

Running head: DECISION MAKING IN SCHOOL PSYCHOLOGY

SCHOOL PSYCHOLOGY STUDENTS' DECISION MAKING: IMPLICATIONS FOR TRAINING

A DISSERTATION  
SUBMITTED TO THE FACULTY  
OF  
THE GRADUATE SCHOOL OF APPLIED AND PROFESSIONAL PSYCHOLOGY  
OF  
RUTGERS,  
THE STATE UNIVERSITY OF NEW JERSEY  
BY  
DINA CHAYA SIEGEL  
IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE  
OF  
DOCTOR OF PSYCHOLOGY

NEW BRUNSWICK, NEW JERSEY

OCTOBER 2020

APPROVED:

\_\_\_\_\_  
Nancy S. Fagley, PhD

\_\_\_\_\_  
G. Terence Wilson, PhD

DEAN:

\_\_\_\_\_  
Francine Conway, PhD

# DECISION MAKING IN SCHOOL PSYCHOLOGY

Copyright 2020 by Dina Siegel

## ABSTRACT

Although scholars have suggested that cognitive errors may plague school psychologists' decisions, there is little research documenting these effects or the extent to which school psychologists are trained to spot these biases. I conducted a study using experimental vignette methodology (EVM) in conjunction with a survey of knowledge, attitudes, and training in clinical judgment heuristics and biases. Different versions of a vignette formed a 2 (referral concern) x 2 (assessment scores) between subjects design, and participants were randomly assigned to one of the cells of the design. Participants were recruited from NASP-approved, APA-accredited, and/or locally accredited school psychology programs. Although 310 participants provided informed consent for the survey, 175 were included for analysis in the final sample, as only they indicated they were enrolled in a school psychology program. The first goal of the study was to determine school psychology trainees' knowledge, attitudes, and training related to topics in clinical judgment. The second goal of the study was to assess whether trainees exhibit judgmental biases when identifying Specific Learning Disabilities (SLD). I hypothesized that trainees would have received minimal formal training in limitations of clinical judgment, and they would demonstrate bias in classifying SLD based on the wording of the referral concern (i.e., whether or not the referral mentioned SLD). Results indicated that most participants reported having training in at least one area of clinical judgment and demonstrated some knowledge of clinical judgment. Overall attitudes toward understanding clinical judgment limitations within school psychology were positive ( $M = 1.78$ ,  $SD = .40$ ). There was a significant main effect of assessment score discrepancy on SLD classification,  $F(1, 171) = 4.43$ ,  $p = .037$ , meaning that participants based their classification decision, at least in part, on the assessment scores. Partial eta squared = .025. There was a non-significant main effect of referral concern on

SLD classification,  $F(1, 171) = 0.009, p > .05$ , meaning that participants were not biased by the framing of the referral concern. The interaction effect was also non-significant,  $F(1, 171) = 0.271, p > .05$ . Limitations and implications for future training of school psychologists are discussed.

## ACKNOWLEDGEMENTS

Nancy, thank you for helping me get this dissertation off the ground and providing helpful insights at each step. At times, writing a dissertation can feel daunting and knowing that you were just as excited about the topic helped me to push through. Thank you for sharing your expertise on the topic of clinical judgment and setting up a Sakai site to direct me toward resources. Congratulations on your hard-earned retirement! Terry, thank you for introducing me to the clinical judgment literature during my first year of graduate school. The ideas that you expressed in your courses continue to shape my professional identity and I feel very fortunate to have had the opportunity to take your courses. I am grateful to the GSAPP school psychology faculty who expanded my view of the profession and fostered a sense of community.

Thank you to the following people who each invested in my progress on this dissertation:

To GSAPP's Greatest Gals for encouraging me every tiny step along the way of this dissertation! Who else would spend hours listening to my ideas, setting up motivational challenges, and rooting for me with the best emojis?! You are truly the best and I can't wait to share post-grad school adventures together. To Arielle Walzer for sharing her dissertation wisdom with me. To Rachel Langer for piloting my survey and editing an earlier version of this dissertation.

My parents and siblings navigated the fine line of celebrating my accomplishments and refraining from asking about my dissertation. I know it's a tall order and I appreciate your unwavering support! The seeds for this dissertation were sown long before I entered graduate school...As we say in our family, "evidence, shmevidence."

## TABLE OF CONTENTS

|                                                              |           |
|--------------------------------------------------------------|-----------|
| <b>ABSTRACT</b> .....                                        | <b>ii</b> |
| <b>ACKNOWLEDGMENTS</b> .....                                 | <b>iv</b> |
| <b>LIST OF TABLES</b> .....                                  | <b>vi</b> |
| <b>INTRODUCTION</b> .....                                    | <b>1</b>  |
| Clinical Judgment Limitations .....                          | 2         |
| Clinical Judgment in School Psychology .....                 | 4         |
| School Psychology Training .....                             | 10        |
| Present Study .....                                          | 14        |
| <b>METHOD</b> .....                                          | <b>16</b> |
| Participants .....                                           | 16        |
| Procedure.....                                               | 17        |
| Measures.....                                                | 18        |
| Design.....                                                  | 19        |
| <b>RESULTS</b> .....                                         | <b>20</b> |
| Descriptive Statistics .....                                 | 20        |
| Relationship between Knowledge, Attitudes, and Training..... | 23        |
| ANOVA.....                                                   | 23        |
| Additional Vignette Analyses .....                           | 24        |
| <b>DISCUSSION</b> .....                                      | <b>25</b> |
| Limitations.....                                             | 29        |
| Future Directions.....                                       | 31        |
| Impact on School Psychology Training .....                   | 32        |
| <b>REFERENCES</b> .....                                      | <b>36</b> |
| <b>APPENDICES</b> .....                                      | <b>53</b> |

## LIST OF TABLES

| Table                                                | Page # |
|------------------------------------------------------|--------|
| Diagram of the Experimental Design .....             | 44     |
| Demographics .....                                   | 45     |
| Knowledge of Clinical Judgment Errors .....          | 46     |
| Attitudes toward Clinical Judgment Research .....    | 47     |
| Training in Clinical Reasoning .....                 | 48     |
| Correlations between Clinical Judgment Measures..... | 49     |
| Descriptives for Vignette Conditions .....           | 50     |
| ANOVA Summary Table .....                            | 51     |
| Correlations for Vignette Variables .....            | 52     |

### **Introduction**

Since the mid-1950s, a growing body of interdisciplinary researchers have conducted empirical investigations on the reliability and predictive validity of clinical judgments (e.g., Arkes & Hammond, 1986; Baker et al., 1982; Tanner, 2006). In most studies, actuarial judgments, or those calculated based on historical data yielding mathematical probabilities, have matched or outperformed human judgment in reliably predicting behavior (Kahneman, 2011; Meehl, 1986). Paul Meehl, a pioneer of clinical judgment research, remarked, “There is no controversy in social science which shows such a large body of qualitatively diverse studies coming out so uniformly in the same direction as this one...when you can hardly come up with a half dozen studies showing even a weak tendency in favor of the clinician, it is time to draw a practical conclusion” (Meehl, 1986). His conclusion that statistical prediction yields more accurate results continues to be supported by studies with direct comparisons between algorithms and human judgment. Consequently, researchers and clinicians have sought to understand the limitations and find a defensible role for clinical judgment in the decision-making process (Kleinmuntz, 1990).

When applying the clinical judgment method to a professional context, a practitioner may use inductive or deductive methods to form a working hypothesis. Regardless of whether they collect specific observations to draw a broad conclusion (inductive method) or attempt to fit new pieces of information into an existing theory (deductive method), practitioners can rarely sort through all the information. For example, emergency room personnel may have under three minutes to make a diagnosis, auto mechanics must find the faulty part among a complicated set of components, and umpires have less than a few seconds to call a play. Although the consequences of these decisions vary, in each case, factors such as prior experience and training affect which information is noticed, what information is sought out, and how all input is



combined and interpreted. Moreover, these decisions are often made under time constraints and with ambiguous data, resulting in prioritizing speed over accuracy. Distinguishing which cognitive processes facilitate, rather than inhibit, accurate integration is crucial for effective decision-making (Turk & Salovey, 1985).

### **Clinical Judgment Limitations**

#### ***Heuristics***

In an effort to understand – and ultimately improve – decision making, Tversky and Kahneman (1974), conducted a series of experiments that highlighted the cognitive processes involved in estimating probabilities. In a paper summarizing their findings, they explained that people consistently rely on mental shortcuts, or heuristics, to determine odds of uncertain events. In Kahneman and Tversky's (1974) original formulation, the key heuristics were labeled representativeness, availability, and adjustment and anchoring. These heuristics explain the mechanisms people use to make inferences and the reasons they sometimes fail to select the rational choice, or the one predicted based on objective fact. Subsequently, additional heuristics, and their resulting misapplications, have been identified.

Initially, research on heuristics focused on the fallibility of human judgment, but the utility of heuristics has also been examined. Processes that reduce cognitive load are necessary to increase the amount of data that a person can store, making heuristics essential for ordinary functioning (Dumont, 1993). Although imprecise, heuristic thinking, when matched to the environment, often fosters “good enough” decision-making (Gigerenzer & Gaissmaier, 2011). Because humans have increasingly created environments, situations, and tasks with a multitude of outcomes that differ from those encountered by our early ancestors, Gigerenzer and Gaissmaier (2011) proposed a model of heuristics based on ecological rationality. In this model,

which heuristic is selected for a given task depends on the conditions of the environment.

Specifically, when multiple, complex cues are presented, using heuristics reduces the input of information and outperforms linear regression models, leading to the conclusion that “less is more” (Gigerenzer & Gaissmaier, 2011). As the science behind heuristics continues to be explored, more research is needed to understand when heuristics, or simple decision-making models, should be favored over weighted formulas.

### ***Cognitive Biases***

Cognitive biases are the product of systematic reasoning errors which lead to gathering incomplete or unreliable data. Cognitive biases are particularly insidious because people are often unaware that their reasoning process is flawed (Bowes et al., 2020). Confirmation bias is one of the most pervasive, systematic sources of error in reasoning. At times, confirmation bias is the result of an attempt to use newly gathered data to support an a priori hypothesis; in other cases, a person selectively attends to information even when they are not invested in a particular outcome (Nickerson, 1998). Lilienfeld and Lynn (2014) described confirmation bias as “the mother of all biases,” affecting clinicians, researchers, and lay people. Other related biases, such as overconfidence, and blind spot bias may have negative consequences in a clinical context (Lilienfeld & Lynn, 2014).

Because inductive reasoning, or systematically gathering all information relevant to a case, is virtually impossible when dealing with psychological diagnoses, clinicians use deductive methods that often rely on fitting information into preexisting theories or frameworks (Wilcox & Schroeder, 2015). If a clinician approaches a situation with the intention of providing support for their hypothesis, they may neglect to seek evidence that would disconfirm the hypothesis, thereby reinforcing their initial beliefs. Collectively, using judgmental heuristics or exhibiting

cognitive biases can impact the reliability and validity of clinical judgment (Wilcox & Schroeder, 2015). Throughout this dissertation, the term *cognitive errors* refers to all cognitive processes that can negatively affect clinical judgment. Although a review of all cognitive errors is beyond the scope of this work, key cognitive errors and their implications for school psychology are addressed. See Bowes et al., (2020) for a comprehensive overview of cognitive errors in clinical practice.

