The Pennsylvania State University

The Graduate School

College of Education

DECISION-MAKING BY SCHOOL PSYCHOLOGISTS: USE OF THE REPRESENTATIVENESS HEURISTIC AND IMPORTANCE OF ASSESSMENT DATA IN DETERMINATION OF SPECIAL EDUCATION ELIGIBILITY

A Dissertation in

School Psychology

by

Sharise Mavis Wilson

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2010

The dissertation of Sharise Mavis Wilson was reviewed and approved* by the following:

Beverly J. Vandiver
Associate Professor of Education
Dissertation Adviser
Co-Chair of Committee

James C. DiPerna
Associate Professor of Education
Professor in Charge of the Program of School Psychology
Co-Chair of Committee

Rayne Sperling
Associate Professor of Education

Jolynn Carney
Associate Professor of Education

*Signatures are on file in the Graduate School.

Abstract

The purpose of this project was to explore the decision-making approach and types of data that school psychologists use in determining special education classification. There were three research objectives: (a) to investigate the types of conditions and measures needed to test the use of the representativeness heuristic and assessment data, (b) to assess whether school psychologists use the representativeness heuristic and (c) to examine which types of assessment data are used by school psychologist when cases differ in complexity. Participants in Study 1 were 28 alumni from a graduate school psychology program, who were employed outside of Pennsylvania. In Study 2, participants were 272 school psychologists from the Association of School Psychologists of Pennsylvania (ASPP). All participants were certified in school psychology and had either a Master's or Doctoral degree. For both studies, a post-test only analogue design was used. Three student summaries, embedded with assessment data, were systematically manipulated to vary in complexity level (Learning Disability, Normal, and Complicated) and either randomly or systematically assigned to be reviewed by participants. The Decision-Making Questionnaire, developed by the author, was used to gather data about the use of the representativeness heuristic and assessment data by school psychologists. The representativeness heuristic was measured in four ways based on the research design of previous studies. Descriptive statistics, *t*-tests, contingency tables, ANOVAs, and MANOVAs with descriptive discriminant analysis were used to answer a total of eight research questions. Overall, findings suggested that the representativeness heuristic played a role in the decision-making of school psychologists, although depending on how it was measured, its use varied with the complexity of the student data summary. The findings also indicated that school psychologists were relatively consistent in rating the following data important in making eligibility decisions: achievement scores, IQ score, response-to-classroom modification, state criteria, and assessments of behavior. Additionally, focus on progress-monitoring data and referral information appears to influence the propensity to make classification errors. Lastly, the theory of bounded rationality is supported by results regarding the use of the representativeness heuristic and assessment data. Future research is needed to extend the generalizability of findings and to assess reliable ways of measuring heuristics.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Not surprisingly, as I sit here, I have tears in my eyes. My tears stream not only from the pride that I feel, but also from knowing that it is only by the strength that has been bestowed upon me by God, my family and friends, and my dissertation committee that I have reached this summit.

Thank you very much to my committee for your hard work on this project. Dr. Beverly Vandiver, my advisor, it is your wisdom, time, and patience that helped me to turn this "Little D" into a great work. How you managed to work through my very rough drafts and whimpering moments is still beyond me. You have challenged me, molded me, and eventually set me free. Dr. James DiPerna, I appreciate the investment that you have had in me throughout the years. You have been integral in my development from being a first year student, to pursuing a Master's degree, to now being a doctor. Dr. JoLynn Carney, I will never forget your endless "warm wishes", support, and the encouragement that you have sent my way. Dr. Rayne Sperling, thank you for challenging me to think above and beyond what I am used to.

My dissertation is dedicated to my family, who has earned this doctorate with me. To my mommy, Lenise Walker-Wilson, I love you with all of my heart and while I struggled through many aspects of this program, you were there to tell me how much you were proud of me and to remind me that I can do all that I put my mind and heart to. To my sisters, I love you dearly! Shaleah, I thank you for listening to my endless gripes about *everything*. I have always been amazed at how well you were able to keep me abreast of what was going on in the outside world--while keeping the outside world abreast of what was going on with me. Sharell, no matter what we say, your sympathy towards me has been amazing. Thanks for always making me feel like my success was certain. Daddy, Carlos A. Wilson, thanks for the visits, understanding that I couldn't be as hospitable as I would have liked, and still telling me how proud you were-I love you. To all of my family, aunts, uncles, cousins, and godmother thanks for always believing in me and understanding when I couldn't always be around for your miracles and milestones. Do know that the traditions and values that have been instilled in me as a member of the Walker/Wilson family have enabled me to be who I am at this very moment.

Where would I be without my extended family? Dr. Cherise Y. Hamblin, we claimed it while at the Bronx High School of Science, and we did it! Though distance often keeps us apart, you are always "right here" for me. I love you for that as well as for offering me a different perspective of things. Talia Nobbee, my cheerleader, we have been through it all and the best is yet to come. Your "you would never believes" helped me to come up for air when I needed it the most. Dr. Nathaniel Frederick, II, thanks for the continual motivation, check-ins, and endless empathy. You always have just the right words to say, even when I cannot find them myself. Dr. Amy S. Walker, thanks for being a Special Soror as well as a great friend. We have shared many wonderful moments at Penn State, which will never be forgotten. Keisha Johnson, from prayers, to work sessions, to timelines, you have motivated me to work diligently on completing this dissertation-Thanks. Phyllisa Deroze, your camaraderie and support has been utterly appreciated over the past few months. To my cohort (Katie, Maria, Terry, & Kasey), we embarked on this journey together not knowing what was to come. I will never forget the numerous times that we have spent over the past several years studying, commiserating, and developing as school psychologists. Your words of encouragement and enthusiasm about my endeavors have been remarkable. To all of my friends and colleagues, present and past, mentioned or not, none of you will ever be forgotten. From PS 346 to PSU, everyone has left a footprint in my path. I believe that every encounter has purposely been sent from God to impact who I am and who I will become as Dr. Sharise M. Wilson.

Introduction and Literature Review

One of the most important decision-making processes for school administrators, parents, and students is determining whether a student qualifies for special education services (Frey, 2002; Pierangelo & Giuliani, 2007). Special education eligibility determines not only the type of instruction that a student will receive, but also affects school services that families qualify for, as well as the funding that schools receive for programming. The process used to determine special education eligibility involves several steps and a number of people. School psychologists, however, have often been viewed as the gatekeepers to special education and as essential to determining special education eligibility. A challenge faced by school psychologists during this process is to sort through the numerous data that are collected about the referred student. This process can be difficult due to (a) the amount of information, (b) possible inconsistencies among the information, and (c) philosophical differences over the value of the information (e.g., standardized vs. unstandardardized data and parental versus teacher responses). This difficulty is intensified by the changing demands of special education law. One of the more recent changes is the consideration of response to intervention (RTI) when determining the presence of a learning disability.

Prior research has suggested that various practitioners use heuristics to facilitate the processing of cognitively demanding tasks (Cioffi & Markham, 1997; Dewhurst et al., 2007; Garb, 1996). It is likely that school psychologists also use heuristics to manage the complex nature of making special education decisions. It is also possible that certain assessment data will be weighed more heavily than others during the evaluation process. The purpose of the current research was to explore measures of the use of the representativeness heuristic and assessment data and to investigate whether school psychologists (a) use the representativeness heuristic and

(b) differentially weigh assessment data when determining special education eligibility. To contextualize the study, the subsequent literature review covers the following: (a) a brief review of the special education decision process, (b) the underlying theory and description of heuristics, (c) research on the use of heuristics by practitioners, and (d) prior studies about the importance of assessment data.

Special Education Decision Process

Best practice suggests that a team should be involved in the process to determine special education eligibility (Pierangelo & Giuliani, 2007; Reschly, 1996; Salvia & Ysseldyke, 2004). The team should consist of parents, teachers, school psychologists, other services providers, special education administrators or designees, and, when of appropriate age, the student (Catterall, 1972; Mercer, 1975; Salvia & Ysseldyke, 2004). School psychologists, as the team members with expertise in assessment, are often considered to be one of the most influential individuals in multidisciplinary team placement decisions (Gilliam & Coleman, 1981; Salvia & Ysseldyke, 2004).

Federal law (Individuals with Disabilities Education Act [IDEA], 1997, 2004) and their corresponding regulations specify the types of information that should be collected and criteria to be met for a student to be classified as having a special education disability. However, implementation of these regulations is established at the state level (Bellis, 2002). Regardless of the state, the stages of the process required prior to a student receiving special education services are generally the same: (a) prereferral, (b) referral, (c) formal evaluation, (d) eligibility determination, and (e) specification of services, if found eligible. The formal evaluation phase is the most applicable to the current study and is briefly summarized below. For a more comprehensive discussion of the special education decision-making process, see Salvia and

Ysseldyke (2004) or Reschly (1981, 1996).

Following the prereferral phase, where concerns about the student are clarified, and the referral phase, where presenting problems are documented, the formal evaluation phase begins. The formal evaluation is conducted by a school psychologist and individualized to the concerns expressed in the referral (Salvia & Ysseldyke, 2004). Multiple forms of data are gathered to obtain a comprehensive understanding of the student's strengths and weaknesses so that a decision can be made (Pierangelo & Giuliani, 2007; Salvia & Ysseldyke, 2004). School psychologists seek information from school records such as medical history, vision and hearing screenings, disciplinary reports, and the student's grades. Additional information is requested from parents and teachers, such as background information about the student (e.g., developmental milestones, interaction with family and peers, and family history of educational disabilities), input on the student's socio-emotional functioning, and the student's academic history (Hannaford, Simon, & Ellis, 1975; Matuszek & Oakland, 1979; Pierangelo & Giuliani, 2007).

Perhaps the most important data gathered by the school psychologist are the direct assessment of student skills (Salvia & Ysseldyke, 2004), which occurs through classroom observations and the administration of standardized achievement and cognitive tests. Cognitive measures, which usually result in an IQ score, are used to discern the student's current cognitive skill or ability level. Additionally, achievement tests are utilized to determine how well the student is performing academically in relation to same age or grade level peers (Hannaford et al., 1975; Matuszek & Oakland, 1979; Pierangelo & Giuliani, 2007).

Once the formal evaluation is complete, the school psychologist writes a psychoeducational or evaluation report (Salvia & Ysseldyke, 2004). Following the creation of the report, a placement conference (also called a multidisciplinary team meeting, IEP meeting, or staffing) is held. At the meeting, a decision is made about the child's qualification for special education services.

Until the reauthorization of IDEA in 2004, an aptitude-achievement discrepancy approach (assessing the difference between observed achievement and expected achievement based on ability level; Reynolds, 1985) had been the predominant model for determining a learning disability. However, it is also possible for the evaluation team to use the response to intervention model (RTI; Kovaleski, 2007) to identify a child with a learning disability. Under this approach, determination for special education eligibility is based on evaluations of how well the student has progressed after evidence-based interventions have been implemented in the regular education setting. For a more thorough description of RTI, readers are referred to IDEA (2004) and the works of Shinn (2007), Kovaleski (2007), and Batsche, Kavale, and Kovaleski (2006).

Overall, the process used to evaluate a student's need for special education services is straightforward, but not simple. The abundance and integration of data collected in addition to potential discrepancies in the data may present challenges for a school psychologist in making decisions about special education eligibility. This challenge is intensified by inconsistent decision rules, inconsistent theoretical orientations, inconstancy of diagnostic style, inadequate instrumentation, and inconsistent weighing of diagnostic cues (Davidow & Levinson, 1993).

Bounded Rationality

Bounded rationality is a theory that may be useful in explaining what happens when individuals are faced with challenging and cognitively demanding tasks, such as determining eligibility for special education. Bounded rationality refers to how people make decisions quickly and efficiently to meet specific goals in the presence of both internal and external demands (Gigerenzer, 2001). A premise of bounded rationality is that decision makers have information processing limitations and are often limited by time constraints as well as mental resources. Another premise of bounded rationality is that people often weigh the importance of information without much forethought. The construction of preferences depends on the context of the situation (e.g., the order in which information is presented, the format of the information presented, and similarity in decision options) and which factors are more salient to the decision maker.

Some research has provided support for the theory of bounded rationality (Broder, 2000; Gigerenzer & Goldstein, 1996; Gilovich, Griffin, & Kahneman, 2002; Newell & Shanks, 2003). Gigerenzer and Goldstein (1996) used a computer simulation to present participants with questions (e.g., who to treat first in the emergency room: the 80- year-old heart attack victim or the 16-year-old car accident victim) to which they could only respond to by making inferences. The use of bounded rationality was compared to the use of integration algorithms. Bounded rationality was operationalized to occur when participants focused on a subset of cues to help them make an inference. In contrast, the use of integration algorithms was defined by individuals using every possible cue in order to make an inference. It was found that when individuals used bounded rationality versus integration algorithms, their decisions were made faster and with as much accuracy. Others have found that the use of bounded rationality is common, but that it is not used the same way in all contexts (Broder, 2000; Newell & Shanks, 2003).

Heuristics and Decision-Making

The theory of bounded rationality is the foundation for the construct of heuristics (Gigerenzer, 2001; Payne, Bettman, & Luce, 1998; Simon, 1990). Because individuals have limited cognitive resources at any time, they are likely to use heuristics as a result of these limitations. While some researchers have simply called heuristics, “rules of thumb” or “shortcuts,” Simon, the father of heuristics research in judgment and decision-making, described heuristics as methods for arriving at solutions with little computation or effort (Shah & Oppenheimer, 2008; Simon, 1990). Similar to Simon’s definition, heuristics can also be defined as deliberate or unintentional strategies that individuals use to make decisions by relying on perceptions and thoughts (Tversky & Kahneman, 1983). Heuristics can lead to either accurate or inaccurate results (Simon, 1990; Tversky & Kahneman, 1983).

Researchers also have identified when a heuristic is being used and identified various types of heuristics. There are two general patterns when heuristics are used. First, heuristics are used when an individual does not follow a standard decision-making process, but instead skips steps or uses a process that has not formally been accepted by the discipline (Gigerenzer & Goldstein, 1996; Hammond, 1996; Hogarth & Karelaia, 2007; Tversky & Kahneman, 1974; 1983). Second, heuristics are used when the rendered decision does not fall in line with statistical findings or empirically established theory.

The three most prominent heuristics in decision-making research, highlighted by Tversky and Kahneman, (1974) are availability, anchoring and adjustment, and representativeness. The availability heuristic is based on what the clinician has encountered previously and is in use when judgments are based on accessibility of information and ease of recall (Tversky & Kahneman, 1973). This type of heuristic is usually used when individuals are expected to predict the likelihood or frequency of an occurrence or event. For example, school psychologists might overestimate the frequency with which RTI is used if they can easily think of many speakers who recently presented their school district's implementation of RTI at a conference.

In contrast to the availability heuristic, the anchoring and adjustment heuristic refers to judgments that are based on information that occurs first. Earlier estimates are adjusted or altered to yield a final answer based on a salient value. This heuristic is typically used when one is estimating a value, size, or quantity, and is often applied when one is making a series of numerical estimates or predictions (Ellis, Robbins, Schult, Ladany, & Banker, 1990; Tversky & Kahneman, 1974). For example, if a clinician is rating a client's progress after several sessions and uses previous ratings to help determine a rating for the client's current progress, the clinician is considered to be using the anchoring and adjustment heuristic.

The last classic heuristic is the representativeness heuristic, which refers to judgments made based on how similar an individual is to a prototype. A prototype can be considered a model or the most typical example of an object or situation. The representativeness heuristic, like the other heuristics, can be used to make predictions and estimations, but is most useful when an individual is placing an object, item, or person into a category (Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1974; 1983). A limitation of the representativeness heuristic is that base rates are often ignored or underused (Gigerenzer, 2001). Base rates refer to the chance of an outcome occurring or the prevalence in the population. Disregard of base rates can occur when a clinician determines a diagnosis for a client without considering how frequently the selected diagnosis actually occurs in the population. An example of the representativeness heuristic is when a school psychologist uses the prototype of a student diagnosed with autism (Student *A*) to determine the diagnosis of another student (Student *B*).

The school psychologist may assess how similar Student *B's* characteristics are to Student *A*. If Student *B* appears to be similar to Student *A*, then the school psychologist is likely to diagnose Student *B* with autism without consideration of how often autism occurs in school-age children.

Beyond its use in making classification decisions, the representativeness heuristic has also been used more frequently than the other two heuristics (Cioffi & Markham, 1997; Czaczkes & Ganzach, 1996). Given the purpose and prevalence of the representativeness heuristic, the use of this heuristic by school psychologists was a primary focus of the current research.

When discussing the use of heuristics, it is also necessary to consider their effects. The representativeness heuristic, like heuristics in general, may be used to explain both optimal and erroneous decisions (Simon, 1990; Tversky & Kahneman, 1974; 1983). A potential benefit of using the representativeness heuristic is that it can shorten the time spent on making a decision by limiting the number of data that are considered. Another benefit is that consistency in decisions can be improved if the decision maker frequently refers to the same set of prototypes when making decisions. On the other hand, the representativeness heuristic can lead to errors if the decision maker's experiences with members of the population are not representative of that population and the decision maker incorrectly ascribes some characteristic to the entire population. This error would result in decisions that are not consistent with other clinicians or published decision criteria. In the review of literature about the use of the representativeness heuristic by practitioners (e.g., Dumont, 1993; Garb, 1996), such errors have been noted most frequently.

Use of the Representativeness Heuristic by Practitioners

Dumont (1993) described a longstanding error in psychiatric diagnosis, which was due to clinical psychologists succumbing to the representativeness heuristic. For many years clinicians labeled Tourettes Disorder as a psychological disorder and attempted to treat it with psychoanalysis because of its symptoms of compulsive repetition of obscene words, grunts, tics, twitches, and echolalia. Dumont contended that clinical psychologists used the representativeness heuristic by comparing Tourettes Disorder to other psychological disorders. This error and use of the representativeness heuristic was eventually brought to light when Tourettes Disorder was discovered to be a neurological disorder; thus, psychoanalysis was not effective and needed to be replaced with medical treatments.

To investigate clinicians' use of the representativeness heuristic, Garb (1996) presented 67 clinical psychologists and clinical psychology interns with a case study. The case study, based on DSM-III-R criteria, depicted a patient who had some borderline personality features, but with careful examination should have been diagnosed as having antisocial personality disorder. After reading the case study, participants diagnosed the patient in the case. A 10-point likelihood rating scale was used to indicate how likely participants believed the patient to have each of four diagnoses (antisocial personality, histrionic, narcissistic, and borderline personality disorder). Garb assumed that when clinicians made a diagnosis of borderline personality disorder, they were basing their decisions on prototypes, and thereby using the representativeness heuristic. The researcher also inferred that the representativeness heuristic was not used when clinicians attended to explicit criteria listed in the DSM-III-R and diagnosed the patient with the correct diagnosis of antisocial personality disorder. Garb found that only 18 of the 67 clinicians made a diagnosis consistent with the DSM-III-R criteria, suggesting that the representativeness heuristic was, in fact, used by 73% of the clinicians.

In the same study, Garb (1996, p. 273) also examined whether a *high correlation between* *judges' likelihood ratings and similarity ratings* was indicative of the representativeness heuristic. Likelihood ratings indicated the extent to which clinicians believed that the client had one of the four personality disorders. Similarity ratings indicated the extent in which clinicians perceived the patient to be similar to a typical person with each of the four disorders. A similarity rating of 0 indicated that the psychologist or intern thought that the patient was not similar to the "typical" person with the disorder, while a rating of 10 indicated that the psychologist or intern perceived the patient's characteristics to exemplify the disorder. Garb found that likelihood ratings and similarity ratings were highly correlated (r [65] = .92, $p < .05$). Garb used this finding to conclude that the representativeness heuristic occurred.

A year after Garb's (1996) study, Cioffi and Markham (1997) used similar methodology to investigate the occurrence of the representativeness heuristic in the medical field. Participants in this study were 30 midwives who were given simulated patient assessment situations of high and low complexity child births and asked to make a decision about the cases. Low complexity cases had all relevant information available with a predictable relationship between the signs and symptoms, while high complexity cases lacked all relevant information and had an unpredictable relationship between the signs and symptoms. The midwives were instructed to ask questions of the investigator until they were ready to provide a diagnosis. After rendering a diagnosis, the midwives completed a short report form and provided estimates of base rates for the incidence of the clinical conditions (e.g., hemorrhaging) that were presented to them. Midwives reported that their decisions depended on how closely the patients' symptoms resembled archetypes, suggesting the use of the representativeness heuristic. Cioffi and Markham concluded that the midwives were using the representativeness heuristic not only when they used an archetype in diagnosing patients, but also when they disregarded base rates.

Cioffi and Markham (1997) further analyzed their data to (a) compare the use of the representativeness heuristic to the use of the availability heuristic and anchoring and adjustment heuristic, and (b) to discern in which situations the representativeness heuristic emerged the most. Findings based on frequency counts and ANOVA indicated that the representativeness heuristic was used more often than the other two heuristics in diagnosing patients, F (2) = 73.83, $p < .001$. Additionally, the representativeness heuristic was more frequently used in more complex situations than in less complex situations, t (29) = -2.93, $p < .001$.

Dewhurst et al. (2007) provided further support for the use of the representativeness heuristic by practitioners. Participants, students in genetic counseling (N = 89) and genetic counselors (N = 156), were asked to determine the most likely boy-girl birth order sequence among five different families. Participants reviewed descriptions of families that had between 3 and 7 children. The presence of the representativeness heuristic was based on when participants reported specific sequences that would not have been possible had participants attended to rules of probability. Participants were described as not using the representativeness heuristic if they were correct in reporting that the birth order sequence containing the fewest number of children was the most likely to occur. Data analyses indicated that about 35% of the students and 47% of the genetic counselors correctly applied the rules of probability and selected the correct response. Thus, between 53 - 65% of the participants appeared to use the representativeness heuristic.

Overall, prior research (Arkes, Saville, Wortmann, & Harkness, 1981; Bieri, Orcutt, & Leaman, 1963; Detmer, Fryback, & Gassner, 1978; Dumont, 1993; Friedlander & Phillips, 1984; Friedlander & Stockman, 1983; Garb, 1996; Gauron & Dickinson, 1969) indicates that heuristics are used by practitioners when they are presented with a complex situation. When specifically considering the representativeness heuristic, it often is used by practitioners who have the responsibility of making classification decisions. One way of discerning that the representativeness heuristic has been used is when the decision maker reports a correspondence between (a) the likelihood that a person or objects fits a category and (b) the similarity between the person or object and some sort of prototype. Another way that the representativeness heuristic occurs is when the decision maker reports that a prototype was important to the decision. Lastly, a possible indicator of the representativeness heuristic is when the decision maker fails to consider established criteria or base rates while making a decision.

Previous researchers have used decision errors as indicators of the representativeness heuristic. For example, Dumont (1993) used the erroneous classification of Tourettes Disorder as a psychological disorder as an indicator that the representativeness heuristic occurred. However, Simon (1990) noted that, while heuristics can sometimes lead to suboptimal outcomes, there are other times they can result in accurate and satisfactory decisions. Therefore, one limitation of previous research is that it failed to look for indicators of the representativeness heuristic regardless of misclassification.

A second limitation of previous studies is the difficulty in generalizing results to school psychology. Participants in prior research were mostly from the medical field. Practicing school psychologists, however, are unique professionals in that they focus on both mental health as well as academic domains. While some researchers (Davidow & Levinson, 1993; Fagley, 1988) have suggested that the representativeness heuristic is applicable to school psychology, its use has not been tested. Given the lack of research on the use of heuristics in the field of school psychology, Fagley (1988) and Barnett (1988), independently, called for more research to consider the extent to which school psychologists utilize shortcuts in making decisions about special education placements. Although such a call was made again in 1993 (Davidow & Levinson), a search of the school psychology literature indicated that this call has not been answered. It is necessary to assess whether there are conditions that facilitate the use of the representativeness heuristic in the practice of school psychology as it does with other professions.

A final limitation of previous research is in the design of some of the studies. Some researchers (e.g., Dumont, 1993) have not directly assessed the use of heuristics. Instead, they have analyzed archival data, which are gathered and interpreted after the decision has occurred. Thus, conclusions based on archival data need to be viewed with caution because they tend to be based more on recollection than prospective assessment of decision-making. By reviewing prerecorded data, there has been no control over standardization and factors that could have affected the outcome of the study (Gall, Gall, & Borg, 2003). There is a need to explore more appropriate measures of heuristics.

Influence of Assessment Data

No studies that have directly assessed the use of heuristics by school psychologists have been found, but studies have been conducted to investigate other key factors influencing the decision-making of school psychologists. Researchers (e.g., Knoff, 1983; Salvia & Ysseldyke, 1978; Thurlow & Ysseldyke, 1980) have examined the extent to which school psychologists and others who make special education decisions are influenced by data gathered during the assessment process. Such data (also referred to here as assessment data) include intelligence quotient (IQ), achievement, socio-emotional functioning, and sociocultural variables.

IQ and achievement. Prior research has found that the most used data by school psychologists to make decisions regarding special education eligibility are IQ and academic achievement (Salvia & Ysseldyke, 1978; Thurlow & Ysseldyke, 1980; Ysseldyke, Algozzine, Regan, & Mcgue, 1981). Thurlow and Ysseldyke (1980) used four designs to compare the influence of IQ and achievement scores to seven other pieces of information (e.g., sex, attractiveness, socio-economic status, adaptive behavior, visual motor abilities, language proficiency, and behavior ratings). The four designs used were (a) a computer simulation of assessment, (b) self-reports following actual team meetings, (c) self-report based on reflection of prior experiences, and (d) self-report following a presentation on current issues in decision making and assessment. Participants in the study were 17 special education directors, 25 school psychologists, 52 regular education teachers, 50 special education teachers, and 15 support personnel. Thurlow and Ysseldyke indicated that IQ and achievement scores were rated as more important than the personal characteristics (e.g., sex, attractiveness, and socio-economic status) of the student during determination of eligibility for special education services. Follow-up analyses after an ANOVA revealed that sex, socio-economic status, and attractiveness were not linked with special education eligibility. IQ and achievement scores also were found to be of greater importance on placement decisions than language proficiency, adaptive behavior, and visual motor abilities. In addition, achievement scores were considered more important than scores derived from behavioral rating scales.

In contrast, Knoff (1983) reported that IQ scores do not always have a disproportional influence on the educational placement of students who potentially have mental retardation. Knoff surveyed four independent samples of school psychology practitioners (n = 20), special education practitioners (n = 20), school psychology graduate trainees (n = 20), and special education trainees (n = 20). Participants were presented with data one piece at a time, receiving either IQ data before adaptive behavior data or adaptive behavior data before IQ information. After reviewing each piece of information, participants were asked to make a placement decision for the student using a rating scale from 1 (*full time special class*) to 10 (*full time regular class* *with no change in instruction*). Results indicated that regardless of profession or status (e.g., practitioner or graduate student), whether IQ scores or adaptive measures were most influential to the placement of the student depended on which data were presented first, with earlier data being more influential.

Referral question. Another common factor thought to be influential to school psychologists' decision-making is the referral question (Della Toffalo & Pedersen, 2005; DeMesquita, 1987; Ysseldyke et al., 1981). Ysseldyke, Algozzine, Regan and McGue (1981) assessed whether reason for referral influenced diagnostic decisions, using a computer simulation that allowed 159 participants to select their assessment instruments. The referral question was either behavioral or academic in content. When school psychologists (n = 25), regular educators (n = 52), special education teachers (n = 50), administrators (n = 17), and support personnel (n = 15) were given information that differed only on the stated reason for referral, diagnostic decisions were affected. Furthermore, even though students had average behavior as indicated by standardized assessments, students were likely to be diagnosed with an emotional disturbance if the referral reason reflected a behavioral concern.

Similarly, Ysseldyke and Algozzine (1983) suggested that an academic versus a behavioral referral question influenced decision makers' likelihood of recommending a student for an emotional disturbance. Specifically, the chances of a student being diagnosed with an emotional disturbance increased when the student was referred for behavioral difficulties. O'Reilly, Northcraft, and Sabers (1989) also found when school psychologists were provided with information that differed only on whether the student was referred for gifted placement versus support for learning disability, school psychologists' recommendations for placement were completely determined by the referral information.

Socio-emotional status. Socio-emotional status is another factor thought to influence special education decisions. In an analogue study, Della Toffalo and Pedersen (2005) presented 215 school psychologists with cases of students who met eligibility as well as cases of students who did not meet eligibility for special education services. Participants were instructed to use a Likert-type scale to indicate the extent to which the presence of an Oppositional Defiance Disorder diagnosis, and the student's age, gender, IQ score, grades, behavior rating scale scores, and achievement scores influenced their decision regarding eligibility. Multiple regression analyses revealed that the presence of a psychiatric diagnosis predicted recommendations for special education placement, F (1, 213) = 10.52, p = .001. Findings indicated that the order (from strongest to weakest impact) of student data on decisions about emotional disturbance was (a) behavioral rating scales, (b) presence of diagnosis, (c) achievement test data, and (d) IQ. This finding was in contrast with prior research (Salvia & Ysseldyke, 1978; Thurlow & Ysseldyke, 1980; Ysseldyke et al., 1981) that supported IQ and achievement as the most influential determinants for special education eligibility.

