint seeking questions asked, (3) efficiency scores, and (4) integration planning scores.

Number of questions. A problem is solved when the examinee's questions have reduced the number of pictures to two or three. At this point, the number of questions used to solve the problem is counted. Questions that could not be answered yes or no or those questions that the examiner asked the examinee to rephrase because of lack of clarity are not included in the number of question counts. A mean number of questions score is obtained by summing and averaging the number of questions used in solving the three problems. Thus, if the examinee used 7, 9, and 5 questions to solve problems 1, 2, and 3 respectively, his mean score would be seven ($7 + 9 + 5 = 21/3$).

Percent constraint-seeking questions. Acceptable yes/no questions are designated as constraint-seeking questions (CS) or guesses. As stated earlier, CS questions eliminate more than one picture from consideration regardless of whether they are answered yes or no. Guesses can occur in two forms referred to as hypothesis-scanning (HS) and pseudoconstraint (PC) questions. HS questions, also called frank guesses, target a single picture by name (e.g., Is it the horse?); whereas, PC questions target a single picture, but are phrased like CS questions (e.g., Is it the animal with a mane?). The percent of CS questions is determined by dividing the number of CS used to solve the three problems by the total number of questions ($CS/CS + HS + PC$). Thus if an

examinee asked 14 CS questions, 2 HS questions, and 2 PC questions for the test, his % CS score would be .78 ($14/14 + 2 + 2$).

Question efficiency. Efficiency scores are calculated for the first four questions of each of the three problems. A question's efficiency is determined by dividing the smaller of two numerators, pictures targeted or pictures eliminated, by the number of pictures available on the problem solving board when the question is asked. For example, if the target picture were "tennis ball" for the problem solving board shown in Figure 2.1, and the examinee's first question was "Is it a dessert?," the examiner would answer "no" and cover the pictures of the desserts. Since there are eight pictures of desserts, this question's efficiency score would be .50 ($8/32 = .25 \times 2$). There was no smaller numerator for this question since eight pictures were targeted and eight were eliminated. However, had the examinee's first question been "Is it the tennis ball?," the smaller numerator of one would be used to calculate the question efficiency score of .06 ($1/32 \times 2$) to negate the effects a fortunate guess on the question efficiency score. The four efficiency scores are averaged for each problem; the problem efficiency scores are averaged for the test.

Integration planning score. Marshall et al. (2006) introduced the integration planning score (IPS) in a study that compared RAPS performance of subjects with and without severe mental illness. This score was based on the premise that the examinee's ability to ask a highly efficient first question that eliminates approximately half the pictures from consideration reflects a degree of planning on the part of the individual congruent with the goal of identifying the target picture with as few questions as possible.

The IPS score is determined by assigning a value from 1 to 6 to the first question for each problem based on the number of pictures targeted by the examinee's first question: 1 = one picture; 2 = two or three pictures; 3 = four or five pictures; 4 = six or seven pictures; 5 = eight pictures; 6 = nine or more pictures. Integration planning scores are then summed and averaged for the three problems.

Categorization of Questions on the RAPS

Categorizing the types of questions asked in solving problems on the RAPS and determining when different types of questions are asked in the problem solving sequence provides valuable insights into how individuals, with and without neurological damage, go about solving problems. Marshall and Karow (2008) developed explicit definitions to categorize the types of questions asked by normal adults on the RAPS. Using these definitions, the researchers reliably categorized 4842 questions from 373 normal adult subjects as novel, category-focused, narrowing, or inefficient constraint questions or as guesses. Their definitions follow:

Novel questions were defined as those that target nine or more pictures and/or have efficiency scores above 50%.

Category-limited questions were defined as questions that targeted all pictures in one semantic category. Typically, these questions targeted 4, 6, or 8 pictures since these are the sizes of the picture categories on the RAPS, but it is also possible for category-limited questions to target fewer pictures if other pictures have already been eliminated from consideration by prior questions.

Narrowing questions were defined as constraint questions asked after a “yes” answer to a category-limited question. For example, suppose the target picture for the problem solving board in Figure 2.1 was “pie,” and the examinee asked “Is it a dessert?” Here the examiner would respond “yes” and all pictures except the animals would be eliminated. If the examinee asked a subsequent CS question such as “Is it served cold?” this would qualify as a narrowing question because it further reduces (narrows) the number of pictures under consideration and indicates the problem solver’s continued awareness of the goal to solve the problem with as few questions possible.

Inefficient constraint questions were defined as questions with efficiency scores of less than 50% that were not category-focused, novel, or narrowing questions, but still qualified as constraint questions. This category of questions is necessary because it is possible to ask constraint questions that are not efficient.

Guesses were defined as questions that targeted one picture and solved the problem with a “yes” answer and had little effect on problem solving with a “no” answer.

Strategy group assignment

Marshall and Karow (2008) in a normal study using the RAPS assigned 373 adults to one of three strategy groups: novel, category-focused, or mixed. Strategy group assignment was determined by the type of first question asked on each of the three problems. Subjects were assigned to the novel strategy group if their first question was consistently a novel question. Those assigned to the category-focused group consistently asked a category-limited question first while those assigned to the mixed strategy group

asked first questions that were a mixture of novel, category-limited, and occasionally an inefficient question.

Research with the RAPS

Marshall et al. (2003a) introduced the RAPS as a clinical measure to assess problem solving of hard-to-test patients in clinical settings and as a spinoff of the 20Q task. They provided background on the 20Q task, described differences between the 20Q task and the RAPS, described the materials, administration, and scoring of the RAPS, and illustrated its clinical application with three traumatic brain injured clients. In addition, they reported RAPS results for 70 normal subjects. These results indicated that normal individuals predominantly asked constraint questions, solved problems with an average of five questions, and preferred category-limited questions focusing on semantic categories or features. Question efficiency scores tended to increase from question one-to-question four suggesting that as more pictures are eliminated from consideration, question-asking efficiency increases. In keeping with what might be expected on a test of executive function, normal subjects in this study did not perform perfectly on the RAPS, but reflected a range of performance levels. Performance levels were largely found to be related to two components of problem solving: planning and shifting set. Planning difficulties were reflected in the differences in number of pictures targeted by the first questions asked by the normal subjects. Differences in the ability to shift sets were seen in the variability with which normal subjects asked narrowing questions after the target picture category was known.

Marshall and colleagues (Marshall, Karow, Morelli, Iden, & Dixon, 2003b) compared the verbal problem solving abilities of 21 neurologically intact (NI) and 21 traumatically brain injured adults (TBI). The TBI and NI groups were matched for age, gender, and education. All subjects were administered the RAPS on a single occasion. Results revealed no differences between the groups in the number of questions used to solve problems. NI subjects asked significantly more CS questions and TBI subjects did significantly more guessing, but the NI subjects also did some guessing. Over 70% of the time subjects' guesses reflected the inability to ask a narrowing question. For both groups, guesses tended to be a PC question (e.g., Is it the animal with a long neck?) rather than a frank guess (e.g., Is it the giraffe?). This finding was thought to reflect awareness on the part of the subjects that guessing was not appropriate and an inability to switch to a narrowing question. Question efficiency scores were significantly higher for the NI group, but both groups increased efficiency scores from question one to question four. Both groups reflected a range of performance levels on the RAPS; some of the TBI subjects performed as well as the NI subjects, while some of the NI subjects performed below the level of the TBI subjects.

Marshall, McGurk, Karow, Kairy, and Flashman (2006) used the RAPS to examine problem solving by subjects with and without severe mental illness (SMI). This study involved 47 individuals with SMI participating in an urban outpatient clinic treatment program and an equal number of healthy age and gender matched controls. The SMI subjects solved fewer problems on the RAPS, and when they did solve problems, they did so less efficiently. The two groups differed markedly in the types of questions they used to solve problems. In general, the healthy controls took a systematic, organized,

but not always optimal approach to solving problems. The SMI subjects used some of the problem solving strategies of the healthy controls, but they frequently guessed rather than ask constraint questions, particularly if a category-focused question was insufficient to solve the problem.

Ferguson, Marshall, and Olson (2012) compared performance on the RAPS for two groups of brain injured subjects, individuals with traumatic brain injuries and soldiers with documented blast injuries, and age-matched controls. The control subjects had significantly better scores on the number of questions, percent constraint questions, question-asking efficiency, and integration planning measures than the two brain injured groups. However, the two brain injured groups also differed on some of the RAPS scores, particularly the IPS score where the subjects with blast injuries performed superiorly to those with traumatic head injuries.

Three studies have examined the effects of interactive strategy modeling training (ISMT) of performance on the RAPS (Marshall, et al., 2004; Marshall, Capilouto, & McBride, 2007; Maddy & Marshall, 2012). In these studies, training stimuli were problem solving boards of 32 words representing items from a wide range of semantic categories arranged in a 4 x 8 matrix. Matrices of words permitted the investigators to create a variety of word problems that were similar to, yet also different from those used on the RAPS. During ISMT, the examiner/clinician and examinee/patient alternated roles of tester and problem solver. The examiner/clinician modeled and reinforced the use of constraint questions and asking of efficient questions (those that eliminate more pictures from considerations), and provided negative feedback when the examinee/patient

guessed. The objective of the ISMT was to teach the examinee/patient to think strategically and to ask more efficient questions by heightening knowledge of the consequences of the questions asked.

Marshall and colleagues (2004) tested 20 home-dwelling traumatically brain injured survivors with the RAPS before (Pre-test), immediately after (Post-test), and one month after ISMT (Follow-up). Subjects received ISMT three times per week and solved 12 problems in each training session. The examiner/clinician took the role of problem solving for six problems; roles were reversed for the other six problems. Training continued until the subject was able to ask constraint questions 80% of the time during a training session. Once this level of performance was attained, the Post-test RAPS was administered. Training was then discontinued and a follow-up RAPS was administered one month later. Participants significantly improved in solving problems from the Pre – to the Post- and Follow-up tests. They solved problems with fewer questions, asked more constraint questions, and increased their question-asking efficiency scores.

Two studies have examined the effects of ISMT on the performance of individuals with Alzheimer's disease (AD) on the RAPS. Marshall et al., (2007) used a multiple probe design to assess the effects of ISMT in solving problems on the RAPS with three women diagnosed with early stage AD. ISMT was delivered for 12 sessions and 10 word matrix problems were solved per session. To assess the effects of the training, subjects were given problems from the RAPS to solve before (Baseline), after every other training session (Probes), and at two and four weeks after training ended (Maintenance). All subjects increased their use of constraint questions and reduced the

amount of guessing from baseline. Improvements were maintained at two and four weeks after training. Maddy and Marshall (2012) trained four participants with AD, two men and two women, to solve word matrix problems and assessed training effects using the RAPS. Training sessions occurred for one hour, three times per week. These researchers made slight modifications in their version of ISMT. This involved the use of scripted training procedures in which optimal questions were written on cards. The examiner/clinician handed the cards to the examinee/patient one at a time and the patient read the card. The examiner then crossed out the words eliminated by the question. Questions were scripted in a sequence that would allow the problem to be solved with four questions by reducing the number of pictures from 32 to 16, 16 to eight, and so forth. The RAPS was administered to each of the participants before, midway, immediately after, and one month after treatment. Performance was quantified by determining the number of questions, percent constraint questions, question-asking efficiency, and integration planning scores. All participants improved their ability to solve problems on the RAPS. These improvements reflected asking more constraint questions and doing less guessing after training. The participants did not however, improve their use of the questions used for the scripted training procedure, but rather the use of category-focused questions.

Marshall and colleagues (2007) conducted a cross group study in which the RAPS was used to compare problem solving abilities of neurologically intact (NI) individuals and those with diffuse neurologic involvement (DNI). Participants included two groups of NI subjects, older and younger and three groups of DNI subjects. One group of DNI subjects had relatively acute traumatic brain injuries (ATBI; less than 1 year post onset);

a second had relatively chronic traumatic brain injuries (CTBI; more than 2 years post onset); and a third DNI group included persons with a diagnosis of schizophrenia (SWS). Each group had 20 subjects. All subjects were administered the RAPS on a single occasion. Findings revealed significant differences on all the objective scores of the RAPS, the types of questions asked, and the strategies used to solve the problems. These differences were seen between the NI groups and two of the three DNI groups, ATBI and SWS, but the CTBI group performed similarly to the NI groups. Findings suggested problem solving on the RAPS by subjects with and without DNI could be distinguished using selected components of Scholnick and Friedman's (1993) developmental theory of planning, specifically decision to plan, strategy choice, strategy execution, and monitoring the effects of prior actions.

In 2008, Marshall and Karow published an update on the RAPS. This expansion of their earlier normative study (Marshall et al., 2003a) included 373 adult subjects (213 women and 160 men) ranging in age from 18 to 87 years of age ($M = 41.8$ years; $SD = 18$) with 8 to 20 years of education ($M = 14.5$ years; $SD = 2.2$). Some subjects were administered the RAPS once. However, 203 subjects were tested twice to assess test-retest stability. A smaller group of 74 subjects were tested twice, once with the examiner crossing out the pictures eliminated by the questions and once not doing this. Test results were further analyzed to examine the impact of demographic, psychometric, and other factors on performance on the RAPS. In addition, subjects were assigned to strategy groups, novel, category-focused, or mixed based on the types of first questions asked. Findings revealed good test-retest stability for the RAPS, no differences in test scores for subjects tested with and without the cross out procedure, and a modestly significant

correlation between the efficiency score for the RAPS and scores on a non-verbal measure of problem solving, the Raven Colored Progressive Matrices (RCPM; Raven, Court, & Raven, 1984; Raven, 2000). Participants in the novel strategy group were found to perform significantly better on the RAPS than those in the other two strategy groups.

Recently, Marshall and Karow (2013) developed a rubric to score the RAPS. A rubric is a scoring tool that divides a task into component elements and provides a description of levels of performance for each element. The components of the six-element rubric for the RAPS included: planning, strategy choice, strategy execution, awareness of category size, use of narrowing questions, and number of questions. To score the RAPS with the rubric, each element was scored 2, 1, or 0. The researchers compared sensitivity (probability of identifying abnormal functioning in an impaired individual) and specificity (probability of identifying normal functioning in a healthy individual with the test in question) for the RAPS for groups of neurologically intact (NI) and neurologically compromised (NC) subjects matched for age, gender, and education using a rubric scoring system and the traditional RAPS scores, number of questions, % constraint questions, and question asking efficiency. Rubric scores successfully identified 87% of the NC subjects. Traditional scores identified far fewer subjects. Specificity of the RAPS was not improved with rubric scoring. Administration and scoring time for the RAPS was reduced with the use of the rubric. The authors concluded that the rubric scoring method balances clinical observation and measurement and may help time-conscious clinicians develop more efficient ways to quantify performance on multi-component executive function tasks such as the RAPS.

To summarize, the RAPS has been used successfully to examine the problem solving abilities of normal adult subjects across the life span, to compare problem solving of neurologically intact and neurologically compromised adults, and to assess the effects of different problem solving interventions with adults. The game-like format of the RAPS is appealing and well tolerated by adults, particularly adults with brain damage who may be hard to test, particularly in clinical settings. While the RAPS is a deceptively simple test, it does qualify as a test of executive function. Normal subjects do not perform flawlessly on the RAPS, but reflect a range of performance levels on the test. This has been the case in normative studies using the RAPS (Marshall et al., 2003a; Marshall & Karow, 2008) and in those studies comparing performance of normal and brain damaged subjects on the RAPS.

Chapter 3: Methods

This cross sectional study examined the performance of typically developing children on the Rapid Assessment of Problem Solving test (RAPS; Marshall et al., 2003a). The study was approved by the University of Kentucky Institutional Review Board (IRB #14-0351-X1B).