### **Clinical Judgment in School Psychology**

In 2010, the National Association of School Psychologists (NASP) created the Model for Comprehensive and Integrated School Psychological Services (Practice Model) to reflect changing practices within the profession. The 10 domains that comprise the Practice Model were established to guide school psychological training and practice. According to this model, the first practice domain is Data-Based Decision Making and Accountability, defined in part as “a systematic and comprehensive process of effective decision making and problem solving that permeates all aspects of service delivery” (NASP, 2010). Despite NASP’s emphasis on embedding data driven techniques within practice, there is a paucity of research on the methods school psychologists use to execute decisions (Andrews & Syeda, 2017; Gammelmo, 2015; McGill et al; 2016, Woodward, 2020). Moreover, school psychologists may be unaware of the sources of data that influence their decisions (Aspel et al., 1998).

As schools shift toward a more comprehensive service delivery model, which includes a greater emphasis on linking assessment and intervention practices, school psychologists require data driven solutions for developing evidence-based recommendations and evaluating their effectiveness (NASP, 2010). Statistical tools have been developed to aid school psychologists’ clinical reasoning, but few studies have examined their use or impact on decision making within

schools (e.g., Pendergast et al., 2018). More commonly, school psychologists use a combination of data sources with varying levels of validity (e.g., teacher reports, prior experience, informal observations of behavior during test administration, psychological test results) to determine special education classification and subsequent intervention. Several studies have demonstrated that school psychologists prioritize personal experience, knowledge, and beliefs over evidence from research literature when designing interventions (e.g., Forman et al., 2009; Lilienfeld et al., 2012).

Stoddard (2017) conducted a survey of school psychologists to understand which methods are most often used to assess students' progress in counseling. He found that participants rarely relied exclusively on evidence-based measures to assess progress (Stoddard, 2017), indicating their perceptions of students' need for continued services was biased by external factors. Possible factors, such as prior experience working with similar students (representativeness heuristic) or the session frequency outlined in the student's IEP (anchoring), may have influenced decisions. Further complicating the study of school psychologists' data-based decision making is the lack of a universally accepted method for interpreting assessment data, thereby reducing the generalizability of classification decisions.

### ***Conceptual Models of Cognitive Errors in the Decision-Making Process***

Data on school psychologists' everyday decision making are limited; conceptual models based on cognitive science are useful in increasing awareness of effective decision-making strategies and potential pitfalls. Lilienfeld et al., (2012) developed a list of recommendations for utilizing science to inform school psychologists' decisions. They stressed the importance of following a scientific process, or "a commitment to rooting out errors in our web of beliefs," to reduce the effects of cognitive errors and promote accurate reasoning. Andrews and Syeda

(2017), outlined an approach to clinical reasoning that is embedded in learning from the modeling of experienced practitioners. Because the ratio of school psychologists to students is high, school psychologists may not regularly interact and modify their own reasoning by learning from a fellow practitioner. Wilcox and Schroeder (2015) proposed using inductive and deductive methods to promote optimal decision making. While in the initial phases of learning, trainees tend to rely on deductive methods because they have not had sufficient training and experience to develop heuristics, or mental shortcuts. They may not have the necessary context to disregard some data sources as irrelevant for the current task. Careful consideration of whether assessment data or the absence of expected data support a specific claim or diagnosis often leads to accurate conclusions but the process can be time consuming. Professors can facilitate awareness of clinical reasoning through direct instruction and collaborative discussion, with the added benefit of providing immediate feedback for error correction. Supervisors can play a role in coaching trainees to explain their thought process behind selecting a diagnosis (Wilcox & Schroeder, 2015).

Within the field of decision-making, several conceptual models have been proposed to explain how people make decisions and avoid potential errors in reasoning (e.g., Arkes, 1991; Reason, 1990). Watkins' (2009) suggested applying a three-part model to classify the various types of errors school psychologists make. Accordingly, errors can be skills-based, knowledge-based, or rule-based, and a deficit in any area can lead to poorer decision making. Consistent with Watkins' (2009) umbrella of rule-based errors, Fagley (1988) highlighted possible cognitive errors that could plague school psychologists' decision making. As school psychologists are primary members of special education teams, they can shape the course of a student's education. However, school psychologists often operate under strict deadlines, thereby necessitating

efficient decision-making that can result in biased cognition (Fagley, 1988). At each phase in the special education referral process, clinicians may seek and attend to information that confirms their prior beliefs about the case (Wilcox & Schroeder, 2015). Wilcox & Schroeder (2015) elaborated on the ways that cognitive errors may lead to poor decision making. Prior to conducting an evaluation, clinicians may adjust their expectations based on the existing framework of the referral concern (anchoring) or attend to incomplete information worded to focus on gains versus losses (framing effect). When interpreting assessment data, several heuristics and biases may cause the school psychologist to overestimate the likelihood that a particular student has a disability (base rate neglect), focus on irrelevant aspects of data to create a coherent narrative (illusory correlation), or seek out or prioritize data that concur with the initial hypothesis (confirmation bias).

### ***Empirical Studies of Cognitive Errors in the Decision-Making Process***

In the 1980s and 90s several empirical investigations of clinical judgment within school psychology provided mixed support for the conceptual models of decision making and cognitive errors (e.g., Gnys et al., 1995; Kennedy, et al., 1997; Ysseldyke & Algozzine, 1981). Notably, when school psychologists were asked to rank order data used to determine a student's special education classification, their subjective impressions of which sources most influenced their choice of classification differed from the sources of data that they used in practice. School psychologists lacked awareness of the factors that contributed to their decisions (Aspel et al., 1998) which may increase their susceptibility to bias. Findings from several experiments indicated that school psychologists' special education classification decisions were biased by referral concerns (Ysseldyke & Algozzine, 1981), influenced by irrelevant factors within the testing data (Gnys et al., 1995), and impacted by clinical information more than base rates

(Kennedy, et al., 1997). Other concurrent findings failed to find evidence of cognitive errors among school psychologists (e.g., Heubner, 1991). Fagley, et al., (1999) found that school psychology students, overall, were significantly influenced by framing of alternative options. The decision problems involved decisions regularly faced by school professionals, but the options between which they were asked to choose were framed either in terms of potential losses or gains. Importantly, the outcomes in the two frames were objectively identical (e.g., 20 of 30 children would read at grade level vs 10 of 30 would continue to read below grade level). Notably, Fagley et al., (1999) failed to find evidence of a framing effect when the problem was about special education placement, which is a routine activity for school psychologists. Consequently, further study is needed to understand the susceptibility of school psychologists' decision making to cognitive error.

**Recent Peer-Reviewed Studies.** In the decade since the reauthorization of the Individuals with Disabilities Education Act in 2004 (renamed the Individuals with Disabilities Education Improvement Act [IDEA], 2004) and the development of the NASP Practice Model in 2010, there have been few peer-reviewed studies of school psychologists' clinical judgment. Studying the judgment process of school psychologists, who are tasked with identifying students with a broad range of disabilities, is increasingly relevant (Jordan, 2019). Della Toffalo and Pedersen (2005) studied whether a student's having a psychiatric diagnoses biases school psychologists' classification of students. When the student in the vignette had a psychiatric diagnosis, school psychologists were more likely to classify the student with an Emotional Disturbance, regardless of whether eligibility criteria for special education were met. Although Della Toffalo and Pedersen (2005) did not explicitly discuss cognitive factors, their findings indicate that school psychologists' judgments are susceptible to bias. Knowingly or

unknowingly, participants relied on the student's diagnosis instead of strictly adhering to classification criteria, suggesting that the representativeness heuristic caused them to ignore base rates for ED classification and classify a student because of their diagnosis.

**Recent Dissertations.** Several unpublished dissertations have examined school psychologists' decision-making during the assessment process. In contrast to previous studies' findings that school psychologists focus on irrelevant data in determining special education classifications (e.g., Gnys et al., 1995, Ysseldyke & Algozzine, 1981), Gammelmo (2015) found that trainees relied heavily on cognitive and academic achievement test scores and referral questions to make classifications. Gammelmo (2015) analyzed the extent to which school psychology trainees relied on different types of data, including referral question, demographics, and assessment scores in determining classifications for special education. She concluded that trainees were not overly influenced by sources of bias, such as race, age, or sex, such that classifications were not determined by these demographic characteristics. Her analyses did not consider the cognitive strategies school psychology students use when collecting and interpreting information, leaving out a key source of potential bias. Baker (2000), noted that age and gender have not been found to bias classification decisions and varied the type of referral (academic or behavioral) to determine other sources of bias. She found that regardless of whether the referral concern depicted a student with an academic or behavioral problem, participants made appropriate classification decisions based on the background information, observation, and cognitive and academic test results. Additionally, participants' judgments did not differ depending on whether the referral concern was presented before or after the assessment findings.

Sadeh (2016) and Woodward (2020) modified case vignettes in their dissertations to control for level of data ambiguity, such that the vignettes differed in terms of level of detail



and/or clarity of assessment results. Sadeh (2016) failed to find a difference in the rate of classification of students based on race when specifically looking at students considered for ED. Her results remained consistent when controlling for level of data ambiguity, suggesting that her sample of school psychologists relied on valid predictors of ED. Woodward (2020) directly assessed cognitive errors within the assessment process. He found that when considering sources of data for identifying SLD, school psychologists were not biased by the referral concern or level of ambiguity, suggesting they did not use a confirmatory approach. Collectively, these dissertation studies found that school psychologists applied accurate reasoning to make special education classification decisions.

### **School Psychology Training**

#### ***Traditional Approaches to Learning Disability Identification***

Although administering and interpreting cognitive assessments has been a foundational practice of school psychology, the purpose of administering them has evolved to include diagnosing and remediating SLD (Braden & Shaw, 2008). As outlined by IDEA (2004), “a Specific learning disability means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations, including conditions such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia.” Based on the federal definition, understanding of a person’s psychological process is necessary for diagnosing SLD. Early theories suggested that a person demonstrated a deficit in psychological processing if there was a severe discrepancy between their intelligence, or ability level, and academic functioning. Although the severe discrepancy model was accepted based on face validity, it failed to yield adequate specificity or

sensitivity, creating a need for alternative approaches to SLD identification. Despite limited evidence that cognitive assessments can reliably differentiate SLD, they continue to form the basis of SLD identification (Fletcher & Miciak, 2017).

### ***Recent Advances in Learning Disability Identification***

Researchers from differing traditions within school psychology have advanced several approaches to more accurately identify SLD. The development of Patterns of Strengths and Weaknesses (PSW) approach to SLD identification, which attempted to revitalize the science behind learning disabilities, relies on clinical interpretation of patterns among assessment scores (Flanagan, 2017). According to Flanagan's (2017) definition, "A diagnosis of SLD is a clinical judgment...based on a convergence of data sources that appear to be consistent with the SLD construct." PSW models provide an alternative to relying on a single discrepancy score while still conceptualizing SLD as a cognitive processing deficit. In contrast to cognitive methods, which define SLD as an underlying cognitive processing problem, RTI uses an ecological behavioral lens to hypothesize the reasons for a student's failure to make progress within their academic environment (Greenwood & Kim, 2012). RTI utilizes screening assessments to gauge the academic needs of the entire student population and considers students who fall below a threshold or have a slower than expected rate of progress following remediation to have SLD (Clemens, et al., 2016). Proponents of using comprehensive cognitive assessment within an RTI framework suggest that a merging of both perspectives leads to the most accurate identification (Hale et al., 2006).