Similar to psychiatric information, emotional indicators also have been found to influence eligibility for emotional disturbance. Knoff (1983) administered questionnaires to a total of 80 school psychologists and special educators, who were asked to rate the importance of several pieces of information when making a decision about a student's need for special education. The four most important pieces of information the practitioners and trainers reported using were (a) classroom observations, (b) assessment of language proficiency, (c) interview with child, and (d) emotional indicators. Contrary to Knoff's findings, Brooks (1979) contended in a position paper that emotional and affective information do not receive much consideration when assessing children and adolescents with learning and adjustment problems. Brooks' assertion came from examining psychoeducational assessment reports and noticing that specialists never mentioned either the child's emotional status or its impact on the evaluation. Instead, specialists tended to focus on standardized cognitive test scores.

Sociocultural variables. Sociocultural variables, such as socioeconomic status, gender, and age (Craig, Kaskowitz, & Malgoire, 1978; Frey, 2002) have been found to be important when assessing students with an emotional disturbance. Craig et al. (1978) looked at the characteristics of 7,000 students recommended for special education by using indicators derived from parent and teacher recommendations, medical examinations, school behaviors, test scores, and developmental histories. Variations were investigated for disabilities, including emotional disturbance and mental retardation. Relevant findings were that teachers were more likely to place students who were from lower socio-economic status, African American, and male in classes for students with emotional disturbance or mental retardation.

Frey (2002) also investigated the influence of sociocultural characteristics on disability determination by surveying 350 special educators who read one of four descriptive case studies. Each case study described a fourth-grade, male student, who differed on socio-economic status (high or low) and race (African American or Caucasian). The special educators were then instructed to determine the student’s educational placement. Findings of the logistic regression analysis were that recommendations from special educators who read descriptions of African American students were not statistically significant from that of special educators who read descriptions of Caucasian students. Therefore, with socio-economic status controlled for, race did not influence teachers’ recommendations. However, students from low socio-economic status were five times more likely to be recommended for more restrictive placements than students from high SES.

Conclusion. Overall, results from prior research has supported that certain assessment data, such as IQ, achievement, referral question, socio-emotional indicators, and sociocultural factors, might be more important than other assessment data. This is similar to research about heuristics that has suggested that certain assessment factors (e.g., established criteria, base rates) are often ignored by professionals. However, one of the most obvious drawbacks of the studies reviewed about the influence of assessment data on special education decisions is that they are dated (e.g., Knoff, 1983; Salvia & Ysseldyke, 1978; Thurlow & Ysseldyke, 1980; Ysseldyke et al., 1981), with the notable exceptions of the work of Della Toffalo and Pedersen (2005), and Frey (2002). During the last decade, changes have occurred within special education policy and practice, such as the increased emphasis on response to intervention, instructional methods, accountability, and parental involvement (Salvia & Ysseldyke, 2004; Seltzer, 1998; Washburn-Moses, 2005). Therefore findings obtained more than a decade ago about the influence of certain student characteristics on special education decisions should be interpreted with caution because contemporary special education personnel are trained differently, which is shifting how they practice. Given the changes, it has yet to be seen whether factors, such as IQ and achievement, continue to be as influential as they were.

Besides being dated, findings from prior research about how assessment data influence special education decisions were often based on reports or perceptions of undergraduate students, graduate students, learning support teachers, and regular education teachers, rather than practicing school psychologists (e.g., Frey, 2002; Knoff, 1983; Salvia & Ysseldyke, 1978; Thurlow & Ysseldyke, 1980; Ysseldyke et al., 1981). Given the credentialing requirements for practicing school psychologists, results based on how college and graduate students' decisions were influenced by characteristics of students referred for special education might not be applicable to school psychologists' decisions.

Another limitation of prior research is the exclusion of some of the most readily available and common types of collected assessment data. In prior research, predictors have included IQ, standardized achievement scores, report card grades, and standardized behavioral ratings (Della Toffalo & Pedersen, 2005, Knoff, 1983; Salvia & Ysseldyke, 1978; Thurlow & Ysseldyke, 1980; Ysseldyke et al., 1981). However, not all of these variables were studied simultaneously. Furthermore, some important assessment data (e.g., response to modification data and parents' descriptive comments) have not been included in studies examining the influence of assessment data on special education decision-making.

Current Project

Purpose of the Research

Prior research on practitioners' use of heuristics has primarily focused on physicians, nurses, and clinical psychologists (Detmer, Fryback, & Gassner, 1978). Additionally, measures of heuristics used by researchers vary greatly. As such, the purpose of the current project was to explore the measures and conditions used to assess heuristics and to examine their use in the field of school psychology. In this research, heuristics were defined as deliberate or unintentional strategies or shortcuts that individuals use to make estimations or predictions by relying on perceptions and thoughts (Tversky & Kahneman, 1983).

According to previous investigators (e.g., Cioffi et al., 1997; Garb, 1996; Payne et al., 1998), heuristics are used when one needs to reduce the cognitive demand of a complex task (Davidow & Levinson, 1993). Cioffi and colleagues described a complex situation as one in which ambiguity exists and decisions are not easily made. Complex conditions also have been described as situations where several possible solutions exist (Payne et al., 1998). For the current project, complexity was defined as the ease with which it could be decided that a given student has a certain classification. Is it easy to determine that the student does not have a disability (low complexity) or is it easy to classify the student as having a disability when in fact it is not true (high complexity)? In the current project, the terms classification and identification are used interchangeably and refer to the 13 categories (e.g., nonexceptional, mental retardation, & specific learning disability) recognized by the Pennsylvania State Board of Education as areas of school-age special education qualification (State Board of Education, 2008). These categories are described in further detail in the Method section of Study 1.

For the purpose of the research, the focus was on the representativeness heuristic. The representativeness heuristic operates when judgments are based on how *similar* an individual is to the *prototype* that the assessor has in mind (Garb, 1996; Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1983). The representativeness heuristic also involves disregard for established criteria and base rates (Cioffi & Markham, 1997; Dewhurst et al., 2007) and is most likely used during complex tasks when an individual is placing an object, item, or person into a category (Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1983). For these reasons, the representativeness heuristic was thought to fit well into the special education decision-making process, in which school psychologists often have the responsibility of sorting through a plethora of assessment data to determine a classification for a student who is referred. While the applicability of the representativeness heuristic to school psychology has been suggested (Davidow & Levinson, 1993; Fagley, 1988), its use has not been tested previously. Plausibility of the representativeness heuristic operating during the special education decision-making process comes from prior research with practitioners from professions outside of school psychology (Cioffi & Markham, 1997; Dewhurst et al., 2007; Garb, 1996).

Somewhat related to the use of heuristics is the reported under- and over-weighing of certain assessment data. Therefore, another purpose of the current project was to examine which assessment data are most influential in school psychologists' classification decisions. Prior researchers have indicated that school psychologists collect an abundance of assessment data in determining a student's eligibility for special education; however, some information, such as IQ, achievement, referral question, socio-emotional, and sociocultural factors, might be more important to special education eligibility decisions than other assessment data (Knoff, 1983; Salvia & Ysseldyke, 1978; Thurlow & Ysseldyke, 1980; Ysseldyke et al., 1981). However, prior researchers have excluded some of the most available and common types of assessment data (e.g., parent's comments about student's behavior, hearing results, & vision results) when studying antecedents to special education decisions. Over the past 10 years research on the influence of assessment data has been rare. Recent changes have occurred in special education practice and law, including increased acceptance of response to intervention when determining the need for special education services, frequent critiques of instructional techniques, and increased demands that parents become involved in their children's education (Salvia & Ysseldyke, 2004; Seltzer, 1998; Washburn-Moses, 2005). In light of these recent changes and commonly collected assessment data, the importance of typical assessment data was also examined in addition to the use of the representativeness heuristic.

Definition of Key Terms

For the remainder of this document, the following terms are defined as follows:

1. Accuracy of classification: The extent in which school psychologists selected the correct classification for the student.
2. Assessment data: Information about a student collected by school personnel, especially school psychologists, during a psychodeducational assessment, which includes information such as IQ, academic achievement, grades, and medical history.
3. Classification: Identification of a student for special education services; the 13 categories (e.g., nonexceptional, mental retardation, specific learning disability) recognized by the Pennsylvania State Board of Education as areas of school-age special education qualification.
4. Client-prototype similarity: The extent in which an individual believed that a target was like a prototype.
5. Complexity: The level of ambiguity and ease in which decisions could be made (Cioffi &

Markham, 1997) or where several possible solutions exist (Payne et al., 1998). For the purpose of this research, complexity was the ease in which it could be decided that the student described in a given student summary had a certain classification.

6. Data summary: A description and summary of assessment data for a hypothetical student referred to a school psychologist for a psychoeducational assessment.
7. Likelihood of accurate classification: The extent in which a school psychologist believed that the student in a student summary had the classification that the school psychologist selected.
8. Prevalence: The chance of an outcome occurring; a base rate; how likely a situation was in the population.
9. Prototype: A model or the most typical example of an object or situation.
10. Representativeness heuristic: A method for arriving at a solution with diminished effort in which judgments were based on how similar an individual was to a prototype.

Project Overview

The current research involved two studies. Study 1 was conducted to establish a viable research design and strong measures to tap the representativeness heuristic and decision-making questions. Based on the findings from Study 1, Study 2 was designed and conducted. To increase clarity of the studies, the research questions are not presented at the end of this Literature Review as is typically done. Rather, the research questions are presented with each study.

Study 1

Purpose

Study 1 was a preliminary study. The purpose was to assess whether the data summaries were functioning as intended. Furthermore, Study 1 was conducted to check the clarity and adequacy of the questions created to tap the representativeness heuristic and the importance of assessment data.

Research Questions

1. How could data summaries be understandable to school psychologists yet vary in ambiguity of data presented?
2. What types of conditions and questions were needed to measure the use of the representativeness heuristic by school psychologists?
3. What types of assessment data do school psychologists typically use when determining special education eligibility?

Method

Research Design

Study 1 was a post-test only experimental analogue design (Gall et al., 2003). Three student summaries with assessment data were systematically manipulated to vary in complexity level (Learning Disabled [LD], Normal, and Complicated). School psychologists were randomly assigned to review one of the three student summaries.

Participants

The Study 1 sample was recruited from 60 alumni from a graduate school psychology program in Pennsylvania. All participants were employed outside of Pennsylvania at the time of data collection. A random number table was used to assign an equal number (20) of individuals to one of three conditions. Of the surveys distributed, 25% (*n* = 15) of the emails bounced back because the email address was not recognizable, 46% (*n* = 28) of those emailed responded to the survey, and 75% (*n* = 21) of the respondents completed the survey in its entirety. In the Normal condition, 55% (*n* = 11) of those emailed responded to the survey, with 81% (*n* = 9) of these respondents completing the entire survey. In the LD condition, 35% (*n* = 7) of those emailed responded to the survey, with 57% (*n* = 4) of these respondents completing the entire survey. Lastly, in the Complicated condition, 46% (*n* = 10) of those emailed responded to the survey, of which 80% (*n* = 8) completed the entire survey.

The final sample was 28 school psychologists who responded to the survey. The sample consisted of five males and 23 females, with ages ranging from 28 to 49 (*M* = 33.21, *SD* = 10.11). All of the participants identified as White or Caucasian and were certified as school psychologists in their respective states. On average, participants had been working in the field for approximately 12 years (*M* = 12.22, *SD* = 11.93). Most participants had a doctorate degree (92.9%, *n* = 26) and two had a master's degree (7.1%). Half of the participants reported working in a suburban area (*n* = 14, 50.0%). However, participants also reported working in rural (*n* = 11, 39.3%) and urban (*n* = 3, 10.7%) regions. In terms of their training, almost all participants reported being trained under the scientist-practitioner model (*n* = 27, 95.4%) with one participant reporting being trained under the practitioner-scholar model (*n* = 1, 3.6%). The demographics of Sample 1 are provided in Appendix A.

Materials

The materials used in Study 1 were the cover letter, informed consent, and three student data summaries

Cover letter and informed consent. The cover letter, provided in Appendix B, invited the school psychologists to participate in a graduate student's research project investigating the decision-making process of school psychologists. The informed consent addressed appropriate ethical codes relevant to conducting research as it pertained to the study. Participants were informed that all responses to the survey would be anonymous, and that completion and submission of the survey implied the participants' agreement to the terms of the study. The informed consent is provided in Appendix C.

General student summary information. Three student summaries were used in this study. Each of these summaries had the same basic description of a 9-year old child, whose gender was not provided. Besides gender, other sociocultural variables (e.g., ethnicity and socio-economic status) also were excluded to eliminate the effects that these types of variables could have on classification decisions. On each student summary, all quantitative data were presented in a psychometric table while parent and teacher comments were presented descriptively. In the psychometric summaries, standard scores, confidence intervals, and percentile ranks (when appropriate) were reported. This format allowed participants to interpret the scores in the fashion that they typically would (Della Toffalo & Pedersen, 2005). Assessment data for each summary were standardized and presented in the following order: (a) referral question from teacher, (b) parent comments about student's behavior, (c) response-to-classroom modifications, (d) student's grades, (e) student's vision, hearing, and medications, (f) parent's standardized behavior ratings, (g) teacher's standardized behavior ratings, (h) visual motor abilities, (i) classroom observation, (j) standardized intelligence scores, and (k) standardized achievement test scores. The referral question indicated that a student had been referred to determine eligibility for special education services because the student was having difficulties in school (e.g., not completing schoolwork and frequently not being on task). Parent comments and standardized behavior rating scales indicated that the student was not having behavioral problems at home. Visual motor ability scores were reported as average, and hearing and vision were indicated as within normal range. A list of classroom modifications was included to convey that three interventions had been implemented with the student.

Normal data summary. This student summary was designed to be the least complex. It presented a student who clearly did not meet eligibility criteria for special education services. The response-to-classroom modifications data indicated that the student had made progress with three types of interventions. The student's grades were mostly 80% out of 100%. The teacher's standardized behavior ratings indicated that overall the student's behavior was average. There were no Clinically Significant behaviors, and only atypicality was endorsed as At-Risk. The student's Full Scale IQ and all standardized achievement scores were in the Average range. The Normal student summary was included as a baseline condition, as this summary was expected to be easy for school psychologists to assess.

Learning disabled (LD) data summary. This student summary was designed to be of medium complexity. It presented a student that was easily determined to have a Specific Learning Disability. The response-to-classroom modifications data indicated that the student had made no progress with three types of interventions. The student's grades mostly fell in the 60% range. The teacher's standardized behavior ratings indicated that the student's most problematic behaviors were learning and school problems. The student's Full Scale IQ was Average and the standardized achievement scores were in the Low Average range. As students are often referred for learning disabilities, this student summary was included to explore whether school psychologists would typically use the representativeness heuristic and which assessment data would be rated as important to their classification decisions.

Complicated data summary. The Complicated student summary was designed to be the most complex in comparison to the other student summaries. It depicted a student whose eligibility for special education services was not readily apparent. The response-to-classroom modifications data indicated that the student had made minimal progress with three types of interventions. The student's grades were mostly in the 70% range. The teacher's standardized behavior ratings indicated that the student's hyperactivity, attention problems, and social skills were extremely problematic. The student's Full Scale IQ score was in the Low Average range and standardized achievement scores were in the Borderline and Low Average ranges. This student summary was included to examine whether school psychologists were likely to use the representativeness heuristic and which assessment data were rated important when a diagnosis is unclear. Some researchers (Cioffi & Markham, 1997; Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1983) have found that individuals are more likely to use the representativeness heuristic when situations are more complex.

Measures

The measures used were a demographic questionnaire and the Decision-Making Questionnaire (Author, 2009).

Demographic questionnaire. The demographic questionnaire was used to elicit information on several respondent variables: age, ethnicity, level and background of training, years of experience as a school psychologist, and location of employment. The demographic questionnaire is provided in Appendix D.

Decision-making questionnaire. After reviewing the student summary, participants were asked to make decisions about the case and then answer questions about their decisions. Specifically, participants were asked to select a classification from a list of identification labels (e.g., nonexceptional, specific learning disability, & emotional disturbance) recognized by the Pennsylvania Department of Education. Five choices for a placement decision were provided as well. The choices were presented in order from the least restrictive to the most restrictive placement and described as follows: (a) Full time regular class with no basic change in teaching program or support services; (b) Full time regular class with specialists available in the school for consultation; (c) Regular class plus additional resource room or learning support; (d) Part-time special class, resource room, or learning support; (e) Full time special class or special school.

The Decision-Making Questionnaire, developed by the researcher, was based on prior surveys and questions used by previous researchers (e.g., Cioffi & Markham, 1997; Garb, 1996; Knoff, 1983; Thurlow & Ysseldyke, 1980), who have investigated either the use of the representativeness heuristic or the importance of assessment data. The questionnaire contained questions designed to elicit information about the school psychologists' decision-making process. A Likert-like scale from 1 (*not at all*) to 5 (one of three descriptions: *extremely, exclusively, or critical*) was used to measure each of the following: (a) school psychologists' confidence in their classification decision, (b) the likelihood that the student has the classification that the school psychologists selected, (c) the use of prototypes to determine classification, (d) similarity between the student in the summary and a prototype, and (e) the importance of student assessment variables to the school psychologists' decisions. The Decision-Making Questionnaire also contained four open-ended questions about the steps used to determine a classification, the identification of target student characteristics that resembled a prototype, and reasons why assessment data were described as "critical." Space was provided so that participants could also provide feedback about the survey. All participants received questions that were equivalent, except for the name of the student.

Procedures

All materials were administered online via SurveyMonkey, a secure web-based survey program. During March 2009, 60 school psychologists received emails describing the general purpose of the study. Each email contained one of three weblinks that had been randomly assigned to one of three email address sets by using a random number table. Weblinks directed participants to the Survey Monkey website, where materials were presented in the same sequence for all participants. Participants filled out the demographic questionnaire, read one of the student summaries, and finally, completed the Decision-Making Questionnaire. After each question, participants were asked to provide feedback on the format of the material and the clarity of the questions and to recommend revisions. At the end of the survey, participants were once again asked to provide feedback about the overall survey as well as the types of assessment data included in the data summaries. Revised materials based on the feedback from Study 1 were used in Study 2.

Data Analyses

For Study 1, data analyses were conducted to discern how participants responded to each question in the Decision-Making Questionnaire. Cross-tabulation analyses, independent samples *t*-tests, calculation of means, and a review of qualitative comments were used to analyze the data.

Results

Findings about school psychologists' responses to the original Decision-Making Questionnaire are presented by the individual questions on the survey.

Question 1

"Based on the information provided how would you classify Casey, Logan, or Addison: (a) Nonexceptional, (b) Autism, (c) Specific Learning Disability, (d) Severe Emotional Disturbance, (e) Mental Retardation, (f) Traumatic Brain Injury, (g) Other Health Impairment, or (h) Multiple Disabilities?"

Question 1 was used to assess the type of classification that school psychologists would select for the student presented to them. The question was designed to test whether the manipulation of the conditions was effective--the classification selected would vary based on the condition (Normal, LD, or Complicated). Classification accuracy should be highest for those in the Normal condition and lowest for those in the Complicated condition. Classification accuracy for the LD condition was expected to fall in between the Normal and Complicated conditions. Cross-tabulation analyses were conducted on complexity level (3 types; Normal, LD, Complicated), classification selected (e.g., nonexceptional, specific learning disability, other health impairment), and the accuracy of classification (accurate or inaccurate) to check whether the classification and accuracy of classification varied with the conditions. The findings indicated that in the LD and Normal conditions, the students were classified with 100% accuracy. In contrast, identification of the student in the Complicated condition varied and resulted in 40% misclassification. Thus, as intended, the students in the Normal and LD conditions appeared to be easier for the school psychologists to classify than the student in the Complicated condition. However, the manipulation of the conditions did not result in a discernable difference between the complexity level of the Normal and LD conditions. A summary of the observed counts is provided in Table 1. No suggestions, concerns, or questions were noted about Question 1.

Table 1

Accuracy of Classification by Complexity Level (N = 28)

Condition	Accurate		Inaccurate	
Normal (*n* = 11)	Nonexceptional	11 (100%)	___	___
Learning Disabled (*n* = 7)	SLD	7 (100%)	___	___
Complicated (*n* = 10)	Nonexceptional	6 (60%)	SLD	2 (20%)
	___	___	OHI	2 (20%)

Note. SLD = Specific Learning Disability; OHI = Other Health Impairment.

Question 2

"Select ONE type of placement that you would recommend for Casey, Logan, or Addison: (a) Full time regular class with no basic change in teaching program or support service, (b) Full time regular class with specialists available in the school for consultation, (c) Full time regular class with the addition of resource room or learning support, (d) Part-time special class, resource room, or learning support, (e) Full time special class or special school."

Question 2 was included in the Decision-Making Questionnaire to explore whether school psychologists with different student data summaries would vary in their placement recommendations. Cross-tabulation analyses were conducted on complexity level (3 types;

Normal, LD, and Complicated), recommendations for education placement (e.g., regular education, regular education with a specialist, part-time learning support), and accuracy of classification (accurate or inaccurate) to check whether recommendations and accuracy of classification varied with the conditions. Findings indicated that school psychologists in the Normal condition (n = 11), who all correctly classified the student, recommended the student for either Regular Education (n = 7; 63%) or Regular Education with consultation of a specialist (n = 4; 36%). In terms of the LD condition (n = 7), participants were also all accurate, with a majority recommending the student for Regular Education with Learning Support (n = 5; 71%). In terms of the Complicated condition, a review of the data revealed that the same participants (n = 6; 60%) who classified the student as Nonexceptional, or normal, recommended that the student remain in Regular Education with a Specialist. Overall, the recommendations for placement from school psychologists in the Normal condition varied less than in the LD and Complicated conditions. However, the manipulation did not result in a notable difference between the number of placement recommendations for the LD and Complicated conditions. A summary of the observed counts is provided in Table 2. No suggestions, concerns, or questions were noted for Question 2.

Question 3

"Indicate how confident you are in your classification: (a) not at all, (b) a little, (c) moderately, (d) very, or (e) extremely."

Question 3 was included to gather information about the perceived complexity levels of the student data summaries. Independent samples *t*-tests were conducted between the conditions on confidence in classification decision. A statistically significant difference was found in confidence in decisions between the Normal condition and the other two conditions: LD t (16) =

Table 2

Placement Decisions by Complexity Level (N = 28)

	Placement			
	Reg Ed	Reg Ed with Specialist	Reg Ed with Learning Support	Part-Time Learning Support
Condition	*n* (%)	*n* (%)	*n* (%)	*n* (%)
Normal (*n* = 11)	7 (63%)	4 (36%)	0	0
Learning Disabled (*n* = 7)	0	1 (14%)	5 (71%)	1 (14%)
Complicated (*n* = 10)	0	6 (60%)	2 (20%)	2 (20%)

Note. Reg Ed = Regular Education.

3.34, $p = .004$, and Complicated t (19) = 6.70, $p < .001$. Participants in the Normal condition were on average more confident in their classification (1 = *not confident*; 5 = *extremely confident*; M = 4.45) versus the LD (M = 3.42) and Complicated (M = 2.71) conditions. There was no statistically significant difference in confidence of classification for those in the LD and Complicated conditions, t (15) = 2.05, $p = .06$. A summary of the mean ratings is presented in Table 3. In terms of accuracy, no statistical analysis could be conducted due to the small sample size. Participants in the Complicated condition seemed to be less than *moderately* confident in their decisions regardless of the accuracy level (Inaccurate $M = 2.75$; Accurate $M = 2.66$). No suggestions, concerns, or questions were noted about Question 3.

Table 3

Confidence in Classification by Complexity Level and Accuracy of Classification (N = 28)

	Accurate			Inaccurate		
Condition	*n*	*M (SD)*	Min-Max	*n*	*M(SD)*	Min-Max
Normal	11	4.45 (.52)	4 - 5	__	___	___
Learning Disabled	7	3.42 (.78)	2 - 4	__	___	___
Complicated	6	2.66 (.81)	2 - 4	4	2.75 (.50)	2 - 3

Question 4

"What's the likelihood that Casey, Logan, or Addison has the classification that you listed in Question 1? (1) 0 to 20%, (2) 21-40%, (3) 41-60%, (4) 61 to 80%, (5) 81-100%"

Question 4 was included in the Decision-Making Questionnaire for two reasons. One reason was to gather additional information about the perceived complexity level of the data summaries. The second reason was to measure the representativeness heuristic based on Garb's (1996) framework. Namely, the representativeness heuristic should be measured by rating the perceived likelihood of an accurate classification and the similarity between the client and a prototype.

Independent samples *t*-tests were conducted on likelihood of accurate classification between the complexity levels (3 types; Normal, LD, and Complicated). A statistically significant difference was found in reported likelihood of accurate classification between the Normal and Complicated conditions, *t* (18) = 6.31, *p* < .001. Participants in the Normal (1 = 0-20%; 5 = 81-100%; *M* = 4.81) condition were more likely than those in the Complicated (*M* = 3.40) condition to think that their student fit the classification they assigned. Likelihood of accurate classification did not differ between the LD condition and the other conditions: Normal *t* (16) = 2.64, *p* = .018 and Complicated *t* (14) = 2.75, *p* = .016. Overall, these findings indicated that although participants in the Normal and Complicated conditions differed in how accurate they perceived their classification, there was no discernable difference with the LD condition. Therefore, the manipulation was only partially working as intended. A summary of the mean ratings is provided in Table 4. No suggestions, concerns, or questions were noted for Question 4.

Table 4

Likelihood of Accurate Classification by Complexity Level and Accuracy of Classification (N = 28)

	Accurate			Inaccurate		
Condition	*n*	*M (SD)*	Min-Max	*n*	*M(SD)*	Min-Max
Normal	11	4.81 (.40)	4 - 5	___	___	___
Learning Disabled	7	4.14 (.69)	3 - 5	___	___	___
Complicated	6	3.40 (.89)	2 - 5	4	2.75 (.50)	2 - 3

Question 5

"When you read the student Data Summary describing Casey, Logan, or Addison did a prototypical student with a specific classification come to mind? Yes or No?"

Question 5 was used to measure the use of prototypes—a measure of the representativeness heuristic—by school psychologists in the three conditions. Cross-tabulation analyses with complexity level (3 types; Normal, LD, and Complicated), thinking of a prototype (yes or no), and accuracy of classification (accurate or inaccurate) indicated that in comparison to the other two groups, participants in the Normal condition were least likely to think of a prototype (n = 2; 18%), even when giving an accurate classification of the student in the data summary. Over half of the participants in the LD condition reported thinking of a prototype (n = 57; 4%) and were still able to provide an accurate classification. In the Complicated condition, 40% (n = 4) of the participants reported thinking of a prototype. However, when this condition was separated into those who had accurate and inaccurate classifications, 75% of those who inaccurately classified the student reported thinking of a prototype. Only 16% of those who correctly classified the student in the Complicated condition reported thinking of a prototype. Overall, it appeared that the representativeness heuristic, as measured by consideration of a prototype, was used more by school psychologists in the LD condition than in the Normal and Complicated conditions. However, school psychologists in the Complicated condition who were inaccurate used the representativeness heuristic as a prototype more than those in the Complicated condition who were accurate as well as those in the Normal and LD conditions. Observed counts are provided in Table 5. No suggestions, concerns, or questions were noted about Question 5.

Table 5

Prototype use by Complexity Level and Accuracy of Classification (N = 28)

	Accurate		Inaccurate	
Classification	*n*	*Prototype use (%)*	*n*	*Prototype use (%)*
Normal	11	(18%)	___	___
Learning Disabled	7	(57%)	___	___
Complicated	6	(16%)	4	(75%)

Question 6

"How similar is Casey, Logan, or Addison to your prototype? Choose one: (a) not at all, (b) a little, (c) moderately, (d) very, or (e) extremely."

Question 6 was used in conjunction with Question 4 to measure the representativeness heuristic. As stated above, research (Garb, 1996) has suggested that the representativeness heuristic can be measured by tapping into the perceived likelihood of an accurate classification along with the client-prototype similarity. Independent samples *t*-tests were conducted to assess whether client-prototype similarity varied by complexity level (3 types; Normal, LD, and Complicated). No statistically significant differences were found among the three conditions on client-prototype similarity. The Normal condition did not differ from the other conditions: LD $t(4) = 2.58$, $p = .061$ and the Complicated $t(4) = .96$, $p = .391$. Additionally, the LD condition did not differ from the Complicated condition, $t(6) = .93$, $p = .390$. These findings indicated that client-prototype similarity did not vary by condition as intended. A summary of the mean ratings is provided in Table 6. No suggestions, concerns, or questions were noted for Question 6.

Table 6

Client-Prototype Similarity by Complexity Level and Accuracy of Classification (n = 10)

	Accurate			Inaccurate		
Condition	*n*	M (SD)	Min-Max	*n*	M(SD)	Min-Max
Normal	2	3.50 (.70)	3 - 4	___	___	___
Learning Disabled	4	2.25 (.50)	2 - 3	___	___	___
Complicated	1	4.00 (---)	4 - 4	3	2.33 (.57)	2 - 3

Question 7

"To what extent did you use your prototype to determine a classification for Casey, Logan, or Addison? Choose one: (a) not at all, (b) a little, (c) moderately, (d) a lot, or (e) exclusively."

Question 7 was intended to measure how much school psychologists reported using prototypes. Independent samples *t*-tests were conducted on the extent that prototypes were used by complexity level (3 types; Normal, LD, and Complicated). No statistically significant differences were found between the Normal condition and the other two conditions: LD t (4) = 1.55, p = .196 and Complicated t (4) = 1.55, p = .196. Additionally, no statistically significant differences were found between the LD and Complicated conditions, t (6) < .01, p = 1.00. Overall, the findings did not support that the conditions differed on the use of prototypes. A summary of the mean ratings is in provided in Table 7. No suggestions, concerns, or questions were noted about Question 7.