Participants

Two-hundred-seventy-five typically developing children residing in the State of Kentucky (127 males and 148 females) participated in the study. Most of the participants were Caucasian with less than 5% of the sample representing African American, Hispanic, Asian or other cultures. Participants ranged in age from 5.1 to 17.11 years ($M = 11.5$ years; $SD = .5$ years). No participant was considered to have or reported any sensorimotor (i.e., visual, hearing, or physical impairments) or cognitive-communicative impairments (i.e., speech/language, cognitive, or learning problems) that would interfere with them being assessed with the RAPS.

Screening Tasks

Before being administered the RAPS, each child passed two screening tests. The first, a picture matching task, required the child to orally name or to identify 30 of the 126 pictures on the RAPS. This was done to ensure that stimuli from the RAPS were familiar to children as the RAPS has only been used with adults. The 30 pictures (see Figure 3.1) were selected randomly by choosing one or two pictures from each of the 18 picture categories of the RAPS. The same set of pictures was used for each child. Naming

responses were scored right or wrong. Alternative responses such as descriptive responses (e.g., yellow flower for “zinnia”), categorical names (e.g., tool for “wrench”), semantically related responses (e.g., cone for “ice cream cone”) and other responses indicating the child recognized the picture were considered as correct. If a picture was misidentified or not named by the child, picture recognition was assessed with a word-to-picture matching task. This was done by presenting the missed picture in an array with three other pictures and asking the child to point to the misidentified picture (e.g., “point to the ship”). The child passed this screening task if he or she named or recognized 80% (24/30) of the pictures.

The second screening task was used to ensure the child was able to ask yes/no questions. For this task, two 12-picture problem solving boards, similar to the larger 32-item boards of the RAPS, were constructed. None of these pictures were from the RAPS. Six pictures were colored and six were black and white. The pictures represented three categories (e.g., hats, dogs, and games). Each board had one category of 6, 4, and 2 pictures respectively and no two pictures from the same category appeared in adjacent positions (see Figure 3.1). For this task, the child was told, “I am thinking of one of these pictures. I want to hear you ask me some questions that I can answer “yes” or “no” to try to figure out the picture I’m thinking of.” If the child asked a yes/no question, it was answered “yes” or “no,” then the child was encouraged to ask another question. If the child did not ask a yes/no question, further instruction was given, e.g., “You need to ask me a question I can answer yes or no, try it again.” To pass this screening test, the child needed to ask two consecutive yes/no questions. All 275 children passed both screening tests.

Figure 3.1 RAPS Screening Protocol for Children

Rapid Assessment of Problem Solving—Screening

Picture Recognition & Oral Naming

Instructions: *“I am going to show you some picture cards. You tell me what the picture is called. You get to keep each picture you get right and I keep any other ones. If you have more than me at the end, then you win!”*

For any pictures the child cannot name, place the unknown picture in an array of four and ask “Can you show me the _____?” Make note of any pictures that were correctly identified with a check mark (✓) in the +/- column.

	Colored Pictures	+/-	Black & White Pictures	+/-	Comments
01	horse		16	elephant	
02	football		17	soccer ball	
03	top		18	puzzle	
04	guitar		19	trumpet	
05	hotdog		20	table	
06	zinnia		21	tree	
07	ladybug		22	spider	
08	blender		23	toaster	
09	shot		24	eye	
10	parrot		25	cardinal	
11	wrench		26	saw	
12	lawn mower		27	water hose	
13	popsicle		28	octopus	
14	shirt		29	pants	
15	ship		30	car	
	Subtotal		Total*		

*Score of 24 (80%) or higher passes.

Yes/No Question Formulation

Instructions: *“I want to hear you ask me some questions that I can answer yes or no. Look at this board (point to Screening board A or B) and try to figure out which picture I’m thinking of by asking me some questions that I can answer yes or no. Go ahead and ask me a question when you are ready.”*

The child passes this portion if he/she can independently formulate 2 to 4 questions that can be answered yes or no. Write their questions below.

Question 1: _____

Question 2: _____

Question 3: _____

Question 4: _____



Procedures

Graduate students in a research methods class recruited the child participants, performed the screening tests, and administered the RAPS. Before doing any of these tasks, the students completed two modules of the Collaborative Institutional Training Initiative (CITI) required of entry level investigators and participated in two 90-minute training sessions.

Training

For the first training session, the students read and critiqued two peer-reviewed articles on problem solving, one related to adults (Marshall et al., 2003a) and one related to children (Winsler & Naglieri, 2003). During the training session itself, the investigator discussed strategies for critical analysis of research based on a paper by Locke, Silverman, and Spiruso (2010). The students also worked in small groups to determine how they could apply this information to the article they had read and to the literature review process in general.

The materials, administration, and scoring of the RAPS (Marshall et al., 2003a; Marshall & Karow, 2008) were introduced in the second training session. The RAPS was presented as a novel problem solving test that had been used successfully with adults but not children. Students were trained to administer and score the RAPS. They practiced giving the test to each other and recording the necessary information on the recording forms for the RAPS to score the test. This “hands on” training was carried out under the direction and supervision of the investigator. The investigator gave the students feedback on their administration of the test, recording of responses, and scoring, and answered

students' questions as they came up. At the end of the training session, the students were asked to participate as testers in a study examining the performance of typically developing children on the RAPS. They were also informed that their participation was voluntary and if they chose not to participate, another assignment would be provided. No student declined to participate. Finally, student volunteers were asked to recruit children between the ages of 7 and 17.11 years of age, from as diverse backgrounds as possible, screen children for inclusion in the study, administer the RAPS to 4-8 children individually, and to videotape two of their tests.

Administration of the RAPS to Children

The 275 child participants were seen in the summer of 2013 and fall of 2014. Student testers met with children individually in convenient locations (e.g., homes, schools, and churches) in a quiet room. After ensuring that the child had no sensorimotor or cognitive-communication problems that would exclude them from the study, the students gave the screening tasks and administered the RAPS test.

All 275 children were administered the RAPS individually in single sessions. Twenty-two children (12 male, 10 female; mean age 12 years) were administered the RAPS a second time to assess test-retest reliability data on use of the RAPS with children. For these children, the second test followed the first after a short break and included different problem solving boards and target pictures.

When administering the RAPS, the student testers followed guidelines proposed by Marshall et al., (2003b) and described in Chapter 2. However, since some of the study

participants were young children, an apriori decision was made to permit the testers to repeat the instructions “as needed.” As needed was defined as any indication of confusion by the child expressed verbally, via facial expression, body language, signs of inattention (i.e., looking away, responding to environmental noise), or a direct request from the child to repeat the instructions. To further standardize test administration, one third of graduate student testers used problem solving boards 1-3, 4-6, or 7-9 respectively. The target picture for each board was stipulated in advance by the investigator and different pictures were used for each board. To begin the test, the student tester placed the first problem solving board on the table in front of the child and gave the following instructions:

We are going to play a question-asking game. I am thinking of one of these pictures (tester gestures to the pictures) and your job is to figure out which one it is. The way to do this is to ask me questions that I can answer “yes” or “no.” You can ask me any question you want so long as I can answer it “yes” or “no.” Try to ask as *few* questions as possible. When you are ready, go ahead and ask your first question.

After each question, the tester (a) answered “yes” or “no,” (b) covered the pictures eliminated by the question with small blank cards, (c) wrote the question down the recording form, (d) designated the number of pictures targeted and eliminated by the question, (e) designated if the question was a constraint question or a guess, and (f) classified each question as a novel, category-limited, inefficient, narrowing constraint-seeking question or as a guess. The child then asked his or her next question. This process continued until the participant’s questions had reduced the number of pictures on the 32-item problem solving board to two or three or the target picture was identified by name.

The second and third problems were then presented for solving in a similar manner. Participants were permitted to ask questions until each problem was solved or until they had asked 10 questions. If the participant had not solved the problem by question 10, the tester defaulted to a “yes” answer and terminated the problem solving effort. Because a problem is considered to be solved on the RAPS when the examinee’s questions have reduced the number of pictures to two or three, there were occasions when the target picture was not identified by name. In these instances, the child was allowed to ask an additional question or two to obtain “closure” but the extra questions were not included in the total question count. After the third problem was solved and the testing with the RAPS had been completed, the graduate student tester concluded the session and provided the participant with general praise and encouragement.

Follow-up activities

After completing the testing, the graduate students met with the investigator for two follow-up sessions. In the first session, they worked in small groups under the guidance of the investigator to ensure the information entered on recording forms was correct. They performed an initial review of all questions asked by the child participants to determine the following:

1. Did the examiner accurately record the answer to the question as yes or no?
2. Did the examiner accurately record the number of pictures available when the question was asked, how many pictures were targeted by the question, and how many were eliminated by the question?

3. Did the examiner accurately classify the question as a novel, category-limited, narrowing, or inefficient constraint-seeking question or as a guess?

After an initial review of the 275 tests, it was necessary to exclude the tests of 46 participants. Seven tests were confounded by unresolvable recording errors. Thirteen tests were excluded because they were mistakenly administered to children under the age of seven. Finally, 26 tests had to be excluded because the participant's first question identified the target picture with a guess (e.g., Is it a doll?) or asked a lucky constraint question that identified the target picture's category (e.g., Is it a toy?). These tests were excluded because the individual solved a problem with too few questions, which made it impossible to calculate all scores of the RAPS.

Following exclusion of the 46 tests, the graduate students, again under direction of the investigator, did the necessary calculations needed to score the 229 remaining tests and entered the data on the recording forms. These calculations involved (a) counting the number of questions needed to solve each problem, (b) counting the number of constraint-seeking questions needed to solve each problem, and (c) calculating question-efficiency scores for the first four questions for each problem. The graduate students also classified all questions asked on each problem as a novel, category-limited, or inefficient constraint question or as a guess, and tallied the numbers for each type of question.

At the second follow-up session, pairs of student testers watched their respective videotaped administrations of the RAPS. This served as a quasi-procedural reliability check to ensure the RAPS was administered properly. Each member of the pair was given a 12-point checklist (see Table 3.1) delineating steps to be followed when giving the

RAPS. Each graduate student tester used the check list to determine if their partner had followed all the steps. Fifty-one videotapes were reviewed resulting in a total of 612 (51 x 12) checks. Review of the quasi-procedural reliability check revealed that there were only 11 occasions where a step in the checklist was omitted or not followed appropriately.

Table 3.1 Procedural Review Checklist for Students

Checklist Item	Yes	No	N/A
1 Gave correct and complete instructions.			
2 Responded appropriately (yes/no) to questions asked.			
3 Covered eliminated pictures.			
4 Redirected appropriately when a question could not be answered yes/no.			
5 Followed up appropriately when tester was unsure of pictures targeted by question.			
6 After 10 questions, appropriately brings task to a close.			
7 Records question on protocol verbatim.			
8 Categorizes question.			
9 Records number of pictures available.			
10 Records number of pictures targeted.			
11 Records number of pictures eliminated.			
12 Ends task when only 2 or 3 available pictures remain.			

Scoring

Scoring of the RAPS was done by the investigator and a trained research assistant (RA). This entailed calculating and averaging scores for each problem to derive mean scores for each participant. To facilitate explanation of how the scoring was done, a completed RAPS test for a participant is provided in Figure 3.2. The top-most segment

(RAPS Scoring Summary) of Figure 3.2 shows all the scores computed for the RAPS for the participant. The bottom three segments (recording form + problem boards) show that the participant solved problems associated with boards 1, 2, and 3 and that the target pictures for these boards were the items cookie, shirt, and plane respectively. The Figure 3.2 recording forms show (1) the questions asked by the participant for each problem, (2) whether the question was answered yes (Y) or no (N), (3) if the question was designated as a constraint-seeking question (C) or a guess (G), (4) the number pictures on the problem solving board when the question was asked (pictures considered), (5) the number pictures targeted by the question (pictures targeted), and (6) the number of pictures eliminated by the question (pictures eliminated). Finally, Figure 3.2 problem boards indicate which pictures on the problem solving board were eliminated by the child's first (blue), second (green), third (purple), fourth (yellow), and fifth (pink) questions.

Mean number of questions (M#Q). This score was obtained by averaging the number of questions needed to solve each problem. Figure 3.2 shows that the participant used four, four, and five questions to solve problems 1, 2, and 3 respectively. This resulted in a M#Q score of 4.33 ($4 + 4 + 5 = 13/3$).

Percent constraint-seeking questions (%CS). This score represents the percentage of CS questions asked in solving the three problems. It is obtained by dividing the number of CS questions by the total number of questions (CS questions + Guesses). Figure 3.2 shows that the participant used a total of 13 questions to solve the three

problems and that 11 of 13 questions were CS questions. The %CS score for this participant as shown in Figure 3.2 was .85 (11/13).

Figure 3.2 Sample Scoring for the RAPS

RAPS Scoring Summary

ID Number	Types of Questions	Total # of Questions	Total # of Constraint Seeking Questions	Integration Planning Score	Problem Solving Efficiency Score
027	Category Limited	3	4	5	.61
Problem Number 1	Constraints				
	Narrowing	1			
	Novel Inefficient Constraint				
	Guesses				
	Frank Guess				
	Pseudo-Constraint				
Comments/Notes					
<hr/>					
Types of Questions	Total # of Questions	Total # of Constraint Seeking Questions	Integration Planning Score	Problem Solving Efficiency Score	
Category Limited	3	4	4	4	.56
Constraints					
Narrowing	1				
Novel Inefficient Constraint					
Guesses					
Frank Guess					
	Pseudo-Constraint				
Comments/Notes					
<hr/>					
Types of Questions	Total # of Questions	Total # of Constraint Seeking Questions	Integration Planning Score	Problem Solving Efficiency Score	
Category Limited	2	5	3	6	.56
Constraints					
Narrowing					
Novel Inefficient Constraint	1				
Guesses					
Frank Guess	2				
	Pseudo-Constraint				
Comments/Notes					
<hr/>					
Combined Total Types of Questions for All Problems	Total # of Questions*	Total # of Constraint Seeking Questions*	Integration Planning Score*	Overall RAPS Efficiency Score*	
Category Limited	8	13 / 4.33	11 / .85	5	.58
Constraints					
Narrowing	1	*Total for All Problems		*Mean for All Problems	
Novel Inefficient Constraint	2				
Guesses					
Frank Guess	2				
	Pseudo-Constraint				
Question Sequence Efficiency Score*					
Problem Number Total	QE1	.50	.38	1.00	.62
	QE2	.50	.46	.50	.49
	QE3	1.00	.40	.50	.63
	QE4	.44	1.00	.22	.55

Copyright © Janice Carter Smith 2012

Figure 3.2 Sample Scoring for the RAPS (continued)