Evaluating students for special education and related services continues to be a core function of school psychologists (Larson & Choi, 2010); nonetheless, methods for identifying SLD vary widely (Benson et al., 2020). Maki, et al., (2017) found that school psychologists'

decision to identify a student with SLD, regardless of which identification model used, was impacted by the conclusiveness of the evaluation data. Their study surveyed practicing school psychologists and asked them to use one of three identification models to interpret a student's evaluation data. The models that participants used were IQ-Achievement Discrepancy, PSW, and RTI. The cases were manipulated such that in one scenario the student clearly met criteria for SLD, in another scenario the student clearly did not meet criteria for SLD, and in the last scenario the results were inconclusive as to whether the student met criteria for SLD. When the student's scores were close to the established cut score for SLD, and therefore inconclusive, school psychologists' accuracy tended to decrease. Although they did not specify which other information school psychologists used, the authors acknowledged that participants relied on individual judgment to supplement established criteria, which biased their decisions. The results of this study support that heuristics are likely to lead to decreased accuracy when data are inconclusive.

### ***Developing Clinical Reasoning***

Numerous studies in related health fields have demonstrated the presence of incidental and systematic biases in clinical judgment, yet application of these findings to training professional psychologists has lagged. Although researchers have demonstrated the presence of bias across multiple contexts, they have failed to devote equal attention to developing and teaching techniques to reduce it (Lilienfeld & Lynn, 2014). Educating trainees on cognitive errors may be a prerequisite for improving decision making (Watkins, 2009). However, there is limited information on the extent to which mental health professionals are exposed to ideas about cognitive errors (Rock, 1994; Turk & Salovey, 1985). In Rock's (1994) survey of mental health clinicians, participants were mostly unfamiliar with the literature on the limitations of clinical

judgment. If school psychologists are unaware of potential biases that could arise in their job, they are unlikely to avoid them or correct them if they occur.

School psychology students who are still forming their professional identity and practice are well positioned to benefit from training in clinical judgment errors. Clinicians routinely engage in the same types of activities and are, therefore, likely to commit the same errors, in the absence of explicit training. Arkes (1993) reviewed findings from several of his earlier studies in which he, along with colleagues, found that repetition of a statement increases the perception of its veracity. The validity of the statement was judged based on its familiarity, rather than its accuracy. This phenomenon poses a significant challenge to school psychologists, as they may form an inaccurate belief about a problem formulation and then continue to conceptualize similar cases from that framework because they are familiar with it. Engaging in the same processes over time may increase clinicians' confidence, regardless of the validity of their methods. Furthermore, psychologists may become fond of a specific technique, even if it lacks evidence, because they use it repeatedly and are familiar with it. Once they have established a preferred method or technique, they are unlikely to receive immediate feedback on the accuracy of their judgments from supervisors and are, therefore, less likely to assess for bias (Dawes, et al., 1989).

Studies spanning the past 40 years have noted that school psychologists use a combination of data collection and interpretation methods, but these studies have generally excluded trainees. Trainees represent a critical group for study, as their clinical judgment is actively developing and there is some evidence that they formulate problems differently from experienced practitioners (e.g., Spengler et al., 2009; Wilcox & Schroeder, 2015). Spengler et al., (2009) conducted a meta-analysis to understand the role of experience level in clinical judgment among mental health practitioners. Spengler and Pilipis (2015) updated the original meta-

analyses by including articles from 1970-2010. Across both meta-analyses, there was a small effect of experience, meaning that clinicians with more experience were more likely to make accurate decisions. The authors cautioned that although statistically significant, the effect was small. Added training and experience accounted for a small percentage of the variance in clinical judgment, meaning that other factors contribute to accurate problem formulation and recommendations. Nonetheless, based on Spengler et al., (2009) and Spengler and Pilipis (2015) meta-analyses, trainees may exhibit poorer judgment than experienced professionals when formulating a problem and developing intervention recommendations.

### **Present Study**

Previous articles have indicated that research on clinical judgment in applied psychology is limited, and even more limited when focusing exclusively on school psychology (Andrews & Syeda, 2017; Lilienfeld et al., 2012, Wilcox & Schroeder, 2015). However, school psychologists make decisions that can affect individual students, as well as overall school system operations. In the past two decades, standards for school psychological practice have changed substantially, but few empirical studies have evaluated the extent to which school psychologists rely on clinical judgment in addition, or to the exclusion of, practice standards. Most of the research that was conducted on cognitive errors in school psychology predates the passage of IDEA and the creation of the NASP Practice Model. Evidence from earlier studies suggests that school psychologists' decisions are based on multiple indicators, including clinical judgment, and they are often unaware of the factors that affect their decision making (e.g., Aspel et al., 1998; Maki, 2017). Further study of school psychologists' judgment would enhance understanding of which biases, if any, are most prominent within their practice.

To my knowledge, no studies have directly assessed school psychologists' knowledge, attitudes, and training related to clinical judgment and its limitations. Furthermore, studies that did examine clinical judgment often focused on practicing school psychologists and excluded trainees. As training standards in school psychology have changed, there is a need to assess whether school psychology trainees experience training consistent with these standards. In the absence of step by step practice guidelines and/or professional training, school psychologists are likely to be unaware of these biases in their practice (Watkins, 2009). This study will help to validate a measure of clinical judgment among school psychology trainees to understand their level of engagement with the topic.

The goal of the present study is twofold; I will document school psychology trainees' knowledge, attitudes and training related to cognitive errors in clinical judgment and evaluate their susceptibility to cognitive errors when making special education classification decisions.

### ***Hypotheses***

Based on a review of the literature, there appear to be several circumstances in which cognitive errors can affect the practice of school psychology. Without knowledge of these errors, trainees are unlikely to be able to anticipate, avoid, or correct them (Lilienfeld et al., 2012; Watkins, 2009).

1. I hypothesize that trainees will have received minimal exposure to the topic of clinical judgment errors in their formal training.
2. I hypothesize that participants will rate the likelihood that the hypothetical student has a SLD as higher when the referral concern explicitly states that the teacher suspects a SLD.

### ***Research Questions***

The proposed study aims to answer the following questions:

1. How much training do school psychology graduate students receive in cognitive errors and or limitations of clinical judgment?
2. How much do school psychology graduate students know about the cognitive limitations of clinical judgment?
3. To what extent do school psychology graduate students think that clinical judgment research is relevant to their practice?
4. Do school psychology graduate students engage in biased reasoning when making decisions about special education classification?

### **Method**

#### **Participants**

Participants were drawn from a convenience sample of school psychology graduate students from NASP, APA, and locally accredited programs. A total of 310 participants provided informed consent for the study. Of these, 67 declined to answer any survey questions, 10 indicated that they were not current school psychology students, and 58 did not complete the demographic information and could not be verified as school psychology students. The remaining 175 responses were analyzed (56.5%). The sample of 175 participants was predominantly female (86.3%), White (71.4%), and between the ages of 22 and 31 (83%). Ages ranged from 22 to 50 ( $M=27.7$ ,  $SD=5.47$ ). Participants had completed from 1-7 years of their program ( $M=3.3$ ,  $SD=1.275$ ). About half were enrolled in a school psychology doctoral program (52.0%), about a third in an educational specialist school psychology program (33.1%), and the rest were in a masters school psychology program (14.9%) at the time of the study. No other exclusion criteria were applied.

**Procedure**

The study was designed and administered using Qualtrics online survey software. An initial request with a link to the study survey was sent to 202 NASP-approved and/or APA-accredited school psychology program chairs or coordinators, as well as to a listserv for school psychology trainers. Participants received an email from their program contact and could click the link to access the informed consent page and survey. Participants were notified of informed consent policies and their rights as study participants. All participant responses were recorded anonymously and identifying information was not collected. See Appendix A for a copy of the email request sent to training directors and graduate students.

Participants were presented with a vignette about a hypothetical student who was referred for special education and related services. The vignette included a brief summary of the student's academic performance and the teacher's reason for referral. After reading the vignette, the participants were presented with cognitive and achievement assessment results for the same hypothetical student. They were asked three follow up questions about the case, including the likelihood that the student met criteria for a SLD, their confidence in their decision, and the similarity of the case to other students they have encountered.

The current study used a 2x2 factorial design with two levels for referral concern and assessment data. The referral concern was manipulated to be representative of a student with SLD or representative of a student with general academic problems. In the first condition, the referral concern included a statement that suggested the student had a SLD; whereas in the second condition, the student was described without reference to SLD. The two assessment data conditions were either scores that are consistent with SLD or inconsistent with SLD. For the purpose of this study, scores were considered consistent with a classification of SLD if the



student demonstrated a significant discrepancy between IQ and achievement. See Appendix B for the vignette text.

In accordance with the aim of documenting school psychology trainees' formal training, knowledge, and attitudes toward clinical judgment and its limitations, participants responded to questions that assessed knowledge of clinical judgment heuristics and biases, attitudes toward clinical judgment heuristics and biases, and training regarding clinical judgment heuristics and biases. The questions in these three areas were scored as three instruments assessing knowledge, attitudes, and training. These questions were administered at the end of the survey to prevent biasing responses to the vignette. Participants also responded to demographic questions about their age, gender, race/ethnicity, year in graduate program, and type of program.

## **Measures**

### ***Demographic Questions***

The author wrote questions that assessed personal and professional characteristics, such as education and training, age, race, and gender. These questions are provided in Appendix C.

### ***Clinical Judgment Measures***

Rock (1994) developed the Clinical Judgment Survey of Mental Health Professionals to assess knowledge, attitudes, and beliefs of mental health practitioners in general. I constructed the knowledge of clinical judgment questions after reviewing a multiple-choice exam from a graduate course on cognitive and affective psychology (Fagley, 2005) and adapting items to include school psychology specific terminology. I also adapted many of Rock's (1994) items to assess attitudes toward clinical judgment heuristics and biases. Because few studies have explicitly measured knowledge of this construct, there are no known validated instruments for

use with school psychology students. The survey contains three scales to assess knowledge, attitudes, and training related to clinical reasoning and judgment errors. See Appendix D.

**Knowledge of Clinical Judgment Errors Scale.** A 10-item scale was used to assess trainees' familiarity with clinical judgment topics. Nine were multiple-choice questions with four response options (only one correct choice) and one question was a true-false item. Responses were scored as 1 or 0 for correct or incorrect responses, respectively. A total knowledge score was computed by adding the scores for each item. An internal consistency reliability coefficient was calculated ( $\alpha=.53$ ) and was found to be poor. A total scale score was computed, and individual item responses were analyzed.

**Attitudes toward Clinical Judgment Research Scale.** The eight-item attitude scale was modeled based on Rock's (1994) survey and wording was adapted to apply to school psychology. Attitudes toward clinical judgment were assessed using a 5-point Likert scale. Response options ranged from strongly agree (1) to strongly disagree (5). One item was reverse coded to prevent biased responding. An internal consistency reliability coefficient was calculated and was found to be adequate ( $\alpha=.75$ ).

**Training in Clinical Reasoning Scale.** Training was assessed by asking participants whether they had been exposed to clinical judgment limitations, cognitive biases/heuristics, and diagnostic reasoning in a required course. An internal consistency reliability coefficient was calculated ( $\alpha=.55$ ) and was found to be poor. Participants also noted which methods influence their decision making.