Table 7

Extent of Prototype use by Complexity Level and Accuracy of Classification (n = 10)

	Accurate			Inaccurate		
Condition	*n*	*M (SD)*	Min-Max	*n*	*M(SD)*	Min-Max
Normal	2	1.50 (.70)	1 - 2	___	___	___
Learning Disabled	4	2.25 (.50)	2 - 3	___	___	___
Complicated	1	2.00 (---)	2 - 2	3	2.33 (.57)	2 - 3

Question 8

"Rate how important each of these factors was to your decision regarding Casey, Logan, or Addison in particular. For each factor, choose only ONE rating: (a) not at all, (b) a little, (c) a little, (d) moderately, or (e) critical."

Question 8 was used to explore which assessment data school psychologists in the three conditions considered to be important to their classification decisions. Some components of Question 8 (i.e., importance of state criteria and prevalence ratings) were also used to measure the representativeness heuristic. Descriptive statistics were run to determine the top five assessment data per condition. Means for the Normal condition indicated that the most important factors were (a) response to modification, (b) classroom observations, (c) state criteria, (d) achievement score, and (e) implications for the student. For the LD condition, the most important factors were (a) state criteria, (b) IQ score, (c) achievement score, (d) response to modifications, and (e) classroom observations. Ratings differed for those school psychologists in the Complicated condition who had inaccurate and accurate classifications. When the student was classified correctly in the Complicated condition, the most important factors were (a) state criteria, (b) achievement score, (c) IQ score, (d) BASC-2 teacher report, and (e) BASC-2 parent report. For the Complicated condition in which the student was misclassified, the most important factors were (a) achievement score, (b) implications for the student, (c) referral, (d) BASC-2 teacher report, and (e) IQ score. Mean ratings for each factor are provided by condition in Table 8.

The ratings of importance of assessment data also were examined to discern the agreement between conditions. The school psychologists in the Complicated condition who misclassified the case had an 80% overlap with those in the Normal condition, which is the greatest amount of overlap between any of the conditions. The two conditions overlapped on four factors: achievement score, implications for the student, referral, and IQ score. There was a 60% overlap on the five most important factors between the following conditions: (a) Normal and LD, (b) Normal and Complicated with correct classification, and (c) LD and Complicated with correct classification. Across all of these comparisons, the same factors were rated as important: state criteria, IQ score, and achievement score. The school psychologists in the Complicated condition who misclassified the case had a 40% overlap with those in the LD condition. Both conditions rated achievement score and IQ score as important assessment data.

Table 8

Importance of Assessment Data by Complexity Level and Accuracy (N = 28)

	Accurate				Accurate		Inaccurate	
	Normal (*n* = 11)		LD (*n* = 7)		Complicated (*n* = 6)		Complicated (*n* = 4)	
	M (SD)	Min-Max	M (SD)	Min-Max	M (SD)	Min-Max	M (SD)	Min-Max
RefQuest	2.80 (.91)	1 - 4	2.60 (.89)	1 - 3	3.00 (1.41)	1- 4	3.50 (.57)	3 - 4
H&V	3.00 (.81)	2 - 4	2.80 (1.48)	1 - 5	3.00 (1.15)	2 - 4	3.00 (.81)	2 - 4
VMI	2.60 (.51)	2 - 3	2.20 (.83)	1 - 3	1.50 (.57)	1 - 2	2.00 (.00)	2 - 2
Observat	3.70 (.94)	2 -5	3.40 (.54)	3 - 4	3.00 (.81)	2 - 4	3.25 (1.25)	2 - 5
IQ	3.20 (.78)	2 - 4	3.80 (.44)	3 - 4	3.50 (1.00)	3 - 5	3.25 (.50)	3 - 4
Achieve	3.40 (.96)	2 - 5	3.60 (.54)	3 - 4	3.75 (.95)	3 - 5	4.00 (.00)	4 - 4
RTM	4.10 (.56)	3 - 5	3.40 (.54)	3 - 4	2.75 (.50)	2 - 3	3.00 (.81)	2 - 4
Grade	3.30 (1.05)	2 - 5	2.80 (.83)	2 - 4	3.00 (.81)	2 - 4	3.00 (.81)	2 - 4

Table 8 (continued)

Importance of Assessment Data by Complexity Level and Accuracy (N = 28)

	Accurate				Inaccurate			
	Normal (*n* = 11)		LD (*n* = 7)		Complicated (*n* = 6)		Complicated (*n* = 4)	
	M (SD)	Min-Max	M (SD)	Min-Max	M (SD)	Min-Max	M (SD)	Min-Max
PNarrativ	2.90 (.56)	2 - 4	2.80 (.44)	2 -3	3.00 (.00)	3 - 3	3.00 (.81)	2 - 4
BASC-T	3.20 (.63)	2 - 4	3.20 (.83)	2 -3	3.00 (.00)	3 -3	3.25 (.50)	3 - 4
PA Criter	3.45 (1.29)	1 - 5	4.00 (.00)	4 - 4	3.80 (1.09)	2 - 5	3.25 (.50)	3 - 4
DSMIV	1.90 (1.13)	1 - 4	1.33 (.51)	1 - 2	2.40 (1.14)	1 - 4	2.25 (.95)	1 – 3
SimProt	2.11 (.92)	1 - 3	2.25 (.95)	1 - 3	1.75 (1.50)	1 - 4	2.50 (.57)	2 - 3
Prevalence	1.60 (.96)	1 - 3	1.25 (.50)	1 - 2	1.00 (.00)	1 - 1	1.00 (.00)	1 - 1
Student	3.33 (1.50)	1 - 5	2.75 (1.50)	2 - 5	2.75 (1.25)	1 - 4	3.75 (.95)	3 - 5
School	2.11 (1.26)	1 - 4	1.50 (1.00)	1 - 3	1.00 (.00)	1 - 1	1.00 (.00)	1 - 1

Note. RefQuest = referral questions; H&V = hearing and vision; VMI = visual-motor index; Observat = classroom observations; Achieve= achievement score; RTM = response-to-classroom modification; PNarrativ = parents comments; PA Criter= state criteria; DSMIV = DSM criteria; SimProt = similarity with prototype; Prevalence = prevalence of classification; Student = Implications for the student; School = Implications for the school. Possible scores for the ratings of assessment data ranged from 1 (not at all important) to 5 (critical to decision).

Those in the Complicated condition who misclassified the student also rated implications for student, referral, and BASC-2 Teacher results as important, whereas those in the LD condition rated state criteria, response to classroom modifications, and classroom observations as important factors. However, those in the Complicated condition who misclassified versus all of those who correctly classified in the three conditions agreed on only one of the five (20% overlap) assessment data as important (achievement score). This same factor was the only one rated as important by school psychologists in every condition.

The use of state criteria and prevalence rates also was used to measure the representativeness heuristic (Cioffi & Markham, 1997; Dewhurst et al., 2007; Garb, 1996). Independent samples *t*-tests were conducted to assess whether school psychologists in the three conditions of complexity differed in their use of state criteria to determine special education eligibility. Results indicated that the Normal condition did not differ from the other two conditions in using state criteria: LD t (15) = 1.02, p = .325, and Complicated t (18) = .20, p = .844. Likewise, no statistically significant difference was found in using state criteria between the LD and Complicated conditions, t (13) = 1.22, p = .245. School psychologists in each condition, on average, reported that state criteria were *moderate* to *very important* to their decisions. This finding did not support that use of the representativeness heuristic, measured as state criteria, varied with the complexity level of the case.

Independent samples *t*-tests were also run to determine whether school psychologists in the three conditions of complexity differed in their use of prevalence rates. Results indicated that the Normal condition did not differ from the other conditions in using prevalence rates: LD t (10) = .51, p = .624 and Complicated t (14) = .26, p = .802. Likewise, no statistically significant difference was found in the use of prevalence rates between the LD and Complicated conditions,

t (10) = .22, p = .833. On average, prevalence of classification was rated as *not at all important* or *a little important* across all three conditions. This finding did not support that the use of the representativeness heuristic, measured as disregard of prevalence rate, varied with the complexity level of the case. No questions, comments, or suggestions were made for Question 8.

Question 9

"You may use this section to list any additional information that you would liked to have to make a psychoeducational decision or to provide any comments or suggestions that could improve this study."

Question 9 was used to gain information about the overall quality of the research materials. Ten of the 28 participants responded to this question. Half of the respondents were from the Complicated condition (n = 5), while three were from the LD condition and two were from the Normal condition. All suggestions were for more detailed data summaries. For example, participants reported that they would have preferred more detailed information on RTI, work samples, and reasons for referral. Participants also requested the inclusion of a student interview, a measure of executive functioning, and projective testing. However, no suggestion was expressed by more than one participant. Overall, findings suggested that the survey was clear and, although not completely comprehensive, included assessment factors that school psychologists typically encounter in practice.

Discussion

The purpose of Study 1 was to assess whether the data summaries were operating as designed. Furthermore, Study 1 was conducted to check the clarity and adequacy of the questions created to tap the representativeness heuristic and the importance of assessment data. Three research questions guided this study. Limitations of the findings and implications and recommendations for Study 2, are discussed.

Research Question 1: How Could Data Summaries be Understandable yet Vary in Ambiguity of Data Presented?

Based on a lack of corrections or requests for clarity by participants, it can be concluded that the data summaries were clear. Nonetheless, there were some concerns about the lack of detail and missing types of assessment data. However, there was no single type of assessment data that was noted as being preferred by more than one participant. Additionally, the study included all of the assessment data types that have been found in previous work (Della Toffalo & Pedersen, 2005; DeMesquita, 1987; Ysseldyke et al., 1981) plus more. The findings suggest that the data summaries presented information typically received by school psychologists.

In addition to being clear and representative, assessment data summaries were designed to differ in terms of level of complexity. Four factors (accuracy in classification, confidence in classification, number of placement recommendations, and perceived likelihood of accurate classification) were considered indicators of complexity level. The first indicator of complexity level was accuracy in classification. In the Normal and LD conditions, school psychologists were more accurate than those in the Complicated condition, suggesting that, as intended, the Normal and LD conditions were less complex than the Complicated condition. However, no discernable difference was found in the accuracy in classification for school psychologists in the

Normal and LD conditions. There are two possible reasons for the lack of delineation between the Normal and LD conditions. First, the qualification of a student might not be the primary factor in determining complexity level. Second, the Normal and LD conditions might have been more similar than intended. Nonetheless, interpretations should be taken with caution due to the small sample size.

Although no difference was found between the Normal and LD conditions on the accuracy of classification, these conditions did differ in level of confidence and placement recommendations. In the Normal condition, recommendations for educational placement were almost the same and school psychologists were more confident in their decisions. Taken together, these findings suggest that to some extent school psychologists differed in their responses to the Normal and LD conditions, with the Normal condition leading to more certainty about classification than the LD condition.

Like the LD condition, participants in the Complicated condition were less confident than in the Normal condition. This finding suggests that, as anticipated, the Complicated condition was more complex than the Normal condition. However, no difference was found in confidence in classification between the LD and Complicated conditions. Similarly, no difference was found in the number of placement recommendations for the LD and Complicated conditions. Three factors may account for the lack of differentiation between the LD and Complicated conditions on confidence in classification and in placement recommendations. First, it is possible that school psychologists in the Complicated condition were overconfident in their decisions. Second, school psychologists in the LD condition might have been less confident in their decisions. Finally, the LD and Complicated conditions might have been more similar than intended.

Findings suggest that the Normal and Complicated conditions differed on perceived likelihood of accurate classification; school psychologists in the Normal condition believed that their classification was more accurate than those in the Complicated condition. Therefore, this difference in response may be indicative of complexity level of the case summaries. The school psychologists in the Normal condition may have been more confident about their classification because the case summary was straightforward, whereas the case summary in the Complicated condition wasn't clear cut. However, no difference was found between the LD condition and either the Normal or the Complicated condition. One possible explanation for this finding is that perceived likelihood of accurate classification might not be a reliable measure of complexity level. Another possible explanation is that the small sample size obscured significant findings.

In summary, school psychologists responded differently to the three conditions, although the delineation was not as clear as intended. There was a greater distinction between the Normal and Complicated conditions, but less distinction was not notable between the LD condition in comparison to the Normal and Complicated conditions. Thus, data summaries need to be revised to create the necessary contrast between the LD and other conditions.

Research Question 2: What Types of Conditions and Questions were Needed to Measure the use of the Representativeness Heuristic by School Psychologists?

Based on a lack of corrections or requests for clarity by participants, it can be concluded that the questions contained in the Decision-Making Questionnaire were comprehensible. It was also anticipated that the questions would adequately tap the representativeness heuristic. Specifically, it was expected that respondents in the three conditions would respond differently to the measures of the representativeness heuristic (use of prototypes, likelihood of accurate classification, client-prototype similarity, importance of state criteria, and importance of prevalence of classification).

Two questions were used to determine the use of prototypes. One question asked whether or not prototypes were considered while the school psychologists were determining classification. Data revealed that school psychologists in the three conditions differed in their consideration of prototypes. Overall, prototypes were thought of more in the LD condition than in the Complicated and Normal conditions. They also were thought of more in the Complicated condition than in the Normal condition. Additionally, when separated by accuracy of classification, school psychologists in the Complicated condition who were inaccurate thought of prototypes more than those in the same condition who were accurate; as well as those in the Normal and LD conditions. The association between accuracy of decisions and use of prototype has not been explored in previous literature. However, that prototypes were thought of the most in the LD condition conflicts with previous research findings (Cioffi & Markham, 1997). Cioffi and Markham found that midwives who reviewed cases without ambiguous data were less likely to base decisions on archetypes than when the cases were more complex. Therefore, the LD condition, which was presumed to be less complex than the Complicated condition, should have had fewer participants using prototypes than in the Complicated condition. Difference in findings between the two studies could be due to differences in profession (e.g., school psychologists versus midwives) or methodology (e.g., having participants self report the use of prototypes versus inferring that archetypes were used).

The second question used to assess the use of prototypes asked the extent in which participants used prototypes to make classification decisions. Overall, the findings did not support that the three conditions (Normal, LD, and Complicated) differed on the use of prototypes. On average, in each condition, participants reported that the extent in which they used prototypes ranged from *not at all* to *moderately*. This range is low considering that these school psychologists had already reported thinking of prototypes during their decision-making. Therefore, the repeated prompting about prototypes throughout the survey might have given away the purpose of the study, resulting in distorted ratings of the use of prototypes when participants finally answered the second prototype question. School psychologists might have minimized using prototypes to avoid suggesting their use of an unethical approach to decision-making. In contrast to basing decisions on prototypes, ethical decision-making is often described as being clearer, deliberate, and structured (Barnett & Johnson, 2008; Dieterly, 1979).

The second measure of the representativeness heuristic examined in Study 1 was the use of state criteria. Findings indicated that there was no difference between the conditions in their use of state criteria. In each of the conditions, on average, state criteria was rated as *moderate* to *very important* in determining classification decisions. School psychologists' adherence to state criteria despite the complexity of the case suggests that their classifications can be consistent regardless of the level of ambiguity in a data summary. However, the lack of disparity in use of state criteria between the conditions is in contrast to prior research that found disregard of established criteria as an indicator of the representativeness heuristic, which is expected to vary across the complexity level of situations (Garb, 1996; Tversky & Kahneman, 1983). The difference in findings between the current study and previous research could be due to (a) differences in profession (e.g., school psychologists versus clinical psychologists), (b) differences in methodology (e.g., self report of use of established criteria versus inference of use of established criteria), or (c) the use of established criteria as an unreliable measure of the representativeness heuristic.

Based on Garb's (1996) conceptualization of the representativeness heuristic, a third measure was the association between (a) the likelihood of accurate classification and (b) client-prototype similarity. However, the small sample size of the study precluded the use of correlational analyses so the relationship between client-prototype similarity and likelihood of accurate classification could not be calculated. Therefore, data were analyzed to assess whether school psychologists in the three conditions would respond differently to questions tapping into these two domains. School psychologists in the Normal condition were more likely to believe that their student fit the classification assigned than those in the Complicated condition. However, there was no difference between the LD condition and the Normal and Complicated conditions on this measure. An additional study with a larger sample is needed to better assess whether the representativeness heuristic, measured as the association between client-prototype similarity and likelihood of accurate classification, varies with complexity level of the condition.

The last measure of the representativeness heuristic was the use of prevalence of classification to make a decision. It was unexpectedly found that prevalence of classification did not vary across the three conditions. Specifically, on average participants rated prevalence of classification as *not at all important* or *a little important*. This finding was inconsistent with previous research (Cioffi & Markham, 1997). Cioffi and Markham found that a measure of the representativeness heuristic was when midwives provided low base rates for clinical conditions and overdiagnosed patients with the same clinical condition. Cioffi and Markham also found that the representativeness heuristic was more frequently used in complex situations. Therefore, it is perplexing that school psychologists across all conditions barely used base rates. However, several explanations may account for this finding. First, Study 1 had a small sample size, which limits interpretations of this finding. Second, the difference in findings between the current study and that of Cioffi and Markham might be due to differences in profession. Third, the difference might be due to methodological differences. For example, participants in the current study were asked how important prevalence rates were to their decisions, whereas Cioffi and Markham looked for a pattern between participants' diagnostic decision and estimated base rates of conditions. Fourth, it is possible that all school psychologists were using a form of the representativeness heuristic, which was measured as a disregard of prevalence rates.

In summary, there was preliminary support that the measures used in Study 1 were sufficient to measure the representativeness heuristic. Participants did not request clarification of the questions, suggesting that the measures were clear and easy to use. Additionally, the use of the representativeness heuristic, as measured in Study 1, varied to some extent with the complexity level of the data summaries. When the representativeness heuristic was measured as a prototype, differences were observed among the conditions. Additionally, although, it was not possible to assess the relationship between likelihood of accurate classification and client-prototype similarity, some differences in the conditions were observed for reported likelihood of accurate classification. However, there are concerns that participants in the three conditions did not respond differently to questions about (a) the extent in which prototypes were used, (b) the use of prevalence of classification, (c) the use of state criteria to determine classification, and (d) reported similarity between the student and a prototype.

Research Question 3: What Types of Assessment Data do School Psychologists Typically use when Determining Special Education Eligibility?

As expected, it was found that IQ score, achievement score, and behavior measures were important to school psychologists' decisions. This finding is consistent with that of previous research (Della Toffalo & Pedersen, 2005; Salvia & Ysseldyke, 1978; Thurlow & Ysseldyke, 1980). However, school psychologists did not limit themselves to these three forms of data.

Other assessment data that school psychologists frequently used were state criteria, classroom observations, referral information, implications for the student, and response to classroom modifications. Therefore, it seems logical for researchers investigating the use of assessment data to consider multiple forms of assessment data, including, but not limited to cognitive and achievement scores and behavior measures.

No other study has looked at whether the importance of assessment data varies with the complexity of data summaries. Findings of this study suggested that there was 60-80% overlap in the top five assessment data used by school psychologists in the three conditions. However, the only factor rated as important by school psychologists in every condition was achievement score. Thus, regardless of the condition of the student summary, achievement score was considered vital to decisions about special educational eligibility.

Mean ratings were also inspected to assess the level of overlap in ratings from school psychologists who classified correctly versus those who classified incorrectly. The Complicated condition was the only condition that resulted in both accurate and inaccurate classification decisions. Only achievement score fell in the five most important assessment data for those in the Complicated condition who were inaccurate as well as those who were accurate across all conditions. Unlike the other conditions, those in the Complicated condition who misclassified the student also rated referral information as important. This finding suggests that this factor might have been involved in the classification error. Why this would happen is not fully understood. It is possible that this finding would have been understood better had there also been school psychologists who incorrectly classified the student in the LD and Normal conditions. A larger sample size might have resulted in more inaccurate classifications in these conditions.

Limitations

There were several limitations to Study 1. A major limitation was the sample size; there were only 28 participants in the study. Due to the small sample size, it often was not possible to run inferential statistical analyses. Thus, it was not possible to test whether there was a relationship between likelihood of accurate classification and client-prototype similarity. This relationship has been listed as an indicator of the representativeness heuristic (Garb, 1996) and is a measure in Study 2.

Another limitation is the lack of generalizability of the findings. Because all of the participants came from the same program, it is likely that their beliefs and practice as school psychologists have been similarly influenced. Therefore, it is possible that results would have been different had the participants come from varied backgrounds and training experiences.

The last two limitations of the study concern the study design. First, the student data summaries did not include all of the data typically available to school psychologists when making classification decisions. In an attempt to keep the survey manageable for participants, the protocols only included data identified as important in prior research (e.g., Cioffi & Markham, 1997; Garb, 1996; Knoff, 1983; Thurlow & Ysseldyke, 1980). For example, as mentioned by participants, detailed RTI information, projective test results, and a student interview were not included in the data summaries.

Finally, participants in the study were not aware of the purpose of the study. Therefore, their feedback about data summaries and the questions in the survey might have been more extensive had they known the exact purpose of the study. For example, participants might have provided insight into whether or not the repeated mention of prototypes made it too obvious that the survey was intended to measure the use of prototypes.

Implications and Revisions for Study 2

The purpose of Study 1 was to examine the manipulation of data summaries and the clarity of the Decision-Making Questionnaire prior to their use in Study 2. Therefore, several findings from Study 1 were of critical importance to Study 2. First, the finding that there were very few suggestions, concerns, or comments about the data summaries and questions in the Decision-making Questionnaire suggested that in their original state, these materials were understood by participants. Therefore, it seemed reasonable to maintain the format of materials and measures used in Study 1 for Study 2.

Second, that there was a lack of consistency in suggestions for which types of assessment data should be added to the data summaries suggested that overall the data summaries contained enough assessment data that are typically reviewed by school psychologists. Additionally, it was important not to overwhelm participants with an abundance of assessment data as this could provoke participants to drop out. Therefore, the types of assessment data reviewed in Study 2 remained as they were in Study 1.

Third, that the Normal, LD, and Complicated conditions did not differ as much as intended implied the need for additional manipulation of the data summaries. Therefore, to further distinguish the Complicated condition from the other conditions the Complicated condition was made more ambiguous by lowering the report card grades into the borderline range (e.g., math grade from 70 to 65). Additionally, the LD condition was revised by lowering the report card grades and indicating that the student had even less response to classroom modifications. Such changes to the LD condition reflected greater academic difficulty for this student than in the original LD data summary, thereby possibly creating greater distinction in responses between the LD and Complicated conditions. The Normal condition remained as it was because there are few reasonable ways to manipulate this case so that the student would continue to be easily classified as nonexceptional.

Fourth, there were two questions in the Decision-Making Questionnaire that addressed the use of prototypes, but resulted in conflicting responses by participants, indicating a need to revise the measurement of prototype as the representativeness heuristic. Analysis of the survey indicated that the redundancy in asking about prototypes possibly gave away the purpose of the study, resulting in lower ratings on using prototypes in the last question. Therefore, the question asking the extent in which participants used prototypes to make classification decisions (see Question 7 in Study 1) was removed from the Decision-Making Questionnaire in Study 2.

Finally, because school psychologists who were accurate versus inaccurate in the Complicated condition differed in their use of prototypes and use of assessment data, it was important that the association between accuracy of classification and decision-making be considered. Although accuracy of classification might be associated with the use of the representativeness heuristic as well as the use of certain assessment data, this relationship was barely addressed in Study 1. In Study 2 these relationships were examined: between accuracy and both the use of the representativeness heuristic and the use of assessment data. This type of investigation has not been found in previous research about decision-making (e.g., Cioffi & Markham, 1997; Della Toffalo & Pedersen, 2005; Frey, 2002; Garb, 1996)

Study 2

Purpose

Based on the findings of Study 1, Study 2 was designed to examine whether school psychologists use the representativeness heuristic in making decisions about a student's special education eligibility. Furthermore, the focus was on which assessment data school psychologists used when the complexity level of the case varied.

Research Questions

1. Was school psychologists' accuracy of classification of students associated with the complexity level of the student data summary?
2. Did school psychologists use the representativeness heuristic more frequently as the complexity level of the student data summary varied?
3. Did the use of the representativeness heuristic vary as a function of the complexity level of the data summary and the accuracy in classification of the student?
4. Did the rating of importance of assessment data by school psychologists vary as a function of the complexity of the student summary?
5. Did the rating of the importance of assessment data by school psychologist vary as a function of the complexity of the student summary and the accuracy of the school psychologists' classification of the student in the student summary?

Hypotheses

Hypothesis 1. School psychologists in the Normal condition were expected to accurately classify the student more than school psychologists in the LD and Complicated conditions. School psychologists in the LD condition were expected to accurately classify their case more than those in the Complicated condition.

Hypothesis 2. It was expected the representativeness heuristic would be used most frequently in the Complicated condition in comparison to the LD and the Cleary Normal conditions. Also, the representativeness heuristic would be used more frequently in the LD condition in comparison to the Normal condition.

Hypothesis 3. It was expected that use of the representativeness heuristic would vary as the student summary changed in complexity and would result in more misclassifications as complexity of the case varied. Specifically, the representativeness heuristic would be used more by school psychologists in the Complicated condition who misclassified their case in contrast to those in the other conditions (LD & Normal) who were also inaccurate in classification. School psychologists in the LD condition who misclassified their case were expected to use the representativeness heuristic more versus those in the Normal condition who also erred in classification.

Hypothesis 4. It was hypothesized that school psychologists' ratings of importance of assessment data would vary based upon the complexity level of the student summary. In the Normal and LD conditions school psychologists would rate the top five assessment data as more influential in their decision making than those in the Complicated condition. School psychologists in the Normal condition, would also rate the influence of these assessment data as higher than in the LD condition.

Hypothesis 5. It was expected that the ratings of importance of the top five assessment data would vary based upon the complexity of the student summary and the accuracy of classifications selected by school psychologists. No specific direction was hypothesized.

Method

Research Design

Study 2 was a post-test only quasi-experimental analogue design (Gall et al., 2003). Three student summaries with assessment data were systematically manipulated to vary in complexity level (LD, Normal, and Complicated). School psychologists were systematically assigned to review one of the three student summaries.

Participants

The sample to participate in Study 2 was recruited from 602 practicing school psychologists listed in the Membership Directory of the Association of School Psychologists of Pennsylvania (ASPP). Approximately 20% (n = 133) of the emails sent to the ASPP members bounced back because the email address was not recognizable. Of the remaining population (N = 469), 68% (n = 319) responded to the survey.

Further inspection of the data resulted in the removal of 38 more cases from the sample of 319 participants. A systematic error in the survey format sent to eight participants resulted in their exclusion from the sample. Data for 30 participants (eight in the Normal condition, 10 in the LD condition, and 12 in the Complicated condition) were removed due to missing responses on more than 50% of their survey (e.g., no response to questions regarding the importance of assessment data). Univariate and multivariate outliers were detected using the z critical value of 3.29 ($p < .001$) and Mahalanobis distance test, respectively (Tabachnick & Fidell, 2001). No univariate outliers were found, but the multivariate scores were extreme in one case, resulting in its removal. Thus, the remaining 272 participants had completed on average most of the survey (99.7%) and were used in all subsequent statistical analyses, when possible. Prior to completing data collection, power analyses, using the GPower 3.1 software program (Erdfelder, Faul, &

Buchner, 2009) were conducted to identify the sample size that would be necessary to detect statistical significance at an alpha level of .01, power of .80, and effect size of .30 (Gall et al., 2003). The findings were that a minimum sample size of 155 would be sufficient to run a 3 x 2 contingency table, 159 was needed to conduct a 3 x 2 ANOVA, and 44 would suffice to run a 3 x 2 MANOVA.

The final sample consisted of 69 males and 203 females, with ages ranging from 23 to 74 (*M* = 42.21; *SD* = 12.52) years old. A majority of the school psychologists identified as White or Caucasian (96.7%; *n* = 263), while only 9 participants used other racial or ethnic self identifications: 1.1% Black or African American (*n* = 3), 1.1% Mixed Heritage (*n* = 3), .7% Hispanic or Latino (*n* = 2), and .4% Asian (*n* = 1). All participants were certified as school psychologists and on average had been working in the field for approximately 13 years (*M* = 13.29; *SD* = 10.73). Each participant had either a master's (73%; *n* = 199) or doctoral (27%; *n* = 73) degree in school psychology. All participants were practicing school psychologists in Pennsylvania at the time of data collection, with most working in schools (*n* = 216; 79.4%). However, participants also reported working for intermediate units (*n* = 42; 15.4%), themselves (*n* = 13; 4.8%), or residential treatment facilities (*n* = 1; .4%). More participants reported working in a suburban area (*n* = 127; 46.7%) than in rural (*n* = 96; 35.3%), urban (*n* = 42; 15.4%), or multiple (*n* = 7; 2.6%) regions. On average, school psychologists reported completing 55.58 (*SD* = 32.15) evaluations per year, though the number varied greatly (3 to 200).

In terms of professional orientation, most participants reported typically using the discrepancy model *often* (46%, *n* = 125) or *always* (43.8%, *n* = 119), while the remaining 15% reported using it *sometimes* (8.5%, *n* = 23), or *never* (1.8%, *n* = 5). Most participants reported *never* using the RTI model for psychoeducational evaluations (54%, *n* = 147), while others reported using it *sometimes* (32.7%, *n* = 89), *often* (10.7%, *n* = 29), and *always* (2.7%, *n* = 7). A paired samples *t*-test indicated that there was a statistically significant difference between participants' reported use of the discrepancy model (t (271) = 22.68, $p < .001$; M = 3.32; SD = .71) in comparison to the RTI model (M = 1.61; SD = .78), with school psychologists being almost twice as likely to usually use the discrepancy model over the RTI model. In regards to training orientation, almost 50% (46%; *n* = 126) reported being trained in the scientist-practitioner model, with other school psychologists almost split between the practitioner-focused (28%; *n* = 77) and practitioner-scholar (25%; *n* = 69) models. A full summary of the demographic description of the entire sample is provided in Appendix A.