RAPS

Robert C. Marshall & Colleen M. Karow


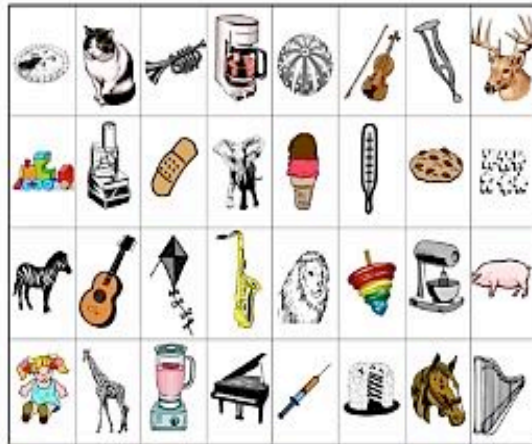
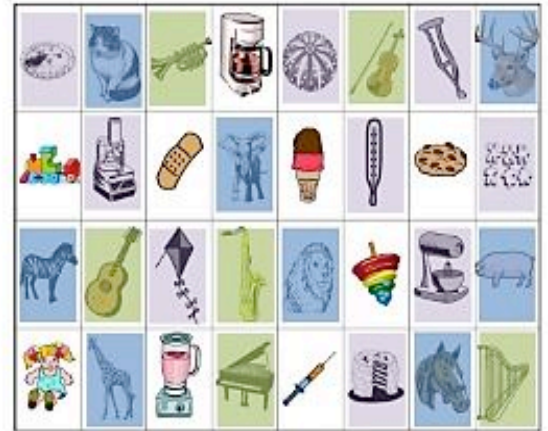
Board #1 Recording Form						
Patient Name: <u>027</u>			Date of Testing: 			
D.O.B:		(Age) <u>9:8</u>	Target: <u>cookie</u>	Problem: <input checked="" type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3		
Question Asked (write out)	Y/N	IC/Constraint / IG/guess	# Pictures Considered	# Pictures Targeted	# Pictures Eliminated	Question Efficiency
1. <u>Is it an animal?</u>	N	C (catlim)	32	8	8	.5
2. <u>Is it an instrument?</u>	N	C (catlim)	24	6	6	.5
3. <u>Is it in color?</u>	Y	C (Novel)	18	9	9	1.0
4. <u>Is it something you eat?</u>	Y	C (catlim)	9	2	7	.44
5. <u>Is it the cookie?</u>	Y	G (FG)	2			
6.						
7.						
8.						
9.						
10.						
				Scores:		
				1. Total Number of Questions		4
				2. Total Number of Constraint Seeking Questions:		4
				3. Integration Planning Score:		5
				4. Questions Asking Efficiency Score:		
				Q1	Q2	Q3
				.5	.5	1.0
						.44
Comments:						Mean = .61
Tally Types of Questions						
Constraints				Guesses		
<i>Category-Limited</i>	<i>Narrowing</i>	<i>Novel</i>	<i>Inefficient Constraint</i>	<i>Frank Guess</i>	<i>Pseudo-constraint</i>	
3		1				
# ITEMS	CATEGORY	BLACK & WHITE		COLOR		
4	Kitchen Items	Food Processor, Mixer		Coffee Maker, Blender		
4	Medical Equipment	Thermometer, Crutch		Shot, Band-Aid		
4	Desserts	Pie, Cake		Cookie, Ice Cream Cone		
6	Musical Instruments	Trumpet, Harp, Piano		Violin, Guitar, Saxophone		
6	Toys	Ball, Puzzle, Kite		Train, Top, Doll		
8	Animals	Elephant, Lion, Zebra, Giraffe		Deer, Horse, Pig, Cat		

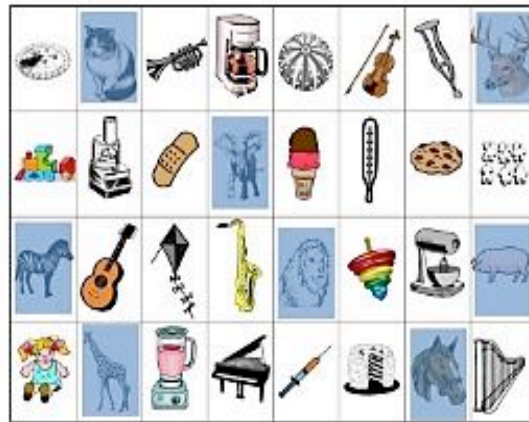
Figure 3.2 Sample Scoring for the RAPS (continued)



RAPS Board 1—Before



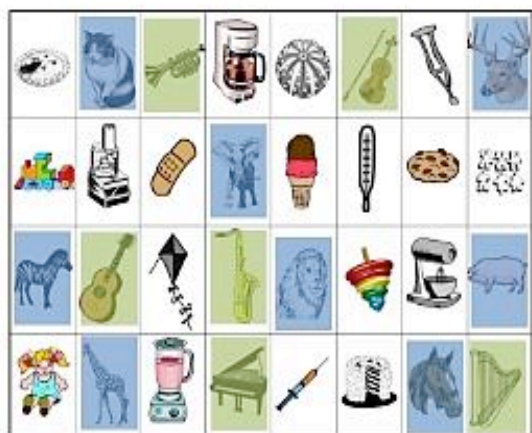
Question 3—Is it in color? (Y)



Question 1—Is it an animal? (N)



Question 4—Is it something you eat?
(Y) *Problem considered "solved" at this point.



Question 2—Is it an instrument? (N)



Question 5—Is it the cookie? (Y)

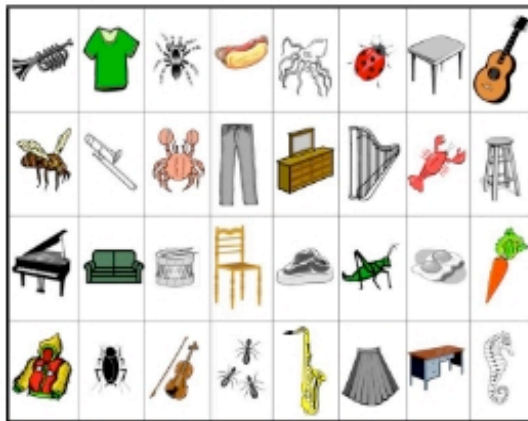
Figure 3.2 Sample Scoring for the RAPS (continued)

RAPS

Robert C. Marshall & Colleen M. Karow

Board #2 Recording Form						
Patient Name: <u>027</u>			Date of Testing:			
D.O.B:		(Age) <u>9:8</u>	Target: <u>shirt</u>	Problem: <input type="checkbox"/> 1 <input checked="" type="checkbox"/> 2 <input type="checkbox"/> 3		
Question Asked (write out)	Y N	[C]onstraint /[G]uess	# Pictures Considered	# Pictures Targeted	# Pictures Eliminated	Question Efficiency
1. Is it a bug?	N	C (catlim)	32	6	6	.38
* 2. Is it in a house?	N	C (catlim)	26	6	6	.46
3. Is it something you wear?	Y	C (catlim)	20	4	16	.4
4. Does it have green?	Y	C (Nar)	4	2	2	1.0
5. Is it the jacket?	N	G (FG)	2			
6. Is it the shirt?	Y	G (FG)	—			
7.						
8.						
9.						
10.						
				Scores:		
				1. Total Number of Questions		4
				2. Total Number of Constraint Seeking Questions:		4
				3. Integration Planning Score:		4
				4. Questions Asking Efficiency Score:		
				Q1	Q2	Q3
				.38	.46	.4
				Q4	1.0	
Comments:						Mean = .56
* He indicated furniture here.						
Tally Types of Questions						
Constraints				Guesses		
<i>Category-Limited</i>	<i>Narrowing</i>	<i>Novel</i>	<i>Inefficient Constraint</i>	<i>Frank Guess</i>	<i>Pseudo-constraint</i>	
3	1					
# ITEMS	CATEGORY	BLACK & WHITE		COLOR		
4	Clothing	Pants, Skirt		(Shirt) Jacket		
4	Food	Steak, Eggs		Hot dog, Carrot		
4	Sea Creatures	Seahorse, Octopus		Crab, Lobster		
6	Insects	Spider, Ant, Roach		Bee, Lady bug, Grasshopper		
6	Furniture	Table, Couch, Stool		Desk, Dresser, Chair		
8	Musical Instruments	Harp, Trombone, Piano, Trumpet		Guitar, Saxophone, Violin, Drum		

Figure 3.2 Sample Scoring for the RAPS (continued)



RAPS Board 2—Before



Question 3—Is it something you wear? (Y)

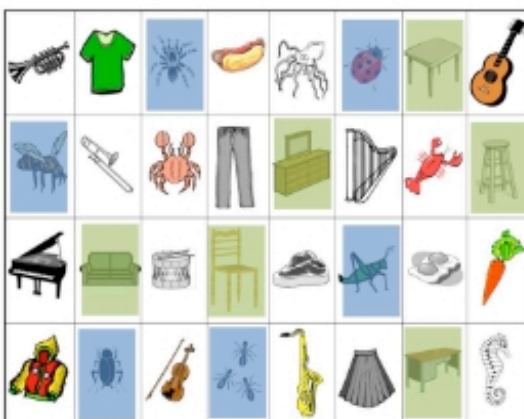


Question 1—Is it a bug? (N)



Question 4—Does it have green? (Y)

*Problem considered solved at this point.



Question 2—Is it in a house? * (N)

*child indicates furniture



Question 5—Is it the jacket? (N)

Figure 3.2 Sample Scoring for the RAPS (continued)

RAPS

Robert C. Marshall & Colleen M. Karow


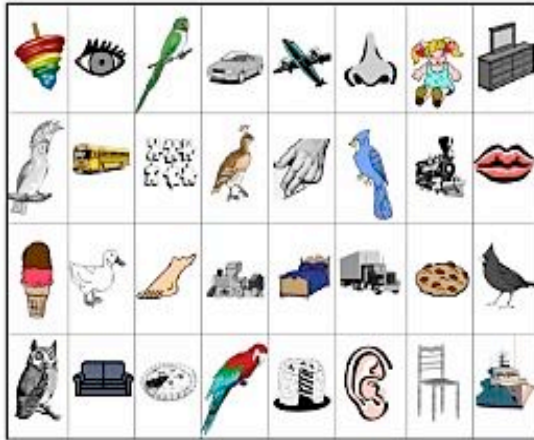
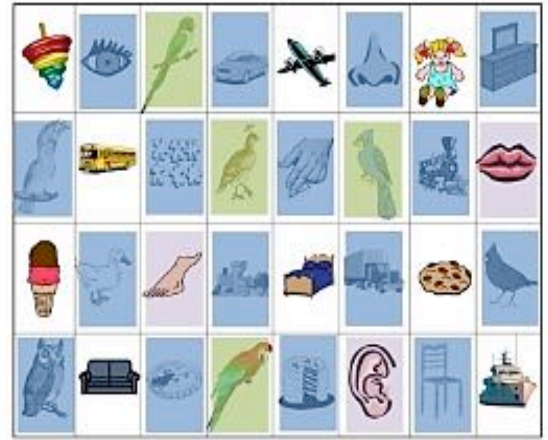
Board #3 Recording Form						
Patient Name: <u>027</u>			Date of Testing: 			
D.O.B:		(Age) <u>9:8</u>	Target: <u>plane</u>	Problem: <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input checked="" type="checkbox"/> 3		
Question Asked (write out)	Y/N	(C)onstraint / (G)uess	# Pictures Considered	# Pictures Targeted	# Pictures Eliminated	Question Efficiency
1. Is it in color?	Y	C (Novel)	32	16	16	1.0
2. Is it a bird?	N	C (catlim)	16	4	4	.5
3. Is it on my body?	N	C (catlim)	12	3	3	.5
4. Is it the bus?	N	G (FG)	9	1	1	.22
5. Is it the plane?	Y	G (FG)	8			
6.				Scores:		5
7.				1. Total Number of Questions		
8.				2. Total Number of Constraint Seeking Questions:		3
9.				3. Integration Planning Score:		6
10.				4. Questions Asking Efficiency Score:		
				Q1	Q2	Q3
				1.0	.5	.5
						.22
Comments:						Mean = .56
Tally Types of Questions						
Constraints				Guesses		
<i>Category-Limited</i>	<i>Narrowing</i>	<i>Novel</i>	<i>Inefficient Constraint</i>	<i>Frank Guess</i>	<i>Pseudo-constraint</i>	
2		1		2		
# ITEMS	CATEGORY	BLACK & WHITE		COLOR		
4	Toys	Train, Puzzle		Top, Doll		
4	Furniture	Dresser, Chair		Bed, Couch		
4	Desserts	Pie, Cake		Ice cream, Cookie		
6	Body Parts	Eye, Nose, Hand		Ear, Lips, Foot		
6	Transportation	Car, Truck, Train		Ship, <u>Plane</u> , Bus		
8	Birds	Cockatoo, Duck, Owl, Cardinal		Blue jay, Parrot, Parakeet, Quail		

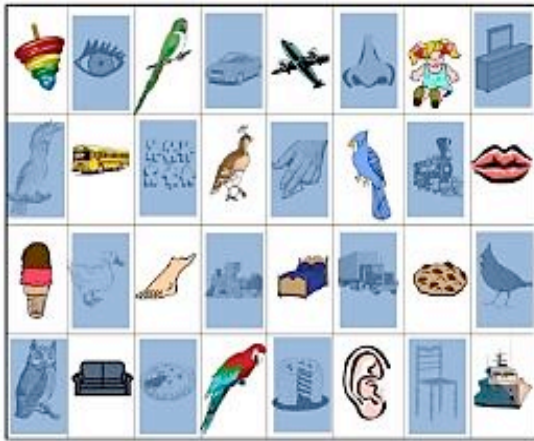
Figure 3.2 Sample Scoring for the RAPS (continued)



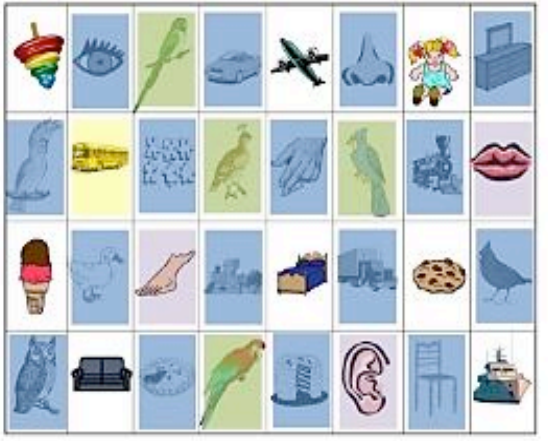
RAPS Board 3—Before



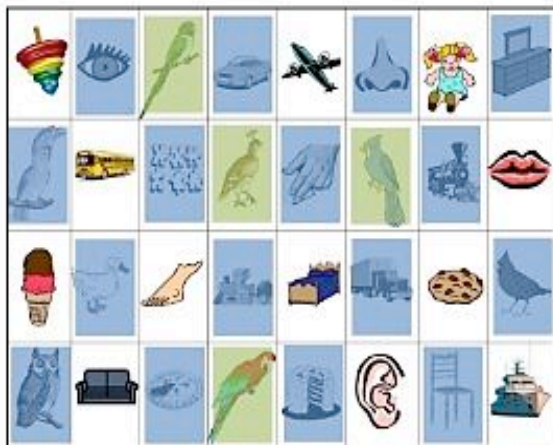
Question 3—Is it on my body? (N)



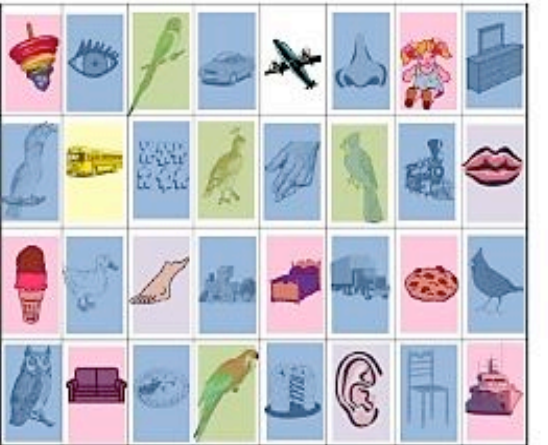
Question 1—Is it in color? (Y)



Question 4—Is it the bus? (N)



Question 2—Is it a bird? (N)



Question 5—Is it the plane? (Y)

Mean integration planning score (MIPS). An integration planning score (IPS) was determined for the first question of each problem. To do this, a score ranging from 1 to 6 was assigned to the question based on the number of pictures targeted: 1 = one picture; 2 = two or three pictures; 3 = four or five pictures; 4 = six or seven pictures; 5 = eight pictures; 6 = nine or more pictures. The IPS scores on the recording forms for each problem indicate that the first questions of the participant targeted 8, 6, and 16 pictures for problems 1, 2, and 3 respectively, and that these resulted in IPS scores of 5, 4, and 6. The RAPS scoring summary show that when the IPS scores were averaged, the resulting MIPS score was 5 ($5 + 4 + 6/3$).