## **Design**

Experimental Vignette Methodology (EVM) has been used in psychology and education research to approximate real world scenarios (e.g. Huebner, 1991; Woodward, 2020). EVM uses

carefully crafted vignettes, which differ only in respect to the variables of interest, to engage participants with a decision-making task. This design can be constructed as within or between subjects, depending on whether participants view each vignette condition. EVM allows researchers to achieve a level of experimental control that cannot always be attained in the natural environment, thereby increasing internal validity (Aguinis & Bradley, 2014). EVM is a useful, often underutilized, approach for decision making research (Huebner, 1991).

Although the ecological validity of EVM has been criticized, Huebner (1991) suggested that there is ample evidence to support its use. Criticism of vignette research is based on the assertion that clinician's reserve judgment until accessing more information and, therefore, case vignettes do not mirror what happens in actual practice. Nonetheless, there is no evidence that clinician's sort data differently in real life (Dawes, 1986). Aguinis and Bradley (2014) provided recommendations to increase the ecological validity of EVM by simulating the contextual variables associated with the decision-making process. This study incorporates EVM into survey methodology and uses a between subjects 2 x 2 design. See Table 1 for a breakdown of the experimental design conditions.

## **Results**

### **Descriptive Statistics**

Prior to analyzing the data, survey responses were screened for inclusion criteria and incomplete data to determine the final sample. Of the total survey responses, 175 participants indicated that they were current school psychology students at the time of the survey administration (June-July 2020). These responses did not contain any missing data and were included in the final sample.

### ***Demographics***

The gender and racial breakdown of the sample was consistent with school psychologist demographic characteristics from the most recent NASP member survey (Walcott & Hyson, 2018). In the current sample, 86.3% of participants were female and 71.4% were White. Over half of the participants in the present study were enrolled in doctoral programs, as compared with 25.2% of participants in the NASP study who were doctoral level school psychologists (Walcott & Hyson, 2018). See Table 2 for sample characteristics.

### ***Knowledge of Clinical Judgment Errors***

Knowledge of clinical judgment errors was assessed via 10 items: nine multiple-choice items and one true-false item. These items were intended to survey knowledge of specific cognitive biases and heuristics, as applied to the practice of school psychology. This 10-item scale had poor internal consistency ( $\alpha = .53$ ). Therefore, in addition to the overall descriptive statistics, descriptive statistics are reported for each item separately. Participants' scale scores, reflecting overall knowledge of clinical judgment errors, ranged from 2-10 ( $M = 6.61$ ,  $SD = 1.81$ ), reflecting the number of items answered correctly. For items one, two, four, five, eight, nine, and 10, most participants correctly identified the clinical judgment topic. Less than half of the sample answered correctly for items three (44.6%), six (14.9%), and seven (44.0%). Item six, which assessed participants' belief that conducting a comprehensive assessment leads to more accurate integration of data, had the lowest correct response rate. See Table 3 for a breakdown by individual item.

### ***Attitudes toward Clinical Judgment Research***

Attitudes were assessed using eight multiple-choice items that were rated on a five-point Likert scale, with *strongly agree* corresponding to a score of 1. Consequently lower scores

indicated more positive attitudes. The average rating by a participant across all eight items was computed. The mean of these averages indicated participants had favorable attitudes toward applying clinical judgment training in school psychology ( $M = 1.78, SD = .40$ ). Item 5, which stated that literature suggests clinical judgment is often unreliable and inaccurate, had the lowest agreement rating ( $M = 3.06, SD = .81$ ). The attitude item means and overall mean are listed in Table 4.

### ***Training in Clinical Reasoning***

Participants answered three yes or no questions about types of training they had received. A response of yes was coded as one and a response of no was coded as 0. A total score was computed by summing the responses to the three items. Scores ranged from 0-3 ( $M = 2.05, SD = 1$ ). These items were found to be significantly correlated with each other and the three-item scale had poor internal consistency ( $\alpha = .55$ ). Most participants reported having training in at least one area of clinical judgment (90.9%), with diagnostic reasoning being the least common (56.0%). Training in diagnostic reasoning did not significantly differ based on the number of years in a school psychology program,  $t(173) = 1.72, p = .09$ , though this difference was trending toward significance. Whereas students who had received training in this area had completed an average of 3.45 years of their program ( $SD = 1.25$ ), students who had not received training in this area had completed an average of 3.12 years of their program ( $SD = 1.29$ ).

Participants reported using a variety of methods to determine SLD classification, with Response to Intervention as the most common (33.7%), followed by Patterns of Strengths and Weaknesses (29.7%), IQ-Achievement Discrepancy (26.9%), and Other (9.7%). Participants also reported using additional methods (e.g. personal experience, supervisor recommendations, research literature) to make classification decisions. Cognitive and academic assessment results

were the most commonly endorsed, followed by evidence from research literature, and supervisor recommendations based on their experience. See Table 5 for a detailed breakdown of training in clinical judgment and SLD identification.

### **Relationship between Knowledge, Attitudes, and Training**

Pearson correlations were computed to determine the relationships among the three clinical judgment scales. A total score was computed for the Knowledge of Clinical Judgment Errors and Training in Clinical Reasoning Scales, where higher scores indicated greater knowledge and training, respectively. Overall Attitudes toward Clinical Judgment Research was assessed using a mean score, where lower scores indicated more favorable attitudes. There was a significant, negative relationship between knowledge and attitudes ( $r = -.24$ ), such that greater knowledge of clinical judgment errors was associated with more favorable attitudes toward clinical judgment research. Other correlations between scales were non-significant (see Table 6).

### **ANOVA**

I conducted a 2 (Referral Concern) x 2 (Assessment Scores) ANOVA to determine the effects of referral concern wording and IQ-achievement discrepancy on likelihood of endorsing SLD classification. Levene's Test was non-significant,  $p = .18$ , indicating that the assumption of homogeneity of variance was not violated. There was a significant main effect of assessment score discrepancy on SLD classification,  $F(1, 171) = 4.43, p = .037$ . This finding indicates that perceived likelihood of the student qualifying for a SLD was higher for the large IQ-achievement discrepancy condition ( $M = 48.94, SD = 26.79$ ) than the small IQ-achievement discrepancy condition ( $M = 39.87, SD = 27.96$ ). As indexed by partial eta squared, assessment scores accounted for 2.5% of the variance in the likelihood rating, which represents a small effect. There was a non-significant main effect of referral concern wording on SLD classification,  $F(1,$

171) = 0.009,  $p > .05$ . This finding indicates that mean likelihood did not differ depending on whether the referral source mentioned concern about SLD ( $M = 44.42$ ,  $SD = 26.35$ ) or not ( $M=44.56$ ,  $SD=29.44$ ). The interaction effect was also non-significant,  $F(1, 171) = 0.271$ ,  $p > .05$ . See Table 7 for descriptive statistics for each vignette condition and Table 8 for a summary of the ANOVA results.

### **Additional Vignette Analyses**

#### ***Confidence***

Participants reported their level of confidence in their decision on a scale from 0-100. Mean confidence rating was 65.29% across conditions. A 2 (Referral Concern) x 2 (Assessment Scores) ANOVA yielded a non-significant effect on participants' confidence level in their decision. Participants' confidence level did not vary based on the wording of the referral concern,  $F(1, 171) = 0.292$ ,  $p > .05$  or the degree of discrepancy between IQ and achievement scores,  $F(1, 171) = 0.032$ ,  $p > .05$ .

#### ***Similarity***

Participants reported the degree of similarity between the student described in the vignette and students they had encountered in their training. Similarity was rated on a scale from 0-100, and the full range of possible scores was evidenced in this sample. Mean similarity rating was 58.90 % across conditions. Participant's similarity ratings did not significantly differ by condition,  $p > .05$ .

#### ***Pearson Correlations***

There was a significant, weak relationship between similarity and likelihood ( $r = .36$ ). Participants who rated the student as more similar to students they had previously encountered, were more likely to endorse a higher probability that the student met criteria for SLD. This

relationship was varied when broken down by assessment scores condition. When the scores indicated SLD, the relationship weakened, but remained significant. When the scores did not indicate SLD, the relationship strengthened to a moderate correlation, meaning participants in this condition relied on similarity even when it contradicted the assessment scores. The relationship between similarity and confidence was significant ( $r = .31$ ), such that as the degree of similarity between the hypothetical student and students that participants had encountered in their training increased, so did their confidence in their judgment. This study failed to find a significant relationship between confidence and likelihood ( $r = .03$ ). See Table 9 for a correlation matrix.

### **Discussion**

The hypothesis that students would have minimal exposure to topics within clinical judgment through their graduate programs was not supported (H1). Participants reported that limitations of clinical judgment and cognitive biases and heuristics were covered in their required coursework. The study did not assess the degree to which these topics were covered in depth or applied to school psychology practice. Although these findings are encouraging, more research is needed to understand which training models are most effective in safeguarding against cognitive errors. Moreover, training in diagnostic reasoning was less common than other forms of clinical judgment. Because school psychologists spend a significant amount of time in assessment related tasks (Larson & Choi, 2010), refining diagnostic reasoning should be prioritized.

When presented with knowledge-based definition and scenario questions, participants demonstrated knowledge of cognitive biases, heuristics, and clinical judgment. Participants were less familiar with actual prediction than with specific cognitive errors. Notably, most participants



reported the belief that conducting a comprehensive assessment leads to more accurate interpretation of data. Although this claim may appeal to school psychologists as having face validity, it has not been supported by research. Conducting a comprehensive assessment can lead to more instances of bias without a systematic or actuarial method for interpreting the results (e.g. Kahneman, 2011, p. 224-225; Kennedy et al., 1997). Possible explanations for this fallibility in interpreting results are the false notion that connections between data points represent meaningful relationships and the tendency to overestimate the likelihood of rare events based on an exceptional feature in the data (Kahneman, 2011, p. 224-227). The fact that most participants agreed with the superiority of interpreting comprehensive assessment with clinical judgment highlights that, despite having knowledge of cognitive errors, they are not aware of how bias may impact interpretation of assessment results. Awareness of biases that prevent systematic data collection and interpretation is a prerequisite for following best practices for assessment.

Participants demonstrated overall positive attitudes toward applying clinical judgment literature to school psychology. Trainees expressed that further training in training in clinical judgment limitations would enhance their capacity to perform the functions of a school psychologist. Although attitudes were favorable overall, most participants' agreement ratings were lower for the statement that the based on the literature, clinical judgment is often unreliable. As Kahneman (2011) explained, many psychologists have failed to accept that their judgment and predictive abilities in regard to clinical matters cannot outperform those of resulting from a combination of statistical cues. This finding suggests that school psychology students may not be familiar with the literature on the distinction between actuarial and clinical prediction methods,

which is consistent with the relatively lower knowledge score (44.6% correct) for the definition of actuarial prediction.

Exploratory analyses revealed a significant correlation between knowledge of clinical judgment errors and attitudes toward clinical judgment research. Correlations failed to indicate a significant relationship between knowledge of clinical judgment errors and training in clinical reasoning. These correlations suggest that training may not be sufficient to increase knowledge of clinical judgment errors. Or, this may reflect the limitations of the measures, both in terms of content as well as psychometric properties. However, when people have more knowledge of these constructs, they may express greater interest in learning about the applications of the research to clinical practice. Even though attitude response options ranged from 1-5, only three items obtained ratings across the entire range, item means ranged from 1.38 to 3.06, and the observed overall means ranged from 1-3. Similarly, the participants in the present sample endorsed high levels of knowledge, with an observed score range from 2-10. Because of the somewhat restricted range, the modest reliability, and number of items, the correlation may not capture the extent of the relationship between the two constructs and should be interpreted with caution.