The demography of school psychologists in this sample was representative of school psychologists nationwide. In a national survey on the recent demographics of the profession of school psychology (*N* = 1,748; Curtis et al., 2008), 26% of school psychologists were males and 74% were females. The average age of school psychologists was 46.2 years old (SD unavailable), and they primarily identified as Caucasian (92.6%), with the remaining 7.4% reflecting other racial/ethnic groups: 3% Hispanic, 1.9% African American, and less than 1% Native American/Alaskan Native, Asian/Pacific Islander, or Other. Approximately two-thirds had a master's degree (67.5%) and one-third had a doctoral degree (32.4%). School psychologists at the national level reported almost 15 years of work experience (M = 14.8 years; SD unavailable), with 28.6% working in urban areas, 49.9% in suburban areas, and 28.8% in rural areas (responses totaled more than 100% because multiple choices were possible).

Materials

The materials used in Study 2 were the same as in Study 1: cover letter, informed consent, and three student data summaries. As result of the findings in Study 1, minor revisions were made to the data summaries in attempt to further differentiate responses to the three conditions. First, the report card grades in the Complicated condition were lowered closer to the borderline range (e.g., math grade from 70 to 65). Additionally, so that the LD condition would reflect a student with even greater academic difficulty, the report card grades were lowered and the student's response to classroom modifications were reduced in the LD condition. The Normal data summary remained as it was in Study 1. The three data summaries are contained in Appendix E (Normal data summary), Appendix F (LD data summary), and Appendix G (Complicated data summary).

Measures

The measures used in Study 2 were also the same as used in Study 1: a demographic questionnaire and the Decision-Making Questionnaire (Author, 2009). The demographic questionnaire remained the same as in Study 1. However, the Decision-Making Questionnaire was revised to reduce redundancy. Specifically, one of the questions about the use of prototypes (see Question 7 in Study 1) was removed. The revised Decision-Making Questionnaire is in Appendix H.

Procedures

The research packet was administered online via the same web-based survey program used in Study 1, SurveyMonkey. During May and June of 2009, 602 school psychologists, who were members of the Association of School Psychologists of Pennsylvania (ASPP), received emails describing the general purpose of the study. Each email contained one of three weblinks that had been systematically assigned to one of three email address sets (Gall et al., 2003). To create email address sets, a list of school psychologists was provided by ASPP. The list was then alphabetized by last name and divided into thirds. Random assignment was not used in order to keep school psychologists who might have been listed in the directory more than once from being placed into different conditions. Weblinks directed participants to the Survey Monkey website, where research packets were presented in the same sequence for all participants. First, participants were given a cover letter and informed consent, which they had to accept the terms of in order to participate in the study. Participants were able to print a copy of the informed consent form, if desired. Subsequently, participants filled out the demographic questionnaire, read one of the student summaries, and finally, completed the Decision-Making Questionnaire. Participants required about 20 minutes to complete the questionnaires. As an incentive, participants had the option of entering into a drawing to receive a gift card valued at thirty dollars. Twenty-five (almost 14% of those who entered the drawing) participants were randomly selected and received their prize by the end of August 2009.

Which participants completed the survey was tracked by the survey program, while keeping responses anonymous. Following web-based survey guidelines, reminders were sent to nonrespondents during the second, third, and fourth weeks of data collection (Archer, 2007). Input validation (e.g., respondents were allowed to enter only numerical values for certain questions) was used when possible to improve data quality. Additionally, to avoid respondents answering the survey several times, only one response was allowed per IP address.

Data Analyses

SPSS 16.0 for Windows software was used to analyze the data. First, descriptive statistics (means, standard deviations, and correlations) of major variables (e.g., decision confidence, likelihood ratings, and similarity ratings) were conducted for the overall sample and by experimental group. Second, the data were tested to determine whether the appropriate assumptions for the statistics of a two-way contingency table, ANOVA, and MANOVA were met. Third, preliminary analyses were conducted to determine whether the experimental conditions were equivalent or proportional on the relevant demographic variables (e.g., age, gender, years as certified school psychologists, and location of employment).

Statistical analyses were selected based on the nature of the research question, the number of dependent variables, the number of independent variables, the scale of measurement of the variables, and the degree of the relationship among the dependent variables.

1. For the first hypothesis, a 3 x 2 contingency table analysis was conducted to determine whether a statistically significant association would be found between the complexity of the student summary and the accuracy of classification, with less accuracy occurring in the LD and Complicated conditions, in comparison to the Normal condition and less accuracy occurring in the Complicated condition in comparison to the LD condition . Accuracy of classification was based on whether the school psychologist selected a classification (see Appendix H, Question 1) that was consistent with the data presented.
2. For the second hypothesis, it was expected that school psychologists in the Complicated condition would use the representativeness heuristic more than those in the LD and Normal conditions, while those in the LD condition would use the representativeness heuristic more than those in the Normal. Several statistical analyses were used to assess whether the use of the representativeness heuristic differed based on the complexity level of the student summary, as multiple dependent variables and measures of representativeness heuristic were used.

    (a) When the dependent variable was the use of a prototype (yes or no; see

Appendix H, Question 9) when assessing a student summary, a 3 x 2 contingency analysis was conducted.

(b) When the dependent variables were both (a) the likelihood of classifying accurately (see Appendix H, Question 6) and (b) client-prototype similarity (see Appendix H, Question 9-A), a one-way MANOVA was used.

(c) When the dependent variable was the rating of importance of state criteria for classification of cases (see Appendix H, Question 10-L), a one-way ANOVA was used.

(d) When the dependent variable was the use of the prevalence of classification to make a classification decision (see Appendix H, Question 10-O), a one-way ANOVA was used.

3. For the third hypothesis, several statistical analyses were used to assess whether the emergence of the representativeness heuristic differed with the complexity level of the student summary and the accuracy of school psychologists' classification. Specifically, the representativeness heuristic would be used more by school psychologists in the Complicated condition who misclassified their case in contrast to those in the other two conditions. In turn, school psychologists in the LD who misclassified their case were expected to use the representativeness heuristic more than those in the Normal condition who erred in classification. Once again, there were multiple dependent variables and measures of the representativeness heuristic.

(a) When the dependent variable was the use of a prototype (yes or no; see Appendix H, Question 9) while reviewing the student summary, a 3 x 2 x 2 contingency analysis was planned, although a 2 x 2 x 2 contingency analysis was used.

(b) When the dependent variables were both (a) the likelihood of accurate classification (see Appendix H, Question 6) and (b) client–prototype similarity (see Appendix H, Question 9-A), a 3 x 2 MANOVA was planned, although a 2 x 2 MANOVA was conducted.

(c) When the dependent variable was the rating of how important state criteria were in classifying the case (see Appendix H, Question 10-L), a 3 x 2 ANOVA was planned, but a 2 x 2 ANOVA was conducted.

(d) When the dependent variable was the rating of how important the prevalence of the classification was in making a decision (see Appendix H, Question 10-O), a 3 x 2 ANOVA was planned, but a 2 x 2 ANOVA was run.

4. For the fourth hypothesis, a one-way MANOVA, followed by a descriptive discriminant analysis, was run to examine whether the five assessment data that were reported as most important to school psychologists' decisions (see Appendix H, Question 10) differed based on the complexity level of the student summary, with the top five assessment data being rated as more important to school psychologists in the Normal condition in comparison to the other two conditions, and more important to school psychologists in the LD condition in comparison to the Complicated condition.
5. For the fifth hypothesis, a 3 x 2 MANOVA was proposed, but a 2 x 2 MANOVA was run to explore how the five assessment data reported as most important (see Appendix H, Question 10) differed based on the complexity level of the student summary and the accuracy of school psychologists' classifications.

Due to the extensive number of statistical analyses conducted, Bonferroni adjustment resulted in the statistical significance level set at .01 for all analyses conducted in this study. This statistical significance level minimized Type I error while balancing Type II error (Howell, 2007). An a priori salient effect size was set at $r > .30$.

Results

Preliminary Analyses

A series of preliminary analyses were conducted (a) to provide a descriptive summary of the data, (b) to test whether the data met the assumptions necessary for using inferential statistics, and (c) to examine whether participants across experimental conditions were comparable on demographic features. Descriptive statistics, including means, standard deviations, bivariate correlations, percentages, skew, and kurtosis were calculated for major variables (accuracy of classification, likelihood of classification, client-prototype similarity, and importance ratings of assessment data).

Accuracy of classification. The school psychologists varied in their accuracy of classifying the student summary based on the status of their experimental condition. Most of the school psychologists in the Normal condition (97.7%; n = 86) classified the student accurately, which was approximately 14% higher than the percent of school psychologists in the LD condition (86%; n = 80) and 51% higher than those in the Complicated condition (46.2 %; n = 42). That school psychologists in the Normal condition were the most accurate and those in the Complicated condition the least accurate demonstrated that the manipulation of the independent variable (complexity level of student summaries) was effective.

Likelihood of accurate classification. In the Normal condition, on average school psychologists reported a *moderate* to *high* (1 = *not likely*; 5 = *extremely likely*; M = 3.67) likelihood that the student had the classification that they selected. Like the accuracy rating, school psychologists in the LD (*moderately likely*; M = 3.08) and Complicated (a *little* to *moderately likely*; M = 2.91) conditions had on average lower likelihood ratings than those in the Normal condition. The mean scores of likelihood of classification per experimental group are provided in Appendix I, Table I-1.

Confidence in classification. School psychologists in the Normal condition were on average *very* to *extremely* (1 = *not confident*; 5 = *extremely confident*; M = 4.13) confident in their classification of the student in the summary reviewed. School psychologists in the LD and Complicated conditions were less confident in their classifications (*moderately* to *very*; M = 3.48 & 3.03, respectively). The mean scores of confidence in classification per experimental group are provided in Appendix I, Table I-1.

Client–prototype similarity. Only school psychologists who indicated that they thought of a prototype (45.9%; n = 125), while reviewing the student summary were routed to questions asking them how similar the student in the summary was to their prototype. The mean rating of client–prototype similarity was 3.33 for the Normal condition, 3.40 for the LD condition, and 3.53 for the Complicated condition; all scores were reflective of the *moderate* to *very similar* range (1 = *not similar*; 5 = *extremely similar*). The mean scores for client–prototype similarity by condition are provided in Appendix I, Table I-2.

Intercorrelations. For each condition, there were 6 intercorrelations between the following variables: accuracy of classification, likelihood of accurate classification, confidence in classification, and client–prototype similarity. Half of the intercorrelations for both the Normal and the LD conditions were statistically significant at the .01 level and in the Complicated condition, two correlations were statistically significant. Intercorrelations ranged from $|.01|$ to $|.68|$ (Mdn = .32) for the Normal condition; from $|.03|$ to $|.77|$ (Mdn = .15) for the LD condition; and from $|.06|$ to $|.39|$ (Mdn = .27) for the Complicated condition. In each condition, confidence in classification had a positively statistically significant correlation with accuracy of classification (Normal r = .31, LD r = .23, & Complicated r = .27) and with likelihood of accurate classification (Normal *r* = .68, LD *r* = .77, & Complicated *r* = .39). School psychologists who expressed more confidence in their decisions were more accurate in their classification. Additionally, those who were more confident were more likely to report that their classification was accurate. Accuracy of classification had a positively statistically significant correlation with likelihood of accuracy in both the Normal (*r* = .29) and LD conditions (*r* = .43), but not in the Complicated condition (*r* = -.08). Only in the Normal condition was there a statistically significant correlation between client–prototype similarity and any other variable (confidence; *r* = .57). Intercorrelations between accuracy of classification, likelihood of accurate classification, confidence in classification, and client–prototype similarity are provided in Appendix I, Table I-3.

Assumptions. The following assumptions of parametric statistics were met for the scores of the dependent variables: (a) independent observations, (b) normality (visual examination of histograms & boxplots), and (c) linearity (visual examination of scatterplots). The assumptions of homogeneity of variance and covariance are presented with the specific statistical analyses. For the non-parametric chi-square tests, the assumption of sufficient cell frequencies (expected count $\geq$ 5) was met, except for the data associated with hypothesis 3a. Thus, one experimental condition (Normal) had to be excluded from the analyses for hypothesis 3a to meet the assumption.

Assessment data. Descriptive statistics for the ratings of how important assessment data (e.g., referral question, IQ score, achievement score, and response-to-classroom modifications) were to school psychologists are presented by experimental condition (Normal, LD, and Complicated) in Appendix J, Tables J4 to J6, respectively. Possible scores for the ratings of all assessment data ranged from 1 (*not at all important*) to 5 (*critical to decision).*

Comparison of the conditions on the five types of assessment data with the highest mean ratings of importance indicated similarity in ratings. The LD and Complicated conditions had 100% overlap on the five assessment data that were rated as most important, although the order of ratings varied. The Normal condition had 80% overlap with the other two conditions but differed from them in the order of importance and that classroom observation was rated as one of the top five assessment data instead of BASC-2 Teacher report. Mean ratings for the Normal, LD, and Complicated conditions are in Appendix J, Tables J-4 through J-6, respectively. The mean ratings of assessment data collapsed across the conditions are provided in Appendix I, Table I-7.

Demographic comparison of conditions. ANOVAs and Chi-square analyses were conducted to check whether participants in the three groups (Normal, LD, and Complicated) were comparable on demographic variables. No statistically significant differences were found between the participants across the three conditions on the major demographic variables. Findings of a series of one-way ANOVAs indicated that the three groups were not statistically significant different on (a) mean age of participants, F (2,267) = .07, p = .93, partial eta squared < .01; (b) number of years as certified school psychologists in Pennsylvania, F (2, 269) = .03, p = .97, partial eta squared < .01; (c) number of evaluations conducted in the prior year, F (2, 269) = 2.81, p = .06, partial eta squared = .02; (d) frequency of typically using the discrepancy model during assessment, F (2, 269) = .72, p = .49, partial eta squared = .01; and (e) frequency of using the Response to Intervention model during assessment, F (2, 269) = 2.32, p = .10, partial eta squared = .02.

Findings of the two-way contingency analyses indicated that the experimental conditions did not differ in the number of males to females, X^2 (6, N = 272) = 2.25, p = .33, Cramer's V =

.09 and in the distribution of ethnicities, X^2 (6, *N* = 272) = 8.18, *p* = .43, Cramer's *V* = .12. The three groups also did not differ in the number of school psychologists employed in various systems (school district, residential treatment center, independent or private practice, & intermediate unit) X^2 (6, *N* = 272) = 5.28, *p* = .51, Cramer's *V* = .10 or in the residential region of employment (urban, rural, suburban), X^2 (6, *N* = 272) = 9.43, *p* = .15, Cramer's *V* = .13. Additionally, participants across experimental conditions were equivalent in the number who had attained a master's or doctoral degree, X^2 (6, *N* = 272) = 10.72, *p* = .03, Cramer's *V* = .14. Finally, no statistically significant differences were found in participants across conditions in the program's orientation of training (scientist-practitioner, practitioner-scholar, and practitioner focused), X^2 (6, *N* = 272) = 2.66, *p* = .62, Cramer's *V* = .07. Because the participants were comparable in demographic features across the conditions, it was not necessary to consider demographic features as covariates in subsequent analyses. Summaries of the ANOVAs on the demographic comparisons of the experimental groups are presented in Appendix J.

Analyses of the Hypotheses

Hypothesis 1. It was hypothesized that a statistically significant association would be found between the complexity of the student summary and the number of school psychologists who accurately classified the student in each case, with less accuracy occurring for those in the LD and Complicated conditions, respectively, in comparison to those in the Normal condition. A 3 (complexity of student summary) x 2 (accuracy of classification; yes or no) contingency table was used to test the hypothesis. The finding of the contingency analyses was statistically and practically significant, X^2(2, *N* = 272) = 73.3, *p* < .001; Cramer's *V* = .52. Most of the school psychologists with the Normal case (97.7%) and LD case (86%) accurately classified their student as such, but fewer than 50% of the school psychologists with the Complicated case accurately classified their student (Nonexceptional). Thus, as student summaries varied in complexity, there was a decrease in the number of school psychologists who accurately classified the student in the summary. Specifically, there was approximately a 14% decrease in accuracy between the Normal case and the LD one, a 51% decrease in accuracy between the Normal and Complicated case, and a 40% decrease in accuracy between the LD and Complicated cases. Hypothesis 1 was supported. A summary of the observed counts is presented in Table 9.

Table 9

Summary of 3 x 2 Contingency Analysis of Complexity Level and Accuracy in Classifying Student (N = 272)

Complexity		Accurate	Inaccurate	Total
Normal	Count	86.0	2.0	88.0
	Percent	97.7	2.3	100.0
LD	Count	80.0	13.0	93.0
	Percent	86.0	14.0	100.0
Complicated	Count	42.0	49.0	91.0
	Percent	46.2	53.8	100.0
Total	Count	208.0	64.0	272.0
	Percent	76.5	23.5	100.0

Note. LD = Learning Disability.

Hypothesis 2. Hypothesis 2 was composed of four components, all associated with testing the presence of the representativeness heuristic: (a) the use of a prototype, (b) the likelihood of an accurate classification and client-prototype similarity, (c) the importance of state criteria, and (d) the importance of the prevalence of the classification. In general, the representativeness heuristic was expected to be used more in the Complicated condition in comparison to the other conditions, and to be used more in the LD condition than in the Normal condition.

Hypothesis 2a. It was expected that a greater number of school psychologists in the Complicated condition would use a prototype in assessing their case in comparison to the other conditions, and more in the LD condition were expected to use a prototype than school psychologists in the Normal condition. A 3 (complexity of student summary) x 2 (thought of prototype; yes or no) contingency analysis was conducted to test this hypothesis. The findings were statistically significant: an association was found between the complexity level of the student summary and the number of school psychologists who used a prototype to classify the student, X^2 (2, N = 272) = 18.36, p < .001; Cramer's V = .26. Approximately 7% of the variance in using a prototype was accounted for by the complexity level of the student summary. Twice as many school psychologists in the LD (56%), and in the Complicated (54%) conditions reported using a prototype for their case in comparison to those in the Normal condition (27%). No delineation was found in the use of prototypes between the LD and Complicated conditions. Hypothesis 2a was partially supported. A summary of the observed counts for Hypothesis 2a is provided in Table 10.

Hypothesis 2b. In comparison to the other experimental conditions, school psychologists in the Complicated condition were expected to express a higher likelihood to have classified accurately and to have viewed the case as being similar to a prototype. The same pattern was expected to occur between school psychologists in the LD and the Normal conditions, respectively. Following the protocol of prior research (Garb, 1996), a combination of the likelihood of classifying accurately and the client-prototype similarity was treated as a linear

Table 10

Summary of 3 x 2 Contingency Analysis of Complexity Level and use of Prototype (N = 272)

Complexity		Didn't think of a prototype	Thought of a prototype	Total
Normal	Count	64.0	24.0	88.0
	Percent	72.7	27.3	100.0
LD	Count	41.0	52.0	93.0
	Percent	44.1	55.9	100.0
Complicated	Count	42.0	49.0	91.0
	Percent	46.2	53.8	100.0
Total	Count	147.0	125.0	272.0
	Percent	54.0	46.0	100.0

Note. LD = Learning Disability.

composite reflective of the representative heuristic. A one-way MANOVA was used to test this hypothesis, with complexity level of the student summary as the independent variable. The two dependent variables were the reported likelihood of accurately classifying the case and the client-prototype similarity. Only participants who reported thinking of a prototype were routed to questions about client-prototype similarity, so for this hypothesis, there was a reduction in sample size (n = 125) from the original sample size (N = 272). The assumption of homogeneity of covariance was met (Box's Test p = .26).

The one-way MANOVA was not statistically significant, F (4, 242) = 1.03, p = .40, partial eta squared = .02. School psychologists in the three experimental conditions (Normal, LD, and Complicated) did not rate their cases differently on the linear composite of likelihood of accurate classification and client-prototype similarity. These findings did not support that the representativeness heuristic was used more in the Complicated case than in the other conditions or used more in the LD condition in comparison to the Normal condition. Hypothesis 2b was not supported.

***Hypothesis 2c*.** In the Complicated condition, school psychologists were expected on average to rate state criteria for classification as less important in making a classification decision in comparison to the other conditions. School psychologists in the LD condition were expected to rate state criteria as less important than those in the Normal condition. A one-way ANOVA was used to test this hypothesis. The independent variable was complexity level of the student summary (3 types; Normal, LD, and Complicated) and the dependent variable was the rating of how important state criteria were in classifying the student. The assumption of homogeneity of variance was not met based on the Levene's Test ($p < .001$). However, research (Lindman, 1974, p. 33; Tinsley & Brown, 2000) has indicated that the F statistic is quite robust against violations of this assumption unless sample sizes are unequal. In this case, the sample sizes were unequal and thus any statistically significant findings should be interpreted with caution.

The one-way ANOVA was statistically significant, $F(2, 269) = 11.27$, $p < .001$, partial eta squared = .08, an indication that school psychologists in the three complexity conditions differed in rating the importance of state criteria in classification decisions. Approximately 8% of the variance in the ratings of importance of state criteria was accounted for by the complexity level of the student summaries. A summary of the ANOVA on the importance of state criteria is presented in Table 11.

Table 11

Summary of ANOVA on the Importance of State Criteria to Classification Decisions (N = 272)

Source	*df*	Mean Square	*F*	Sig.	Partial Eta Squared
Complexity	2	9.75	11.27	< .001	.08
Error	269	.87			
Total	272				

To counter the violation of homogeneity of variance, the Games-Howell post hoc test was used to compare the mean differences between the conditions (Morgan, Leech, Gloeckner, & Barrett, 2004). Two of the post-hoc comparisons were statistically significant: individuals who reviewed the Complicated student summary rated state criteria as less important than individuals who reviewed the other student summaries. Thus, the representativeness heuristic was used more so by those in the Complicated case than school psychologists in the other cases but no difference was found in the use of the representativeness heuristic between those in the Normal and LD conditions. Hypothesis 2c was partially supported. A summary of the Games-Howell findings and statistical significance levels is presented in Table 12.

***Hypothesis 2d*.** A final measure of the representativeness heuristic was the use of prevalence of the classification to make a classification decision. School psychologists were expected to rate the prevalence of the classification as less important in the Complicated condition in comparison to the other conditions. Furthermore, those in the LD condition were

Table 12

Post Hoc Analyses (Games Howell) of Importance of State Criteria to Classification Decisions (N = 272)

Post Hoc Analysis	Complexity (I)	Complexity (J)	Mean Difference (I-J)	Std. Error	Sig.	95% C.I Lower Bound	Upper Bound
Games-	Normal	LD	.18	.12	.268	-.09	.46
Howell	Normal	Complicated	.64	.14	< .001	.30	.98
	LD	Complicated	.46	.15	.008	.10	.81

Note. LD = Learning Disability.

also expected to rate the prevalence of classification as less important than those in the Normal condition. A one-way ANOVA was conducted to test this hypothesis. The independent variable was complexity level of the student summary. The dependent variable was the rating of importance of prevalence of classification. With an a priori alpha level of .01, the assumption of homogeneity of variance was not met, (Levene's Test p = .02), but, as indicated earlier, research (Lindman, 1974, p. 33; Tinsley & Brown, 2000) has indicated that the F statistic is quite robust against violations of this assumption unless sample sizes are unequal. In this case, the sample sizes were unequal, and thus, any statistically significant findings should be interpreted with caution. The one-way ANOVA was not statistically significant, F (2, 269) = 1.28, p = .28, partial eta squared = .01. Regardless of the complexity level of the student summaries, school psychologists did not differ in their ratings of how important prevalence of classification was to their decision. Therefore, the representativeness heuristic, as measured by disregard of prevalence of the classification, did not differ among the conditions. Hypothesis 2d was not supported. A summary of the ANOVA for Hypothesis 2d is presented in Table 13.

Table 13

Summary of ANOVA on the Importance of Prevalence of Classification (N = 272)

Source	*df*	Mean Square	*F*	Sig.	Partial Eta Squared
Complexity	2	1.07	1.28	.28	.01
Error	269	.84			
Total	272				

Hypothesis 3. It was hypothesized that the representativeness heuristic would vary based on the complexity level of the student summaries and the accuracy of classification. As with hypothesis 2, the representativeness heuristic was measured in four ways (a) school psychologists' use of a prototype, (b) the combination of participants' reported likelihood of classifying accurately and client-prototype similarity, (c) the importance of state criteria, and (d) the importance of prevalence of classification.

Hypothesis 3a. It was expected that a prototype would be used more by school psychologists who misclassified their case in the Complicated condition rather than in the other two conditions. Additionally, school psychologists were expected to use a prototype more when they misclassified their case in the LD condition rather than in the Normal condition. The initial plan was to run a 3 (complexity level of student summary) x 2 (accuracy of classification; yes or no) x 2 (thought of a prototype; yes or no) contingency analysis to test this hypothesis. However, this analysis was not possible as the sample size for the table cell of the Normal condition with inaccurate classifications (*n* = 2) was not sufficient to run the analysis and the expected count was only 1(instead of the minimum requirement of 5). Thus, the analysis was reduced to a 2 x 2 x 2 design, with only the LD and Complicated conditions used for complexity level. Thus, only one facet of the hypothesis could be tested.

With an a priori alpha level set at .01, the findings were not statistically significant, indicating that the number of school psychologists who thought of a prototype was not associated with the function of the complexity level (LD versus Complicated) of the student summary and the inaccuracy of classification, X^2 (1, N = 184) = 4.77, p = .03; Cramer's V = .23. Hypothesis 3a was either not supported or could not be tested. A summary of the observed counts is provided in Table 14.

Table 14

Summary of 2 x 2 x 2 Contingency Analysis of Complexity Level and Misclassification and Thinking of a Prototype (n = 184)

Complexity	Accuracy		Didn't think of a prototype	Thought of a prototype	Total
LD	Inaccurate	Count	10.0	3.0	13.0
		Percent	76.9	23.1	100.0
Complicated	Inaccurate	Count	21.0	28.0	49.0
		Percent	42.9	57.1	100.0
Total	Inaccurate	Count	31.0	31.0	62.0
		Percent	50.0	50.0	100.0

Note. LD = Learning Disability.

Hypothesis 3b. In comparison to the other experimental conditions, school psychologists in the Complicated condition were expected to be more likely to misclassify their case, express a higher likelihood to have classified accurately, and express higher client-prototype similarity. The same pattern was expected to occur between school psychologists in the LD and the Normal conditions, respectively. Initially, a 3 x 2 MANOVA was planned but due to the insufficient sample size for the cell of the Normal condition with inaccurate classifications (n = 2), a 2 x 2 MANOVA was used to test this hypothesis. Thus, the hypothesis could not be fully tested. The independent variables were the condition of the student summary (two types; LD & Complicated) and the accuracy of the classification (two types; yes or no). The dependent variables were the likelihood of accurate classification and the client-prototype similarity. All assumptions of MANOVA were examined and met, including the assumption of homogeneity of covariance (Box's Test p = .22).

The 2 x 2 MANOVA was not statistically significant, F (2, 96) = 1.12, p = .33, partial eta square = .02. The linear combination of expressed likelihood of accurate classification and reported client-prototype similarity did not differ on misclassification by school psychologists in the two experimental conditions. Thus, the ratings of likelihood of accuracy and client-prototype similarity did not support that the representativeness heuristic was used more by school psychologists with inaccurate classification in the Complicated condition in comparison to the Normal or the LD condition. Hypothesis 3b was either not supported or could not be tested. A summary of the MANOVA for hypothesis 3b is provided in Table 15.

Table 15

Summary of Multivariate Analysis of Variance for Complexity Level on Likelihood and Similarity (n = 184)

Effect	Wilks' Lambda	*F*	Hypotheses *df*	Error *df*	Sig.
Complexity	.99	.63	2.00	96.00	.535
Accuracy	1.00	.13	2.00	96.00	.880
Complexity*Accuracy	.98	1.12	2.00	96.00	.332

***Hypothesis 3c*.** It was hypothesized that a greater number of school psychologists in the Complicated condition would misclassify their case and provide a low rating of importance for state criteria for classification in comparison to the other conditions. Furthermore, school psychologists in the LD condition were expected to be more likely than those in the Normal condition to misclassify their case and express a low importance rating for state criteria for classification. A 3 x 2 ANOVA was planned to test this hypothesis. However, a 2 x 2 ANOVA was conducted due to the insufficient sample size for the cell of the Normal condition with inaccurate classifications ($n = 2$), which resulted in data from the Normal condition being excluded. Thus, the hypothesis could not be fully tested. All assumptions of ANOVA were examined and met, except for the assumption of homogeneity of variance (Levene's test $p < .001$), but research (Lindman, 1974, p. 33; Tinsley & Brown, 2000) has indicated that the F statistic is quite robust against violations of this assumption unless sample sizes are unequal. In this case, the sample sizes were unequal and thus any statistically significant findings should be interpreted with caution. The two-way ANOVA was not statistically significant, F (1, 180) =

4.11, *p* = .04, partial eta squared = .02. The use of state criteria for determining classification did not differ on misclassification by school psychologists in the experimental conditions. Thus, the ratings of the importance of state criteria for classification did not support that the representativeness heuristic was used more by school psychologists with inaccurate classification in the Complicated condition in comparison to the LD condition. Hypothesis 3c was either not supported or could not be tested. A summary of the ANOVA is provided in Table 16.