Efficiency scores. Question-efficiency (QE) scores were calculated for the first four questions for each problem (12 scores total). A question's efficiency was determined by dividing the smaller of two numerators, pictures targeted or pictures eliminated, by the number of pictures available when the question was asked, and multiplying the result by 2. Figure 3.2 shows the 12 QE scores for the participant, four for each problem. The RAPS Scoring summary shows that these QE scores were averaged in different ways to obtain three different efficiency scores for the participant, problem solving efficiency (PSE), overall RAPS efficiency (ORE), and question sequence efficiency (QSE).

Problem solving efficiency (PSE). PSE scores were obtained by averaging the four QE scores for each problem. This resulted in three PSE scores per participant, one per problem. Figure 3.2 shows that the PSE scores for this participant were .61 ($.50 + .50 + 1.0 + .44 = 2.44/4$), .56 ($.38 + .46 + .40 + 1.0 = 2.24/4$) and .56 ($.62 + .40 + .63 + .55 = 2.20/4$) for problems 1, 2, and 3 respectively.

Overall RAPS efficiency (ORE). The ORE score was obtained by averaging the three PSE mean scores. Figure 3.2 shows that the participant's overall efficiency score was .58 ($.61 + .56 + .56 = 1.73/3$).

Question sequence efficiency (QSE). Figure 3.2 shows that four QSE scores were obtained for the participant. The scores were derived by averaging the QE scores for the first question of each problem, the second question for each problem and so forth. Figure 3.2 shows that the QSE scores for this participant were .62 ($.25 + .37 + .50 = 1.12/3$), .49 ($.37 + .42 + .50 = 1.29/3$), .63 ($.67 + .67 + .78 = 2.12/3$), and .55 ($1.00 + .75 + .25 = 2.00/3$) for questions 1, 2, 3, and 4 respectively.

Strategy Group Assignment

Similar to what was done in the Marshall and Karow (2008) study with adults, children were assigned to novel or category-focused strategy groups if their first questions were always novel or category-focused questions respectively. In this study, however, children placed in the mixed strategy group differed from the adults in the Marshall and Karow study because they asked a mixture of all types of questions and made some guesses. In contrast, the adults in the Marshall and Karow study mixed strategy group just asked novel and category-limited questions and an occasional inefficient constraint question. It was also necessary to create a fourth group designated for the child subjects. This was deemed to be a "no strategy" group and it included children who only guessed or on their first questions.

Grouping of Participants by Age

To analyze the test results from the RAPS, the participants were separated into three age groupings, young (7-9 years; N = 73), early-adolescent (10-13 years; N = 79), and adolescent (14-17 years; N = 77). Table 3.2 provides demographic information on the three groups.

Table 3.2 Characteristics of the participants included in the data analysis

	All	Young	Early Adolescent	Adolescent
	7.0-17.11	7:0-9:11	10:0-13:11	14:0-17:11
	(N=229)	(N=73)	(N=79)	(N=77)
Age (years)				
Mean	11.70	7.92	11.48	15.52
SD	3.25	.82	1.08	1.18
Min, Max	7.0, 17.11	7.0, 9.11	10.0, 13.11	14.0, 17.11
Gender				
Male	109	32	38	39
Female	120	41	41	38

Chapter 4: Results

All child participants solved three problems and completed the RAPS without difficulty. Results of the study are presented in three sections. Section One provides information on reliability of question classification, computation of scores, and test-retest stability for the RAPS with children. Section Two summarizes parametric and non-parametric analyses to examine gender differences as well as within and between group differences for the objective scores of the RAPS using an alpha level of .05 for all comparisons. Section Three summarizes the results of a series of descriptive analyses related to the types of questions asked by the children in solving problems on the RAPS, when the questions were asked in the question-asking sequences leading to solving the problem, and components of problem solving considered to be important to performance on the RAPS.

SECTION 1

Reliability and Test Stability

Classification of Questions

The 229 participants asked a total of 3367 questions. Each question was classified as a novel, category-limited, inefficient constraint, narrowing question or as a guess using explicit definitions from Marshall and Karow (2008). Question classifications, initially made by the graduate student testers, were checked for accuracy by the investigator or RA. Only 93 questions (.03%) were found to be misclassified. These errors were

corrected by rechecking the question classifications with the explicit definitions from Marshall and Karow (2008).

Question-Efficiency Scores

Question efficiency (QE) scores were calculated by the graduate students for the first four questions of each problem. In most cases, 12 QE scores were calculated (4 scores x 3 problems) per participant. Approximately 20% (549) of the QE scores were re-calculated and checked for accuracy several weeks after completion of the initial scoring by the RA or the investigator. Only two of the 549 scores were found to be in error.

Problem Solving Efficiency

QE scores for each problem were summed and averaged to provide a problem solving efficiency score (PSE) for each problem. There were three PSE scores per participant and 687 PSE for all participants. Twenty percent (142) of these scores were re-calculated by the investigator several weeks after the initial calculations. Only five PSE scores were found to be in error.

Overall RAPS Efficiency Scores

The three PSE scores were averaged to obtain a single overall RAPS efficiency score (ORE) for each participant by the RA. Approximately 20% (24) of the ORE scores were re-calculated by the investigator and only one score was found to be in error.

Test-Retest Stability

To assess test-retest stability on the RAPS with children, 22 children were administered the test a second time approximately one hour after their first test. Pair-wise comparisons of scores on the first and second tests were made for the following: mean total questions, mean questions per problem solved, percent constraint questions, RAPS integration planning score, and overall RAPS efficiency scores. Mean scores for the 22 children for the first and second test for all of these measures reflected minimal differences and none of the comparisons between the mean scores for the first and second tests indicated statistically significant differences with $p < .05$.

SECTION 2

Gender Differences

Though no gender differences on the RAPS have been reported for adult normal subjects (Marshall et al., 2003a; Marshall & Karow, 2008), and because this study focused on children, gender effects were examined for three primary scores for the test: mean number of questions (M#Q), percent constraint seeking questions (%CS), and problem solving efficiency (PSE). Separate two-way ANOVAs were conducted for each of these scores. For all three analyses evidence was not strong enough to conclude that differences between groups are different for males and females. There was no evidence of significant group by gender interactions.

Within and Between Group Differences

Do typically developing young, early-adolescent, and adolescent children differ in the number of questions (M#Q) needed to solve problems on the RAPS?

A problem was considered solved when the participant's questions had reduced the number of pictures under consideration to two or three. Question counts were made to calculate the mean number of questions needed to solve the three problems and the mean number of questions per problem solved. Table 4.1 shows that on the average, younger, early-adolescent, and adolescent participants used 15.6 (SD = 3.9), 14.1 (SD = 2.78), and 14.4 (SD = 2.67) questions to solve the three problems. Table 4.1 also shows the mean number of questions per problem solved for the young, early-adolescent, and adolescent participants were 5.2 (SD = 1.3), 4.7 (SD = 0.9), and 4.8 (SD = 0.9) respectively.

ANOVA results revealed that the young group asked more total questions and more questions per problem than the early-adolescent and adolescent groups with a statistical significance of $p = 0.03$.

Do typically developing young, early-adolescent, and adolescent children differ in the percentage of constraint-seeking questions (%CS) used to solve problems on the RAPS?

Two forms of yes/no questions are acceptable on the RAPS: constraint-seeking (CS) and guesses. The percentage of CS questions (%CS) was determined for each participant by summing the number of CS questions for the entire test and dividing by the total number of questions. Table 4.2 shows the mean percentage of constraint questions for each group. On the average, young, early-adolescent, and adolescent participants asked constraint questions 62% (SD = 0.26), 79% (SD = 0.22), and 80% (SD = 0.18) of the time

respectively. ANOVA results revealed statistical differences between the groups ($p < .0001$) and that young subjects asked fewer CS questions than the early-adolescent ($p < .0001$) and adolescent groups ($p < .0001$). Kruskal-Wallis tests also revealed differences between the groups at similar levels.

Table 4.1. Total number of questions needed to solve 3 RAPS problems and mean number of questions (M#Q) per problem solved.

	<i>M</i>	Total Questions			M#Q		
		<i>M</i>	Median	SD	<i>M</i>	Median	SD
Young 7:0-9:11	7:9	15.6	15	3.9	5.2	5.0	1.3
Early Adolescent 10:0-13:11	11:5	14.1	14	2.78	4.7	4.7	0.9
Adolescent 14:0-17:11	15:5	14.4	14	2.67	4.8	4.7	0.9

Table 4.2. Mean percentage of constraint questions for each age group.

	<i>M</i>	Median	SD
Young	0.62	0.64	0.26
Early Adolescent	0.79	0.86	0.22
Adolescent	0.80	0.85	0.18

Do typically developing young, early adolescent, and adolescent children exhibit differences in mean integration planning scores (MIPS) on the RAPS?

Planning is an integral component of problem solving (Lezak et al., 2004; Scholnick & Friedman, 1993). Marshall et al., (2003a; 2006) have suggested that developing a plan to solve problems on the RAPS with as few questions as possible requires the examinee to integrate information from the problem solving board in order to ask an optimal first question. For example, one might determine that half of the pictures are in color and half black and white, that pictures belong to different semantic categories, some larger than others, and that pictures are arranged in rows and columns. The integration planning score, developed exclusively for the RAPS (Marshall et al., 2006) was intended to capture “goodness” or “poorness” of the examinee’s first question. This score ranges from 1 to 6 and it is based on the number of pictures targeted by the question (see Chapters 2 and 3). Integration planning scores for each problem were summed (see Chapter 3) and averaged to derive a mean integration planning score (MIPS) for each participant and to calculate group means. Table 4.3 shows that the MIPS for the young, early-adolescent, and adolescent groups were 3.7, (SD = 1.6), 4.5 (SD = 1.1), and 4.8 (SD = 1.0) respectively. ANOVA results revealed statistical differences between the groups ($p < 0.0001$) and that the young group had a significantly lower MIPS score than the other two groups. Kruskal-Wallis tests also revealed differences, but resulted in a lower p-value when younger and early-adolescent groups were compared ($p < 0.003$).

Table 4.3. Mean integration planning scores for each group.

	<i>M</i>	Median	SD
Young	3.7	4.0	1.6
Early Adolescent	4.5	4.7	1.1
Adolescent	4.8	4.7	1.0

Do typically developing young, preadolescent, and adolescent children exhibit differences in problem solving efficiency scores (PSE) on the RAPS?

Question-efficiency (QE) scores are based on the premise that the optimal way to solve problems on the RAPS is to ask questions that target larger, and then smaller numbers of pictures so as to reduce the number of pictures under consideration in near-50% increments (Hartley & Anderson, 1983; Marshall & Karow, 2008). These scores, described in Chapters 2 and 3, range from 0 to 1.00 and reflect the degree to which the examinee’s questions conform to this optimal. As described in Chapter 3, QE scores for individual questions were summed and averaged in different ways to calculate three different efficiency scores, problem solving efficiency (PSE), overall RAPS efficiency (ORE), and question sequence efficiency (QSE).

To obtain PSE scores, QE scores for each problem were averaged. The PSE score permits the investigator to determine if problem solving is improving his/her performance from problem-to-problem and assess possible learning effects. Normative studies with the RAPS with adults have not shown evidence of learning effects (Marshall et al., 2003a; Marshall & Karow, 2008). To assess possible learning effects in this study, which dealt

exclusively with children, PSE scores were derived for each participant by averaging the QE scores for the first four questions for each problem. Due to having repeated measurements from each subject, a mixed model using Kenward and Roger (1997) degrees of freedom was fit. Results showed that there was not strong enough evidence to support that the differences in means between these efficiency scores differed for the three groups ($p = 0.68$). Removing this interaction from the model, there was not strong enough evidence to conclude that the means were any different for these three variables ($p = 0.81$). However, there was strong enough evidence to conclude that mean efficiency scores differ by group ($p < 0.0001$). Specifically, the young group tended to have lower efficiency scores on average ($p = 0.0001$ vs. early adolescent group; $p < 0.0001$ vs. adolescent group).

Table 4.4 Mean problem solving efficiency scores for young, early-adolescent, and adolescent groups for the three RAPS problems.

	Problem 1		Problem 2		Problem 3	
	<i>M</i>	SD	<i>M</i>	SD	<i>M</i>	SD
Young	0.44	0.21	0.43	0.21	0.45	0.22
Early Adolescent	0.54	0.20	0.54	0.16	0.56	0.18
Adolescent	0.57	0.15	0.57	0.16	0.56	0.14

Do typically developing young, early-adolescent, and adolescent children differ in the overall RAPS efficiency (ORE) scores?

Overall RAPS efficiency scores were obtained for each participant by averaging the three PSE scores; individual scores were then averaged to obtain a mean for each group. Table 4.5 shows that the mean ORE scores for the young, early adolescent, and adolescent groups were 0.44 (SD = 0.19), 0.55 (SD = 0.15), and 0.57 (SD = 0.11) respectively. ANOVA results indicated there were statistical differences between the groups ($p < 0.0001$) and that the young group had significantly lower efficiency scores than the early adolescent ($p = 0.0001$) and the adolescent group ($p < 0.0001$). Kruskal-Wallis tests indicated similar differences.

Table 4.5 Overall RAPS efficiency scores.

	<i>M</i>	Median	SD
Young	0.44	0.46	0.19
Early Adolescent	0.55	0.57	0.15
Adolescent	0.57	0.58	0.11

Do typically developing young, early-adolescent, and adolescent children exhibit differences in question sequence efficiency scores on the RAPS?

Normal adults and the children in this study usually solved problems on the RAPS with five or fewer questions. Question sequencing efficiency (QSE) scores, as described in Chapter 3, assess performance from question-to-question. These scores are derived by

summing and averaging the QE scores for the first, second, third, and fourth questions of each of the three problems. QSE scores were used to examine the possibility that reducing information load by covering eliminated pictures after each question would improve question-asking efficiency. If this were found to be the case, a participant would reflect higher QSE scores from question-to-question. Table 4.6 (upper portions) shows the mean scores for each group on a problem by problem basis and the mean scores averaged for the three problems (lower portion) used for the statistical analyses.

Repeated measurements were obtained from each participant, therefore a mixed model with Kenward and Roger (1997) degrees of freedom was utilized. Results showed that the means of the average efficiency scores across the four questions differed for the three groups ($p = 0.006$). Table 4.6 (lower portion) revealed that young participants evinced minimal differences in their QSE scores across questions ($p = 0.26$).