The hypothesis that participants would be biased to interpret assessment findings consistent with the wording of the referral concern was partially supported (H2). Although there were no significant differences between the referral wording conditions, there was a lot of within condition variance. Participants' likelihood scores for identifying SLD ranged from 0 to 100 across conditions. When broken down by specific SLD identification method, the standard deviations remained large, indicating that participants using the same reported SLD identification method provided different estimates for the likelihood that the student met criteria for SLD.

Furthermore, the assessment scores in the SLD condition were designed to reflect a greater than 22-point difference between IQ and Achievement, a common cut point for determining SLD using an IQ-Achievement discrepancy formula. Nonetheless, participants in the large discrepancy condition who reported using an IQ-Achievement discrepancy model did not uniformly rate the probability of the student meeting criteria for SLD as high. This finding suggests that participants relied on factors in addition to the supplied data and did not adhere strictly to one SLD identification method. Correlational findings provide further support for this conclusion. There was a significant, weak correlation between similarity and likelihood of SLD. This relationship became stronger when looking exclusively at the participants in the non-SLD assessment scores condition. This relationship suggests that when the scores did not clearly indicate the presence of SLD, participants who believed the wording of the referral concern in the vignette to be more similar to students they had encountered based their decision in part on the representativeness heuristic. Their judgement that the student in the case was similar to past students, or a prototypical student, may have led them to minimize the contribution of the assessment scores and rate the likelihood that the student met criteria for SLD as higher, leading to an inaccurate judgment.

Despite correlational evidence of possible bias among trainees, this study failed to find a causal impact of referral concern wording on SLD classification. The current study findings are part of a trend from other recent dissertations which found that school psychology trainees and practitioners used valid indicators to make classification decisions (e.g. Gammelmo, 2015; Sadeh, 2016; Woodward, 2020). There are several possible explanations which should be considered to address the discrepancy between evidence of cognitive errors in published studies and failure to replicate those findings in unpublished dissertations. None of the dissertation

studies compared clinical to actuarial judgment. Participants in the dissertation samples used valid indicators for determining student classification at one time point. These studies failed to look at the reliability of the same clinician's judgment across time points, a situation in which actuarial methods are particularly advantageous. Additionally, although the sample in the current study reflected sufficient power to detect a medium effect size using ANOVA, the sample is still relatively small in comparison to the population of school psychology trainees from which it was drawn. Moreover, the breakdown of participants per condition was not evenly distributed, potentially affecting overall power. Because this is a study on cognitive error, any explanation that failed to consider the "belief in small numbers," or tendency to believe that a random sample mirrors the characteristics of the larger population from which it was drawn, would be remiss (Tversky & Kahneman, 1971). Because of the relatively few studies of clinical judgment in school psychology and small sample sizes, the samples may not be accurate representations of the larger populations of school psychology trainees and practitioners. More studies are needed to determine whether school psychologists' judgment is negatively impacted by cognitive bias. Lastly, the discrepant findings between peer-reviewed articles and unpublished dissertations may reflect a filing drawer problem, or the bias toward publishing only studies with significant findings in peer-reviewed journals (e.g., Rosenthal, 1979).

### **Limitations**

The initial survey request was sent in June 2020 which may have reduced the likelihood that program chairs received and forwarded the email to their students. The survey may not have reached a wide cross section of school psychology students. Students who received the link may have been less likely to participate in the survey over summer vacation, especially as the study occurred during the global Covid-19 pandemic. Although participants were asked initially if they

were a school psychology graduate student, the only options provided were *yes* or *no*. It is possible that recent graduates selected *no*. Participants did not complete other demographic information until the end of the survey and, therefore, participants who dropped out prior to completing the survey could not be verified as school psychology students and were not included in the final sample. A response option for recent graduates would have allowed for more accurate representation of participants' standing in their programs. Because the survey was administered from June-July, recent graduates may have been excluded. Nonetheless, based on the collected demographic information, the sample was comparable to NASP membership in terms of race and gender.

There are several limitations that may affect interpretation of the findings. The versions of the vignette were designed to have similar sentence length and detail across conditions. The only difference between the conditions was whether the teacher expressed general concern about the student or specific concern of a SLD. Although this study did not find a direct indication that participants were biased by the referral information, the vignette conditions may have been too similar. The assessment score profiles were intended to be similarly designed to control for the effect of overall FSIQ. The assessment score conditions may have been too similar, or lacked detail, because mean SLD likelihood scores for both conditions were below 50%, indicating participants did not view a classification of SLD as probable in either condition.

Additionally, in a naturally occurring school setting, classroom observation and interview data may also be available to supplement a referral concern. Whereas in a school setting, school psychologists can view multiple sources of data simultaneously, participants were not able to switch back and forth between the vignette and the assessment scores, which may reduce the ecological validity of the study. Participants were asked to rate the similarity of the student

description to students they had encountered as another ecological validity measure. Responses were recorded on a “slider” scale from 0-100 and the question rating was initially set to 74 as default setting, rather than zero, and participants had to adjust their response accordingly.

Reported similarity ratings may have been positively skewed as a result.

### **Future Directions**

This was the first study that directly assessed knowledge of clinical judgment errors among school psychology trainees. Additional studies are needed to refine and validate the measures used in this study to more accurately capture school psychology trainee’s knowledge of specific domains within the clinical judgment literature. The clinical judgment scales used in this survey should be further refined and validated with samples of school psychologists and adapted based on their feedback and validation results.

A natural extension of the present study would evaluate the effects of educating students on specific clinical judgment topics and their limitations. Participants in the present study demonstrated the ability to recognize cognitive errors in common school-based scenarios. Although cognitive errors have been studied for decades, there are few validated ways to correct them or prevent them in practice. Some have suggested that training in cognitive errors could be one path toward improving clinical judgment (e.g. Andrews & Syeda, 2017; Lilienfeld et al, 2012). Training, although a theoretically compelling first step, may come with potential unintended consequences. Nisbett & Ross (1980) suggested that people may learn to recognize their errors, but despair about their ability to correct them. Future studies should evaluate the effects of providing direct education to school psychology students, in the form of a course or brief workshop.

This study adds to the literature indicating that SLD identification varies by practitioner, reducing inter-rater reliability. In order to be a clinically meaningful diagnosis, a greater degree of consensus on what constitutes a diagnosis is necessary. In the current study, assessment score results, particularly score discrepancy, informed participants' decision. Nonetheless, participants did not differentiate between the conditions, as the likelihood ratings were below 50% for both conditions, meaning that neither condition was representative of a student with SLD. One of the barriers to accurate identification of SLD is the lack of agreement on the SLD criteria and best practice for identification. Although school psychological practice has changed considerably in the past decades, the federal definition of SLD has remained consistent since it was first introduced in 1970. Future research is needed to refine the SLD construct.

### **Impact on School Psychology Training**

Evidence from previous studies suggests graduate students and experienced practitioners engage in biased decision making (e.g. Dumont, 1993). Some findings indicate that psychology students are more likely than experienced practitioners to make errors in their reasoning (Spengler et al., 2015; Waltman et al., 2013). These biases affect the clinician's ability to accurately make diagnoses, select treatment options, and provide evidence-based services. Although participants in the present study were not biased by whether the teacher stated a concern about SLD, they demonstrated use of clinical judgment, rather than adhering to a specific SLD identification model. Because clinicians do not receive consistent, immediate feedback on their judgments, these biases are unlikely to correct themselves in the absence of training and policy changes. Psychology training programs are well positioned to increase awareness of clinical judgment errors and methods to reduce them.

More studies are needed to understand the extent to which incorporating clinical judgment limitations into coursework on data-based decision-making yields better knowledge of cognitive biases, and ultimately, reduces them. Teaching students the tools for recognizing biases may not be sufficient in counteracting their effects (Larrick, 2004). As programs design their training sequence in accordance with the NASP Practice Model, they should consider a variety of strategies for improving decision-making. Larrick (2004) provided an overview of cognitive, motivational, and technological strategies that have been studied in the context of counteracting the negative effects of cognitive errors, or “debiasing.” Courses should incorporate cognitive strategies, which have been shown to be effective in reducing several types of biases, such as “consider the opposite,” or teaching students to systematically identify reasons that support an alternative hypothesis (Arkes, 1991). Cognitive strategies can be further divided into training on the application of statistical concepts to applied settings and training on the cognitive processes that result in error when left uncorrected (Larrick, 2004). Motivational strategies, such as providing incentives for following normative rules, may be less effective for students, as they do not have the knowledge of the normative rules. Technological strategies, which target the environmental context in which decisions are made, are particularly relevant to school psychology. School psychologists often make decisions as part of multi-disciplinary special education teams, but school psychology training programs rarely interact with these related disciplines. Whereas a substantial body of literature, albeit dated, exists on the cognitive errors that affect school psychologists’ judgment, more research is needed to identify suitable ways to reduce bias.

Given the low level of interest in clinical judgment among researchers, best practices for training school psychology students in safeguards against these errors have not been developed



(Watkins, 2009). Watkins (2009) identified preliminary recommendations for minimizing clinical judgment errors among school psychologists. According to his model, school psychologists should study the literature on clinical judgment, as well as developments in the field of school psychology to become informed practitioners. Despite being informed, school psychologists should avoid conflating experience with expertise; expertise comes from applying rules to make decisions and receiving immediate feedback and error correction. In addition to acquiring knowledge, practitioners should use methods that increase consistency among practitioners and across time points. In the case of school psychologists, decision making may include checklists, computerized scoring for assessments, and actuarial models formed from key predictors. Lastly, school psychologists should approach their practice from a scientific lens and use falsifiable, working hypotheses when decision rules are not available (Watkins, 2009).

School psychology trainees need models of how to effectively use clinical reasoning to interpret the science behind assessment and intervention. To date, this process has not been well publicized for consumers (e.g. students, school personnel) and trainees (e.g. school psychology students). Because of the absence of published literature on training in limitations of clinical judgment, school psychology students may struggle to translate knowledge of assessment and intervention using science-based clinical decision making (Andrews & Syeda, 2017). Wilcox and Schroeder (2015) recommend explicit training in clinical reasoning as a safeguard against errors.

Lastly, SLD classification, specifically, and special education classification more broadly is determined through a combination of methods, which were reported by participants in this study. Trainees in this study reported using different identification methods, as well as sources of data to inform their decisions. This lack of uniformity within special education classification can lead to negative consequences for students, such as removal of services when transferring

between school districts. School psychology would benefit from clear guidelines for conducting a multi-source evaluation and determining SLD classification to aid clinical decision making.