Hypothesis 3d. Based on complexity level of student summary complexity and accuracy in classification, school psychologists were expected to differ in how important prevalence of classification was to making classification decisions. Participants in the Complicated condition who had misclassified the student were expected to rate prevalence of classification as less important in determining classification than the other conditions. In turn, those in the LD condition who had misclassified the student were expected to rate prevalence of classification as

Table 16

Summary of ANOVA on Importance of State Criteria to Classification by Complexity and Accuracy of Classification (n = 184)

Source	*df*	Mean Square	*F*	Sig.	Partial Eta Squared
Complexity	2	1.50	1.95	.144	.01
Accuracy	1	11.65	15.17	< .001	.05
Complexity*Accuracy	2	2.20	2.89	.057	.02
Error	266	.77			
Total	272				

less important in determining classification than the Normal condition. To test hypothesis 3d, a 3 x 2 ANOVA was planned, but reduced to a 2 x 2 ANOVA due to the insufficient number (*n* = 2) of participants in the Normal condition who misclassified the case. Thus, the hypothesis could not be fully tested. The assumption of homogeneity of variance was not met (Levene's test; $p = .04$), and any statistically significant findings should be interpreted with caution when sample sizes are unequal as they were in this case (Lindman, 1974, p. 33; Tinsley & Brown, 2000). The two-way ANOVA was not statistically significant, F (1, 180) = .01, $p = .96$, partial eta squared < .01. Regardless of the experimental condition and the level of accuracy in classification, school psychologists did not differ in their regard of the prevalence of the classification. Hypothesis 3d was either not supported or could not be tested. A summary of the ANOVA is presented in Table 17.

Table 17

Summary of ANOVA on Importance of Prevalence to Classification by Complexity Level and Accuracy of Classification (n = 184)

Source	*df*	Mean Square	*F*	Sig.	Partial Eta Squared
Complexity	1	1.82	2.33	.129	.01
Accuracy	1	.17	.22	.639	< .01
Complexity*Accuracy	1	.01	.01	.961	< .01
Error	180	.78			
Total	184				

Hypothesis 4. It was hypothesized that school psychologists in the three experimental conditions (Normal, LD, and Complicated) would differ on their importance ratings for assessment data. The linear composite of the five most important assessment data across all conditions was expected to be rated as more important to school psychologists in the Normal condition in comparison to the other conditions and more important to those in the LD condition in comparison to those in the Complicated condition. A one-way MANOVA was conducted, with the condition of the student summary (3 types; Normal, LD, and Complicated) as the independent variable and the ratings of importance of assessment data by school psychologists as the dependent variables. The dependent variables were the assessment data that had received the five highest ratings from the school psychologists across the sample: (a) achievement score, (b) state criteria, (c) IQ, (d) response-to-classroom modification, and (e) BASC-2 Teacher report.

All assumptions of MANOVA were met, except homogeneity of covariance (Box's Test of Equality of Covariance Matrices ($p < .001$). The decision to proceed with the analysis was based on several reasons. One, statistical significance of the Box's M test can be easily obtained due to its sensitivity to large sample sizes (Huberty & Petoskey, 2000; Tabachnick & Fidell, 2001). Two, large sample sizes ($n > 50$; Finch & Davenport, 2009) and large group sizes ($n > 6$ times number of DVs; Hubert & Petoskey, 2000) make the MANOVA robust to violation of assumptions. Three, Pillai's Trace criterion was used as the multivariate statistic instead of the traditional Wilks' lambda because its use increases robustness of the MANOVA when homogeneity of covariance has been violated and sample size of the cells are unequal (Tabachnick & Fidell, 2001). Finally, statistically significant findings were interpreted with caution.

The MANOVA was statistically significant, F (10, 532) = 4.65, $p < .001$, partial eta squared = .08, indicating that the school psychologists in the three conditions, Normal, LD, and Complicated, differed in their mean ratings on the linear composite of achievement score, state criteria, IQ, response-to-classroom modification, and BASC-2 Teacher report. The effect size indicated that 8% of the variance in ratings of the importance of the five assessment data was accounted for by the complexity level of the student summaries.

To interpret how the three conditions (Normal, LD, and Complicated) differed on the linear composite of the importance of the five assessment data, a descriptive discriminant analysis was conducted. As expected a maximum of two discriminant functions emerged, with the first discriminant function accounting for 12% of the variance and the second function accounting for almost 4% of the variance. Based on the effect size (salient $R \geq |.30|$) and the statistical significance of the dimension reduction analysis, only the first function was interpreted: Functions 1 to 2 Wilks' lambda = .84, X^2 (10) = 45.34, p < .001; Function 2 Wilks' lambda = .96, X^2 (4) = 10.31, p = .04. A summary is presented in Table 18 of the standardized coefficients and correlation of the functions to the five most important assessment data. The standardized coefficients for the first discriminant function ranged from -.44 to .59, with the importance of response-to-classroom modification data contributing the most to the function, followed by the importance of state criteria, and the importance of achievement scores. The importance of BASC-2 Teacher report had a negative contribution to the function. The structure coefficients for the first function ranged from .01 to .77, with the importance of state criteria sharing approximately 59% of its variance to the function, followed by importance of response-to-classroom modification (50%), and importance of achievement score (22%). In contrast, the importance of IQ score (6%) and BASC- 2 Teacher report (< 1%) shared little variance to the function. Thus, ratings of the importance of response-to-classroom modification and state

Table 18

Standardized Coefficients and Structure Coefficients for Descriptive Discriminant Analyses for the Five most Important Assessment Data (N = 272)

Assessment Data	**Function 1**		Function 2	
	B	***s***	*B*	*s*
State Criteria	**.56**	**.77**	-.23	-.19
Response-to-classroom modification	**.59**	**.71**	.74	.60
Achievement Score	**.39**	**.47**	-.62	-.51
IQ score	**-.11**	**.24**	-.25	-.44
BASC- 2 Teacher Report	**-.44**	**.01**	.40	.22
	R^2_{c1} = .12		R^2_{c2} = .04	

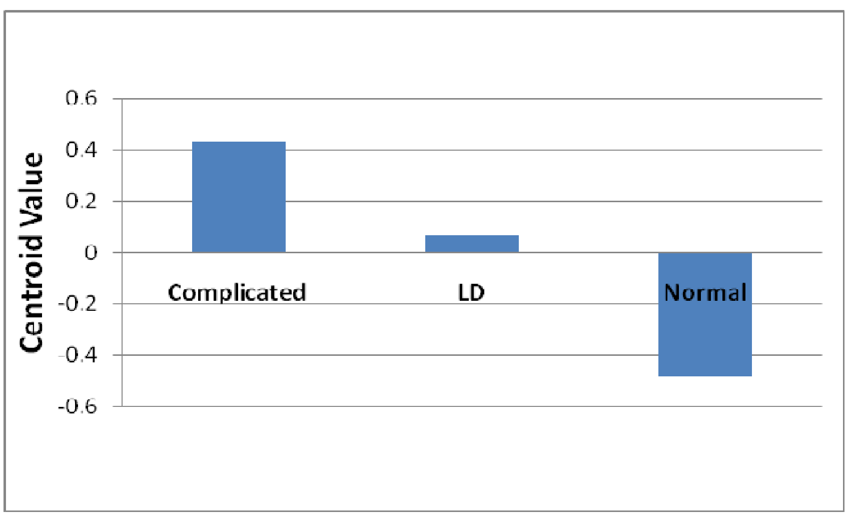
Note. Statistically significant functions are bolded. *B* = standardized coefficient; *s* = structure coefficient.

criteria best described the discriminant function, followed by importance of achievement score. Response-to-classroom modification, state criteria for identification, and academic achievement scores are all associated with the monitoring of progress and academic standing. Additionally, these assessment data have been considered by the Pennsylvania Department of Education as essential to the identification of students with special educational disabilities (IDEA, 2004). Based on the high correlations of academic achievement and progress monitoring to the function, the variate was named *Academic Monitoring.*

A comparison of the centroids on the variate for the experimental conditions showed that on average school psychologists in the Normal condition (.43) were most likely to rate the variables of Academic Monitoring as important in determining a classification for the case in comparison to the other school psychologists. School psychologists in the Complicated condition were least likely to rate (-.48) Academic Monitoring criteria as important in their classification decision, whereas school psychologists in the LD condition fell in the middle (.06) on the set of criteria. A pictorial representation of the centroids on Academic Monitoring is provided in Figure 1.

Hypothesis 4 was supported in that school psychologists in the three conditions differed in how they rated the importance of the linear composite of achievement score, state criteria, IQ, response-to-classroom modification, and BASC-2 Teacher report, labeled as Academic Monitoring criteria. Academic Monitoring was rated as more important to those in the Normal condition, less important to the LDcondition, and then even less important to those in the Complicated condition.

Figure 1. Display of group separation on the discriminant variate of academic monitoring.



Note. *N = 272.* Low scores indicate that participants rated Academic Monitoring as less important. High scores indicate that Academic Monitoring was rated as more important.

Hypothesis 5. It was expected that the average response of school psychologists to the top five ratings of importance of assessment data would vary based on the interaction between the complexity level of the student summary (Normal, LD, and Complicated) and the accuracy of classification (yes or no). No specific direction was hypothesized. Initially a 3 x 2 MANOVA was to be conducted to explore the hypothesis; however, data from the Normal condition could not be included in this analysis due to the unusually low number of school psychologists in the Normal condition who misclassified the student ($n = 2$). While the sample size was adequate to run the MANOVA, the minimum sample size of the cell was violated ($n >$ number of DVs, in this case 5 DVs; Tabachnick & Fidell, 2001). Therefore, only the LD and Complicated conditions could be used in the analysis, resulting in a 2 x 2 MANOVA to test hypothesis 5. Thus, the hypothesis could not be fully tested.

The two independent variables were the complexity level of the student summary (two types; LD & Complicated) and the accuracy in classification (two levels; yes or no). The dependent variables were the five most important assessment data rated by all school psychologists. All assumptions were met with the exception of homogeneity of covariance (Box's M Test $p < .001$). Based on the same rationale for hypothesis 4, despite the violation of this assumption, the one-way MANOVA was considered robust (Finch & Davenport, 2009; Huberty & Petoskey, 2000; Tabachnick & Fidell, 2001). The main effect of accuracy of classification on the linear composite of the five assessment data rated as most important (achievement score, state criteria, IQ score, response-to-classroom modifications, & BASC-2 teacher report) was statistically significant, F (5, 176) = 4.99, $p < .001$, partial eta squared = .12. The interaction effect of complexity level of student summary and accuracy of classification on the linear composite was also statistically significant, F (5, 176) = 3.91, $p = .002$, partial eta squared = .10. Because the interaction effect of the MANOVA encompassed the statistically significant main effect as well, it is recommended to interpret the interaction effect (Nichols, 1993). The effect size of the interaction indicated that approximately 10% of the variance in the importance of assessment data was accounted for by the interaction between complexity level of the student summaries and accuracy of classification. In Table 19 a summary is provided of the results of the two-way MANOVA.

An examination of the univariate statistics indicated that the interaction of complexity level and accuracy of classification was statistically significant on the perceived importance of response-to-classroom modifications, F (1, 180) = 12.73, $p < .001$, partial eta squared = .07. The interaction effect was plotted for this outcome variable and is depicted in Figure 2. The figure showed that school psychologists in the Complicated condition who were accurate in classification rated response-to-classroom modifications more important than those in the LD condition who were accurate in classification. Furthermore, participants in the LD condition

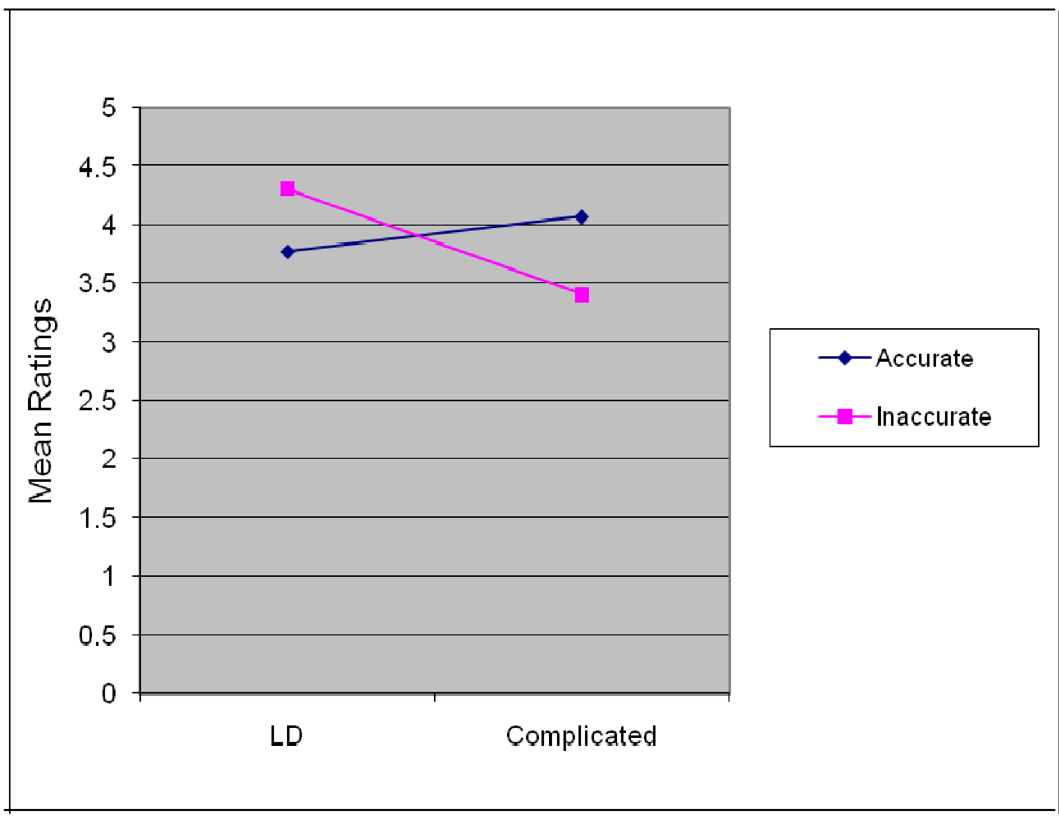
Table 19

Summary of 2 x 2 Multivariate Analysis of Variance for Complexity Level and Accuracy on the Importance of top Five Assessment Data (n = 184)

Effect	Pillai's Trace	*F*	Hypotheses *df*	Error *df*	Sig.
Complexity	.03	1.11	5.00	176.00	.359
Accuracy	.12	4.99	5.00	176.00	< .001
Complexity*Accuracy	.10	3.91	5.00	176.00	.002

whose classification of the student were inaccurate reported using response-to-classroom modifications more so than those in the Complicated condition whose classification was inaccurate. Thus, the general premise of hypothesis 5—an expected interaction effect of complexity level of the student summary and accuracy of classification on the top five assessment data rated as important in decision-making—was partially supported, as the Normal condition could not be included in the design.

Figure 2. Display of means for the rating of importance of response-to-classroom modifications data.



Note. *n = 184*

Discussion

The purpose of Study 2 was to examine whether Pennsylvania school psychologists use the representativeness heuristic in decision-making and which assessment data are considered important in determining special education classification. Specifically, five hypotheses were tested. One was supported, three were partially supported, and one was not supported. Implications of these findings, as well as the limitations and recommendations for future research, are discussed.

Hypothesis 1: Classification of Data Summaries

The findings provide support for Hypothesis 1. The complexity of the student data summaries was associated with the accuracy in classifying the respective student. Among the three complexity conditions, school psychologists in the Complicated condition were most likely to misclassify the student, those in the LD condition were less likely to misclassify, and those in the Normal condition were least likely to misclassify the student. Assessment data summaries were designed to differ in terms of level of complexity. The Normal data summary was designed to be straightforward and described a student who was average in all assessed domains. The LD data summary, which was developed to be slightly more difficult, described a student with low grades, discrepant IQ and achievement test scores, and minimal response to classroom modifications. The Complicated data summary was created to be the most complex and described a student whose status bordered between specific learning disability and nonexceptional. The referral information and reports from the teacher suggested learning problems, but the IQ and achievement scores were commensurate, and grades were mostly in the satisfactory range. Two conclusions can be drawn from the findings. One, the manipulation of creating different psychoeducational student data summaries worked the way it was intended.

Two, school psychologists do respond differently to cases based on the level of complexity of the case.

These finding replicate the results of previous research (e.g., Cioffi & Markham, 1997). Cioffi and Markham found that vagueness increases the complexity level of situations. Thus, as planned, the manipulation of ambiguity level of the student summaries resulted in these cases varying in complexity level. However, unique from prior research (Cioffi & Markham, 1997), this study also assessed whether the complexity level of situations is associated with the accuracy of decisions. Post hoc inspection of the data revealed that the most complex case (nonexceptional) led to 49 of 91 cases (54%) being misclassified as describing a student with either a specific learning disability (n = 40) or an other health impairment (n = 9). It is possible that these errors can be attributed to school psychologists (a) focusing on assessment data that suggests a learning problem (e.g., below average grades and negative teacher comments), (b) concentrating on assessment data from teachers, or (c) overlooking the consistency between ability and achievement levels.

Almost one-third of the variance (27%) in accuracy of classification was accounted for by the complexity level of the student summaries. That school psychologists would make such errors is a concern because their decisions have significant implications for students and schools. For example, it is possible to hinder the academic development of a student by recommending a placement that is more restrictive than necessary (Fuchs, D., Fuchs, & Fernstrom, 1992). It is also possible for over identification of students to lead to a misallocation of funds for school programming (National Coalition on Workforce and Disability, 2004). Like Study 1, school psychologists in Study 2 were exceptionally accurate in their classification of the student in the Normal condition. Therefore, when the case was relatively simple, school psychologists were able to discern that the student did not have a disability. This finding suggests that over identification of students is not universal and depends on the context of the student summary.

Hypothesis 2: Complexity Level and Representativeness Heuristic

Hypothesis 2 was partially supported. This hypothesis stated that the representativeness heuristic would be used more often in the Complicated condition in comparison to the other two conditions and more by school psychologists in the LD condition in comparison to the Normal condition. To increase the reliability and validity of measuring the representativeness heuristic, it was operationalized in four ways: (a) the use of a prototype, (b) the likelihood of an accurate classification and client-prototype similarity, (c) the importance of state criteria, and (d) the importance of the prevalence of the classification. These methods have been used previously in measuring the representativeness heuristic (Cioffi & Markham, 1997; Garb, 1996).

As expected, school psychologists in the LD and Complicated conditions used prototypes almost two times more than participants in the Normal condition. This result is consistent with findings of Study 1, where school psychologists in the Normal condition considered prototypes less than in the LD and Complicated conditions. The result is also consistent with Cioffi and Markham's (1997) findings. Cioffi and Markham found that midwives who reviewed cases without ambiguous data were less likely to base decisions on archetypes than when the cases were more complex. Cioffi and Markham credited this finding to midwives attempting to reduce the cognitive demand in the more difficult tasks. Thus, the parallel in findings suggests that school psychologists may also be using prototypes to reduce the cognitive demands of complex cases. For example, basing a decision on client-prototype similarity can reduce the number of assessment data considered.

Approximately 7% of the variance in using a prototype was accounted for by the complexity level of the student summary. Thus, two issues appear to be present. First, there is some degree, although modest, of using a prototype that is related to complexity, indicating that the representativeness heuristic may influence decision-making. Second, factors other than complexity level are also contributing to the use of prototypes. Potential factors might be the ability to think of a suitable prototype, recency of prototype exposure, and expectation that the data might be indicative of a disability.

Another expected finding is that when presented with a complicated case versus a less ambiguous case (Normal or LD student summaries), school psychologists rate state criteria as less important to their decisions. While this finding is not consistent with results of Study 1, in which state criteria was rated as *moderate* to *very important* across all three conditions, the finding is consistent with research that disregard of established criteria is an indicator of the representativeness heuristic (Garb, 1996; Tversky & Kahneman, 1983). Garb found that clinical psychologists and clinical psychology interns were likely to disregard DSM-III criteria for mental health disorders when diagnosing patients with personality disorders. Thus, both groups of psychologists may disregard criteria when trying to determine classification for a complex case. The disregard of state criteria when reviewing complicated cases is a concern because these guidelines have been established to ensure that school psychologists are consistent in their classification decisions. Therefore, the lack of consideration of state criteria may increase the risk that decisions will vary from one practitioner to another and may in turn reduce the consistency in services provided to students.

Regarding the meaningfulness of findings, 8% of the variance in using state criteria was accounted for by the complexity level of the student summary. Thus, while modest, there is some indication that the use of the representativeness heuristic, measured as disregard of state criteria, is associated with the complexity level of the case. A possible reason for this association is that state criteria may not provide enough direction for classifying ambiguous cases, leading decision makers to rely on other information or strategies. As indicated previously, making decisions without regard for established criteria results in inconsistency in working with students.

When considering the meaningfulness of the association between the importance of state criteria and the complexity of cases, it is also necessary to account for competing explanations. An alternative explanation of the finding is that the heuristic measured may not have been representativeness, but was reflective of a general heuristic or specifically the base rate heuristic (Uhlmann, Brescoll, & Pizarro, 2007). It is known that general heuristics are used when the decision maker does not follow a standard decision making process. Instead the decision maker uses a process that has not formally been accepted by the discipline (Gigerenzer & Goldstein, 1996; Hammond, 1996; Hogarth & Karelaia, 2007; Tversky & Kahneman, 1974; 1983). Thus, the disregard of established criteria by school psychologists regardless of the student data summary reviewed may be indicative of a general or another heuristic, not representativeness (Garb, 1996). In essence, disregard of established criteria may not be the best stand-alone measure of the representativeness heuristic, but of another form of heuristic.

Three unexpected findings surfaced about the complexity level of the student summaries and use of the representativeness heuristic. According to Cioffi and Markham (1997), archetypes are more likely to be used in situations of greater complexity to reduce the cognitive demand. Therefore, as expected school psychologists in the Complicated condition used prototypes. However, they were also expected to use prototypes more than school psychologists in the LD condition. This finding did not emerge in either Study 1 or 2. The prevalence rates of specific learning disabilities need to be considered to understand the lack of delineation between the two groups (LD and Complicated conditions). On average, school psychologists reported that, of the total evaluations completed in the prior year, almost half of the students met criteria for a specific learning disability (46%), 16% met the criteria for other health impairments, and 30% were determined to have no disabling condition. This finding is similar to statistics that show that the largest percentage (approximately 50%) of children who receive special education services in the United States is composed of students with a specific learning disability (Pastor & Reuben, 2008). Thus, it is possible that the use of a prototype not only depends on the complexity of a case but also on how easily school psychologists can recall and use prototypes of students with the suspected classification**.** Further examination is necessary to determine the conditions in which the use of prototypes depends on the complexity of cases, as suggested by Cioffi and Markham (1997), versus the ease in which prototypes come to mind, as suggested by the finding just reported.

The second unexpected finding is that school psychologists in the Complicated condition, in comparison to the other conditions, did not express a higher likelihood to have classified accurately and to have viewed their case as being similar to a prototype. This pattern also did not occur between school psychologists in the LD and the Normal conditions. These two findings are unexpected given that clinical psychologists' and psychology interns' reported likelihood of accurate classification was correlated with client-prototype similarity (Garb, 1996). The selection of statistical analysis may be one reason for the difference in findings between Garb and this study. Garb operationalized the representativeness heuristic by using bivariate correlations between likelihood of accurate classification and client-prototype similarity. In the current study, a multivariate approach, MANOVA, was used to create a linear composite of the two variables. However, an examination of the bivariate correlations within each condition of the current study (Normal, LD, and Complicated) indicated that the findings did not change. Although in the Normal condition, there was a higher correlation between likelihood of accurate classification and client-prototype similarity (r = .32) than what was found in the other conditions (LD r = .05; Complicated r = -.06), none of the correlations were statistically significant, even at the .05 level, or meaningful ($r \geq .30$). Thus, using Garb's method to assess the representativeness heuristic would have not changed the current findings. It is more likely that other issues, rather than the statistical method used, accounted for the difference in findings.

A more likely reason for the difference in findings between Garb's (1996) study and the current study is that the relationship between perceived likelihood of accurate classification and client–prototype similarity might not be a reliable measure of the representativeness heuristic. In the current study participants were presented with cases of various complexities, whereas complexity was not a variable in Garb's study. Therefore, it could be that the ambiguity presented in the cases of the current study might be a factor in measuring and detecting the representativeness heuristic. The last explanation for the unexpected finding in the current study is that there might be differences between the participants of the two studies (e.g., field, age, and era) that may have influenced responses to questions about likelihood of accurate classification and client-prototype similarity.

The third unexpected result regarding the emergence of the representativeness heuristic is that school psychologists in the three conditions did not differ in considering prevalence of classification as important to their decision-making. When looking at the mean ratings of importance of assessment data, on average each condition reported prevalence rates as being *not important* to *a little important* to their decision, which is inconsistent with previous research

(Cioffi & Markham, 1997). Cioffi and Markham found that a measure of the representativeness heuristic was when midwives provided low base rates for clinical conditions and overdiagnosed patients with the same clinical condition. Cioffi and Markham also found that the representativeness heuristic was more frequently used in complex situations. Therefore, it is perplexing that all groups of school psychologists barely used base rates. However, it should be noted that this finding is consistent with results in Study 1. In Study 1 it was found that school psychologists not working in Pennsylvania rated prevalence of classification as *not important* to a *little important* to their decisions across all conditions. The comparable findings of Study 1 and Study 2 further support that (a) prevalence ratings may be used differently in different professions and (b) the disregard of prevalence rates might be a common heuristic all school psychologists use.

In summary, the use of the representativeness heuristic, as measured in four ways, did not differ among the three conditions to the extent hypothesized. In fact, school psychologists in the Normal condition only differed from the other groups when the representativeness heuristic was measured using either prototypes or state criteria. Furthermore, the LD and Complicated conditions did not show a statistically significant difference in their use of the representativeness heuristic, except when measured as state criteria. The lack of delineation between the LD and Complicated condition suggests that the process used to determine special education eligibility might be very similar when a student clearly has a learning disability and when a student is borderline of having two diagnoses. It is possible that the need to rule-out or confirm a learning disability in both of these situations makes the decision-making process more similar than when it is obvious that the student does not have a disability, such as in the Normal condition. The similarity of the LD and Complicated conditions to each other than either to the Normal condition is supported by other findings reported previously. For example, on average school psychologists in the LD and Complicated conditions agreed 100% on which five assessment data were most important to their decisions. However, those in the Normal condition were on average 80% in agreement with the LD and Complicated conditions.

Hypothesis 3: Representativeness Heuristic and Misclassification

The third hypothesis was that the use of the representativeness heuristic, which was measured in four ways, would vary between conditions (Normal, LD, and Complicated) and misclassification of cases. Due to an unexpectedly low number of school psychologists who misclassified the student after reviewing the Normal summary, the comparison was made only between the LD and Complicated conditions. The use of the representativeness heuristic, regardless of how it was measured (use of prototypes, disregard of state criteria, disregard of prevalence rates, perceived likelihood of classification and client-prototype similarity), did not vary based on the complexity level of the LD and Complicated student summaries and the accuracy of classification. The findings of Hypothesis 3 are not surprising given the results of Hypothesis 2. As reported, school psychologists in the LD and Complicated conditions functioned the same when using the representativeness heuristic as measured in three of the four ways (use of prototype, disregard of prevalence rates, and the linear composite of likelihood of accurate classification and client-prototype similarity). Therefore, the use of the representativeness heuristic, as measured by these three ways, should not have varied much in terms of accuracy for the two conditions.

In Hypothesis 2, the only situation in which use of the representativeness heuristic varied between school psychologists in the LD and Complicated conditions was when state criteria was used as the measure. However, for the current hypothesis, not only was the interaction effect of complexity level and accuracy on state criteria not statistically significant, but the main effect of accuracy of classification was not statistically significant either. These finding are unexpected given that prior researchers (Cioffi & Markham, 1997; Dumont, 1993; Garb, 1996) have all found that the representativeness heuristic is related to errors in diagnostic decisions. There are several potential explanations for the differences in findings between the current study and previous research. First, misclassification of students may not be directly associated with the use of the representativeness heuristic. Second, use of established criteria was determined through inference in previous research, but measured via self-report in the current study, which could have affected the outcomes. Third and finally, to measure the use of established criteria other studies have used correlational designs and have ignored the effect of complexity level. In the current study, the design allowed for established criteria to be measured by the combined effects of accuracy of classification and complexity of condition. Therefore, the research design and statistical analyses employed might have also influenced the outcomes of the studies.

Conclusions could not be drawn about whether the use of the representativeness heuristic varied between the Normal condition and the other conditions, and the extent of misclassification of the respective student summaries. It is possible that had these analyses been viable that statistical significance would have emerged. This possibility is based on the assumption that the decision-making process used by school psychologists in the Normal condition differs from what is used by school psychologists in the other two conditions. The differentiation of decision making by school psychologists in the Normal condition is assumed based on findings that on average these school psychologists (a) are more accurate than in both the LD and Complicated conditions, (b) show greater difference in their use of the representativeness heuristic, and (c) have greater distinction in their use of assessment data.