There were, however, significant differences in mean QSE scores for the older groups. For the early adolescent group, differences in mean QSE scores were statistically significant when comparing questions 1 and 2 ($p = 0.037$), questions 1 and 4 ($p = 0.020$), questions 2 and 3 ($p < 0.001$), and questions 2 and 4 ($p < 0.001$). For the adolescent participants, the mean QSE scores were statistically different for question 2 relative to all other questions ($p < 0.001$). Question-to-question differences in the QSE scores for the three groups are depicted graphically in Figure 4.1. Here it can be seen that the young reflected much less fluctuation in their QSE scores across the questions than the other two groups. In contrast, the early adolescent and adolescent groups had relatively high QSE scores for question 1. Both groups had substantially lower QSE scores on question

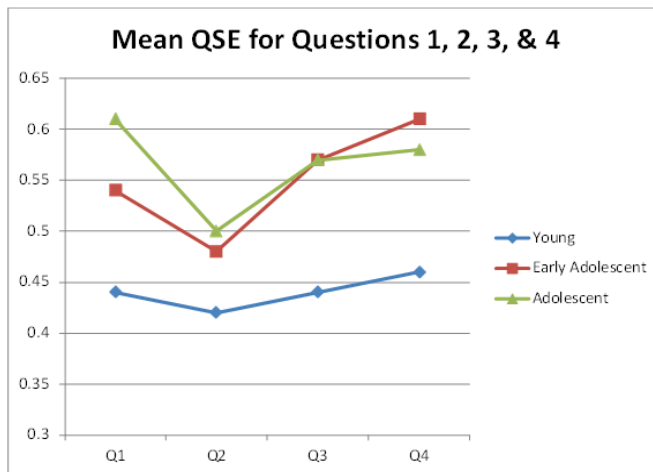
2 than question 1, and both groups had significantly higher mean QSE scores for questions 3 and 4 than question 2.

Table 4.6. Mean and standard deviations for question sequence efficiency scores for young, preadolescent, and adolescent participants on the RAPS.

	Question 1		Question 2		Question 3		Question 4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Problem 1								
Young	0.45	0.33	0.43	0.25	0.45	0.29	0.45	0.28
Early Adolescent	0.53	0.29	0.47	0.24	0.57	0.27	0.60	0.28
Adolescent	0.61	0.31	0.50	0.21	0.57	0.24	0.56	0.23
Problem 2								
Young	0.42	0.29	0.41	0.26	0.43	0.26	0.48	0.27
Early Adolescent	0.56	0.31	0.46	0.21	0.57	0.29	0.58	0.29
Adolescent	0.58	0.31	0.52	0.23	0.59	0.26	0.59	0.23
Problem 3								
Young	0.45	0.33	0.43	0.25	0.43	0.25	0.46	0.30
Early Adolescent	0.52	0.28	0.50	0.24	0.58	0.27	0.64	0.29
Adolescent	0.63	0.30	0.48	0.23	0.54	0.22	0.60	0.22
Mean QE*								
Young	0.44	0.28	0.42	0.20	0.44	0.20	0.46	0.21
Early Adolescent	0.54	0.25	0.48	0.17	0.57	0.18	0.61	0.21
Adolescent	0.61	0.26	0.50	0.13	0.57	0.16	0.58	0.14

Note. Mean QE is the average question efficiency for problem 1, 2, and 3 by question.

Figure 4.1 Mean QSE scores for questions 1 through 4 by age group.



SECTION 3

Descriptive Analysis

Do typically developing young, early-adolescent, and adolescent children ask different types of questions to solve problems on the RAPS?

The 229 participants asked 3367 questions, which were comprised of 2392 CS questions and 975 guesses. Constraint questions were classified as novel, category-limited, narrowing, or inefficient questions using definitions from Marshall and Karow (2008) given in Chapter 2. Table 4.7 shows the numbers and percentages of the types of questions and guesses for each group for the first four questions asked. These data suggest young, early-adolescent, and adolescent participants ask different types of questions when solving problems on the RAPS. When group differences were examined with respect to the probability of asking specific types of questions there were no group differences with respect to the probability of asking inefficient constraint questions ($p = 0.07$), but there were group differences for all other question types. Groups differed with respect to the probability of asking novel questions ($p = 0.02$); young participants asked fewer novel questions than early adolescent ($p = 0.01$) and adolescent participants ($p = 0.02$). While Table 4.7 shows the three groups asked high percentages of category-limited questions, there were significant group differences in the use of these questions ($p = .002$). The probability that young ($p = 0.003$) and early adolescent ($P = 0.02$) would ask a category-limited questions was lower than that for adolescent children. Although narrowing questions constituted a small portion of participant's total questions, there were group differences with respect to the probability of these questions ($p = .005$). The

probability of young children asking a narrowing question ($p = 0.0003$) was lower than that for early adolescent ($p = 0.002$) and adolescent ($p = 0.02$) children. Finally, there was strong evidence to show that the groups differed with respect to the probability of guessing ($p = 0.0001$). The probability that young children would guess was higher than that for early adolescent ($p = 0.0003$) and adolescent ($p = 0.0001$) children.

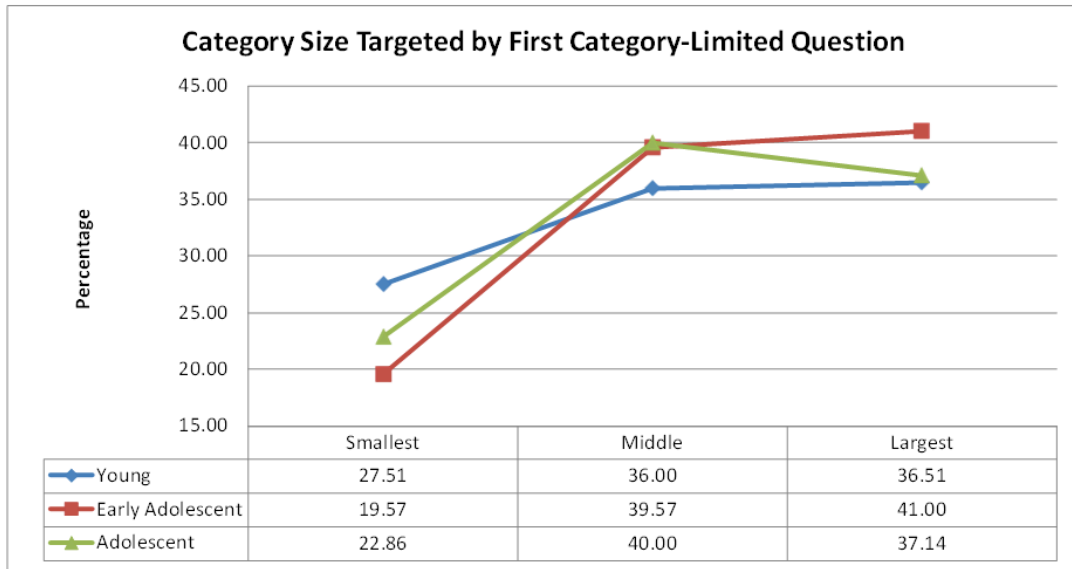
Table 4.7. Number and percentage of novel, category-limited, narrowing, inefficient-constraint questions and guesses for young, early adolescent, and adolescent participants on the RAPS for the first four questions. (Total questions for group)

	Novel	Category-limited	Narrowing	Inefficient Constraint	Guesses
Young (851)	11.9%	46.7%	2.9%	8.1%	30.4%
Early Adolescent (902)	19.5%	51.8%	6.4%	6.3%	16.0%
Adolescent (864)	17.9%	59.1%	5.4%	4.5%	13.0%

Because category-limited questions were asked frequently by each group (See Table 4.7) and the groups differed with respect to the probability of asking these questions, a post-hoc analysis was conducted to determine if participants' first category-limited questions targeted the largest (8 pictures), next-largest (6 pictures), or smallest (4 pictures) picture category when the question was asked. This information is provided in Figure 4.2. This figure shows that the percentage of questions that target the smallest

picture category of four pictures is slightly higher for the young group, but the groups differ minimally with respect to which their questions target the next-largest and largest categories.

Figure 4.2 Percentage of times the first category-limited question targeted each picture category size available by age group.



Do typically developing young, early-adolescent, and adolescent children ask different types of questions at early (questions 1 and 2), middle (questions 2 and 3), and late (questions 5 and 6) points in the problem solving sequence?

While young, early adolescent, and adolescent participants differed with respect to the probability of asking certain types of questions on the RAPS, overall question counts do not provide information about when, in the question asking sequence one

Table 4.8. Percentages of types of questions (novel, category-focused, narrowing, inefficient-constraint, and guesses) asked by young, preadolescent and adolescent participants for the first six questions of the RAPS. (number of questions).

	Questions					
	1	2	3	4	5	6
<i>Young</i>	(219)	(219)	(219)	(194)	(128)	(75)
Novel	23.2%	11.8%	7.8%	3.6%	<1%	0%
Category-Limited	48.4%	57%	47.9%	31.4%	18.8%	22.6%
Narrowing	-----	<1%	4.1%	7.2%	0%	2.6%
Inefficient Constraint	9.6%	11.4%	5.9%	5.2%	0%	2.6%
Guess	18.7%	18.7%	34.2%	52.5%	75.8%	69.3%
<i>Early Adolescent</i>	(237)	(237)	(237)	(191)	(126)	(55)
Novel	32.5%	19.4%	14.*%	9.4%	3.9%	0%
Category-Limited	56.5%	60.3%	51.4%	40.8%	38%	25.4%
Narrowing	-----	<1%	13.9%	17.2%	13.4%	5.4%
Inefficient Constraint	8.4%	7.6%	3.8%	5.2%	2.4%	5.4%
Guess	6.8%	11.8%	20.2%	27.2%	42%	63.6%
<i>Adolescent</i>	(231)	(231)	(229)	(182)	(131)	(75)
Novel	40.5%	14.7%	7.4%	4.9%	2.5%	<1%
Category-Limited	48.2%	69.7%	65.5%	52.7%	54.2%	29.3%
Narrowing	----	3.5%	9.6%	9.3%	3.8%	10.6%
Inefficient Constraint	7.6%	8.6%	2.2%	2.7%	1.5%	2.7%
Guess	3.4%	6.5%	16.2%	30.2%	38.2%	56%

Note. Question count varies across the six questions as all participants did not ask six questions in solving problems. Some participants asked more and some asked less than six questions. Additionally, it is impossible to ask a narrowing question as a first question so that space is left blank for all groups.

question is asked in lieu of another. Since this information would shed light on how children go about solving problems, percentages of the various types of questions were calculated for the first through the sixth question for each group. These data are presented in Table 4.8.

Table 4.8 indicates that all three groups tended to ask certain types of questions at particular times. For example, all three groups of participants asked most of their novel questions on question one. Young participants however asked far fewer novel first questions (23.2%) than early adolescent (32.5%) or adolescent participants (40.5%). Marshall and Karow (2008) reported their adult normal subjects frequently asked “Is it a black and white picture?” or “Is it a colored picture?” as a novel first question. A post-hoc examination of the number of times the child participants asked these novel first questions revealed that the young, early adolescent, and adolescent participants did this 20.5%, 25.7%, and 33.3% of the time.

In contrast to asking novel first questions, guesses on early questions on the RAPS, are usually counterproductive. Table 4.8 shows that the three groups differed markedly in terms of the frequency of guessing on early questions. Young participants guessed on the first two questions nearly 20% of the time. First and second question guessing for early adolescent and adolescent participants was substantially less.

Finally, Table 4.8 shows that when participants asked narrowing questions, they tended to do so at question three and beyond. The young group asked the lowest proportion of narrowing questions at this point. Early adolescent participants asked the most whereas the adolescent group fell in between.

Do young, early-adolescent, and adolescent children use different strategies when solving problems on the RAPS?

As described in Chapter 3, participants were assigned to one of four strategy groups based on the types of first questions asked. Novel strategy users asked a novel question first on all three problems; those assigned to the category-focused strategy group always began with a category-limited question. The children in the mixed strategy group, however, asked novel, category-limited, inefficient constraint, and guesses as first questions. A few children were placed in a fourth strategy group designated as a “no strategy” group because they always guessed on the first question.

Table 4.9 gives the number and percentages of young, early adolescent, and adolescent participants in the four strategy groups. These data exclude one adolescent participant deemed to be an outlier because he asked an inefficient constraint questions first on all three problems. Table 4.9 shows that fewer participants in the young group fell into the novel and category-focused strategy groups and more fell into the mixed and no strategy groups. Fisher’s Exact Test was used to examine differences between the groups in strategy group placement. There is not quite strong enough evidence ($p = 0.07$) to conclude the groups differed in terms of strategy use, but the data shown in table 4.9 do appear to reflect that the young group tended to be placed in the no strategy and mixed strategy group more often.

Table 4.9 Problem solving strategies for each age group.

Strategy*	ALL		Young		Early Adolescent		Adolescent	
	N	%	N	%	N	%	N	%
Novel	42	18.4	8	10.96	14	17.72	20	26.32
Category	56	24.56	14	19.18	21	26.58	21	27.63
Mixed	122	53.51	46	63.01	42	53.16	34	44.74
No Strategy (Guesses)	8	3.51	5	6.85	2	2.53	1	1.32

Note. Results reflect 228 of 229 participants. One participant was excluded as an outlier based on the use of inefficient constraint questions first on all three problems.

Do typically developing young, early-adolescent, and adolescent children exhibit differences in their ability to shift strategies to solve problems on the RAPS?

Narrowing questions are a special type of constraint question that the problem solver has an opportunity to ask after getting a “yes” answer to a category-limited question (e.g. Is it an animal?). Narrowing questions (e.g., Does it live in the jungle?) further constrain (narrow) the number of pictures under consideration and are congruent with the goal of solving the problem with as few questions as possible. Narrowing questions can only be asked when the opportunity arises and require a shift from a category-focused approach to a more abstract approach. The alternative to asking a narrowing question is to guess. To ascertain how narrowing questions were utilized by the three groups, all narrowing opportunities were identified, and the percentage of times members of each group used narrowing questions was calculated. Participants in the

young group had 112 narrowing opportunities, asked narrowing questions 34.5% of the time. Those in the early adolescent group had 118 opportunities to ask a narrowing question and asked a narrowing question 66.1% of the time. Adolescent participants had 114 opportunities to narrow and asked a narrowing question 54.4% of the time. Thus the younger participants shifted to narrowing questions much less often than the two older groups.

Examination of the first questions of the 46 young, 42 early adolescent, and 34 adolescent participants in the mixed strategy groups also provides a means to examine strategy shifting. An example of a positive strategy shift might be starting off with a guess on problem one, and then asking a more efficient question (e.g., category-limited question) on problem two. Accordingly, the number of positive strategy shifts was counted for the child subjects in the mixed strategy groups from problem one to two and problem two to three. Participants in the young, early adolescent, and adolescent groups made positive shifts on 38.8% (33/92), 41.7% (35/84), and 38.2% (26/68) of opportunities.

Chapter 5: Discussion

This study assessed problem solving of typically developing children with the Rapid Assessment of Problem Solving test (RAPS; Marshall et al. 2003a). The RAPS is a modification of Mosher and Hornsby's 20 Question Task (1966). The test involves verbally solving three problems. Each problem requires the examinee to ask yes/no questions to identify a target picture from a 32-picture array with the goal being to solve the problem with as few questions as possible. The RAPS has been used successfully to assess problem solving in neurologically intact (Marshall et al. 2003a; Marshall & Karow, 2008) and neurologically compromised adults (Marshall et al., 2003b, 2006, 2007; Furgeson, Marshall, & Olson, 2012), but published data on the test are not available for children. This cross sectional study sought to establish a normative database for the RAPS for children, examine differences in problem solving skills of children across selected age ranges, and hopefully provide clinicians with a useful tool to identify and assess the effects of problem solving interventions in children with disabilities who are sometimes difficult to test with conventional measures.

Trained graduate student testers administered the RAPS to 279 child participants. All children completed the RAPS testing without difficulty. The graduate students' ensured information from the RAPS was recorded accurately on the test's recording forms, calculated question efficiency scores, and entered information from the recording forms on the scoring summary to calculate various scores for the test as described in Chapter 3. Ultimately, a total of 229 tests were available for inclusion in data analysis. Final scoring of these tests, scoring reliability, and checks for accuracy of classification

of questions were then performed by the investigator and a trained research assistant (RA). The investigator and the RA assessed scoring accuracy, reliability, and the accuracy with which yes/no questions on the test were classified. In general, there were very few errors in scoring or classification accuracy. Errors identified were resolved through discussion by the investigator and the RA. Prior to beginning the data analysis, the 229 participants were divided into three groups: young (ages 7-9; N = 73), early adolescent (ages 10-13; N = 79), and adolescent (ages 14-17; N = 77).