## References

- Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods, 17*(4), 351-371. doi:10.1177/1094428114547952
- Andrews, J.J.W. & Syeda, M.M. (2017). Clinical reasoning in school psychology: From assessment to intervention. *Canadian Journal of School Psychology, 32*(1), 3-15. <https://doi.org/10.1177/0829573516659935>
- Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin, 110*(3), 486. doi:10.1037/0033-2909.110.3.486
- Arkes, H. R. (1993). Some practical judgment and decision-making research. In N.J. Castellan (Ed.), *Individual and group decision making: Current issues* (pp. 3-18). Lawrence Erlbaum Associates.
- Arkes, H. R., & Hammond, K. R. (1986). *Judgment and decision making: An interdisciplinary reader*. Cambridge University Press.
- Aspel, A. D., Willis, W. G., & Faust, D. (1998). School psychologists' diagnostic decision-making processes: Objective-subjective discrepancies. *Journal of School Psychology, 36*(2), 137-149. [https://doi.org/10.1016/S0022-4405\(98\)00002-8](https://doi.org/10.1016/S0022-4405(98)00002-8)
- Baker, J. P., Detsky, A. S., Wesson, D. E., Wolman, S. L., Stewart, S., Whitewell, J., Langer, B., & Jeejeebhoy, K. N. (1982). Nutritional assessment: A comparison of clinical judgment and objective measurements. *New England Journal of Medicine, 306*(16), 969-972. doi:10.1056/NEJM198204223061606

- Baker, N. (2000). Referral information and test observation effects on the classification decision of school psychologists. [Unpublished doctoral dissertation]. Rutgers, The State University of New Jersey.
- Benson, N. F., Maki, K. E., Floyd, R. G., Eckert, T. L., Kranzler, J. H., & Fefer, S. A. (2020). A national survey of school psychologists' practices in identifying specific learning disabilities. *School psychology, 35*(2), 146–157. <https://doi.org/10.1037/spq0000344>
- Bowes, S.M. Ammirati, R.J., Costello, T.H., Basterfield, C., & Lilienfeld, S.O. (2020). Cognitive biases, heuristics, and logical fallacies in clinical practice: A brief field guide for practicing clinicians and supervisors. *Professional Psychology: Research and Practice*. Advance online publication.
- Braden, J. P., & Shaw, S. R. (2009). Intervention validity of cognitive assessment: Knowns, unknowables, and unknowns. *Assessment for Effective Intervention, 34*(2), 106-115. <https://doi.org/10.1177/1534508407313013>
- Clemens, N. H., Keller-Margulis, M. A., Scholten, T., & Yoon, M. (2016). Screening assessment within a multi-tiered system of support: Current practices, advances, and next steps. In S.R. Jimerson, M.K. Burns, & A.M. VanDerHeyden (Eds.) *Handbook of response to intervention* (pp. 187-213). Springer.
- Dawes, R. M. (1986). Representative thinking in clinical judgment. *Clinical Psychology Review, 6*(5), 425–441. [https://doi.org/10.1016/0272-7358\(86\)90030-9](https://doi.org/10.1016/0272-7358(86)90030-9)
- Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*(4899), 1668-1674. <https://doi.org/10.1126/science.2648573>
- Della Toffalo, D. A., & Pedersen, J. A. (2005). The effect of a psychiatric diagnosis on school psychologists' special education eligibility decisions regarding emotional disturbance.

*Journal of Emotional and Behavioral Disorders*, 13(1), 50-63.

<https://doi.org/10.1177/10634266050130010501>

Dumont, F. (1993). Inferential heuristics in clinical problem formulation: Selective review of their strengths and weaknesses. *Professional Psychology: Research and Practice*, 24(2), 196-205. <https://doi.org/10.1037/0735-7028.24.2.196>

Fagley, N.S. (2005). *Cognitive and Affective Psychology* [Course Exam]. Sakai. [Sakai.rutgers.edu/portal](http://Sakai.rutgers.edu/portal).

Fagley, N. S. (1988). Judgmental heuristics: Implications for the decision making of school psychologists. *School Psychology Review*, 17(2), 311-321.

Fagley, N.S., Miller, P.M., & Jones, R.N. (1999). The effect of positive or negative frame on the choices of students in school psychology and educational administration, *School Psychology Quarterly*, 14(2), 148-162. <https://doi.org/10.1037/h0089002>

Flanagan, D.P. (2017). *Case study applications of the WISC-V in cross-battery assessment and SLD using X-BASS* [PowerPoint Slides]. Retrieved from: <https://downloads.pearsonclinical.com/videos/092717-wisc-v/WISC-V-and-X-BASS-Webinar-Handout-092917.pdf>

Fletcher, J. M., & Miciak, J. (2017). Comprehensive cognitive assessments are not necessary for the identification and treatment of learning disabilities. *Archives of Clinical Neuropsychology*, 32(1), 2-7. <https://doi.org/10.1093/arclin/acw103>

Forman, S. G., Fagley, N. S., Steiner, D. D., & Schneider, K. (2009). Teaching evidence-based interventions: Perceptions of influences on use in professional practice in school psychology. *Training and Education in Professional Psychology*, 3(4), 226-232. <https://doi.org/10.1037/a0016874>

- Gammelmo, M. M. (2015). *Validity of diagnostic decision making in a school psychology training clinic* [Unpublished doctoral dissertation]. Penn State University.
- Gigerenzer, G. E., & Gaissmaier, W. (2011). Heuristic decision making. *The Annual Review of Psychology*, 62, 451-482. doi: 10.1146/annurev-psych-120709-145346
- Gnys, J. A., Willis, W. G., & Faust, D. (1995). School psychologists' diagnoses of learning disabilities: A study of illusory correlation. *Journal of School Psychology*, 33(1), 59-73. [https://doi.org/10.1016/0022-4405\(94\)00023-2](https://doi.org/10.1016/0022-4405(94)00023-2)
- Greenwood, C. R., & Kim, J. M. (2012). Response to intervention (RTI) services: An ecobehavioral perspective. *Journal of Educational and Psychological Consultation*, 22(1-2), 79-105. doi:10.1080/10474412.2011.649648
- Hale, J. B., Kaufman, A., Naglieri, J. A., & Kavale, K. A. (2006). Implementation of IDEA: Integrating response to intervention and cognitive assessment methods. *Psychology in the Schools*, 43(7), 753-770. <https://doi.org/10.1002/pits.20186>
- Huebner, E. S. (1991). Bias in special education decisions: The contribution of analogue research. *School Psychology Quarterly*, 6(1), 50-65. <https://doi.org/10.1037/h0088240>
- Individuals with Disabilities Education Act, 20 U.S.C. § 1401 (2004).
- Jordan, J.S., Harvey, J.H., & Weary, G. (1988). Attributional biases in clinical decision making. In D. Turk & P. Salovey (Eds.), *Reasoning, inference, and judgment in clinical psychology* (pp. 99-106). Free Press.
- Jordan, S. (2019). *The use of clinical judgment in differentiating symptoms of autism spectrum disorder from those of other childhood conditions: A delphi study* [Unpublished doctoral dissertation]. University of Denver.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

- Kennedy, M. L., Willis, W. G., & Faust, D. (1997). The base-rate fallacy in school psychology. *Journal of Psychoeducational Assessment, 15*(4), 292-307.  
<https://doi.org/10.1177/073428299701500401>
- Kleinmuntz, B. (1990). Why we still use our heads instead of formulas: Toward an integrative approach. *Psychological Bulletin, 107*(3), 296-310. <https://doi.org/10.1037/0033-2909.107.3.296>
- Larrick, R.P. (2004). Debiasing. In D.J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 273-285). Blackwell Publishing.
- Larson, J. P., & Choi, H. S. (2010). The effect of university training and educational legislation on the role and function of school psychologists. *Journal of Applied School Psychology, 26*(2), 97-114. <https://doi.org/10.1080/15377900903433336>
- Lilienfeld, S. O., Ammirati, R., & David, M. (2012). Distinguishing science from pseudoscience in school psychology: Science and scientific thinking as safeguards against human error. *Journal of School Psychology, 50*(1), 7-36. <https://doi.org/10.1016/j.jsp.2011.09.006>
- Lilienfeld, S. O., & Lynn, S. J. (2014). Errors/biases in clinical decision making. *The Encyclopedia of Clinical Psychology, 1*-9. <https://doi.org/10.1002/9781118625392.wbecp567>
- Maki, K. E., Burns, M. K., & Sullivan, A. (2017). Learning disability identification consistency: The impact of methodology and student evaluation data. *School Psychology Quarterly, 32*(2), 254-267. <https://doi.org/10.1037/spq0000165>
- McGill, R.J., Kennedy, K.S., & Busse, R.T. (2016). Data-based decision making in school counseling: Utilizing multiple single-case indicators to evaluate interventions. *The Practitioner Scholar: Journal of Counseling and Professional Psychology, 5*(1), 104-121.

- Meehl, P. E. (1986). Causes and effects of my disturbing little book. *Journal of Personality Assessment, 50*(3), 370-375. [https://doi.org/10.1207/s15327752jpa5003\\_6](https://doi.org/10.1207/s15327752jpa5003_6)
- National Association of School Psychologists ([NASP] 2010). Model for comprehensive and integrated school psychological services.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175-220. <https://doi.org/10.1037/1089-2680.2.2.175>
- Nisbett, R. E., & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgment. Prentice Hall.
- Pendergast, L. L., Youngstrom, E. A., Ruan-Lu, L., & Beysolow, D. (2018). The nomogram: a decision-making tool for practitioners using multitiered systems of support. *School Psychology Review, 47*, 345–359. <https://doi.org/10.17105/SPR-2017-0097.V47-4>.
- Reason, J. (1990). *Human error*. Cambridge University Press.
- Rock, D. L. (1994). Clinical judgment survey of mental-health professionals: I. An assessment of opinions, ratings, and knowledge. *Journal of Clinical Psychology, 50*(6), 941-950. [https://doi.org/10.1002/1097-4679\(199411\)50:6<941::AID-JCLP2270500620>3.0.CO;2-7](https://doi.org/10.1002/1097-4679(199411)50:6<941::AID-JCLP2270500620>3.0.CO;2-7)
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638-641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Sadeh, S. (2016). *School psychologists' decision making in evaluations for emotional disturbance* [Unpublished doctoral dissertation]. University of Minnesota.
- Spengler, P. M., & Pilipis, L. A. (2015). A comprehensive meta-reanalysis of the robustness of the experience-accuracy effect in clinical judgment. *Journal of Counseling Psychology, 62*(3), 360-378. doi: 10.1037/cou0000065



- Spengler, P. M., White, M. J., Ægisdóttir, S., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C.N., Lampropoulos, G.K., Walker, B.S., Cohen, G.R., & Rush, J. D. (2009). The meta-analysis of clinical judgment project: Effects of experience on judgment accuracy. *The Counseling Psychologist, 37*(3), 350-399.  
<https://doi.org/10.1177/0011000006295149>
- Stoddard, C. P. (2017). *Progress monitoring in school-based counseling: A survey of school psychologists' current practices and awareness of feedback systems* [Unpublished doctoral dissertation]. William James College.
- Tanner, C. A. (2006). Thinking like a nurse: A research-based model of clinical judgment in nursing. *Journal of Nursing Education, 45*(6), 204-211.  
<https://doi.org/10.3928/01484834-20060601-04>
- Turk, D. C., & Salovey, P. (1985). Cognitive structures, cognitive processes, and cognitive-behavior modification: II. Judgments and inferences of the clinician. *Cognitive Therapy and Research, 9*(1), 19-33. <https://doi.org/10.1007/BF01178748>
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76*(2), 105. <https://doi.org/10.1037/h0031322>
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131. doi:10.1126/science.185.4157.1124
- Walcott, C. M., & Hyson, D. (2018). Results from the NASP 2015 membership survey, part one: Demographics and employment conditions. *NASP Research Reports, 3*(1), 1-17.
- Waltman, S. H., Williams, A., & Christiansen, L. R. (2013). Comparing student clinician and licensed psychologist clinical judgment. *Training and Education in Professional Psychology, 7*(1), 33. <https://doi.org/10.1037/a0032072>