Hypothesis 4: Complexity Level and Rating of Assessment Data

Hypothesis 4 was partially supported. It was expected that a combination of the five most important assessment data (achievement score, state criteria, IQ, response-to-classroom modifications, and BASC-2 Teacher report) would weigh more on decision-making in the Normal condition in comparison to the other conditions. The combination was also expected to be more important to decision-making in the LD condition in comparison to the Complicated condition. These expectations aligned with other hypotheses that school psychologists in the conditions would vary on their use of important assessment factors. Response-to-classroom modifications, achievement score, and state criteria formed a statistically significant linear composite, labeled as Academic Monitoring, that clearly separated the three sets of school psychologists. As expected, Academic Monitoring was more important to those in the Normal condition, less important to those in the LD condition, and even less important to those in the Complicated condition in making classification decisions.

Prior work (Della Toffalo & Pedersen, 2005; Salvia & Ysseldyke, 1978; Thurlow & Ysseldyke, 1980) has found IQ score, academic achievement, and behavioral data to be highly important to classification decisions. However, other than Study 1, no prior research has investigated whether the importance of assessment data varies with the complexity level of the case. Additionally, no studies have investigated the importance of state criteria and progress-monitoring data (e.g., response-to-classroom modification and RTI) to classification decisions. When the assessment data that were considered to be most important to school psychologists in Study1 and Study 2 are compared, there are very few congruencies, an indication that the importance of assessment data is not as simple as conceptualized by prior research (e.g., Della Toffalo & Pedersen, 2005; Salvia & Ysseldyke, 1978). For example, when school psychologists review cases that have conflicting or ambiguous information, some assessment data (e.g., academic progress) become less important. A possible explanation is that these school psychologists are likely to disregard important information and instead focus on subsets of data that are in concurrence or data that confirm a diagnosis that was initially expected. This explanation is supported by findings in the current study that school psychologists use general heuristics (e.g., ignoring important information and not following standard decision-making procedures) during decision-making and their use at times depends on the context of the situation.

In terms of the meaningfulness of findings, there are two considerations. First, 12% of the variance in rating of importance of Academic Monitoring was accounted for by complexity level of the student summaries, indicating that some of the influence of data related to Academic Monitoring is associated with the complexity level of the student summary. Second, variance in the rating of Academic Monitoring remains unaccounted. Other factors that may be contributing to the weight given to the Academic Monitoring composite are clarity of data and familiarity with the assessment procedures listed in the student summary (Della Toffalo & Pedersen, 2005; Heubner & Cummings, 1986; LaRocco & Murdica, 2009).

Hypothesis 5: Rating of Assessment Data & Misclassification

Hypothesis 5 was partially supported, as it could not be tested fully. It was expected that the importance of the five most influential assessment data (achievement score, state criteria, IQ, response-to-classroom modifications, and BASC-2 Teacher report) would vary based on an interaction between complexity level of the student summary and accuracy of classification. As with the third hypothesis, the Normal group was excluded from this analysis due to only two school psychologists misclassifying the student. As expected, there was an interaction effect of complexity level of the student summary and accuracy of classification on the ratings of importance of the top five assessment data. The interaction effect was mostly due to the ratings of importance of response-to-classroom modifications, showing an inverse relationship between the LD and Complicated conditions. In regards to correct classification, school psychologists in the Complicated condition rated response-to-classroom modifications as more important than those in the LD condition. However, in regards to misclassification, school psychologists in the LD condition rated response-to-classroom modifications as more important than those in the Complicated condition.

Based on a search of the extant literature, no studies have examined the variables of complexity level, classification accuracy, and the use of assessment data at the same time. The current study seems to be the only one that has assessed whether the reported use of assessment data varies with the complexity level of cases and the accuracy of classification. Approximately 7% of the variance in ratings of importance of response-to-classroom modifications was accounted for by the complexity level of the student summaries and accuracy of classification. It is possible that the reported uncertainty of school personnel regarding effective use of progress-monitoring data (Bender & Shores, 2007; Samuels, 2008) may explain this observed association.

Limitations

Several factors may limit generalizability of the findings. First, as in Study 1, the student data summaries did not include all of the data typically available to school psychologists when making classification decisions. In an attempt to keep the survey manageable for participants, the protocols only included data identified as important in Study 1 and other prior research (e.g., Cioffi & Markham, 1997; Garb, 1996; Knoff, 1983; Thurlow & Ysseldyke, 1980). For example, the student summaries contained brief descriptions of classroom modifications and overall progress within 3 weeks. However, many school psychologists use additional data (e.g., trends of progress) and as a result, some participants expressed discomfort with being asked for a classification without more detailed response to intervention information. Their discomfort is appropriate given recent legislature (IDEA, 2004) that allows for the inclusion of response to intervention data as a tool for assessing specific learning disabilities. Due to this discomfort, ratings of the importance of response-to-classroom modification to participants' decisions might under-represent the relative importance of response to intervention data.

Second, although the number of participants in Study 2 was larger than in Study 1 and adequate for most of the analyses, at times the sample size within cells was too small for some of the proposed analyses (e.g., 3 x 2 contingency table and 3 x 2 ANOVA). There were only two participants in the Normal group who misclassified the student in the summary— a confirmation that the manipulation of the Normal condition worked as designed. However, a cell size of two is not large enough to run most statistical analyses and made it impossible to test whether those who misclassified students differed in the three conditions on variables such as importance of assessment data and use of the representativeness heuristic. Using partial eta square for effect size and violating homogeneity of variance also may have influenced the meaningfulness of the findings. Partial eta square is typically associated with a higher effect size than other measures such as eta square (George & Mallery, 2001). Additionally, despite steps taken to reduce the robustness of statistical analyses, the assumption of homogeneity of variance was violated, indicating that some of the results should be interpreted with caution.

Third and finally, technological problems occurred during data collection, despite attempts to debug the online process during the pilot study. One notable difficulty was that at times participants were unable to proceed even though the survey was not complete. These technological difficulties lead to some missing data and reduction of sample size. Missing data and reduced sample sizes lower the power of analyses and potentially introduce ambiguity into inferences that can be drawn from the study (Tabachnick & Fidell, 2001).

Future Research

The findings and limitations of the study stimulate questions for future research about the decision-making of school psychologists. One useful research objective is to further explore the use of prototypes. Cioffi and Markham (1997) found that midwives were more likely to use prototypes when cases were more complex. In the current study, it was suggested that school psychologists not only use prototypes when cases are complex but also when cases are relatively simple and familiar to school psychologists. Future research should use multiple measures to probe clinicians about the situations in which prototypes are more and less likely to be used.

Another potential objective for future research is to identify reliable measures of the representativeness heuristic. Two measures of the representativeness heuristic used in previous work were not effective in the current study. It is unknown whether differences in results are due to differences in sample or method. It is also unknown whether some of the measures tap only the representativeness heuristic or heuristics in general. Therefore, replication studies are warranted with samples of various practitioners. Thus, a useful research objective is to consider whether the following are useful measures of the representativeness heuristic: (a) disregard of base rates, (b) disregard of established criteria, and (c) an interaction of likelihood of accuracy with client-prototype similarity.

Future research could also assess whether school psychologists use heuristics other than representativeness. Two other common heuristics are the availability and the anchoring and adjustment heuristics (Tversky & Kahneman, 1974); other practitioners have been found to use these heuristics (Bieri et al., 1963; Friedlander & Stockman, 1983). A focus of future studies could be on whether school psychologists use these other heuristics in determining classification.

A fourth direction for future research is replication of this study with school psychologists from states other than Pennsylvania. Participants in this study were recruited exclusively from Pennsylvania, and while their demographics are otherwise representative of the national population of school psychologists, methods of determining special education eligibility vary from state-to-state criteria (IDEA, 2004). For example, states differ in their support of RTI and discrepancy models as well as the level of data required for a student to receive special education services (Bellis, 2002).

A fifth direction for future research is to examine under what conditions heuristics may or may not be useful. In the current study, the use of the representativeness heuristic, regardless of how it was measured, did not vary based on the complexity level of the LD and Complicated data summaries and the accuracy of classification. However, the Normal student summaries were not included in this analysis. It is possible that if the Normal condition was included that a difference would have emerged between the Normal condition and the other conditions on accuracy of classification. Prior researchers have suggested that heuristics, in general, may be used to explain both optimal and erroneous decisions (Simon, 1990; Tversky & Kahneman, 1974; 1983). Some questions for future studies are as follows: When does using heuristics increase or decrease the consistency of decisions? What are other potential benefits and drawbacks of using heuristics? In what situations does the use of heuristics lead to greater accuracy than when it is not used?

Finally, it is recommended that a study be conducted with a sufficient sample size (e.g., $N = 3{,}000$) to generate a large enough subsample of school psychologists who misclassify students in low complexity cases. It has been estimated that there are 32,300 school psychologists in the United States (Jimerson, Skokut, Cardenas, Malone, & Stewart, 2008). It would be important to know whether the findings obtained in this study would be found in a national sample. For example, approximately 2% of school psychologists in Study 2 misclassified a student as having Other Health Impairment when the student should have been easily classified as nonexceptional. If this finding was supported in a national sample of school psychologists, it would mean that about 640 of them would misclassify a nonexceptional case and would have significant implications about ameliorating this assessment error as well as in training future school psychologists. More information is needed about the decision-making process of school psychologists who misclassify simple cases.

Implications

The results of this study have several applications to the training and practice of school psychologists. First, findings that school psychologists' accuracy of classification varies with the type of case imply that an intervention is needed to prevent misclassification in certain situations. Because school psychologists appear to be more likely to make classification errors when cases have ambiguous data, school psychologists should consider consulting with colleagues or collecting additional data when assessing such complex cases. Claims have been made that using a team approach to decision-making and collecting data from multiple sources facilitates accurate decisions and appropriate educational placement of children (Pierangelo & Giuliani, 2007; Reschly, 1996; Salvia & Ysseldyke, 2004).

Second, the fact that almost all school psychologists in the Normal condition accurately classified their student underscores that school psychologists are competent decision-makers and practitioners. It is favorable to recognize that a student does not have a disability when it does not exist (Algozzine & Ysseldyke, 1986). Such recognition limits placing students in restrictive environments unnecessarily. Ensuring that students are in the least restrictive environment in order for them to benefit from instruction has been the foundation of educational legislations (Fuchs, D., Fuchs, & Fernstrom, 1992; IDEA, 2007).

Third, it is important that school psychologists acknowledge their propensities for using heuristics and shortcuts. Until the present study, there has been no empirical evidence on whether school psychologists use specific heuristics. However, the findings of this study indicate that school psychologists use the representativeness heuristic. Given the tendency of school psychologists to use the representativeness heuristic, trainers of school psychologists should inform students about the ways that heuristics are used. School psychologists should also discuss whether it is acceptable and ethical for heuristics to be used in the field. Following these discussions, school psychologists might want to consider alternatives to relying on the representativeness heuristic, such as using checklists, flowcharts, practice guidelines, and scientific approaches (Watkins, 2009). Additionally, awareness of the use of heuristics might reduce any influences associated with the use of heuristics.

Fourth, based on importance ratings of assessment data, it appears that school psychologists consistently focus on certain data (e.g., IQ & achievement scores) when making classification decisions. Overweighing certain assessment data, in essence, represents the use of a heuristic in general. As a result, practicing school psychologists must be mindful of this type of heuristic and are encouraged to avoid overlooking other information that should also be gathered from other assessments, such as hearing and vision results, and visual motor ability. Graduate training programs and professional development coordinators should reinforce how important it is to conduct comprehensive evaluations and consider multiple sources of information during the process to determine special education eligibility (Reschly, 1996; Salvia & Ysseldyke, 2004).

Finally, a systematic misunderstanding of RTI might be affecting the accuracy of classification depending on the condition of the case (Fuchs, Mock, Morgan, & Young, 2003), as the use of response-to-classroom modifications data was likely to be associated with the accuracy of classification in the Complicated condition and misclassification in the LD case. To combat this misunderstanding, school psychologists are encouraged to advance their knowledge and understanding of progress-monitoring data. Given that numerous school districts are using the RTI model for school evaluations (IDEA, 2004), further professional development and training about RTI is needed in the field of school psychology to prevent classification errors in all conditions. Currently, research on RTI and its effectiveness is in its infancy and as a result, more research is needed in regards to decision-making and the use of RTI (Fuchs & Deshler, 2007). It is possible that, in the absence of prerequisite skills to appropriately implement RTI, the use of heuristics will increase as pressures to use RTI intensify. It is also possible that as RTI use becomes more prevalent misclassification will increase.

Conclusion

Overall, the findings of the current project add to the breadth of knowledge about heuristic use. Specifically, the representativeness heuristic appears to play a role in the decision-making of school psychologists, as it does with other practitioners. The emergence of the representativeness heuristic differs based on the context of the student data summary as well as how it is measured and who is using it. First, the representativeness heuristic was less noticeable when it was relatively obvious that the student did not have a disability. Second, the use of the representativeness heuristic differed with the condition of the student data summary when the measure is based on a prototype. Third, the use of the representativeness heuristic, measured as state criteria, varied with the condition of the data summary when school psychologists in Pennsylvania (Study 2) made classification decisions. Additionally, findings from Study 1 and Study 2 are inconclusive as to whether other measures of the representativeness heuristic used in prior research (e.g., disregard of base rates and the interaction of accuracy of classification with client-prototype similarity) are reliable measures or useful in school psychology.

The current study also adds to the literature about the importance of assessment data to special education decisions. It has been found that school psychologists usually consider IQ score, response-to-classroom modification, state criteria, assessments of behavior (e.g., ratings scales or observations), and especially achievement scores as highly important to classification decisions. However, the influence of assessment data is not as simple as conceptualized by previous research (e.g., Della Toffalo & Pedersen, 2005; Salvia & Ysseldyke, 1978). For example, school psychologists are on average less consistent when it comes to their use of academic progress-monitoring data. Data used to monitor academic progress have been found to be more important to school psychologists when cases are straightforward, not complicated.

Additionally, there are mixed findings about whether the focus on certain assessment data (e.g., referral information and progress-monitoring data) influences the propensity to make errors in classification decisions.

A major contribution of the current study is the combined use of two domains of decision-making research (heuristics and assessment data) to arrive at a more comprehensive examination and analysis of both than has been previously found (e.g., Cioffi & Markham, 1997; Garb, 1996; Knoff, 1983; Thurlow & Ysseldyke, 1980). Findings indicate that school psychologists both over- and under-weigh certain assessment data and use the representativeness heuristic during decision-making. The weighing of certain assessment data appears to represent a general form of a heuristic. Therefore, the study indicates that school psychologists use the representativeness heuristic as other practitioners do and might also use more general heuristics as well. The study also suggests that there is a need to continue assessing the validity and reliability of measures designed to tap heuristics.

Together the findings regarding the use of the representativeness heuristic and assessment data not only confirm that heuristics are used, but also provide additional support for the theory of bounded rationality. Bounded rationality refers to how people make decisions quickly and efficiently to meet specific goals in the presence of both internal and external demands (Gigerenzer, 2001). Prior research has operationalized bounded rationality as an individuals' focus on a subset of cues to help them make an inference (Gigerenzer & Goldstein, 1996). That school psychologists might rely heavily on certain assessment data (e.g., IQ score, achievement score, and response-to-classroom modification) and use the representativeness heuristic indicates that school psychologists do in fact focus on a subset of cues and have bounded rationality. Additionally, the use of assessment data and the representativeness heuristic partially appears to depend on the condition of the student data summary, an indication that bounded rationality is not used the same way in all contexts (Broder, 2000; Newell & Shanks, 2003).

Another major contribution of the current study is the development of materials and measures that can be used by other researchers interested in decision-making. Based on the responses of two independent samples of school psychologists, the three data summaries (Normal, LD, and Complicated) differ in levels of complexity. The manipulated data summaries could be adapted and utilized in a wide range of analogue and experimental studies about psychoeducational decision-making. Other researchers might also be interested in the Decision-Making Questionnaire. This survey, created and revised by the author, was easy to administer and use. Therefore it is likely that the Decision-Making Questionnaire could become a useful tool for measuring heuristics and assessment data by school personnel.

It is hoped that the findings of the current study stimulate researchers to expand the generalizability of the findings by continuing research in the area of decision-making. Future research is needed to assess whether the findings of the current study hold true with a national sample of school psychologists. Additionally, research is needed to determine whether the findings are comparable when school psychologists are more comfortable with the quality and quantity of information provided in each case. Future research is also needed to test what methods best measure the use of the representativeness heuristic, not only by school psychologists but by other practitioners as well. Finally, situations in which heuristics are either beneficial, or lead to errors by school psychologists need to be explored.

The current study has important implications for the practice and training of school psychologists. School psychologists need to be trained about the potential of making errors when cases are ambiguous in order to prevent their occurrence. Additionally, school psychologists need to be aware of their natural tendency in cognitive processing to shorten the decision-making process. Awareness of such shortcuts might reduce associated influences and calls into question whether using shortcuts, such as the representativeness heuristic, is acceptable and ethical. Lastly, because it is the responsibility of school psychologists to be competent in their practice (National Association of School Psychologists, 2000), school psychologist as well as school psychology graduate students should seek opportunities to learn how to manage the special education decision-making process. Discovering ways to reduce the complexity of decision-making without relying on heuristics or overweighing assessment data could be helpful. Learning the best ways to interpret and use RTI data during psychoeducational evaluations is also important for reducing classification errors.

References

Algozzine, B., & Ysseldyke, J. E. (1986). The future of the LD field: Screening and diagnosis. *Journal of Learning Disabilities, 19,* 394-398. doi:10.1177/002221948601900707

Archer, T. (2007). Characteristics associated with increasing the response rates of web-based surveys. *Practical Assessment, Research & Evaluation, 12.*

Arkes, H. R., Saville, P. D.,Wortmann, R. L., & Harkness, A. R. (1981). Hindsight bias among physicians weighing the likelihood of diagnoses. *Journal of Applied Psychology, 66, 252*. doi:10.1037/0021-9010.66.2.252

Barnett, D. W. (1988). Professional judgment: A critical appraisal. *School Psychology Review, 17,* 658-672.

Barnett, J. E., & Johnson, W. B. (2008). Ethics desk reference for psychologists. Washington, DC: American Psychological Association.

Batsche, G. M., Kavale, K. A., & Kovaleski, J. F. (2006). Competing views: A dialogue on response to intervention. *Assessment for Effective Intervention, 32,* 6-19. doi:10.1177/15345084060320010301

Bellis, T. J. (2002). *When the brain can't hear*. New York: Simon and Schuster.

Bender, W., & Shores, C. (2007). *Response to intervention: A practical guide for every teacher.* Thousand Oaks, CA: Corwin Press.

Bieri, J., Orcutt, B. A., & Leaman, R. (1963). Anchoring effects in sequential clinical judgments. *Journal of Abnormal Psychology, 67,* 616-623.

Broder, A. (2000). Assessing the empirical validity of the "take-the-best" heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1332–1346. doi:10.1037/0278-7393.26.5.1332

Brooks, R. (1979). Psychoeducational assessment: A broader perspective**.** *Professional Psychology*, 708-722. doi:10.1037/0735-7028.10.5.708

Catterall, C. D. (1972). Special education in transition- implications for school psychology. *Journal of School Psychology, 10,* 91-98. doi:10.1016/0022-4405(72)90035-0

Cioffi, J., & Markham, R. (1997). Clinical decision making by midwives: Managing case complexity. *Journal of Advanced Nursing, 25,* 265-272. doi:10.1046/j.1365-2648.1997.1997025265.x

Craig, P. A., Kaskowitz, D. H., & Malgoire, M. A. (1978). *Teacher identification of handicapped pupils (ages 6-11) compared with identification using other indicators, Volume II.* Menlo Park, CA: Educational Policy Research Center, Stanford Research Institute.

Curtis, M. J., Lopez, A.D., Castillo, J. M., Batsche, G. M., Minch, D., & Smith, J. C. (2008). The status of school psychology: Demographic characteristics, employment conditions, professional practices, and continuing professional development. *Communique, 36(5),* 27-29.

Czaczkes, B., & Ganzach, Y. (1996). The natural selection of prediction heuristics: Anchoring and adjustment versus representativeness. *Journal of Behavioral Decision Making,* 9, 125-139. doi:10.1002/(SICI)1099-0771(199606)9:2<125::AID-BDM221>3.0.CO;2-7

Davidow, J., & Levinson, E. M. (1993). Heuristic principles and cognitive bias in decision making: Implications for assessment in school psychology. *Psychology in the Schools, 30,* 351-361. doi:10.1002/1520-6807(199310)30:4<351::AID-PITS2310300410>3.0.CO;2-X

Della Toffalo, D. A., & Pedersen, J. A. (2005). The effect of psychiatric diagnosis on school psychologists’ special education eligibility decisions regarding emotional disturbance. *Journal of Emotional and Behavioral Disorders, 13*, 53-60. doi:10.1177/10634266050130010501

DeMesquita, P. (1987). *The information processing and diagnostic decision making of school psychologists while solving a computer-simulated diagnostic referral problem.* Paper presented at the meeting of the National Association of School Psychologists, New Orleans.

Detmer, D., Fryback, D., & Gassner, K. (1978). Heuristics and biases in medical decision making. *Journal of Medical Education, 53*, 682-683.

Dewhurst, M. E., Veach, P. M., Lampman, C., Petraitis, J., Kao, J., & LeRoy, B. (2007). Probability biases in genetic problem solving: A comparison of undergraduates, genetic counseling graduate students, and genetic counselors. *Journal of Genetic Counseling, 16,*157-70. doi:10.1007/s10897-006-9050-7

Dieterly, D. L. (1979). The bubble gum school of research. *American Psychologist, 34,* 562-563. doi:10.1037/0003-066X.34.6.562.b

Dumont, F. (1993). Inferential heuristics in clinical problem formulation: Selective review of their strengths and weaknesses. *Professional Psychology: Research and Practice, 24,* 196-205. doi:10.1037/0735-7028.24.2.196

Ellis, M. V., Robbins, E. S., Schult, D., Ladany, N., Banker, J. (1990). Anchoring errors in clinical judgments: Type I error. *Journal of Counseling Psychology, 37,* 343-351. doi:10.1037/0022-0167.37.3.343

Erdfelder, E., Faul, F., & Buchner, A. (2009). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28,* 1-11.

Fagley, N. S. (1988). Judgmental heuristics: Implications for the decision making of school psychologists. *School Psychology Review, 17*, 311-321.

Finch, W. H., & Davenport, T. (2009). Performance of Monte Carlo permutation and approximate tests for multivariate means comparisons with small sample sizes when parametric assumptions are violated. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 5,* 60-70.

Frey, A. (2002). Predictors of placement recommendations for children with behavioral and emotional disorders. *Behavioral Disorders, 27,* 126-136.

Friedlander, M. L., & Phillips, S. D. (1984). Preventing anchoring errors in clinical judgment. *Journal of Consulting and Clinical Psychology, 52*, 366-371. doi:10.1037/0022-006X.52.3.366

Friedlander, M. L., & Stockman, S. J. (1983). Anchoring and publicity effects in clinical judgment, *Journal of Clinical Psychology, 39*, 637–644. doi:10.1002/1097-4679(198307)39:4<637::AID-JCLP2270390433>3.0.CO;2-Q

Fuchs, D., Fuchs, L. S., & Fernstrom, P. (1992). *A conservative approach to special education reform: mainstreaming through transenvironmental programming and curriculum-based measurement.* Washington, DC: US Department of Education.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for learning disabilities construct. *Learning Disabilities: Research & Practice*, *18*, 157-171. doi:10.1111/1540-5826.00072

Fuchs, D., & Deshler, D. D. (2007). What we need to know about responsiveness to intervention (and shouldn't be afraid to ask). *Learning Disabilities Research & Practice, 22,* 129-136. doi:10.1111/j.1540-5826.2007.00237.x

Gall, M. D., Gall, J. P., & Borg, W. R. (2003). *Educational research: An introduction,* (7th ed.). Boston, MA: Pearson Education, Inc.

Garb, H. (1996). The representativeness and past-behavior heuristics in clinical judgment. *Professional Psychology: Research and Practice, 27,* 272-290. doi:10.1037/0735-7028.27.3.272

Gauron, F. E., & Dickinson, J. K. (1969). The influence of seeing the patient first on diagnostic decision making in psychiatry. *American Journal of Psychiatry, 126*, 199-205.

George, D., & Mallery, P. (2001). *SPSS for windows: Step by step a simple guide and reference.* Needham Heights, MA: Allyn & Bacon.

Gigerenzer, G. (2001). The adaptive toolbox. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 37–50). Cambridge, MA: MIT Press.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and fugal way: Models of bounded rationality. *Psychological Review, 103,* 650–669. doi:10.1037/0033-295X.103.4.650

Gilliam, J. E., & Coleman, M. C. (1981). Who influences IEP committee decisions? *Exceptional Children, 47*, 642-644.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment.* New York: Cambridge University Press.

Hammond, K. R. (1996). How convergence of research paradigms can improve research on diagnostic judgment. *Medical Decision Making, 16*, 281-288. doi:10.1177/0272989X9601600313

Hannaford, A. D., Simon, J., & Ellis, D. (1975). Criteria for special class placement of the mildly retarded- multidisciplinary comparisons. *Mental Retardation, 13,* 7-10.

Hogarth R. M., & Karelaia, N. (2007). Heuristic and linear models of judgment: Matching rules and environments. *Review of Psychological Review,114*, 733. doi:10.1037/0033-295X.114.3.733

Howell, D. C. (2007). *Fundamental statistics for the behavioral sciences* (6th ed.). Florence. KY: Wadsworth Publishing.

Huberty, C. J, & Petoskey, M. D. (2000). Multivariate analysis of variance and covariance. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 183-208). New York: Academic Press.

Huebner, E. S. (1990). The generalizability of the confirmation bias. *School Psychology International*, *11*, 281-286. doi:10.1177/0143034390114005

Huebner, E. S. (1991). Bias in special education decisions: The contribution of analogue research. *School Psychology Quarterly*, *6,* 50-65. doi:10.1037/h0088240

Huebner, E. S., & Cummings, J. A. (1986). Influence of race and test data ambiguity upon school psychologists' decisions. *School Psychology Review*, *15,* 410-417.

Individuals with Disabilities Education Act (IDEA) of 1997. (1997). *Pub. L. No. 105-17*, 105th Congress.

Individuals with Disabilities Education Act (IDEA) of 2004. (2004). *Pub. L. No. 108-446,* 108th Congress.

Jimerson, S. R., Skokut, M., Cardenas, S., Malone, H., & Stewart, K. (2008). Where in the world is school psychology?: Examining evidence of school psychology around the globe. *School Psychology International*, *29*, 131-144. doi:10.1177/0143034308090056

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237-251. doi:10.1037/h0034747

Knoff, H. (1983). Effect of diagnostic information on special education decisions. *Exceptional Children, 49,* 440-444.

Kovaleski, J. F. (2007). Response to intervention: Considerations for research and systems change. *School Psychology Review*, *36,* 638-646.

LaRocco, D. J., & Murdica, P. (2009). *Understanding teachers' concerns about implementing RTI: Practical implications for educational leaders.* Paper presented at the 40th Annual Northeast Educational Research Association Conference, Rocky Hill, CT.

Lindman, H. R. (1974). *Analysis of variance in complex experimental designs.* Oxford, England: W. H. Freeman & Co.

Matuszek, P., & Oakland, T. (1979). Factors influencing teachers' and psychologists' recommendations regarding special class placement. *Journal of School Psychology, 17,* 116-125. doi:10.1016/0022-4405(79)90017-7

Mercer, J. R. (1975). Psychological assessment and the rights of children. In N. Hobbs (Ed.), *Issues in the classification of children* (pp. 239-252). San Francisco: Jossey-Bass.

Morgan, G.A., Leech, N.L., Gloeckner, G. W., & Barrett, K.C. (2004). *SPSS for introductory statistics* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

National Association of School Psychologists. (2000). *Professional conduct manual.* Bethesda, MD: Author.

National Coalition on Workforce and Disability (2004). *Special education law enactment, 10,* 1-4.

Newell, R. B., & Shanks, D. R. (2003). Take the best of look at the rest? Factors influencing “one reason” decision making. *Journal of Experimental Psychology, 29,* 5.

Nichols, P. (1993). Testing simple effects in MANOVA. *SPSS keywords*. Chicago, IL: SPSS Inc.

O'Reilly, D., Northcraft, G. B., & Sabers, D. (1989). The confirmation bias in special education disability decisions. *School Psychology Review, 18,* 126-135.

Pastor, P. N., & Reuben, C. (2008). *Diagnosed attention deficit hyperactivity disorder and learning disability, United States, 2004–2006: Data from the National Health Interview Survey*. MD: U.S. Department of Health and Human Services Centers for Disease Control and Prevention National Center for Health Statistics.

Payne, J. W., Bettman, J. R., & Luce, M. F. (1998). Behavioral decision research: An overview. In M. Birnbaum (Eds.), *Measurement, judgment, and decision making* (pp. 303-359). San Diego, CA: Academic Press.

Pierangelo, R., & Giuliani, G. (2007). A *Step-by-step guide for educators: Special education eligibility.* Thousand Oaks, CA: Sage.

Reschly, D. J. (1996). Diagnostic and treatment utility of intelligence tests. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Beyond traditional intellectual assessment: Contemporary and emerging theories, tests, and issues* (pp. 437-456). New York: Guilford.