The discussion that follows will focus on the following: (1) how findings of the present study relate to Mosher and Hornsby's (1966) early study of children's use of question-asking strategies; (2) age related differences in problem solving on the RAPS between typically developing young, early adolescent, and adolescent children; (3) differences in how children and adults perform on the RAPS; and (4) study limitations, clinical applications, and possibilities for further research.

RAPS and the 20Q Task

Two approaches (Mosher & Hornsby, 1966) were identified for asking yes/no questions to solve 20Q problems, "constraint seeking" and "hypothesis scanning." The former involves asking questions that eliminate half the alternatives with each question and assumes all alternative possibilities are equally likely. This approach minimizes the number of questions that need to be asked and yields useful information regardless of whether a question is answered "yes" or "no." The hypothesis scanning strategy is the opposite of the constraint seeking strategy. Here the individual asks questions that test a specific hypothesis and have no relation to the questions that have been asked before.

Each approach has costs and benefits. Constraint seeking provides efficiency in the use of information and likelihood of success in a reasonable amount of time, but at the expense of a higher cognitive workload (i.e. forming a plan and using a strategy). Hypothesis scanning lessens cognitive demands for formulation and use; it also offers the possibility of quick success that is not attainable with constraint seeking.

Mosher and Hornsby (1966) examined the question asking strategies of six, eight, and 11-year-old children using a fixed alternative 20Q task consisting of 42-pictures of black and white line drawings (see Figure 1.1). All subjects were boys and each age group had 30 subjects. While data from their study were not treated statistically, three salient findings emerged from the study. First, six-year-old boys reflected near-exclusive use of the hypothesis scanning approach to solving 20Q problems. Second, the eight and 11-year-old boys used a constraint seeking approach approximately 50% and 80% of the time respectively suggesting the use of constraint seeking questions to solve 20Q problems increase with age. Finally, constraint seeking questions, particularly from the older boys, were predominantly superordinate category questions (e.g., Is it a toy?) referred to in this study as category-limited questions.

The present study included both boys and girls, divided into three groups (young, early adolescent, and adolescent) based on age. Children younger than seven were excluded from this study for the simple reason that Mosher and Hornsby's (1966) study indicated them incapable of asking constraint questions that are necessary to solve problems on tests such as the RAPS. Two outcome measures for this study, mean number of questions (M#Q), and percent constraint questions (% CS) were also used by Mosher

and Hornsby (1966). Mosher and Hornsby reported six, eight, and 11-year-old boys solved the problem (identified the target picture) with an average of 26, 15, and 11 questions respectively. All children in this study reached a solution with fewer questions. On the average, children in the young, early adolescent, and adolescent groups asked 5.2, 4.7, and 4.8 questions per problem solved. This marked difference in the number of questions needed to reach a solution by the participants in the two studies may be due to several factors. First, six-year-old children were not included in this study. Second, Mosher and Hornsby's (1966) 20Q task had 42 pictures whereas the RAPS problem boards had only 32 pictures. Third, the 42-item picture display of Mosher and Hornsby was not organized in any particular manner, whereas the problem solving boards of the RAPS were specifically designed to encourage the examinee to ask certain types of questions (e.g., half the pictures were colored and half were black and white; pictures could be groups in semantic categories of four, six, and eight pictures; and pictures were arranged in a 4 x 8 grid). Finally, on the RAPS, a problem is considered solved when the number of pictures on the problem board had been reduced to two or three because at this point the only alternative left to the examinee is to guess. It is not clear from the information reported by Mosher and Hornsby when their problem solving effort was terminated.

As previously mentioned, Mosher and Hornsby (1966) reported a substantial difference in the use of constraint seeking questions by eight and 11-year-old boys (50% versus 80%). While participants in the young, early adolescent, and adolescent groups asked constraint questions 62%, 79%, and 80% on the average (see Table 4.2), these differences were statistically significant in that young participants asked fewer constraint

seeking questions than the two older groups. Thus, the frequent use of constraint questions and the increase of these types of questions across the three groups support Mosher and Hornsby's assertion that use of these questions increase with age.

Many studies using the fixed alternative, 42-picture, 20Q task shown in Figure 1.1 (Hartley & Anderson, 1983; Denney, 1985; Denney & Denney, 1973, 1982; Drumm, Jackson, & Magley, 1995; Simon, 1975) have reported that normal subjects reflect a propensity for asking certain types of constraint questions, specifically superordinate category questions (e.g., Is it an animal?). This has also been seen in normative studies with adults using the RAPS (Marshall et al., 2003a; Marshall & Karow, 2008). Mosher and Hornsby (1966) also reported, but did not quantify, that their eight and 11-year-old children predominately asked superordinate category questions. Table 4.7 shows that 47.4% and 54% of the questions the early adolescent and adolescent group asked respectively were category-focused questions and that use of these questions increased in relation to age. Findings of this study support those of Mosher and Hornsby and other research with the 20Q task and the RAPS. It would therefore seem safe to conclude that if a fixed alternative 20Q task is composed of pictures of items that can be grouped into semantic categories, it will prompt the examinee to ask category limited or superordinate category questions even though other types of questions might sometimes be more effective at solving the problem.

Group Differences in Problem Solving

Like the 20Q task, the RAPS is presented as a game that challenges the examinee to ask yes/no questions and try to solve the problem (identify the target picture) with as

few questions as possible. In actuality, the number of questions metric provides limited information about how an individual solves problems. Research with the RAPS, reviewed in some detail earlier in this paper, has focused on the test as an executive function measure. Therefore, differences in problem solving on the RAPS between the young, early-adolescent, and adolescent groups will be examined in relationship to components of executive functioning considered to be important in solving problems: planning, strategy selection, strategy execution, and strategy shift.

Planning. Planning is one of several complex cognitive functions subsumed under the general rubric of thinking (Sohlberg & Mateer, 1989). Planning on the RAPS is reflected in two ways. The first involves making a decision to analyze or to act (Scholnick & Friedman, 1993) before asking the first question. A decision to analyze would be evidenced with the asking of early constraint seeking questions, whereas a decision to act would be seen in a first question guess. Table 4.8 indicates that the young participants guessed 18.7% of the time on their first and 18.7% on their second questions in solving RAPS problems. Conversely, early guesses by early adolescent (6.8% on the first question and 11.8% on the second question) and adolescent (3.4% on the first and 6.5% on the second) participants were far less frequent. Although these group differences in early guessing were not examined statistically, non-parametric analyses examining the probability of guessing on questions one-four revealed that the younger subjects had a higher probability of guessing than older groups.

The integration planning scores (IPS) for each problem and the average of these scores for the test (mean integration planning score; MIPS) also provide information

about the ability to plan. In keeping with test instructions to ask yes/no questions to solve the problem with as *few* questions as possible, an ideal first question would target approximately half of the pictures in the array. The MIPS, which can range from 1 to 6, reflects the average number of pictures targeted by the examinees' first questions. Table 4.3 shows that the MIPS for the young, early adolescent, and adolescent subjects were 3.7, 4.5, and 4.8 respectively. These scores were statistically different with the young group having lower MIPS than the two older groups.

Strategy Selection. Strategy selection refers to the development of a plan of action for solving a problem (Scholnick & Friedman, 1993). Marshall and Karow (2008) identified two distinct strategies used by normal adults in solving problems on the RAPS, novel and category-focused. Novel strategists consistently started with a novel question. This question was frequently "Is it a black and white picture?" or "Is it a colored picture?" Since either question would eliminate 50% of the pictures from consideration regardless of whether it was answered "yes" or "no," Marshall and Karow deduced that novel strategists selected an optimal constraint seeking strategy that would reduce alternatives in near 50% increments. Category-focused strategists, on the other hand, consistently started off with a category-limited question (e.g., Is it a sports ball?). Marshall and Karow placed most of their 373 adult participants in a "mixed strategy" group. Mixed strategists' first questions were a combination of novel and category-limited questions, but did not include guesses. As was the case in the Marshall and Karow study, some children in the current study could be assigned to novel or category-focused strategy groups, but most were assigned to the mixed strategy group. Children in the mixed strategy group, however, asked different questions than Marshall and Karow's

adult subjects. Whereas the adults only asked novel and category limited questions first, the child mixed strategists asked inefficient constraint questions and guessed on first questions as well. This suggests adults are trying to decide between two relatively effective strategies for solving RAPS problems, novel and category-focused. Conversely, children are trying to discern effective and ineffective strategies. Table 4.9 shows that a higher proportion of young participants were placed in the mixed and no strategy groups, and fewer young children were in the novel and category-focused groups. ANOVA results approached significance at $p = 0.07$ but were not strong enough to support group differences in strategy group placement at the $p = 0.05$ level. These differences, while not statistically significant, suggest the possibility of age related differences in ability to select effective strategies for solving problems on the RAPS.

Strategy Execution. Question efficiency (QE) scores provide information about how well or how poorly the examinee executes his or her constraint-seeking strategy in solving problems on the RAPS. QE scores are obtained by dividing the number of pictures still showing when a question is asked by the smaller of two numerators, either pictures targeted or pictures eliminated, and multiplying the result by two. These scores (ranging from 0.00 to 1.00) penalize the problem solver for guessing or asking inefficient questions and reward efficiency. The overall RAPS efficiency score (ORE; an average of the 12 QE scores) has been used to determine if the problem solver performed optimally on the test by asking questions that reduce the number of pictures in near 50% increments so as to be able to solve every problem with four questions.

Table 4.5 shows the mean ORE scores for the young, early adolescent, and adolescent groups were .44, .55, and .57 respectively. Group means were statistically different and the young group had a mean ORE score significantly lower than the other two groups. This suggests there are age related differences in how children execute strategies for solving problems on the RAPS. Further support for this assertion can be found in Marshall and Karow's (2008) normative study with adults. In this study ORE scores for 18-19 year-old (N = 34) and 20-29 year-old (N = 86) were 60.7 (SD = 14.3) and 62.6 (SD = 11.4) respectively. The present study suggests what has been found to be true for adult performance was also seen in children's performance on the RAPS. Differences in ORE scores for both children and adults seem to vary according to age.

QE scores were also used to compute problem solving efficiency scores (PSE) to ascertain if participants in the three groups became more efficient at executing their problem solving strategy from problem-to-problem. If this were found to be the case, it would indicate a possible learning effect. Table 4.4 shows the mean PSE scores for problems one, two, and three for each group. The PSE mean values for each group are essentially identical across the three problems. Additionally, in each case, the mean PSE scores are almost identical to the ORE values for each group as shown in Table 4.5. Analyses, to determine if PSE scores for any of the groups improved from problem to problem reflecting improved strategy execution, showed this not to be the case.

Since the examiner covers up pictures eliminated after each question when giving the RAPS, it is possible an examinee could improve his or her QE scores from question-to-question due to a lessening of cognitive information load. This also might be construed

as a possible learning effect. To assess this possibility, question sequencing efficiency scores (QSE) were obtained by averaging QE scores for the first, second, third, and fourth questions for each problem, and averaging means for the three problems. Table 4.6 shows the mean QSE scores for each group on a problem by problem basis as well as the overall means for the three problems. Figure 4.1 shows that the mean QSE scores for the young subjects are remarkably consistent from question to question for each problem as well as when scores are averaged for all three problems. For the early-adolescent and adolescent groups, the mean QSE scores are higher for question one than question two, higher for question three than question two, and higher for question four than question three. ANOVA analyses were carried out only for the overall means. Results revealed no significant differences for the young group, but significant differences for the early adolescent and adolescent groups.

Marshall et al., (2003a) found that QSE scores increased from question to question in their initial study with the RAPS. This suggested subjects improved question asking efficiency as fewer pictures became available. In the current study, young children demonstrated no such tendency; the two older groups had QSE scores for question one equivalent to those of Marshall et al. (2003a) on questions one, three, and four, but demonstrated a decided drop in their QSE scores for question two. One possible explanation for the lower efficiency scores on question two, in spite of the fact that far fewer pictures are available at this time, is that many participants in studies with the RAPS (both adults and the children in this study) begin with the question, “Is it a black and white picture?” or “Is it a colored picture?” This question eliminates half of the pictures regardless of the answer. However, it also reduces the number of pictures in the

semantic categories represented on the RAPS by 50%. This may pose challenges for formulating an equally efficient or more efficient question for question two.

The optimal constraint seeking strategy to solve problems on the RAPS is to ask questions that target larger, then smaller numbers of pictures and reduce the number of items under consideration in near 50% increments. One question that arises from this study is why more people of all ages do not use a constraint seeking strategy that does this. For example, why more adults, and children for that matter, do not ask constraint questions targeting columns and rows of pictures? This is an easy strategy to execute that would ensure every problem was solved with four questions. A possible explanation for failure to execute an optimal strategy might be that the temptation to group pictures from the problem solving board into semantic categories so powerful that potential problem solvers consider asking any type of question other than a category-limited question to be non-elegant. This possibility was also entertained by Mosher and Hornsby (1996) who were surprised that their child subjects did not solve problems with fewer questions. It is also possible that the presence of easily identifiable semantic categories sets up a cognitive expectation that categories are *supposed* to be used in the problem solving process. This idea of performing to an expectation, although attractive to any age group, may be particularly influential with children.

Strategy Shift. Strategy shifting is required on the RAPS when the examinee receives a “yes” answer to a category-limited question (e.g., Is it an animal?). In such cases, the examinee has the opportunity and should shift to a narrowing question. In this case, if the animals pictured were cat, lion, elephant, deer, giraffe, zebra, horse, and pig,

an optimal narrowing question (e.g., Does it live in Africa?) would reduce the choices by 50% and be compatible with the goal of solving the problem with as few questions as possible. Use of narrowing questions can only be examined in relationship to the opportunities to ask them and this varies from individual to individual.

Table 4.9 shows the point in the question-asking sequence where participants asked certain types of questions and reveals that most narrowing questions were asked on questions three or four. A post hoc analysis of the opportunities to ask narrowing questions was completed for participants in the young (112 opportunities), early adolescent (118 opportunities), and adolescent (114 opportunities) groups. Young, early adolescent, and adolescent groups were found to ask optimal narrowing questions on 34.5%, 66.1%, and 54.4% of these opportunities. These differences suggest the groups differ in their ability to shift strategies.

In sum, examination of quantitative and descriptive data from typically developing young, early-adolescent, and adolescent children on the RAPS revealed several age related differences. For the most part, findings suggest that young children solve problems on the RAPS less efficiently than early adolescent and adolescent children. They also tend to ask different types of questions at particular points in the question asking sequence in solving 20Q problems of the RAPS. For all comparisons, there were negligible differences in the scores for the RAPS between early adolescent and adolescent participants. Interestingly, in some cases, the early adolescent group had scores on the test that were slightly better than the adolescent subjects. Though it is difficult to speculate why the older adolescent subjects did not have significantly higher

scores than the early-adolescent subjects on more of the comparisons, discussions (between the investigator, the RA, and the graduate students of the degree to which children in each group engaged in the RAPS assessment) suggested that the early adolescent participants responded more robustly to the challenges of the RAPS, whereas the task was treated with less seriousness by some of the adolescent participants. Environment of testing may have also influenced performance. At least 25 early adolescent participants were assessed in a classroom very similar to a school setting, whereas some of the adolescent participants were assessed in their homes and at other less structured locations (e.g., Vacation Bible School). The classroom structure could have set an academic tone to the RAPS administration for some early adolescents that may have been absent for some adolescent participants.