- Watkins, M.W. (2009). Errors in diagnostic decision making and clinical judgment. In T.B. Gutkin & C.R. Reynolds (Eds.), *Handbook of school psychology 4<sup>th</sup> edition* (pp. 210-229). Wiley.
- Wilcox, G., & Schroeder, M. (2015). What comes before report writing? Attending to clinical reasoning and thinking errors in school psychology. *Journal of Psychoeducational Assessment, 33*(7), 652-661. <https://doi.org/10.1177/0734282914562212>
- Woodward, M.L. (2020). *The role of data ambiguity on confirmation bias in psychoeducational assessment* [Unpublished doctoral dissertation]. Northern Arizona University.
- Ysseldyke, J. E., & Algozzine, B. (1981). Diagnostic classification decisions as a function of referral information. *The Journal of Special Education, 15*(4), 429-435. <https://doi.org/10.1177/002246698101500404>

Table 1

*Diagram of the Experimental Design Conditions*

|                         |                                                                     | <b>Assessment scores</b>         |                                  |
|-------------------------|---------------------------------------------------------------------|----------------------------------|----------------------------------|
|                         |                                                                     | Large IQ-achievement discrepancy | Small IQ-achievement discrepancy |
| <b>Referral Concern</b> | Teacher states concern that student has a learning disability       | A<br>( <i>n</i> = 51)            | C<br>( <i>n</i> = 47)            |
|                         | Teacher states general concern that student is falling behind peers | B<br>( <i>n</i> = 38)            | D<br>( <i>n</i> = 39)            |

*Letters A-D correspond to four conditions of the 2x2 design.*

Table 2

*Demographics*

| <b>Variable</b>                                  | <i>n</i> | %    | <i>M</i> | <i>SD</i> | Range |
|--------------------------------------------------|----------|------|----------|-----------|-------|
| <b>Age</b>                                       | 174      | 99.4 | 27.7     | 5.47      | 22-50 |
| <b>Gender</b>                                    |          |      |          |           |       |
| <b>Female</b>                                    | 151      | 86.3 |          |           |       |
| <b>Male</b>                                      | 20       | 11.4 |          |           |       |
| <b>Non-Binary</b>                                | 1        | .6   |          |           |       |
| <b>Prefer Not to Say</b>                         | 3        | 1.7  |          |           |       |
| <b>Race/Ethnicity</b>                            |          |      |          |           |       |
| <b>American Indian or Alaska Native</b>          | 2        | 1.10 |          |           |       |
| <b>Asian</b>                                     |          |      |          |           |       |
| <b>Black or African American</b>                 | 10       | 5.71 |          |           |       |
| <b>American Hispanic or Latino</b>               | 4        | 2.29 |          |           |       |
| <b>Hispanic or Latino</b>                        | 13       | 7.43 |          |           |       |
| <b>Native Hawaiian or Other Pacific Islander</b> | 0        | 0    |          |           |       |
| <b>White</b>                                     | 125      | 71.4 |          |           |       |
| <b>Multi-racial</b>                              | 16       | 9.14 |          |           |       |
| <b>Prefer Not to Say</b>                         | 5        | 2.9  |          |           |       |
| <b>Program Year</b>                              |          |      |          |           |       |
| <b>Less than one</b>                             | 6        | 3.4  |          |           |       |
| <b>One</b>                                       | 44       | 25.1 |          |           |       |
| <b>Two</b>                                       | 61       | 34.9 |          |           |       |
| <b>Three</b>                                     | 32       | 18.3 |          |           |       |
| <b>Four</b>                                      | 22       | 12.6 |          |           |       |
| <b>Five</b>                                      | 7        | 4.0  |          |           |       |
| <b>Six or more</b>                               | 3        | 1.7  |          |           |       |
| <b>Program Type</b>                              |          |      |          |           |       |
| <b>Masters</b>                                   | 26       | 14.9 |          |           |       |
| <b>Specialist</b>                                | 58       | 33.1 |          |           |       |
| <b>Doctoral</b>                                  | 91       | 52.0 |          |           |       |

Table 3

*Knowledge of Clinical Judgment Errors (N = 175)*

| <b>Item</b>                        | <b>% Correct</b> | <b>Coef. Alpha if item deleted</b> | <b><i>M</i></b> | <b><i>SD</i></b> |
|------------------------------------|------------------|------------------------------------|-----------------|------------------|
| <b>1 Representativeness</b>        | 82.3             | .526                               |                 |                  |
| <b>2 Confirmation Bias</b>         | 93.1             | .503                               |                 |                  |
| <b>3 Actuarial Prediction</b>      | 44.6             | .499                               |                 |                  |
| <b>4 Confirmation Bias Text</b>    | 74.3             | .499                               |                 |                  |
| <b>5 Sunk Cost</b>                 | 67.4             | .479                               |                 |                  |
| <b>6 Comprehensive Assessment</b>  | 14.9             | .507                               |                 |                  |
| <b>7 Base-rate Neglect</b>         | 44.0             | .475                               |                 |                  |
| <b>8 Actuarial Prediction Text</b> | 78.9             | .509                               |                 |                  |
| <b>9 Hindsight Bias</b>            | 80.0             | .522                               |                 |                  |
| <b>10 Representativeness Text</b>  | 81.1             | .472                               |                 |                  |
| <b>Knowledge Total Score</b>       | -                | -                                  | 6.61            | 1.81             |

*Items 1-5 and 7-10 were multiple-choice questions with one correct answer and three incorrect answers. Item 6 was a true or false item.*

Table 4

*Attitudes toward Clinical Judgment Research (N = 175)*

| <b>Item</b>                                                                                                                                                                                                | <i>M</i> | <i>SD</i> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|-----------|
| <b>1 To increase the relevance of judgment research for clinical practice, researchers should study clinical judgment under conditions more similar to actual practice.</b>                                | 1.72     | .70       |
| <b>2 Clinical judgment is a topic that warrants research.</b>                                                                                                                                              | 1.58     | .65       |
| <b>3 Research on clinical judgment is not applicable to the types of activities that school psychologists conduct.</b>                                                                                     | 1.78     | .75       |
| <b>4 Findings from clinical judgment research could have meaningful implications for school psychology practice.</b>                                                                                       | 1.61     | .63       |
| <b>5 Considering the clinical judgment literature in psychology it is appropriate to conclude that clinicians' judgments are unreliable and inaccurate more often than they are reliable and accurate.</b> | 3.06     | .81       |
| <b>6 Studying clinical judgment should be part of school psychology training.</b>                                                                                                                          | 1.66     | .66       |
| <b>7 Understanding cognitive biases would make me a more effective school psychologist.</b>                                                                                                                | 1.47     | .61       |
| <b>8 Training on biases in human judgment and decision making and how to avoid them should be part of graduate programs in school psychology.</b>                                                          | 1.37     | .53       |
| <b>Average Score</b>                                                                                                                                                                                       | 1.78     | .40       |

*Items were rated on a Likert scale. Response options were '1-Strongly Agree' '2-Agree' '3-Neither Agree nor Disagree' '4-Disagree' and '5-Strongly Disagree.' Item three was reverse scored.*

Table 5

*Training in Clinical Reasoning*

| <b>Item</b>                                                    | <i>n</i> | <i>%</i> | <i>M</i> | <i>SD</i> |
|----------------------------------------------------------------|----------|----------|----------|-----------|
| <b>1 Received training in limitations of clinical judgment</b> | 134      | 76.6     |          |           |
| <b>2 Received training in cognitive biases/heuristics</b>      | 127      | 72.6     |          |           |
| <b>3 Received training in diagnostic reasoning</b>             | 98       | 56.0     |          |           |
| <b>Received training in at least one area</b>                  | 159      | 90.9     |          |           |
| <b>Total Score</b>                                             |          |          | 2.05     | 1         |
| <b>SLD identification method</b>                               | -        | -        |          |           |
| <b>IQ-Achievement Discrepancy</b>                              | 47       | 26.9     |          |           |
| <b>Patterns of Strengths and Weaknesses</b>                    | 52       | 29.7     |          |           |
| <b>Response to Intervention</b>                                | 59       | 33.7     |          |           |
| <b>Sources of information used for decision making</b>         | 135      | 77.1     |          |           |
| <b>Research literature</b>                                     | 86       | 49.1     |          |           |
| <b>Personal experience</b>                                     | 171      | 97.7     |          |           |
| <b>Cognitive and academic assessment</b>                       | 128      | 73.1     |          |           |
| <b>School district policies</b>                                | 134      | 76.6     |          |           |
| <b>Supervisor recommendations based on personal experience</b> | 99       | 56.6     |          |           |
| <b>Supervisor recommendations based on research literature</b> | 135      | 77.1     |          |           |
| <b>Other</b>                                                   | 40       | 22.9     |          |           |

Table 6

*Correlations between Clinical Judgment Measures*

| <b>Variables</b> | Knowledge | Attitudes | Training |
|------------------|-----------|-----------|----------|
| <b>Knowledge</b> | -         |           |          |
| <b>Attitudes</b> | -.238**   | -         |          |
| <b>Training</b>  | .072      | -.042     | -        |

*\*\*Significant at the .01 level. Note. Lower attitudes scores indicate more positive attitudes.*



Table 7

*Descriptive statistics for vignette conditions*

| <b>Referral<br/>Concern</b> | Large IQ-Achievement Discrepancy |           |          | Small IQ-Achievement Discrepancy |           |          |
|-----------------------------|----------------------------------|-----------|----------|----------------------------------|-----------|----------|
|                             | <i>M</i>                         | <i>SD</i> | <i>n</i> | <i>M</i>                         | <i>SD</i> | <i>n</i> |
| <b>LD</b>                   | 49.71                            | 25.35     | 51       | 38.70                            | 26.47     | 47       |
| <b>Non-LD</b>               | 47.92                            | 28.93     | 38       | 41.28                            | 29.93     | 39       |

Table 8

*ANOVA Summary Table*

| <b>Source</b>                                                  | <i>df</i> | MS      | <i>F</i> | <i>p</i> | Partial $\eta^2$ |
|----------------------------------------------------------------|-----------|---------|----------|----------|------------------|
| <b>Assessment Score<br/>Discrepancy</b>                        | 1         | 3352.68 | 4.431    | .037     | .025             |
| <b>Referral Concern</b>                                        | 1         | 6.809   | 0.009    | .925     | .001             |
| <b>Assessment Score<br/>Discrepancy x Referral<br/>Concern</b> | 1         | 205.201 | 0.271    | .603     | .002             |
| <b>Within Cell</b>                                             | 171       | 756.661 |          |          |                  |
| <b>Total</b>                                                   | 175       |         |          |          |                  |

*Note.* Dependent variable was perceived likelihood the child had a specific learning disability.

Table 9

*Correlations for vignette variables*

| <b>Variables</b>  | Likelihood | Confidence | Similarity |
|-------------------|------------|------------|------------|
| <b>Likelihood</b> | -          |            |            |
| <b>Confidence</b> | .311**     | -          |            |
| <b>Similarity</b> | .358**     | .034       | -          |

*\*\*Significant at the .01 level*

## Appendix A

**Initial Contact Letter**

Dear program director,

My name is Dina Siegel and I am doctoral student in the APA and NASP approved School Psychology program at the Graduate School of Applied and Professional Psychology at Rutgers University.

I am conducting a study for my dissertation that examines how school psychology graduate students think about – and would respond to – complex professional practice scenarios.

I need your help. I would greatly appreciate if you could copy and paste the message below to the school psychology students in your department.