Reschly, D. J. (1981). Psychological testing in educational classification and placement. *American Psychologist, 36,* 1094-1102. doi:10.1037/0003-066X.36.10.1094

Reynolds, C. R. (1985). Measuring the aptitude-achievement discrepancy in learning disability diagnosis. *Remedial and Special Education, 6,* 37-48. doi:10.1177/074193258500600508

Salvia, J., & Ysseldyke, J. E. (1978). *Assessment in special and remedial education.* New York: Houghton-Mifflin.

Salvia, J., & Ysseldyke, J. E. (2004). *Assessment: In special and inclusive education.* New York: Houghton-Mifflin.

Samuels, C. A. (2008). Response to intervention sparks interest, questions: Critics say approach depends on too many complex factors. *Education Week*, 23.

Seltzer, T. (1998). *A new IDEA: A parent's guide to the changes in special education law for children with disabilities.* Washington, DC: Bazelon Center for Mental Health Law.

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin, 134,* 207-222. doi:10.1037/0033-2909.134.2.207

Shinn, M. R. (2007). Identifying students at risk, monitoring performance and determining eligibility within response to intervention: Research on educational need and benefit from academic intervention. *School Psychology Review, 36,* 601-617. doi:10.1177/0143034307078103

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*, 1-19. doi:10.1146/annurev.ps.41.020190.000245

State Board of Education. (2008). Rules and regulations: Title 22, education. *Pennsylvania Bulletin*, 38.

Tabachnick, B. G., & Fidell, L.S. (2001). *Using multivariate statistics*. Needham Heights, MA: Allyn & Bacon.

Thurlow, M. L., & Ysseldyke, J. E. (1980). *Factors influential on the psychoeducational* *decisions reached by teams of educators: Research report no. 25.* Minneapolis, MN: Institute for Research on Learning Disabilities.

Tinsley, H. E. A., & Brown, S. D. (2000). *Handbook of applied multivariate statistics and mathematical modeling.* San Diego, CA: Academic Press.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207-232. doi:10.1016/0010-0285(73)90033-9

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases, *Science, 185,* 1124-1131. doi:10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293–315. doi:10.1037/0033-295X.90.4.293

Uhlmann, E. L., Brescoll, V. L., & Pizarro, D. (2007). The motivated use and neglect of base rates. *Behavioral and Brain Sciences, 30,* 284-285. doi:10.1017/S0140525X07001938

Washburn-Moses, L. (2005). Roles and responsibilities of secondary special education teachers in an age of reform. *Remedial and Special Education, 26,* 151-158. doi:10.1177/07419325050260030301

Watkins, M. W. (2009). Errors in diagnostic decision making and clinical judgment. In T. B. Gutkin & C. R. Reynolds (Eds.), *Handbook of school psychology* (4th ed.; pp. 210-229). New York: Wiley.

Ysseldyke, J. E., & Algozzine, B. (1983). On making psychoeducational decisions. *Journal of Psychoeductional Assessment, 1,* 187-195. doi:10.1177/073428298300100210

Ysseldyke, J. E., Algozzine, B., Regan, R., & McGue, M. (1981). The influence of test scores and naturally occurring pupil characteristics on psychoeducational decision making with children. *Journal of School Psychology, 19*, 167-177.

doi:10.1016/0022-4405(81)90058-3

APPENDIX A
SUMMARY OF THE DEMOGRAPHIC FEATURES OF THE STUDY SAMPLE & A NATIONAL SAMPLE

Summary of the Demographic Features of the Study Sample
& a National Sample

Table A

Summary of the Demographic Features of Study Sample 1& 2 and a National Sample

Demographic Variable		***Sample 1* (*N* = 20)**	***Sample 2* (*N* = 272)**	***National* (*N* = 1748)**
Age	*M*	33.21	42.21	46.20
	SD	10.11	12.53	--
Gender				
Female	n	23.00	204.00	--
	% of sample	82.00	75.00	74.00
Male	n	5.00	68.00	--
	% of sample	17.90	25.00	26.00
Race/ Ethnicity				
Asian	n	0.00	1.00	--
	% of sample	0.00	.40	< 1.00
Black or African American	n	0.00	3.00	--
	% of sample	0.00	1.10	1.90
Hispanic or Latino	n	0.00	2.00	--
	% of sample	0.00	.70	3.00
White or Caucasian	n	28.00	263.00	--
	% of sample	100.00	96.70	92.60

Table A (continued)

Summary of the Demographic Features of Study Sample 1 & 2 and a National Sample

Demographic Variable		***Sample 1*** **(*N* = 20)**	***Sample 2*** **(*N* = 272)**	***National*** **(*N* = 1748)**
Mixed Heritage	n	0.00	3.00	--
	% of sample	0.00	1.10	--
Years of Certification	*M*	12.22	13.29	14.80
	SD	11.93	10.27	--
Educational Level				
Master's	n	26.00	199.00	--
	% of sample	92.90	72.20	67.50
Doctorate	n	2.00	73.00	--
	% of sample	7.10	26.80	32.40
Training Model				
Scientist- practitioner	n	27.00	126.00	--
	% of sample	96.40	46.30	--
Practitioner- scholar	n	1.00	69.00	--
	% of sample	3.60	25.40	--
Practitioner focused	n	0.00	77.00	--
	% of sample	0.00	28.30	--

Table A (continued)

Summary of the Demographic Features of Study Sample 1 & 2 and a National Sample

Demographic Variable		***Sample 1* (*N* = 20)**	***Sample 2* (*N* = 272)**	***National* (*N* = 1748)**
Number of Evaluations per Year	*M*	60.40	55.58	--
	SD	21.09	32.15	--
Employment Setting				
School District	n	26.00	199.00	--
	% of sample	92.90	72.20	67.50
Intermediate Unit	n	2.00	73.00	--
	% of sample	7.10	26.80	32.40
Independent Practice	n	27.00	126.00	--
	% of sample	96.40	46.30	--
Residential Facility	n	1.00	69.00	--
	% of sample	3.60	25.40	--
Assessment Approach				
Discrepancy Model	*M*	3.70	3.32	--
	SD	.50	.71	--
RTI Model	*M*	1.40	1.62	--
	SD	.75	.78	--

Table A (continued)

Summary of the Demographic Features of Study Sample 1 & 2 and a National Sample

Demographic Variable		***Sample 1* (*N* = 20)**	***Sample 2* (*N* = 272)**	***National* (*N* = 1748)**
Region Employed				
Rural	n	11.00	96.00	--
	% of sample	39.30	35.30	28.80
Urban	n	3.00	42.00	--
	% of sample	10.70	15.40	28.60
Suburban	n	14.00	127.00	--
	% of sample	50.00	46.70	49.90
Multiple areas	n	0.00	7.00	--
	% of sample	0.00	2.60	--

APPENDIX B
COVER LETTER

Appendix B
Cover Letter

Participation Recruitment Notice

Sharise Wilson is a Ph.D. candidate at the Penn State University, and school psychologist at Sugar Valley Rural Charter School in Loganton, Pennsylvania. She is currently conducting a research study to examine the decision-making process of school psychologists during psychoeducational evaluations and assessments.

Your participation is voluntary; you can choose whether or not you would like to participate in this project. You can refuse to participate in this study, and your decision will not have any personal or professional consequences. All certified school psychologists who are currently practicing in Pennsylvania are eligible to participate. You must be 18 years of age or older to participate in this study.

Completion and submission of the survey implies that you have read the information in this form and consent to take part in the research. As a participant in this study you will be asked to complete three components: a demographic questionnaire, a Student Data Summary module, and the follow up decision-making questionnaire. Filling out the forms will take approximately 20 minutes.

You also have the option of entering a random drawing for a giftcard valued at $30. To enter the drawing, click on the second weblink provided in the email that you received.

If you have any questions regarding the study, contact Ms. Sharise Wilson at 917-434-5956 or smw303@psu.edu.

APPENDIX C
INFORMED CONSENT

Appendix C

Informed Consent

Implied Informed Consent Form for Social Science Research
The Pennsylvania State University

Title of Project:
School Psychologists' Psychoeducational Decision-Making

Principal Investigator:
Sharise Wilson, M.S., Doctoral Candidate
302 1st Avenue
Lock Haven, PA 17745
(917) 434-5956; smw303@psu.edu

Advisor:
Dr. Beverly Vandiver
137 CEDAR Building
University Park, PA 16802
(814) 863-2846; bjv3@psu.edu

1. Purpose of the Study: The purpose of this study is to explore the decision-making process of school psychologists.

2. Procedures to be followed: You will be asked to complete a demographic questionnaire about yourself. Subsequently, you will be asked to review a case study about a student. You will also answer questions on a survey related to the case study.

3. Duration: It will take about 20 minutes to complete the survey.

4. Statement of Confidentiality: Your confidentiality will be maintained to the degree permitted by the technology used. Specifically, no guarantees can be made regarding the interception of data sent via the Internet by any third parties. The survey does not ask for any information that would identify who the responses belong to.

Upon completion of the survey you can click on a separate weblink to enter a drawing for a giftcard. The weblink will direct you to another screen where you will provide some identifying information. However, identifying information for the giftcard drawing cannot be linked to responses on the survey. Entrance in the drawing is completely optional.

In the event of any publication or presentation resulting from the research, no personally identifiable information will be shared because your name is in no way linked to your responses.

5. Benefits: By participating in this study you might learn more about yourself or have a better understanding of how you make psychoeducational decisions.

This research might provide a better understanding of the practice of school psychology and how it compares to that of other professionals. This information could help plan trainings and professional development of school psychologists. This information might assist colleagues and parents in better understanding how school psychologists expend resources and impact the special education process.

6. Right to Ask Questions: Please contact Ms. Sharise Wilson at (917) 434-5956 with questions, complaints or concerns about this research.

7. Payment for participation: Participants have the option of entering in a random drawing for 1 of 25 giftcards, each worth $30. The 25 individuals who win each giftcard will be notified via email by August 1, 2009. The giftcards will be mailed immediately afterwards. Responses on the drawing entrance page cannot be connected to responses in the survey.

8. Voluntary Participation: Your decision to be in this research is voluntary. You can stop at any time. You do not have to answer any questions you do not want to answer. Refusal to take part in or withdrawing from this study will involve no penalty or loss of benefits you would receive otherwise.

You must be 18 years of age or older to take part in this research study.

Completion and submission of the survey implies that you have read the information in this form and consent to take part in the research. Please keep this form for your records or future reference.

APPENDIX D
DEMOGRAPHIC QUESTIONNAIRE

Appendix D
Demographic Questionnaire

Directions: Complete the following information about yourself. Your responses are completely confidential.

1. Your current position is:
(a) school psychologist, (b) supervisor of special education

2. Your current status in the field of school psychology is:
(a) intern, (b) active, (c) inactive, (d) retired

3. How many years have you been a certified school psychologist:
____

4. As a certified school psychologist, you work for:
(a) school/ school district,(b) intermediate unit, (c) judicial system,
(d) independent/ private practice,
(e) other (please specify)______

5. List any additional job titles that you hold
____

6. On average, how many psychoeducational assessments/ evaluations do you complete each year?
____

7. Use the following scale to indicate how often you use the discrepancy model during psychoeducational assessments:
Choose one: (a) Never, (b) Sometimes, (c) Often, (d) Always

8. Use the following scale to indicate how often you use the Response to Intervention (RTI, Response to Instruction) model during psychoeducational assessments:
Choose one: (a) Never, (b) Sometimes, (c) Often, (d) Always

9. The region in which you work can best be described as:
(a) Rural, (b) Urban, (c) Suburban, (d) Other (please specify)_____________

10. Indicate your gender:
(a) Male, (b) Female

11. Indicate your age in years:

    ____

12. Indicate your ethnicity:
    (a) Asian, (b) Black of African American, (c) Hispanic of Latino, (d) Native Hawaiian or Pacific Islander, (e) White or Caucasian, (f) Mixed Heritage,
    (g) Other (please specify)__________

13. Select your highest degree obtained:
    (a) Bachelor's degree, (b) Master's Degree or Equivalent, (c)Doctorate or first professional Degree,
    (d) Other (please specify)_______________________

14. From the following select all that describe your graduate training:
    (a) scientist-practitioner, (b) practitioner-scholar, (c) practitioner focused
    (d) Other (please specify)_______________________

15. In addition to a Pennsylvania certification in School Psychology list any additional certifications/ licensures that you hold:

    ____

Please go to the next page and follow the instructions provided.

APPENDIX E
NORMAL STUDENT DATA SUMMARY

Appendix E

Normal Data Summary

DIRECTIONS:

As a certified school psychologist, assume that Logan has been referred to you for a psychoeducational assessment. Contained in the Student Profile is information that you have collected about Logan. Review the information as you normally would with such an assessment. Once you have reviewed the Student Profile, you will be asked several questions regarding making decisions about this case. When you are ready, scroll through to view the assessment data.

REFERRAL INFORMATION:

Logan is a 9 year old student in the 4th grade. Logan 's teacher is concerned because Logan is having difficulties (e.g., not completing school work) in school. This initial referral was made to see whether Logan qualifies for special education services. English is the only language spoken by Logan and Logan's family.

PARENT COMMENTS ABOUT STUDENT BEHAVIOR:

The mother reports that Logan reached all developmental milestones within normal limits. At home, Logan lives with both parents and two siblings (an older and younger one). Logan 's mother reports being unconcerned about Logan 's behavior at home. Logan gets along with peers and family members well.

RESPONSE TO MODIFICATIONS IN THE REGULAR CLASSROOM:

Three modifications were implemented for 3 weeks with the student by school personnel (e.g., teacher and school aide). The following results with the student were found.

1. Preferential seating in the classroom: Logan was observed being on-task (e.g., looking at teacher, looking at board, reading in book) 60% more often than prior to the modifications.
2. Peer tutor to help with reading and writing assignments: Logan showed a 20% increase in the number of completed assignments when compared to before the modifications were implemented.
3. Daily behavior progress reports from teacher to mother: Logan turned in 15% more homework assignments than prior to implementation of the modification.

REPORT CARD GRADES:
Based on 100% criterion, the student received the following grades at the end of the first half of this school year (4th grade):
Math: 85
Reading: 89
Writing: 85
Science: 77
Social Studies: 81
Physical Education: 90
Art: 90

In kindergarten through 3rd grades, numerical grades are not assigned. Children earn grades of Needs Improvement, Satisfactory, or Outstanding. A review of Logan's report card from kindergarten through 3rd grade indicated that Logan's grades were consistently in the Outstanding range with an occasional grade of Satisfactory.

HEARING/ VISION/ MEDICAL INFORMATION:
The student's vision and hearing were assessed in 2008 and considered to be in the normal range. Logan 's general health is considered to be excellent. No major injuries have been reported. Logan does not take any prescribed medications.

PARENT AND TEACHER RATINGS OF STUDENT BEHAVIOR ON THE BASC-2:
The Behavior Assessment System for Children- Second Edition (BASC-2) measures a broad range of problematic behaviors in children. Ratings of Logan were compared to the general population of students aged 8 to 11. T-Scores correspond to a classification of Clinically Significant, At-Risk, and Average. T-Scores in the Clinically Significant range suggest a high level of maladjustment. Scores in the At-Risk range identify a potentially significant problem. Scores in the Average range indicate that the child is similar to same-age peers in this domain.

BASC-2 PARENT RATINGS

Domain	T-Score	%ile	Classification
Hyperactivity	56	79	Average
Aggression	50	54	Average
Conduct Problems	56	78	Average
Externalizing Problems	55	75	Average
Anxiety	39	9	Average
Depression	44	35	Average
Somatization	42	19	Average
Internalizing Problems	39	11	Average
Attention Problems	53	64	Average
Atypicality	44	30	Average
Withdrawal	46	41	Average
Behav. Symptoms Index	52	67	Average
Adaptability	48	43	Average
Social Skills	50	48	Average
Leadership	52	57	Average
Functional Communication	51	46	Average
Adaptive Skills	51	52	Average
Activities of Daily Living	54	64	Average

BASC-2 TEACHER RATINGS

Domain	T-Score	%ile	Classification
Hyperactivity	57	80	Average
Aggression	57	79	Average
Conduct Problems	53	66	Average
Externalizing Problems	59	84	Average
Anxiety	52	61	Average
Depression	47	46	Average
Somatization	49	35	Average
Internalizing Problems	57	60	Average
Attention Problems	53	64	Average
Learning Problems	52	67	Average
School Problems	50	48	Average
Atypicality	63	46	At-Risk
Withdrawal	38	7	Average
Behav. Symptoms Index	50	56	Average
Adaptability	61	85	Average
Social Skills	59	80	Average
Leadership	50	51	Average
Study Skills	44	29	Average
Functional Communication	55	61	Average
Adaptive Skills	51	52	Average

SCORES ON THE DEVELOPMENTAL TEST OF VISUAL-MOTOR INTEGRATION (VMI):
The Beery-Buktenica Developmental Test of Visual-Motor Integration was administered to Logan. The VMI measures visual and motor abilities. On the VMI, Logan earned a Standard Score of 104 (Percentile Rank = 61), which is in the age-appropriate range.

CLASSROOM OBSERVATION:
A paraprofessional observed Logan for 30 minutes during a math lesson. The paraprofessional noted that Logan independently answered 8 of the 20 teacher questions directed to the entire class. Logan answered these questions with 100% accuracy. Logan was prompted by the teacher 2 times to sit at the desk. Logan sat down immediately upon request and appeared to be on-task for at least 8 minutes. The paraprofessional observed Logan saying "please" and "thank you" to peers. While observing Logan and another student, Logan turned in a math worksheet 45 seconds after the other student. Logan turned in 3 out of 3 worksheets.

FACTOR SCORES ON THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN-FOURTH EDITION
The Wechsler Intelligence Scale for Children- Fourth Edition (WISC-IV) was used to obtain an estimate of Logan's cognitive ability. Standard Scores (SS) have a mean of 100 and Standard Deviation of 15. Confidence Interval ranges and Percentile Rank (PR) scores are also reported below.

WISC-IV						
Indices	SS	Classification	90% Confidence Range			Percentile Rank
Full Scale	100	Average	96	-	104	50
Verbal Comprehension	102	Average	95	-	109	55
Perceptual Reasoning	108	Average	101	-	114	70
Working Memory	83	Average	77	-	92	13
Processing Speed	100	Average	92	-	108	50

ACHIEVEMENT SCORES ON THE WIAT-II:

The Wechsler Individual Achievement Test- Second Edition (WIAT-II) measured Logan 's academic achievement and skill level. Standard Scores have a mean of 100 and Standard Deviation of 15. Logan's WIAT-II Standard Scores (SS), Confidence Interval Range, and Percentile Rank (PR) scores are reported below.

WIAT-II						
Subjects	SS	Classification	90% Confidence Range			Percentile Rank
Word Reading	100	Average	96	-	104	50
Reading Comprehension	98	Average	93	-	103	45
Pseudoword Decoding	106	Average	102	-	110	66
Numerical Operations	102	Average	93	-	111	55
Math Reasoning	100	Average	93	-	107	50
Spelling	95	Average	89	-	101	37
Written Expression	99	Average	90	-	108	47
Listening Comprehension	95	Average	84	-	109	37
Oral Expression	101	Average	93	-	109	53

Thank you for reviewing the Student Profile above. Go to the next page and answer the questions based on the information about Logan. You will be asked to determine a classification and rate the importance of certain assessment data.

APPENDIX F
LD STUDENT DATA SUMMARY

Appendix F
LD Data Summary

DIRECTIONS:
As a certified school psychologist, assume that Addison has been referred to you for a psychoeducational assessment. Contained in the Student Profile is information that you have collected about Addison. Review the information as you normally would with such an assessment. Once you have reviewed the Student Profile, you will be asked several questions regarding making decisions about this case. When you are ready, scroll through to view the assessment data.

REFERRAL INFORMATION:
Addison is a 9 year old student in the 4th grade. Addison 's teacher is concerned because Addison is having difficulties (e.g., not completing school work) in school. The initial referral was made to see whether Addison qualifies for special education services. English is the only language spoken by Addison and Addison 's family.

PARENT COMMENTS ABOUT STUDENT BEHAVIORS:
The mother reports that Addison reached all developmental milestones within normal limits. At home, Addison lives with both parents and two siblings (an older and younger one). Addison 's mother reports being unconcerned about Addison 's behavior at home. Addison gets along with peers and family members well.

RESPONSE TO MODIFICATIONS IN THE REGULAR CLASSROOM:
Three modifications were implemented for 3 weeks with the student by school personnel (e.g., teacher and teacher aide). The following results were found.
1. Preferential seating in the classroom: There has been no significant change in Addison’s academic performance (grades continue to be low) or behaviors (on-task behaviors and completion of assignments remain the same).
2. Peer tutor to help with reading and writing assignments: There has been no significant change in Addison’s academic performance (grades continue to be low) or behaviors (on-task behaviors and completion of assignments remain the same).
3. Daily behavior progress reports from teacher to mother: There has been no significant change in Addison’s academic performance (grades continue to be low) or behaviors (on-task behaviors and completion of assignments remain the same).

REPORT CARD GRADES:

Based on 100% criterion, the student received the following grades at the end of the first half of this school year (4th grade):

Math: 60

Reading: 60

Writing: 60

Science: 77

Social Studies: 81

Physical Education: 90

Art: 90

In kindergarten through 3rd grades, numerical grades are not assigned. Children earn grades of Needs Improvement, Satisfactory, or Outstanding. A review of Addison’s report card from kindergarten through 3rd grade indicated that Addison’s grades were consistently in the Needs Improvement range with an occasional grade of Satisfactory.

HEARING/ VISION/ MEDICAL INFORMATION:

The student’s vision and hearing were assessed in 2008 and considered to be in the normal range. Addison 's general health is considered to be excellent. No major injuries have been reported. Addison does not take any prescribed medications.

PARENT AND TEACHER RATINGS OF STUDENT BEHAVIOR ON THE BASC-2:

The Behavior Assessment System for Children- Second Edition (BASC-2) measures a broad range of problematic behaviors in children. Ratings of Addison were compared to the general population of students aged 8 to 11. T-Scores correspond to a classification of Clinically Significant, At-Risk, and Average. T-Scores in the Clinically Significant range suggest a high level of maladjustment. Scores in the At-Risk range identify a potentially significant problem. Scores in the Average range indicate that the child is similar to same-age peers in this domain.

BASC-2 PARENT RATINGS			
Domain	T-Score	%ile	Classification
Hyperactivity	56	79	Average
Aggression	50	54	Average
Conduct Problems	56	78	Average
Externalizing Problems	55	75	Average
Anxiety	39	9	Average
Depression	44	35	Average
Somatization	42	19	Average
Internalizing Problems	39	11	Average
Attention Problems	65	92	Average
Atypicality	44	30	Average
Withdrawal	46	41	Average
Behav. Symptoms Index	52	67	Average
Adaptability	48	43	Average
Social Skills	50	48	Average
Leadership	52	57	Average
Functional Communication	51	46	Average
Adaptive Skills	51	52	Average
Activities of Daily Living	54	64	Average

BASC-2 TEACHER RATINGS			
Domain	T-Score	%ile	Classification
Hyperactivity	57	80	Average
Aggression	57	79	Average
Conduct Problems	53	66	Average
Externalizing Problems	59	84	Average
Anxiety	52	61	Average
Depression	47	46	Average
Somatization	49	35	Average
Internalizing Problems	57	60	Average
Attention Problems	53	64	Average
Learning Problems	72	96	Clin. Significant
School Problems	70	97	Clin. Significant
Atypicality	63	46	At-Risk
Withdrawal	38	7	Average
Behav. Symptoms Index	50	56	Average
Adaptability	61	85	Average
Social Skills	59	80	Average
Leadership	50	51	Average
Study Skills	44	29	Average
Functional Communication	55	61	Average
Adaptive Skills	51	52	Average

SCORES ON THE DEVELOPMENTAL TEST OF VISUAL-MOTOR INTEGRATION (VMI):
The Beery-Buktenica Developmental Test of Visual-Motor Integration was administered to Addison. The VMI measures visual and motor abilities. On the VMI, Addison earned a Standard Score of 104 (Percentile Rank = 61), which is in the age-appropriate range.

CLASSROOM OBSERVATION:
A paraprofessional observed Addison for 30 minutes during a math lesson. The paraprofessional noted that Addison independently answered 1 of the 20 teacher questions directed to the entire class. Addison answered these questions with 0% accuracy. Addison was prompted by the teacher 2 times to sit at the desk. Addison sat down upon request. The paraprofessional observed Addison saying "please" and "thank you" to peers. While being observed simultaneously with another student, Addison turned in a math worksheet 20 minutes after the other student.

FACTOR SCORES ON THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN-FOURTH EDITION:
The Wechsler Intelligence Scale for Children- Fourth Edition (WISC-IV) was used to obtain an estimate of Addison 's cognitive ability. Standard Scores (SS) have a mean of 100 and Standard Deviation of 15. Confidence Interval ranges and Percentile Rank (PR) scores are also reported

below.

WISC-IV						
Indices	SS	Classification	90% Confidence Range			Percentile Rank
Full Scale	100	Average	96	-	104	50
Verbal Comprehension	102	Average	95	-	109	55
Perceptual Reasoning	108	Average	101	-	114	70
Working Memory	83	Average	77	-	92	13
Processing Speed	100	Average	92	-	108	50

ACHIEVEMENT SCORES ON THE WIAT-II:
The Wechsler Individual Achievement Test- Second Edition (WIAT-II) measured Addison's academic achievement and skill level. Standard Scores have a mean of 100 and Standard Deviation of 15. Addison's WIAT-II Standard Scores (SS), Confidence Interval Range, and Percentile Rank (PR) scores are reported below.

WIAT-II						
Subjects	SS	Classification	90% Confidence Range			Percentile Rank
Word Reading	74	Borderline	70	-	78	4
Reading Comprehension	75	Borderline	70	-	80	5
Pseudoword Decoding	86	Low Average	82	-	90	18
Numerical Operations	75	Borderline	66	-	84	3
Math Reasoning	80	Borderline- Low Average	73	-	87	10
Spelling	70	Extremely Low-Borderline	64	-	76	2
Written Expression	78	Borderline	69	-	87	7
Listening Comprehension	90	Low Average- Average	79	-	101	25
Oral Expression	80	Borderline- Low Average	72	-	88	9

Thank you for reviewing the Student Profile above. Go to the next page and answer the questions based on the information about Addison. You will be asked to determine a classification and rate the importance of certain assessment data.

APPENDIX G
COMPLICATED STUDENT DATA SUMMARY

Appendix G

Complicated Data Summary

DIRECTIONS:

As a certified school psychologist, assume that Casey has been referred to you for a psychoeducational assessment. Contained in the Student Profile is information that you have collected about Casey. Review the information as you normally would with such an assessment. Once you have reviewed the Student Profile, you will be asked several questions regarding making decisions about this case. When you are ready, scroll through to view the assessment data.

REFERRAL INFORMATION:

Casey is a 9 year old student in the 4th grade. Casey 's teacher is concerned because Casey is having difficulties (e.g., not completing school work) in school. This initial referral was made to see whether Casey qualifies for special education services. English is the only language spoken by Casey and Casey's family.

PARENT COMMENTS ABOUT STUDENT BEHAVIOR:

The mother reports that Casey reached all developmental milestones within normal limits. At home, Casey lives with both parents and two siblings (an older and younger one). Casey 's mother reports being unconcerned about Casey 's behavior at home. Casey gets along with peers and family members well.

RESPONSE TO MODIFICATIONS IN THE REGULAR CLASSROOM:

Three modifications were implemented for 3 weeks with the student by school personnel (e.g., teacher and teacher aide). The following results with the student were found.

1. Preferential seating in the classroom: Some progress was found with student being on-task (e.g., looking at teacher, looking at board, reading book as directed) 20% more often than prior to the modifications.
2. Peer tutor to help with reading and writing assignments: No change in on-task behaviors or completion of work was found.
3. Daily behavior progress reports from teacher to mother: No change in on-task behaviors or completion of work was found.

REPORT CARD GRADES:
Based on 100% criterion, the student received the following grades at the end of the first half of this school year (4th grade):
Math: 70
Reading: 69
Writing: 70
Science: 68
Social Studies: 82
Physical Education: 90
Art: 80

In kindergarten through 3rd grades, numerical grades are not assigned. Children earn grades of Needs Improvement, Satisfactory, or Outstanding. A review of Casey's report card indicated that Casey's grades were mostly in the Satisfactory range in 3rd grade. In 2nd grade, Casey's grades were mostly in the Needs Improvement range. In kindergarten and 1st grade, Casey's had an equal number of Needs Improvement and Satisfactory grades.

HEARING/ VISION/ MEDICAL INFORMATION:
The student's vision and hearing were assessed in 2008 and considered to be in the normal range. Casey 's general health is considered to be excellent. No major injuries have been reported. Casey does not take any prescribed medications.

PARENT AND TEACHER RATINGS OF STUDENT BEHAVIOR ON THE BASC-2:
The Behavior Assessment System for Children- Second Edition (BASC-2) measures a broad range of problematic behaviors in children. Ratings of Casey were compared to the general population of students aged 8 to 11. T-Scores correspond to a classification of Clinically Significant, At-Risk, and Average. T-Scores in the Clinically Significant range suggest a high level of maladjustment. Scores in the At-Risk range identify a potentially significant problem. Scores in the Average range indicate that the child is similar to same-age peers in this domain.