Differences in Adult and Child Performance on the RAPS

While the intent of this study was to establish a database for the RAPS with children, findings from the study also provide an opportunity to compare how adults and children perform on the test. It appears that normal adult subjects (Marshall et al., 2003a; Marshall & Karow, 2008) and children differ in their ability to integrate and use information available to them to plan, select and execute strategies, and make the necessary strategy shifts to solve problems on the RAPS. Any examinee, after receiving instructions for the test and before asking his or her first question, has the opportunity to scan the problem solving board and deduce that (a) the 32 pictures are arranged in rows and columns; (2) half the pictures are black and white and half are colored; (3) the pictures belong to common semantic categories; and (4) the picture categories vary in

size. They might also consider how to strategically use this information to solve the problem with as few questions as possible. After asking a question or two and seeing which pictures are eliminated by the question, the examinee may or may not revise their plan of action. Throughout the process, the examinee must keep in mind that the goal of the task is to solve the problem with as few questions as possible.

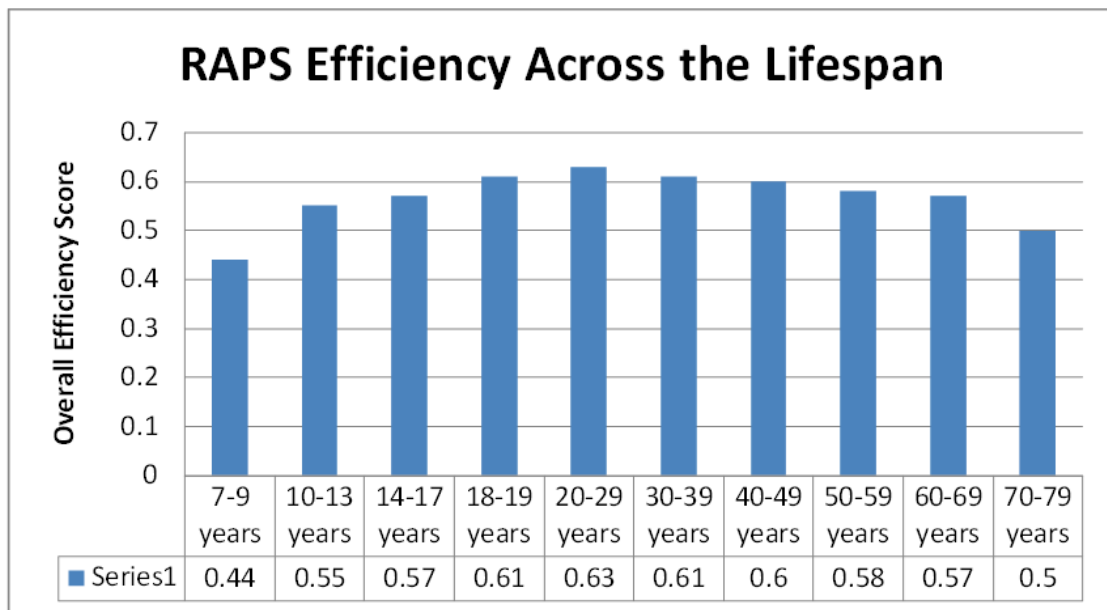
Adults are more likely to start off with questions that target and/or eliminate more pictures from consideration such as “Is it a black and white picture?” Children, particularly, older children do this some of the time, but younger children do so much less often. Young children often guess on early questions. Older children also guess but less often, particularly on early questions. In contrast, adults (Marshall & Karow, 2008) do not guess on first or second questions. While both adults and children ask a preponderance of category-limited questions, (Marshall et al., 2003a; Marshall & Karow, 2008) initial category-limited questions from adults tend to target the largest picture category available. In this study, none of the three child groups reflected a tendency to ask initial category-limited questions targeting the largest category available. This suggests adults are aware that the size of picture categories differ on the RAPS and children are not. Some, but not all, adults are likely to switch to narrowing questions after receiving a yes answer to a category-limited question, but all adults do not do this. The differences in the ability to use narrowing questions (Marshall et al., 2003a; Marshall & Karow, 2008) accounts for some of the variability in adult performance on the RAPS. Children in this study rarely asked narrowing questions, and when they did, they were likely to be older.

Two possible factors may explain why children and adults go about solving problems on the RAPS slightly differently. One is that children, by nature, are more impulsive and prone to seek quick solutions whereas adults are more reflective and may allot more time for planning and selecting a problem solving strategy (Ault, 1973; Zelniker et al., 1977). This possibility could easily be examined by measuring the time elapsing between receiving instructions for the RAPS and asking one's first question. Another possible explanation is that adults and children differ in their ability to scan the problem solving boards to learn information needed to ask efficient questions. To perform well on the RAPS, it is necessary to scan the contents of the problem board before starting to solve and also throughout the question asking process as pictures are eliminated. Adults and children may differ in their visual scanning abilities. To determine if visual scanning abilities are related to performance outcomes on the RAPS, it might be possible to employ visual tracking instrumentation with participants during administration of the test.

Finally, some differences in how children and adults perform on the RAPS are likely the result of cognitive maturation and age related changes. Findings of this study permit some preliminary conclusions to be drawn on performance on the RAPS across the life span. These are easiest to visualize when looking at overall RAPS efficiency (ORE) scores for the children in this study and Marshall and Karow's (2008) adults. Mean ORE scores for the 7-9 (N = 73), 10-13 (N = 79), and 14-17 (N = 77) year old children in this study were .44 (SD=0.19), .55 (SD = 0.15), and .57 (SD = 0.11) respectively. Those for 18-19 (N = 34) and 20-29 (N = 86) year old young adults of Marshall and Karow (2008) were .61 (SD = 0.14) and .63 (SD = 0.11) respectively.

Marshall and Karow’s middle-age subjects, ages 30-39 (N = 56) and 40-49 (N = 67) had mean ORE scores of .61 (SD = 0.11) and .60 (SD = 0.11) respectively. Finally, these researchers’ older adults, 50-59 (N = 60), 60-69 (N = 38) and 70-79 (N = 24) had mean ORE scores of .58 (D = .12), .57 (SD = 0.9), and .50 (SD = 0.16) respectively. Figure 5.1 provides a graphic depiction of these scores. Here it can be seen that from a life span perspective, it appears that the ability to solve problems on the RAPS improves from age seven through young adulthood, flattens out across the middle age, and declines as one gets older (see Figure 5.1).

Figure 5.1 Overall RAPS efficiency scores for children in this study and adults (Marshall and Karow, 2008).



Limitations, Clinical Implications, and Further Research

Limitations

Data for this study were collected in the summer months when the children were not in school. This was convenient, but in retrospect probably limited sample size. Hindsight suggests it would have been advantageous to have had more subjects in the young group who were either seven, eight or nine years of age. Most of the within and between group differences on the measures of interest involve this young group. Having more children at each age level would have permitted an analysis of the performance of the subjects at specific age levels rather than a range of ages. Relatedly, it probably would have helped to have more subjects in the early-adolescent and adolescent groups as well and to balance the groups for gender, i.e., 25 girls and 25 boys at each age. This may have permitted the investigator to identify performance differences between the older groups, which was not the case.

Another limitation of the study was failure to have an alternative plan when the examinee's first question was a lucky guess or fortunate question that permitted the problem to be solved with one or two questions. Twenty-six subjects were lost to data analysis because of this limitation. This could have been prevented by preselecting two alternative target pictures for each problem solving board. A lucky guess or fortunate first question identifying the target picture category would lead the examiner to answer "no" and move to the first back up picture. A lucky guess on question two would lead the examiner to move to the second back up picture. The examiner would then carry on with the test until the problem was solved. This simple administrative modification would

ensure that four questions were asked for each problem and allow for calculation of all RAPS scores.

Assessment of the children in the summer months and concern for their time and activity schedules limited the amount of testing done with the children. It is important to acknowledge that more than 22 children should have been tested twice to ensure test-retest stability. Also noticeably absent in this study is any measure of content validity. To assess content validity of the RAPS with adults, Marshall and Karow (2008) administered both the RAPS and the Coloured Progressive Matrices (CPM; Raven, Court, and Raven, 1984) to 85 of their adult subjects. They reported reasonable high correlations ($p < .01$) between two scores on the RAPS, M#Q and ORE, and CPM scores. While the CPM would not be a good choice for children, there are measures that could have been used to determine content validity. These could include games with unrestricted alternatives such as those used by Mosher and Hornsby (1966). Here the examiner describes a situation, i.e., “A boy goes home from school in the middle of the morning.” and the examinee asks yes/no questions to find out why. Another might include the Wal-Mart task (Marshall, 2008) in which the examinee asks yes/no questions to determine an item the examiner needs to buy at Wal-Mart.

Another study limitation was inability to collect more information on the child participants. Since the RAPS taps executive functions, it would have been beneficial to know more about how children were performing in school, if all were performing at grade level on standardized achievement tests, reading at grade level, etc. This information was not available because data were collected in the summer when children

were not in school and because protection of the children's privacy was a prerequisite to collecting the data in a timely manner.

Clinical Implications

Earlier literature reviews have underscored that the game like format of the 20Q task is well-tolerated by normal and disabled children. In this study, 275 typically developing children completed the RAPS testing without complaint or difficulty. The RAPS has been successfully administered to neurologically compromised adults with diagnoses of traumatic brain injury, dementia, aphasia, and severe mental illness (Marshall et al., 2003b, 2006; 2007). While the RAPS test does not yet have published data for children with disability, the investigator has administered the RAPS with 17 verbal children with autism with minimal difficulty (Smith, Page, & Marshall, 2013). A larger, more representative study with children with disabilities is planned for future research.

Future Research

A number of projects are in the works that involve the RAPS as a dependent variable. First, a project has been approved by the Western Kentucky University IRB that will involve examining problem solving on the RAPS for children with and without autism. Another project, in the planning stages, involves further examination of content validity issues for the RAPS by comparing performance of normal subjects on the RAPS, an unrestricted problem solving task, and the Wal-Mart task (Marshall, 2009).

Research with the RAPS, regardless of whether the test is used with children or adults, needs to address the time it takes to score the test. Marshall and Karow (2013) proposed that the RAPS be scored with a rubric, a scoring tool that assigns specific values to critical elements of executive function tasks. While some of the scores of the RAPS, such as the ORE score, are vital in scoring the test and sensitive to brain damage, a rubric would provide qualitative information about how an individual goes about solving problems on the test, and be an invaluable adjunct to the ORE score.

Finally, we live in an age when many hard copy tests can be administered by computer. With some effort and financing, the RAPS could be administered via computer application, which would speed up administration, facilitate record keeping and storage, and heighten appeal of the test for both patient and clinician. If the RAPS is to be considered a life span test of problem solving and gain any degree of wide spread clinical use, it should be developed into a computer application in the near future.

References

- Agran, M., Blanchard, C., Wehmeyer, M., & Hughes, C. (2002). Increasing the problem-solving skills of students with developmental disabilities participating in general education. *Remedial and Special Education, 23*, 279-288.
- Aitken, S., Chase, S., McCue, M., & Ratcliff, G. (1993). An American adaptation of the Multiple Errands Test: Assessment of executive abilities in everyday life. *Archives of Clinical Neuropsychology, 8*, 212.
- Alderson-Day, B., & McGonigle-Chalmers, M. (2011). Is it a bird? Is it a plane? Category use in problem-solving in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 41*, 555-565.
- Alderson-Day, B., Wright, B., Sims, D., Smart, S., Alwazeer, A., Allgar, V., & Miles, J. (2011). Melatonin versus placebo in children with autism spectrum conditions and severe sleep problems not amenable to behaviour management strategies: A randomised controlled crossover trial. *Journal of Autism and Developmental Disorders, 41*(2), 175-184.
- Allen, G.J., Chinsky, J.M., Larcen, S.W., Lochman, J.E., & Selinger, H.V. (1976). *Community psychology in the schools: A behaviorally oriented multilevel preventative approach*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Ault, R.L. (1973). Problem-solving strategies of reflective and impulsive children. *Child Development, 44*, 259-266.
- Bamdad, M. J., Ryan, L. M., & Warden, D. L. (2003). Functional assessment of executive abilities following traumatic brain injury. *Brain Injury, 17*(12), 1011-1020.
- Barton, J. A. (1988). Problem-solving strategies in learning disabled and normal boys: Developmental and instructional effects. *Journal of Educational Psychology, 80*, 184-191.
- Ben-Yishay, Y., & Diller, L. (1983). Cognitive remediation in traumatic brain injury: Update and issues. *Archives of Physical Medicine and Rehabilitation 74*, 204-213.
- Benson, D., Kuhl, D., Hawkins, R., Phelps, M., Cummings, J., & Tsai, S. (1983). The fluorodeoxyglucose 18F scan in Alzheimer's disease and multi-infarct dementia. *Archives of Neurology, 40*(12), 711-714.
- Borys, S. (1979). Factors influencing the interrogative strategies of mentally retarded and nonretarded students. *American Journal of Mental Deficiency, 84*, 280-288.

- Borys, S.V., & Spitz, H.H. (1976). Short-term retention in retarded adolescents as a function of load, delay, and interpolated activity. *The Journal of Psychology*, 94, 207-216.
- Boyd, T. M., & Sautter, S. W. (1993). Route-finding: A measure of everyday executive functioning in the head injured adult. *Applied Cognitive Psychology*, 7, 171–181.
- Brookshire, R. H. (2007). *Introduction to neurogenic communication disorders*. St Louis, Mo.: Mosby Elsevier.
- Butler, R. W., Rorsman, I., Hill, J. M., & Tuma, R. (1989). Behavioral assessment in neuropsychological rehabilitation: A method for measuring vocational-related skills. *The Clinical Neuropsychologist*, 3, 235–243
- Carroll (1993). *Human Cognitive Abilities: A survey of factor analytic studies*. New York: Cambridge University Press.
- Channon, S., & Crawford, S. (1999). Problem-solving in real-life-type situations: The effects of anterior and posterior lesions on performance. *Neuropsychologia*, 37, 757-770.
- Cornelius, S. W., & Caspi, A. (1987). Everyday problem solving in adulthood and old age. *Psychology and Aging*, 2, 144–153.
- Delis, D.C., Kaplan, E., Kramer, J.H. (2001). *Delis-Kaplan Executive Function System*, San Antonio, Texas: The Psychological Corporation.
- Denney, D. R. (1975). The effects of exemplary and cognitive models and self-rehearsal on children's interrogative strategies. *Journal of Experimental Child Psychology*, 19(3), 476-488.
- Denney, D. R. (1972). Modeling and eliciting effects upon conceptual strategies. *Child Development*, 810-823.
- Denney, D. R., & Denney, N. W. (1973). The use of classification for problem solving: A comparison of middle and old age. *Developmental Psychology*, 9(2), 275.
- Denney, D. R., Denney, N. W., & Ziobrowski, M. J. (1973). Alterations in the information-processing strategies of young children following observation of adult models. *Developmental Psychology*, 8(2), 202.
- Denney, N. (1985). Review of life span research with the Twenty Questions Task: A study of problem-solving ability. *International Journal of Aging and Human Development*, 21(3), 161-173.