Would you please (1) forward this email to your students, (2) reply to let me know that you have done so, and (3) if you know it, please share with me the approximate number of students on your electronic mailing list who will receive the email?

Thank you for your help!

Dina

Dina Siegel, PsyM

Doctoral Candidate

Graduate School of Applied and Professional Psychology

Rutgers, the State University of New Jersey

-----  
Dear fellow graduate student,

I am a fourth-year school psychology graduate student at the Graduate School of Applied and Professional Psychology at Rutgers, the State University of New Jersey.

I am conducting a study for my dissertation that examines how school psychology graduate students think about – and would respond to – complex professional practice scenarios.

The survey presents you with a short vignette about a hypothetical student, and asks you to imagine that you are the school psychologist responsible for working with that student.

The multiple-choice survey should take less than fifteen minutes to complete and is approved by the Rutgers University Institutional Review Board.

Would you kindly help me out by accessing the survey using the following link?[https://rutgers.ca1.qualtrics.com/jfe/form/SV\\_2t55gFiqdAcl3wN](https://rutgers.ca1.qualtrics.com/jfe/form/SV_2t55gFiqdAcl3wN)

Feel free to reach out to myself or my dissertation chair, Dr. Nancy Fagley, with any questions.

[Dina.siegel@rutgers.edu](mailto:Dina.siegel@rutgers.edu) , [Fagley@gsapp.rutgers.edu](mailto:Fagley@gsapp.rutgers.edu)

Thank you for your consideration,

Dina

Dina Siegel

Doctoral Candidate

Graduate School of Applied and Professional Psychology

Rutgers, the State University of New Jersey

### **Follow-up Contact Letter**

Dear School Psychology Program Director:

Thank you for sending the link to my dissertation survey to your school psychology students.

This is a follow-up email regarding my survey on how school psychology students view complex professional practice scenarios.

To help remind those who have not yet completed the survey, would you please send the message below to your enrolled students?

Thank you for your help!

Best,

Dina

Dina Siegel

Doctoral Candidate

Graduate School of Applied and Professional Psychology

Rutgers, the State University of New Jersey

-----  
Dear fellow graduate student,

This is a reminder notice about my dissertation survey. If you have already completed it, I want to let you know that I really appreciate it. If you haven't completed it yet, it isn't too late to help. I am a fourth-year school psychology graduate student at the Graduate School of Applied and Professional Psychology at Rutgers, the State University of New Jersey.

I am conducting a study for my dissertation that examines how school psychology graduate students think about – and would respond to – complex professional practice scenarios.

The survey presents you with a short vignette about a hypothetical student, and asks you to imagine that you are the school psychologist responsible for working with that student.

The multiple-choice survey should take less than fifteen minutes to complete and is approved by the Rutgers University Institutional Review Board.

Below is a link to the online survey:

[https://rutgers.ca1.qualtrics.com/Q/EditSection/Blocks?ContextSurveyID=SV\\_2t55gFiqdAcl3wN](https://rutgers.ca1.qualtrics.com/Q/EditSection/Blocks?ContextSurveyID=SV_2t55gFiqdAcl3wN)

If you have any questions, please feel free to contact me at [dina.siegel@rutgers.edu](mailto:dina.siegel@rutgers.edu)

Thank you in advance for your consideration!

Dina

Dina Siegel, PsyM

Doctoral Candidate

Graduate School of Applied and Professional Psychology

Rutgers, the State University of New Jersey

## Appendix B

**Vignettes**

Please read the following case vignette. Imagine that you are a practicing school psychologist in a public school when reading the vignette.

***Vignette Version 1 - SLD condition***

Avery is a 10-year-old student who was referred for special education and related services by his teacher. He has been struggling in class for the duration of the school year. His teacher is concerned that he has a specific learning disability in reading. On the referral form, the teacher checked off 5/6 academic concerns and 1/6 behavior concerns. She wrote that he rarely participates in class discussions, has poor spelling, and frequently turns in incomplete assignments. The teacher noted that his class performance is well below the class average. Despite informal academic intervention over several months, Avery continues to fall behind his peers.

***Vignette Version 2 - Non-SLD Condition***

Avery is a 10-year-old student who was referred for special education and related services by his teacher. He has been struggling in class for the duration of the school year. His teacher is concerned that he is falling behind his peers. On the referral form, the teacher checked off 5/6 academic concerns and 1/6 behavior concerns. She wrote that he rarely participates in class discussions, has poor spelling, and frequently turns in incomplete assignments. The teacher noted that his class performance is well below the class average. Despite informal academic intervention over several months, Avery continues to fall behind his peers.

After receiving the referral from Avery's teacher, the special education team agreed to conduct a psychoeducational evaluation. Below is a summary table of his results from the evaluation.

*Assessment Score Profile 1 – Large IQ-Achievement Discrepancy*

| WISC-V Score Profile | WJ Achievement Score Profile |
|----------------------|------------------------------|
| FSIQ-105             | Broad Achievement-82         |
| VCI-82               | Broad Reading-79             |
| FRI-115              | Broad Math-86                |
| VSI-109              | Broad Writing-101            |
| WMI-111              |                              |
| PSI-102              |                              |

*Assessment Score Profile 2 - Small IQ-Achievement Discrepancy*

| WISC-V Score Profile | WJ Achievement Score Profile |
|----------------------|------------------------------|
| FSIQ-105             | Broad Achievement-95         |
| VCI-99               | Broad Reading-93             |
| FRI-102              | Broad Math-86                |
| VSI-109              | Broad Writing-101            |
| WMI-111              |                              |
| PSI-102              |                              |



## Appendix C

**Demographic Questions**

Are you a school psychology graduate student?

- Yes
- No

What type of graduate program do you attend?

- Masters
- Specialist
- Doctoral

How many years of your program have you completed?

- Less than one
- One
- Two
- Three
- Four
- Five
- Six+

What is your primary theoretical orientation?

- Behavioral
- Cognitive Behavioral
- Existential
- Family Systems
- Humanistic
- Integrative
- Psychoanalytic
- Psychodynamic
- Other

What percentage of your time do you spend engaged in the following activities per week:

- Administrative Tasks
- Assessment
- Case management
- Consultation
- Counseling
- Other

How old are you?

How do you identify your gender?

- Female
- Male
- Non-Binary
- Prefer Not to Say

Which of the following best represents your racial or ethnic identity? (Choose all that apply)

- American Indian or Alaska Native
- Asian
- Black or African American
- Pacific Islander or Native Hawaiian
- White
- Prefer Not to Say

## Appendix D

**Clinical Judgment Measures*****Knowledge of Clinical Judgment Errors Scale***

The representativeness heuristic is best defined as:

- The degree to which two individuals agree on a topic
- The degree to which a subjective probability is determined based on the similarity and number of salient features between a single case and a larger group
- The number of participants needed in a sample to match the characteristics of the larger population from which the sample is drawn
- The ability to substitute the definition of a word with its visual representation

Confirmation bias is best defined as:

- The tendency to seek and interpret new information in a manner that is consistent with prior hypotheses
- The need to surround oneself with others who share similar beliefs
- The likelihood of believing a position because it is most easily called to mind
- The tendency to evaluate data based on whether it can be explained by a single hypothesis

Actuarial prediction is best defined as:

- Predicting future events based on a person's clinical interpretation of multiple cues about a person's ability, environment, and personality
- Combining predictive cues in a statistical formula to determine the likelihood of an event
- Determining whether a decision will be more beneficial for an individual or an entire group
- Predicting future events based on a person's clinical interpretation of past occurrences

Malonda is a school psychologist in an urban school district. She recently attended a workshop on using Patterns of Strengths and Weaknesses to diagnose learning disabilities and implemented this method in her school. For a particular student, she noticed their test scores seemed to cluster

in the low average range. She interviewed the teacher and asked to see any tests where the student performed below average. Malonda is exhibiting:

- Best practice for assessment
- Deductive reasoning
- Availability heuristic
- Confirmation bias

An elementary school board is considering replacing their current ineffective reading program with a new curriculum. The new, evidence-based, reading intervention program costs \$5,000.

The principal is opposed to the purchase because the school already invested \$3,500 in a reading intervention which has not improved student reading ability. The principal seems to be exhibiting which bias?

- Conjunction
- Sunk cost
- Relevancy
- Availability

Getting a detailed, comprehensive history of a student ensures that the assessment results will be interpreted more accurately.

- True
- False

You are planning a school-wide bullying prevention program. Based on the news articles you have read, you believe there is a strong link between bullying and school shootings. Because rates of bullying are above 25% in this school, you believe the likelihood of a school shooting within the next few years is above 50%. This prediction is an example of:

- Base rate neglect
- Fundamental misattribution bias

- Hindsight bias
- Blind spot bias

Ms. Barnes wants to select two students from her 3rd grade class for demonstrating overall excellence in academics, sportsmanship, and diligence. Which of the following methods of selecting students exhibits the **least** bias?

- Ms. Barnes thinks of the students in her class and choose the two who seem most fitting
- Ms. Barnes reviews the students' records and chooses the two with the fewest behavior incidents
- Ms. Barnes asks the students in the class to vote on each student and chooses the two students who receive the most votes
- Ms. Barnes rank orders each student based on their grades and behavior records

After a high school student on your caseload drops out of school, you say to yourself, "I knew they were never going to graduate." This is an example of:

- Availability heuristic
- Selective perception
- Hindsight bias
- Illusory correlation

According to the **representativeness heuristic**, we judge whether someone is likely to commit suicide by:

- How closely the person fits our prototype of a suicidal person
- How easily we can recall instances of clients who have attempted suicide
- The degree to which the client frames neutral stimuli negatively
- Revising the base rate of suicide attempters based on the number of risk factors exhibited

*Attitudes toward Clinical Judgment Research Scale*

Please rate the extent to which you agree with the following statements:

To increase the relevance of judgment research for clinical practice, researchers should study clinical judgment under conditions more similar to actual practice.

- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

Clinical judgment is a topic that warrants research.

- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

Research on clinical judgment is not applicable to the types of activities that school psychologists conduct.

- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

Findings from clinical judgment research could have meaningful implications for school psychology practice.

- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree

- Strongly Disagree

Considering the clinical judgment literature in psychology it is appropriate to conclude that clinicians' judgments are unreliable and inaccurate more often than they are reliable and accurate.

- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

Studying clinical judgment should be part of school psychology training.

- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

Understanding cognitive biases would make me a more effective school psychologist.

- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

Training on biases in human judgment and decision making and how to avoid them should be part of graduate programs in school psychology.

- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

*Training in Clinical Reasoning Scale*

The following questions ask about some of your training/education experiences in your school psychology program.

Have you received training in the limitations of clinical judgment as part of your graduate program?

- Yes
- No

Was this training part of a required course?

- Yes
- No

Have you received training in cognitive biases/heuristics as part of your graduate program?

- Yes
- No

Was this training part of a required course?

- Yes
- No

Have you received training in refining diagnostic reasoning as part of your graduate program?

- Yes
- No

Was this training part of a required course?

- Yes
- No



What method do you typically use to determine specific learning disability?

- IQ-Achievement Discrepancy
- Patterns of Strengths and Weaknesses
- Response to Intervention
- Other

Please select each of the methods you use to determine students' need for special education:

- Evidence from research literature
- Personal experience working with similar students
- Results from cognitive and/or academic testing
- School district policies
- Supervisor recommendations based on their experience
- Supervisor recommendations based on research literature
- Other