BASC-2 PARENT RATINGS			
Domain	T-Score	%ile	Classification
Hyperactivity	56	79	Average
Aggression	50	54	Average
Conduct Problems	56	78	Average
Externalizing Problems	55	75	Average
Anxiety	39	9	Average
Depression	44	35	Average
Somatization	42	19	Average
Internalizing Problems	39	11	Average
Attention Problems	53	64	Average
Atypicality	44	30	Average
Withdrawal	46	41	Average
Behav. Symptoms Index	52	67	Average
Adaptability	48	43	Average
Social Skills	50	48	Average
Leadership	52	57	Average
Functional Communication	51	46	Average
Adaptive Skills	51	52	Average
Activities of Daily Living	54	64	Average

BASC-2 TEACHER RATINGS			
Domain	T-Score	%ile	Classification
Hyperactivity	75	96	Clin. Significant
Aggression	57	79	Average
Conduct Problems	53	66	Average
Externalizing Problems	59	84	Average
Anxiety	52	61	Average
Depression	47	46	Average
Somatization	61	84	At-Risk
Internalizing Problems	57	60	Average
Attention Problems	72	98	Clin. Significant
Learning Problems	65	83	At-Risk
School Problems	50	48	Average
Atypicality	63	82	At-Risk
Withdrawal	38	7	Average
Behav. Symptoms Index	50	56	Average
Adaptability	61	85	Average
Social Skills	73	95	Clin. Significant
Leadership	50	51	Average
Study Skills	44	29	Average
Functional Communication	55	61	Average
Adaptive Skills	51	52	Average

SCORES ON THE DEVELOPMENTAL TEST OF VISUAL-MOTOR INTEGRATION (VMI):
The Beery-Buktenica Developmental Test of Visual-Motor Integration was administered to Casey. The VMI measures visual and motor abilities. On the VMI, Casey earned a Standard Score of 104 (Percentile Rank =61), which is in the age-appropriate range.

CLASSROOM OBSERVATION:
A paraprofessional observed Casey for 30 minutes during a math lesson. The paraprofessional noted that Casey independently answered 5 of the 20 teacher questions directed to the entire class. Casey answered these questions with 40% accuracy. Casey was prompted by the teacher 2 times to sit at the desk. Casey sat down immediately upon request and appeared to be on-task for at least 3 minutes. The paraprofessional observed Casey saying "please" and "thank you" to peers. While observing Casey and another student, Casey turned in a math worksheet 10 minutes after the other student. Casey turned in 3 out of 3 worksheets.

FACTOR SCORES ON THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN-FOURTH EDITION:
The Wechsler Intelligence Scale for Children- Fourth Edition (WISC-IV) was used to obtain an estimate of Casey 's cognitive ability. Standard Scores (SS) have a mean of 100 and Standard Deviation of 15. Confidence Interval ranges and Percentile Rank (PR) scores are also reported

below.

WISC-IV						
Indices	SS	Classification	90% Confidence Range			Percentile Rank
Full Scale	80	Low Average	76	-	85	9
Verbal Comprehension	83	Low Average	77	-	89	13
Perceptual Reasoning	77	Borderline	72	-	85	6
Working Memory	99	Average	91	-	107	47
Processing Speed	91	Low Average-Average	85	-	100	27

ACHIEVEMENT SCORES ON THE WIAT-II:

The Wechsler Individual Achievement Test- Second Edition (WIAT-II) measured Casey 's academic achievement and skill level. Standard Scores have a mean of 100 and Standard Deviation of 15. Casey 's WIAT-II Standard Scores (SS), Confidence Interval Range, and Percentile Rank (PR) scores are reported below.

WIAT-II						
Subjects	SS	Classification	90% Confidence Range			Percentile Rank
Word Reading	74	Borderline	70	-	78	4
Reading Comprehension	75	Borderline	70	-	80	5
Pseudoword Decoding	86	Low Average	82	-	90	18
Numerical Operations	75	Borderline	66	-	84	3
Math Reasoning	80	Borderline- Low Average	73	-	87	10
Spelling	70	Extremely Low-Borderline	64	-	76	2
Written Expression	78	Borderline	69	-	87	7
Listening Comprehension	90	Low Average- Average	79	-	101	25
Oral Expression	80	Borderline- Low Average	72	-	88	9

Thank you for reviewing the Student Profile above. Go to the next page and answer the questions based on the information about Casey. You will be asked to determine a classification and rate the importance of certain assessment data.

APPENDIX H
DECISION-MAKING QUESTIONNAIRE

Appendix H
Decision-Making Questionnaire

1. Based on the information provided how you would classify Student X?

Please note that throughout the survey "classification" refers to each of the categories listed below.

- Nonexceptional
- Autism
- Specific Learning Disability
- Severe Emotional Disturbance
- Mental Retardation
- Traumatic Brain Injury
- Other Health Impairment
- Multiple Disabilities

2. Select ONE type of placement that you would recommend for Student X:

- Full time regular class with no basic change in teaching program or support services
- Full time regular class with specialists available in the school for consultation
- Full time regular class with the addition of resource room or learning support
- Part-time special class, resource room, or learning support
- Full time special class or special school

3. If you selected "Full time regular class with specialist...", what type of specialist would you consider?

4. Describe the steps and procedures that you used to determine a classification for Student X:

5. Indicate how confident you are in your classification.

Choose one:

- Not at all Confident
- A little Confident
- Moderately Confident
- Very Confident
- Extremely Confident

6. What's the likelihood that Student X has the classification that you listed in Question 1?

- 0-20%
- 21-40%
- 41-60%
- 61-80%
- 81-100%

7. How prevalent in the school-age population of Pennsylvania (not just those referred to you) is the classification that you listed for Student X? Use a percentage.

____%

8. About how many students have you evaluated in the prior year with the classification that you listed for Student X?

# of students:_______________

9. When you read the Student Profile describing Student X, did a prototypical student with a specific classification come to mind.

- Yes
- No

If yes…

9A. How similar is Student X to your prototype? Choose one.

- Not at All Similar
- A Little Similar
- Moderately Similar
- Very Similar
- Extremely Similar

9B. Describe which of Student X's characteristics are similar to your prototype?

9C. To what extent did you use your prototype to determine a classification for Student X? Choose one.

- Not at All
- A Little
- Moderately
- A Lot
- Exclusively

10. Rate how important each of these factors was to your decision regarding Student X, in particular. For each factor choose only ONE rating.

	Not Important	A Little Important	Moderately Important	Very Important	Critical
A. Referral Question					
B. Hearing & Vision Results					
C. VMI Scores					
D. Classroom Observation					
E. IQ Scores					
F. Achievement Scores					
G. Response to Classroom Modifications					
H. Grades					
I. Parent Comments					
J. BASC-2 Teacher Report					
K. BASC-2 Parent Report					
L. State Criteria for Special Education Eligibility					
M. DSM-IV Criteria					
N. Similarity between Student X and Prototype					
O. Prevalence of Classification in PA					

11. For those factors that were “critical” to your decision, briefly describe why you considered them to be of such importance.

Thank you for your participation. You will now be directed to a separate webpage that will allow you to enter the giftcard drawing. If you are not interested in entering the drawing you may close this tab or web browser now.

APPENDIX I

SUMMARY TABLES OF THE RESULTS OF STUDY 2 PRELIMINARY ANALYSES

Appendix I
Summary Tables of the Results of Study 2 Preliminary Analyses

Table I-1

Means and Standard Deviations for Confidence and Likelihood by Condition in Sample 2 (N = 272)

		M	*SD*	Skew	Kurtosis
Normal (*n* = 88)	Confidence	4.13	.81	-.63	.26
	Likelihood	3.67	.62	-1.72	1.75
LD (*n* = 93)	Confidence	3.48	.97	-.68	.85
	Likelihood	3.08	.95	-.55	-.55
Complicated (*n* = 91)	Confidence	3.03	.83	-.31	-.23
	Likelihood	2.91	.94	-.40	-.81

Note. Confidence and likelihood were rated on a Likert-like scale from 1 (*not confident, not likely*) to 5 (*extremely confident, extremely likely*).

Table I-2

Means and Standard Deviations for Similarity by Condition in Sample 2 (n = 125)

	M	*SD*	Skew	Kurtosis
Normal Similarity (*n* = 24)	3.33	.76	-.02	-.22
LD (*n* = 52)	3.40	.60	-.45	.33
Complicated (*n* = 49)	3.53	.50	-.53	-.92

Note. Similarity was rated on a Likert-like scale from 1 (*not similar)* to 5 (*extremely similar*).

Table I-3

Correlations for Accuracy, Confidence, Likelihood, and Similarity by Condition in Sample 2 (N = 272)

		1.	2.	3.	4.
Normal (*n* = 88)	1. Accuracy	--	.31*	.29*	.01
	2. Confidence		--	.68*	.57*
	3. Likelihood			--	.32
	4. Similarity				--
LD (*n* = 93)	1. Accuracy	--	.23*	.43*	.03
	2. Confidence		--	.77*	.07
	3. Likelihood			--	.05
	4. Similarity				--
Complicated (*n* = 91)	1. Accuracy	--	.31*	-.08	-.24
	2. Confidence		--	.39*	-.29
	3. Likelihood			--	-.06
	4. Similarity				--

Note. *$p < .01$.

Table I-4

Correlations, Means, Standard Deviations, Skew, and Kurtosis for Importance of Assessment Data for the Normal Group in Sample 2 (n = 88)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. RefQuest	--	.54*	.39*	.27*	.04	.05	.13	-.07	.32*	.16	.21*	-.12	.09	-.06	.05
2. H&V		--	.41*	.16	.04	.03	.15	.03	.20	.20	.26*	.05	.33*	.07	.23*
3. VMI			--	.12	.19	.19	.25*	.16	.37*	.38*	.40*	.06	.29*	.10	.14
4. Observat				--	.09	.24*	.41*	.05	.14	.36*	.36*	.11	.11	.05	.12
5. IQ					--	.64*	.19	.38*	.30*	.47*	.44*	.21*	.17	-.02	-.09
6. Achieve						--	.24*	.30	.32*	.45*	.37*	.39*	.09	-.02	-.11
7. RTM							--	.39*	.26*	.43*	.43*	.09	.18	.11	.28*
8. Grades								--	.29*	.39*	.36*	.26*	.28*	.12	.07
9. PNarrativ									--	.47*	.57*	.11	.49*	.17	.04
10.BASC-T										--	.91*	.19	.28*	.22*	.10
11. BASC-P											--	.15	.35*	.23*	.11

Table I-4 (continued)

Correlations, Means, Standard Deviations, Skew, and Kurtosis for Importance of Assessment Data for the Normal Group in Sample 2 (n = 88)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
12. PA Criter												--	.11	-.05	-.10
13. DSMIV													--	.31*	.10
14. SimProt														--	.48*
15. Preva															--
M	3.86	3.49	2.59	3.99	4.08	4.35	4.31	3.72	3.50	3.81	3.70	4.38	2.58	2.05	1.69
SD	.76	1.00	.94	.67	.78	.68	.68	.93	.82	.71	.75	.72	1.16	1.07	.99
Skew	-.24	-.28	.32	-.22	-.59	-.57	-.70	-.65	-.13	-.10	.03	-.70	.34	.65	1.32
Kurtosis	-.26	-.49	-.28	.01	.10	-.71	.30	.81	-.46	-.24	-.42	-.75	-.60	-.39	1.21

Note. RefQuest = referral questions; H&V = hearing and vision; VMI = visual-motor index; Observat = classroom observations; Achieve= achievement score; RTM = response-to-classroom modification; PNarrativ = parents comments; PA Criter= state criteria; DSMIV = DSM criteria; SimProt = similarity with prototype; Preva = prevalence of classification. * $p < .01$.

Table I-5

Correlations, Means, Standard Deviations, Skew, and Kurtosis for Importance of Assessment Data for the LD Group in Sample 2 (n = 93)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. RefQuest	--	.33*	.35*	.31*	-.10	-.09	.24*	.29*	.44*	.06	.16	-.01	.11	.07	.12
2. H&V		--	.46*	.22*	.16	.23*	.01	-.01	.46*	.28*	.31*	-.15	.24*	.05	-.06
3. VMI			--	.36*	.19	.07	.13	.15	.44*	.31*	.33*	-.11	.44*	.12	.24*
4. Observat				--	-.10	-.08	.34*	.14	.49*	.36*	.41*	.06	.16	-.01	.17
5. IQ					--	.78*	.02	.10	.04	.46*	.41*	.13	.20	.29*	.15
6. Achieve						--	.12	.20	.02	.49*	.41*	.27*	.16	.29*	.18
7. RTM							--	.28*	.33*	.22*	.23*	.27*	.10	-.11	.15
8. Grades								--	.36*	.27*	.30*	.16	.24*	.10	.25*
9. PNarrativ									--	.37*	.48*	-.04	.36*	-.02	.13
10.BASC-T										--	.89*	.13	.25*	.30*	.33*
11. BASC-P											--	.23*	.34*	.23*	.27*

Table I-5 (continued)

Correlations, Means, Standard Deviations, Skew, and Kurtosis for Importance of Assessment Data for the LD Group in Sample 2 (n = 93)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
12. PA Criter												--	.10	.02	.12
13. DSMIV													--	.37*	.37*
14. SimProt														--	.53*
15. Preva															--
M	3.39	3.59	2.42	3.48	4.16	4.39	3.85	3.52	3.42	3.73	3.57	4.19	2.33	2.13	1.77
SD	1.03	1.09	1.01	.87	.81	.72	.91	.88	.85	.87	.84	.85	1.20	1.10	1.00
Skew	-.36	-.79	.54	.01	-.81	-1.10	-.67	-.44	-.39	-.34	-.06	-.93	.60	.50	1.27
Kurtosis	-.61	.18	-.13	-.63	.28	1.08	.25	-.17	.31	-.47	-.53	.33	-.48	-1.08	1.10

Note. RefQuest= referral questions; H&V = hearing and vision; VMI= visual-motor index; Observat= classroom observations; Achieve= achievement score; RTM = response-to-classroom modification; PNarrativ = parents comments; PA Criter= state criteria; DSMIV = DSM criteria; SimProt = similarity with prototype; Preva = prevalence of classification. * p <.01.

Table I-6

Correlations, Means, Standard Deviations, Skew, and Kurtosis for Importance of Assessment Data for the Complicated Group in Sample 2 (n = 91)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. RefQ	--	.63*	.37*	.63*	.22*	.35*	.56*	.33*	.58*	.23	.44*	.52*	.45*	-.55*	.10
2. H&V		--	.57*	.55*	.23*	.24*	.46*	.26*	.58*	.26*	.60*	.53*	.50*	-.47*	.01
3. VMI			--	.40*	.33*	.27*	.32*	.22*	.32*	.23*	.41*	.37*	.33*	-.23*	-.01
4. Observat				--	.29*	.39*	.61*	.29*	.54*	.31*	.52	.41*	.36*	-.46*	.12
5. IQ					--	.68*	.37*	.29*	.22*	.34*	.27*	.45*	.18	-.14	.03
6. Achieve						--	.43*	.28*	.41*	.37*	.41*	.47*	.32*	-.32*	-.07
7. RTM							--	.50*	.60*	.44*	.49*	.46*	.33*	-.47*	.06
8. Grades								--	.48*	.38*	.37*	.24*	.08	-.31*	-.06
9. PNarrativ									--	.46*	.70*	.38*	.41*	-.51*	-.01
10.BASC-T										--	.70*	.37*	.12	-.17	-.01
11. BASC-P											--	.47*	.33*	-.44*	.01

Table I-6 (continued)

Correlations, Means, Standard Deviations, Skew, and Kurtosis for Importance of Assessment Data for the Complicated Group in Sample 2 (n = 91)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
12. PA Criter												--	.43*	-.58*	-.03
13. DSMIV													--	-.28*	.03
14. SimProt														--	.28*
15. Preva															--
M	3.18	3.09	2.24	3.41	3.93	4.08	3.71	3.37	3.30	3.79	3.53	3.74	2.18	2.31	1.56
SD	1.12	1.22	.97	1.01	.71	.65	1.00	.75	.91	.62	.77	1.16	1.14	1.16	.73
Skew	-.07	-.06	.54	-.37	-.09	-.32	-.23	-.11	-.18	.18	-.02	-.38	.61	.29	.91
Kurtosis	-.91	-1.03	.04	-.42	-.51	.28	-.98	-.41	-.28	-.53	-.31	-1.32	-.71	-1.38	-.56

Note. RefQuest= referral questions; H&V = hearing and vision; VMI= visual-motor index; Observat= classroom observations; Achieve= achievement score; RTM = response-to-classroom modification; PNarrativ = parents comments; PA Criter= state criteria; DSMIV = DSM criteria; SimProt = similarity with prototype; Preva = prevalence of classification. * $p < .01$ level.

Table I-7

Correlations, Means, Standard Deviations, Skew, and Kurtosis for Importance of Assessment Data for the Entire Sample 2 (N = 272)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. RefQuest	--	.51*	.37*	.49*	.06	.14*	.41*	.24*	.47*	.14	.29*	.27*	.26*	-.22*	.10
2. H&V		--	.49*	.36*	.17*	.21*	.24*	.10	.43*	.25*	.39*	.25*	.38*	-.17*	.07
3. VMI			--	.33*	.29*	.19*	.25*	.19*	.38*	.31*	.38*	.17*	.37*	-.02	.15*
4. Observat				--	.10	.20*	.51*	.20*	.43*	.33*	.44*	.29*	.25*	-.19*	.13*
5. IQ					--	.71*	.19*	.26*	.19*	.42*	.37*	.29*	.19*	.05	.05
6. Achieve						--	.28*	.27*	.25*	.43*	.40*	.40*	.20*	-.03	.03
7. RTM							--	.40*	.43*	.33*	.38*	.36*	.23*	-.22*	.16*
8. Grades								--	.39*	.32*	.36*	.24*	.21*	-.03	.11
9. PNarrativ									--	.42*	.58*	.20 *	.20*	-.15*	.06
10.BASC-T										--	.84*	.21*	.22*	.13*	.17*
11. BASC-P											--	.31*	.35*	-.01	.15*

Table I-7 (continued)

Correlations, Means, Standard Deviations, Skew, and Kurtosis for Importance of Assessment Data for the Entire Sample 2 (N = 272)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
12. PA Criter												--	.26*	-.28*	.02
13. DSMIV													--	.12	.19*
14. SimProt														--	.43*
15. Preva															--
M	3.46	3.39	2.42	3.62	4.06	4.27	3.95	3.54	3.40	3.78	3.60	4.10	2.36	2.16	1.68
SD	1.02	1.12	.99	.90	.77	.70	.91	.87	.86	.74	.79	.96	1.18	1.11	.92
Skew	-.41	-.41	.45	-.41	-.49	-.62	-.62	-.35	-.30	-.21	-.12	-.85	.49	.48	1.30
Kurtosis	-.50	-.62	-.21	-.16	-.15	-.01	-.21	.01	-.08	-.21	-.20	-.27	-.66	-1.02	1.29

Note. RefQuest= referral questions; H&V = hearing and vision; VMI= visual-motor index; Observat= classroom observations; Achieve= achievement score; RTM = response-to-classroom modification; PNarrativ = parents comments; PA Criter= state criteria; DSMIV = DSM criteria; SimProt = similarity with prototype; Preva = prevalence of classification. * $p < .01$.

APPENDIX J
DEMOGRAPHIC COMPARISON OF CONDITIONS IN STUDY 2

Appendix J
Demographic Comparison of Conditions in Study 2

A paired samples *t*-test indicated that there was a statistically significant difference between participants' reported use of the discrepancy model (M = 3.32; SD = .71) in comparison to the RTI model (M = 1.61; SD = .78), t (271) = 22.68, p < .001, with school psychologists being almost twice as likely to use the discrepancy model over the RTI model (paired mean difference = 1.70). A summary of the *t*-test results is provided in Table J-1.

Table J-1

Summary of Paired-Samples t-test of Frequency of using the Discrepancy Model Versus RTI model for Study 2 (N = 272)

	M	SD	Std. Error Mean	t	df	Sig. (2-tailed)
DISCREPA - RTIMODE	1.6985	1.23492	.07488	22.684	271	.000

Note. DISCREPA = Frequency of using the discrepancy model; RTIMOD = Frequency of using the RTI model.

Findings for a series of one-way ANOVAs indicated that the three groups were not statistically significant different on (a) mean age of participants, F (2,267) = .07, p = .93, partial eta squared < .01; (b) number of years as certified school psychologists in Pennsylvania, F (2, 269) = .03, p = .97, partial eta squared < .01; (c) number of evaluations conducted in the prior year, F (2, 269) = 2.81, p = .06, partial eta squared = .02; (d) frequency of using the discrepancy model during assessment, F (2, 269) = .72, p = .49, partial eta squared = .01; and (e) frequency of using the Response to Intervention model during assessment, F (2, 269) = 2.32, p = .10, partial eta squared = .02. Summaries of the ANOVAs are provided in Tables K-2 through K-6.

Table J-2

Summary of ANOVA of ages of Participants in Experimental Groups for Study 2 (N = 272)

Source	df	Mean Square	F	Sig.	Partial Eta Squared
Complexity	2	10.89	.07	.933	< .01
Error	267	158.05			
Total	270				

Table J-3

Summary of ANOVA of Participants' Years of Certification Across Experimental Groups for Study 2 (N = 272)

Source	df	Mean Square	F	Sig.	Partial Eta Squared
Complexity	2	3.95	.03	.967	< .01
Error	269	115.85			
Total	272				

Table J-4

Summary of Number of Evaluations Completed per Year by Experimental Groups for Sample 2 (N = 272)

Source	df	Mean Square	F	Sig.	Partial Eta Squared
Complexity	2	2867.36	2.81	.062	.02
Error	269	1020.02			
Total	272				

Table J-5

Summary of ANOVA of Frequency in use of Discrepancy Model by Experimental Groups fro Study 2 (N = 272)

Source	df	Mean Square	F	Sig.	Partial Eta Squared
Complexity	2	.359	.721	.487	.005
Error	269	.498			
Total	272				

Table J-6

Summary of ANOVA of Frequency in use of RTI Model by Experimental Groups for Study 2 (N = 272)

Source	df	Mean Square	F	Sig.	Partial Eta Squared
Complexity	2	1.390	2.316	.101	.017
Error	269	.600			
Total	272				

Findings of the two-way contingency tables indicated that the experimental conditions did not differ in the number of males to females, X^2 (6, N = 272) = 2.25, p = .33, Cramer's V = .09 and in the distribution of ethnicities, X^2 (6, N = 272) = 8.18, p = .43, Cramer's V = .12. The three groups also did not differ in the number of school psychologists employed in various systems (school district, residential treatment center, independent or private practice, & intermediate unit) X^2 (6, N = 272) = 5.28, p = .51, Cramer's V = .10 or in the residential region of employment (urban, rural, suburban), X^2 (6, N = 272) = 9.43, p = .15, Cramer's V = .13. Additionally, participants across experimental conditions were equivalent in the number who had attained a master's or doctoral degree, X^2 (6, N = 272) = 10.72, p = .03, Cramer's V = .14. Finally, no statistically significant differences were found in participants across conditions in the program orientation of training (scientist-practitioner, practitioner-scholar, & practitioner focused), X2 (6, N = 272) = 2.66, p = .62, Cramer's V = .07. Summaries of the expected counts are provided in Tables J7 through J12.

Table J-7

Summary of 3 x 2 Contingency Table of Complexity Level and Gender in Study 2 (N = 272)

		Gender		
Complexity		Male	Female	Total
Normal	Count	27.0	61.0	88.0
	% within	30.7	69.3	100.0
	% of Total	9.9	22.4	32.4
LD	Count	21.0	72.0	93.0
	% within	22.6	77.4	100.0
	% of Total	7.7	26.5	34.2
Complicated	Count	20.0	71.0	91.0
	% within	22.0	78.0	100.0
	% of Total	7.4	26.1	33.5
Total	Count	68.0	204.0	272.0
	% within	25.0	75.0	100.0
	% of Total	25.0	75.0	100.0

Table J-8

Summary of 3 x 4 Contingency Table of Complexity Level and Setting of Employment in Study 2 (N = 272)

Complexity		Setting of Employment: School District	IU	Independent	Residential	Total
Normal	Count	72.0	13.0	3.0	0.0	88.0
	% within	81.8	14.8	3.4	0.0	100.0
	% of Total	26.5	4.8	1.1	0.0	32.4
LD	Count	72.0	17.0	3.0	1.0	93.0
	% within	77.4	18.3	3.2	1.1	100.0
	% of Total	26.5	6.3	1.1	0.4	34.2
Complicated	Count	72.0	12.0	7.0	0.0	91.0
	% within	79.1	13.2	7.7	0.0	100.0
	% of Total	26.5	4.4	2.6	0.0	33.5
Total	Count	216.0	42.0	13.0	1.0	272.0
	% within	79.4	15.4	4.8	0.4	100.0
	% of Total	79.4	15.4	4.8	0.4	100.0

Note. IU= Intermediate Unit

Table J-9

Summary of 3 x 4 Contingency Table of Complexity Level and Region of Employment in Study 2 (N = 272)

		Region				
Complexity		Rural	Urban	Suburban	Multiple	Total
Normal	Count	33.0	14.0	36.0	5.0	88.0
	% within	37.5	15.9	40.9	5.7	100.0
	% of Total	12.1	5.1	13.2	1.8	32.4
LD	Count	37.0	15.0	40.0	1.0	93.0
	% within	39.8	16.1	43.0	1.1	100.0
	% of Total	13.6	5.5	14.7	0.4	34.2
Complicated	Count	26.0	13.0	51.0	1.0	91.0
	% within	28.6	14.3	56.0	1.1	100.0
	% of Total	9.6	4.8	18.8	0.4	33.5
Total	Count	96.0	42.0	127.0	7.0	272.0
	% within	35.3	15.4	46.7	2.6	100.0
	% of Total	35.3	15.4	46.7	2.6	100.0

Table J-10

Summary of 3 x 2 Contingency Table of Complexity Level and Degree Attainment in Study 2 (N = 272)

		Highest Degree Attained		
Complexity		Masters	Doctorate	Total
Normal	Count	60.0	28.0	88.0
	% within	68.2	31.8	100.0
	% of Total	22.1	10.3	32.4
LD	Count	70.0	23.0	93.0
	% within	75.3	24.7	100.0
	% of Total	25.7	8.5	34.2
Complicated	Count	69.0	22.0	91.0
	% within	75.8	24.2	100.0
	% of Total	25.4	8.1	33.5
Total	Count	199.0	73.0	272.0
	% within	73.2	26.8	100.0
	% of Total	73.2	26.8	100.0

Table J-11

Summary of 3 x 3 Contingency Table of Complexity Level and Model of Training in Study 2 (N = 272)

		Model of Training			
Complexity		Scientist-practitioner	Practioner-scholar	Practitoner focused	Total
Normal	Count	42.0	19.0	27.0	88.0
	% within	47.7	21.6	30.7	100.0
	% of Total	15.4	7.0	9.9	32.4
LD	Count	41.0	23.0	29.0	93.0
	% within	44.1	24.7	31.2	100.0
	% of Total	15.1	8.5	10.7	34.2
Complicated	Count	43.0	27.0	21.0	91.0
	% within	47.3	29.7	23.1	100.0
	% of Total	15.8	9.9	7.7	33.5
Total	Count	126.0	69.0	77.0	272.0
	% within	46.3	25.4	28.3	100.0
	% of Total	46.3	25.4	28.3	100.0

VITA
Sharise M. Wilson

EDUCATION

PhD., School Psychology, Pennsylvania State University, University Park, May 2010
M.S., School Psychology, Pennsylvania State University, University Park, December 2005
B.A., Psychology, Pace University, New York, June 2003

CERTIFICATION

Commonwealth of Pennsylvania Professional Certificate: Educational Specialist in School Psychology (NCATE; since 2007)

TEACHING EXPERIENCE

Summer 2009 Frederick Douglass Scholar *Lock Haven University of PA*

CLINICAL EXPERIENCE

Summer 2008- present	Certified School Psychologist	*Sugar Valley Rural Charter School*
Summer 2007- fall 2008	Pre-Doctoral Internship	*Central Intermediate Unit 10*
Fall 2006- present	Behavior Specialist/ Mobile Therapist	*Northwestern Human Services*
Summer 2006- fall 2006	Group Therapist	*Community Services Group*
Summer 2005- summer 2006	Mental Health Advisor	*Community Services Group*

PUBLICATIONS

Watkins, M. W., Wilson, S. M., Kotz, K. M., Carbone, M. C., & Babula, T. (2006). Factor structure of the WISC-IV among referred students. *Educational and Psychological Measurement*, *66*, 975-933.

Wilson, Sharise. (2005). Hair analysis. In J. Neisworth & P. Wolfe (Eds.), *The autism encyclopedia* (p. 95). Baltimore, MD: Paul H. Brookes Publishing Co.

Wilson, Sharise. (2005). Mania. In J. Neisworth & P. Wolfe (Eds.), *The autism encyclopedia* (p. 130). Baltimore, MD: Paul H. Brookes Publishing Co.

ORGANIZATION AFFILIATIONS

National Association of School Psychology
Association of School Psychologists of PA
Delta Sigma Theta Sorority, Inc.

AWARDS RECEIVED

Spring 2009	School Psychology Alumni Fund, Pennsylvania State University
Fall 2007- fall 2009	New York Community Trust Eugen Grabscheid Fund
Spring 2007	Bunton Waller Fellowship
Fall 2004- spring 2005	Burdett E. Larson Graduate Fellowship
Fall 2003-summer 2006	Bunton Waller Fellowship