- Denney, N. W. (1982). Aging and cognitive changes. *Handbook of Developmental Psychology*, 807, 827.
- Denney, N. W., & Connors, G. J. (1974). Altering the questioning strategies of preschool children. *Child Development*, 1108-1112.
- Denney, N. W., & Denney, D. R. (1974). Modeling effects on the questioning strategies of the elderly. *Developmental Psychology*, 10(3), 458.
- Denney, N. W., Jones, F. W., & Krigel, S. H. (1979). Modifying the questioning strategies of young children and elderly adults with strategy-modeling techniques. *Human Development*, 22(1), 23-36.
- Denney, N. W., & Palmer, A. M. (1981). Adult age differences on traditional and practical problem-solving measures. *Journal of Gerontology*, 36(3), 323-328.
- Denney, N. W., & Turner, M. C. (1979). Facilitating cognitive performance in children: A comparison of strategy modeling and strategy modeling with overt self-verbalization. *Journal of Experimental Child Psychology*, 28(1), 119-131.
- Denney, N., & Wright, J. (1976). Cognitive changes during the adult years: Implications for developmental theory and research. *Advances in Child Development and Behavior*, 11, 213-224.
- Donahoe, C. P., & Carter, M. J. Bloem. WD, Hirsch, GL, Laasi, N., & Wallace, CJ (1990). Assessment of interpersonal problem-solving skills. *Psychiatry*, 53, 329-339.
- Drumm, P., & Jackson, D. W. (1996). Developmental changes in questioning strategies during adolescence. *Journal of Adolescent Research*, 11, 285-305.
- Drumm, P., Jackson, D. W., & Magley, V. (1995). The primacy of superordinate-level category questions in the game of 20 questions. *Perceptual and Motor Skills*, 81(1), 271-274.
- D'Zurilla, T.J., & Goldfried, M.R. (1971). Problem solving and behavior modification. *Journal of Abnormal Psychology*, 78(1), 108-126.
- D'Zurilla, T.J., Maydeu-Olivares, A. & Kant, G. L. (1998). Age and gender differences in social problem solving in college students, middle age, and elderly adults. *Personality and Individual Differences*, 25, 241-252.

- Ferguson, A., Marshall, R. C., & Olson, A. E. (2012). Problem solving strategies of subjects with and without brain injuries on a modification of the 20 questions task. Poster Presentation: Undergraduate Research Symposium, University of Kentucky, Lexington, Ky.
- Glago, K., Mastropieri, M., & Scruggs, T. (2008). Improving problem solving of elementary students with mild disabilities. *Remedial and Special Education, 20*(10), 372-380.
- Glosser, G., & Goodglass, H. (1990). Disorders in executive control functions among aphasic and other brain-damaged patients. *Journal of Clinical and Experimental Neuropsychology, 485-501*.
- Gold, J.M., & Harvey, P.D. (1993). Cognitive deficits in schizophrenia. *Psychiatric Clinics of North America, 16*(2), 295-312.
- Goldman, R., & Goldman, M. (1988). Experience-dependent cognitive recovery in alcoholics: A task component strategy. *Journal of Studies on Alcohol, 49*, 142–148.
- Goldstein, F.C, & Levin H.S. (1991). Question-asking strategies after severe closed head Injury. *Brain and Cognition. 17*(1), 23-30. doi: 10.1016/0278-2626(91)90063-E.
- Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in Weigl-type card-sorting problem. *Journal of Experimental Psychology, 38*, 404–411.
- Hartley, A, A, & Anderson, W, J. (1983); Task complexity and problem-solving performance in younger and older subjects. *Journal of Gerontology, 38*, 72-77.
- Heppner, P. P. (1988). *The problem solving inventory: Manual*. Palo Alto, Calif.: Consulting Psychologists Press.
- Horn, G. J. & Cattell, R. B. (1967). Refinement and test of the theory of fluid and crystalized intelligence. *Journal of Educational Psychology, 57*, 253-270.
- Kafer, K. L., & Hunter, M. (1997). On testing the face validity of planning/problem-solving tasks in a normal population. *Journal of the International Neuropsychological Society, 3*, 108–119.
- Keil, K., & Kaszniak, A.W. (2002). Examining executive function in individuals with brain injury: A review. *Aphasiology, 16* (3), 305-335.
- Kenward MG, Roger JH. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics, 53*, 983-997.

- Kesler, M. S., Denney, N. W., & Whitely, S. E. (1976). Factors influencing problem-solving in middle-aged and elderly adults. *Human Development, 19*(5), 310-320.
- Klouda, G. V., & Cooper, W. E. (1990). Information search following damage to the frontal lobes. *Psychological reports, 67*(2), 411-416.
- Knight, C., Alderman, N., & Burgess, P. W. (2002). Development of a simplified version of the multiple errands test for use in hospital settings. *Neuropsychological Rehabilitation, 12*(3), 231-255.
- Kuhn, W., Roebroek, R., Blom, H., van Oppenraaij, D., Przuntek, H., Kretschmer, A., Muller, T. (1998). Elevated plasma levels of homocysteine in Parkinson's disease. *European Neurology, 40*(4), 225-7.
- Laine, M., & Butters, N. (1982). A preliminary study of the problem-solving strategies of detoxified long-term alcoholics. *Drug and Alcohol Dependence, 10*(2), 235-242.
- Levin, H. S., Culhane, K. A., Mendelsohn, D., Lilly, M. A., Bruce, D., Fletcher, J. M., & Eisenberg, H. M. (1993). Cognition in relation to magnetic resonance imaging in head-injured children and adolescents. *Archives of Neurology, 50*(9), 897-905.
- Levin, H. S., Song, J., Scheibel, R. S., Fletcher, J. M., Harward, H., Lilly, M., & Goldstein, F. (1997). Concept formation and problem-solving following closed head injury in children. *Journal of the International Neuropsychological Society, 3*(06), 598-607.
- Lezak, M. D. (1995). *Neuropsychological Assessment*. New York: Oxford University Press.
- Lezak, M.D., Howieson, D.B., Loring, D.W., Hannay, H.J., & Fischer, J.S. (2004). *Neurological assessment* (4th ed.). New York: Oxford University Press.
- Locke, L.F., Silverman, S.J., & Spriduso, W.W. (2010). *Reading and understanding research* (3rd ed.). Thousand Oaks, Calif.: Sage Publications.
- Lubinski, R. (1995). State-of-the-art perspectives on communication in nursing homes. *Topics in Language Disorders, 15*, 1-19.
- MacDonald, S. & Johnson, C.J. (2005). Assessment of subtle cognitive communication deficits following acquired brain injury: A normative study of the functional assessment of verbal reasoning and executive strategies (FAVRES). *Brain Injury, 19*(11), 895-902.

- Maddy, K. M., & Marshall, R. C. (2012). Will treatment facilitate learning of a problem solving strategy by persons with Alzheimer's disease? Presented at the Clinical Aphasiology Conference on May 20-25.
- Marschark, M., & Everhart, V. S. (1999). Problem-solving by deaf and hearing students: Twenty questions. *Deafness & Education International*, 1(2), 65-82.
- Marshall, R. C., Capilouto, G. J., & McBride, J. M. (2007). Treatment of problem solving in Alzheimer's disease: A short report. *Aphasiology*, 21(2), 235-247.
- Marshall, R. C., Dixon, J., Iden, K., Karow, C. M., & Morelli, C. M. (1999). Problem solving efficiency of elderly subjects: Effects of strategy-modeling training. American Speech–Language–Hearing Association Convention, San Francisco.
- Marshall, R. C. (2008). The Wal-Mart Task: a clinical measure of executive function. Paper presented at the American Speech-Language-Hearing Association Convention, Chicago, Ill.
- Marshall, R. C., Hamrick, J., Hill, K., Nietsel, A., Dorton, A., Montgomery, E., & Schroering, C. (2009). Training problem-solving strategies in individuals with Alzheimer's disease. Paper presented at the American-Speech-Language-Hearing Association Convention, New Orleans, La.
- Marshall, R. C., Harvey, S. R., Freed, D. B., & Phillips, D. S. (1996). Question asking strategies of aphasic and non-brain-damaged subjects. *Clinical Aphasiology*, 24, 181-192.
- Marshall, R. C., & Karow, C. M. (2013). Rubric scoring of a clinical test of executive functioning. Presented at the 43rd Clinical Aphasiology Conference. Tucson, Ariz.
- Marshall, R. C. & Karow, C. M. (2008) Update on a clinical measure of problem solving. *American Journal of Speech Language Pathology*, 17, 377-388.
- Marshall, R. C., Karow, C. M., Morelli, C. A., Iden, K. K., & Dixon, J. (2003). Problem-solving by traumatically brain injured and neurologically intact subjects on an adaption of the twenty questions test. *Brain Injury*, 17, 589-608.
- Marshall, R. C., Karow, C. M., Morelli, C. A., Iden, K. K., & Dixon, J. (2003). A clinical measure for the assessment of problem solving in brain-injured adults. *American Journal of Speech-Language Pathology*, 12(3), 333-348.

- Marshall, R. C., Karow, C., Morelli, C., Iden, K., Dixon, J., & Cranfill, T. (2004). Effects of interactive strategy modelling training on problem-solving by persons with traumatic brain injury. *Aphasiology*, *18*(8), 659-673.
- Marshall, R. C., McGurk, S. R., Karow, C. M., & Kairy, T. J. (2007). Problem-solving abilities of participants with and without diffuse neurologic involvement. *Aphasiology*, *21*(6-8), 750-762.
- Marshall, R. C., McGurk, S. R., Karow, C. M., Kairy, T. J., & Flashman, L. A. (2006). Performance of subjects with and without severe mental illness on a clinical test of problem solving. *Schizophrenia Research*, *84*(2), 331-344.
- Mcgonigle-Chalmers, M., & Alderson-Day, B. (2010). Free classification as a window on executive functioning in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *40*(7), 844-857.
- McKinney, J. D. (1973). Problem-solving strategies in impulsive and reflective second graders. *Developmental Psychology* *8*(1): 145.
- Minsheu, N., Siegel, D., Goldstein, G., & Weldy, S. (1994). Verbal problem solving in high functioning autistic individuals. *Archives of Clinical Neuropsychology*, *9*, 31-40.
- Mosher, F. A., & Hornsby, J. R. (1966). On asking questions. In J. S. Bruner et al. (Eds.), *Studies in Cognitive Growth* (pp. 86–102). New York: Wiley.
- Oddy, M. (1984). Pedestrian Accidents (Book). *Journal Of Child Psychology & Psychiatry & Allied Disciplines*, *25*(2), 332. doi:10.1111/1469-7610.ep12807716.
- Peck, M.S. (1978). *The road less traveled: A new psychology of love, traditional values and spiritual growth*. Oxford, England: Simon & Schuster.
- Porteus, S. D. (1965). *Porteus Maze Test: Fifty years' application*. Oxford: Pacific Books.
- Porteus, S. D. (1950). *The Porteus Maze Test and intelligence*. Palo Alto, Calif: Pacific Books.
- Prescott, T., Loverso, F., & Selinger, M. (1992). Microcomputer treatment applications in aphasiology. *Aphasiology*, *6*(2), 155-163.
- Prescott, T., Gruber, J., Olson, M., & Fuller, K. (1987). Hanoi revisited. *The Aphasiology Archive*, 249-259.
- Purdy, M. (2002). Executive functioning in aphasia. *Aphasiology*, *16*, 549-557.

- Raven, J., Raven, J.C., & Court, J.H. (1984) *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, Texas: Harcourt Assessment.
- Raven, J. C. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology*, 1-48.
- Saarnio, P. (1993). Question-asking strategies in abstinent alcoholics: A study of learning potential. *Alcohol and Alcoholism*, 28, 585–592.
- Saykin, A.J., Shtasel, D.L., Gur, R.E., Kester, D.B., Mozley, L.H., Stafiniak, P. & Gur, R.C. (1994). Neuropsychological deficits in neuroleptic naive patients with first episode schizophrenia. *Archives of General Psychiatry*, 51, 124-131.
- Scholnick, E. K., & Friedman, S. L. (1993). Planning in context: Developmental and situational considerations. *International Journal of Behavioral Development*, 16(2), 145-167.
- Shallice, T. (1982). Specific impairment of planning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 298 (1089), 199-209. doi:10.1098/rstb.1982.0082. PMID 6125971.
- Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, 114, 727–741.
- Shattuck, P.T., Roux, A.M., Hudson, L.E., Lounds, T.J., Maenner, M.J., & Trani, J.F. (2012) Services for adults with an autism spectrum disorder. *Canadian Journal of Psychiatry*, 57(5), 284–291.
- Sherr, R. L., Rath, J. F., Langenbahn, D. M., Simon, D., Biderman, D., & Diller, L. (1998). Acquired brain damage: Development of the Problem Solving Role-play Test. *Rehabilitation Psychology*, 43, 184
- Siegler, R. S. (1977). The Twenty Questions Game as a form of problem solving. *Child Development*, 48(2), 395-403. doi:10.1111/1467-8624.ep11228091.
- Simmonds, E.P., (1992). The effects of teacher training and implementation of two methods of improving comprehension skills of students with learning disabilities. *Learning Disabilities Research and Practice*, 7, 194-198.
- Simmonds, E. P. (1990). The effectiveness of two methods for teaching a constraint-seeking questioning strategy to students with learning disabilities. *Journal of learning disabilities*, 23(4), 229-233.

- Smith, J.C., Page, J.L., & Marshall, R.C. (2013). Clinical Measure of Problem Solving in Children with and without Autism Spectrum Disorders. Presented at American Speech Language Hearing Association Convention on November 15, 2013.
- Wade, D. T., Legh-Smith, J., & Hewer, R. (1986). Effects of living with and looking after survivors of a stroke. *British Medical Journal (Clinical Research Edition)*, 293(6544), 418.
- Wilson, B. A., Alderman, N., Burgess, P. et al. (1996). *Behavioral Assessment of the Dysexecutive Syndrome*. Bury St. Edmunds, UK: Thames Valley Test.
- Winsler, A., & Naglieri, J. (2003). Overt and covert verbal problem-solving strategies: Developmental trends in use, awareness, and relations with task performance in children aged 5 to 17. *Child Development*, 74(3), 659-678.
- Zelniker, T., Renan, A., Soror, I., & Shavit, Y. (1977). Effect of perceptual processing strategies on problem solving of reflective and impulsive children. *Child Development*, 48, 1436–1442.

Vita

Janice Carter Smith

Education:

- 2001 **Master of Science in Communication Disorders**, University of Kentucky, Lexington, KY.
- 1999 **Bachelor of Health Sciences: Communication Disorders**, University of Kentucky, Lexington, KY.

Professional Positions Held:

- 2008-Present **Faculty Member**: Western Kentucky University, Bowling Green KY.
- 2005- Present **Clinical Supervisor**: Western Kentucky University, Bowling Green KY.
- 2015- Present **Treasurer**: Kentucky Speech Language Hearing Association, Lexington, KY
- 2014- Present **Publications Chair**: Kentucky Speech Language Hearing Association, Lexington, KY
- 2007-2009 **Speech Language Pathologist**: Commonwealth Health Corporation, Bowling Green, KY.
- 2001- 2007 **Speech Language Pathologist**: Butler County Schools, Morgantown, KY.

Scholastic and Professional Honors:

- 2015 **Johanna B. Hancock Award for Outstanding Dedication and Service**: Kentucky Speech Language Hearing Association, Lexington, KY
- 2013 **RCAP Grant Recipient: \$7897.99** to fund research with children with autism spectrum disorders.

Publications & Presentations:

- Smith, J.C. (*in press*). Assessment of problem solving: What does it tell us about pragmatic language? *KSHA Communicator*, 4-5.
- Smith, J.C. (2011, Winter). Public knowledge of autism: A pilot survey in Kentucky. *KSHA Communicator*, 4-5.
- Smith, J.C. (2015). Clinical Measure of Problem Solving with clinical implications for special populations. Presented at Kentucky Speech Language Hearing Association Convention on February 26, 2015.
- Smith, J.C., Page, J.L., & Marshall, R.C. (2013). Clinical Measure of Problem Solving in Children with and without Autism Spectrum Disorders. Presented at American Speech Language Hearing Association Convention on November 15, 2013.