HOW ACCURATELY CAN RELIGIOUS EDUCATORS PREDICT STUDENT ACHIEVEMENT?

by

Kyle Lyons

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Arts in Education, Curriculum and Instruction

Boise State University

December 2019

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the thesis submitted by

Kyle Lyons

Thesis Title: How Accurately Can Religious Educators Predict Student Achievement?

Date of Final Oral Examination: 25 October 2019

The following individuals read and discussed the thesis submitted by student Kyle Lyons, and they evaluated the student's presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Keith Thiede, Ph.D.	Chair, Supervisory Committee
Richard Osguthorpe, Ph.D.	Member, Supervisory Committee
Michele Carney, Ph.D.	Member, Supervisory Committee

The final reading approval of the thesis was granted by Keith Thiede, Ph.D., Chair of the Supervisory Committee. The thesis was approved by the Graduate College.

DEDICATION

I would like to dedicate this document to my eternal companion, Nicole Lyons. She is my rock and the love of my life. I would not be where I am today without her.

ACKNOWLEDGEMENTS

I would like to acknowledge my committee members, specifically, Keith Thiede, for their help in producing this thesis. I would like to also acknowledge Phil Kelly for his extraordinary teaching. I not only learned from what he taught, but also how he taught it.

ABSTRACT

Teachers make a variety of judgments as they teach. The accuracy of these judgments may influence instruction and student achievement. The present investigation examined (a) how accurately religious educators judge student learning, (b) what cues religious educators report using to judge student learning, and (c) how cue utilization affects the accuracy of judgments of student learning. The research in this study shows the accuracy of judgments for participating teachers is significantly lower than the average judgment accuracy reported in a recent review of teacher judgment literature (Südkamp et al., 2012). The cues participating teachers self-reported using for judging student learning fell into four categories: class performance, personal attributes, external factors, and class behavior. Judgment accuracy is greater for teachers who reported using cues related to class performance than for those who did not. Judgment accuracy is greater for those who did not report using personal attributes as a cue than for those who did. These results are explained in the context of the cue-utilization framework (Koriat, 1997).

Key words: Teacher Judgment Accuracy, Cue Utilization, Links between Cue Use and Accuracy

TABLE OF CONTENTS

DEDICATION iv

ACKNOWLEDGEMENTS v

ABSTRACT vi

LIST OF TABLES ix

LIST OF FIGURES x

CHAPTER 1: STATEMENT OF THE PROBLEM 1

Teacher Judgments in a Religious Education Context 2

CHAPTER 2: LITERATURE REVIEW 6

Measuring the Accuracy of Teacher Judgments 6

Why Teacher Judgments Matter 8

The Cue-Utilization Framework of Judgments 9

Cues Used to Judge Student Learning 10

Gender and Ethnicity 11

Student Behavior 11

Past Performance 12

Teacher Judgments in a Religious Education Context 13

CHAPTER 3: METHOD 15

Participants 15

Measures 16

Knowledge 16

Judgment Accuracy 16

Procedure 16

CHAPTER 4: RESULTS 18

Research Question 1 18

Research Question 2 18

Research Question 3 22

CHAPTER 5: DISCUSSION 27

REFERENCES 32

APPENDIX A 38

LIST OF TABLES

Table 1. Years of Teaching Experience for Teachers 15

Table 2. Individual Cues Classified in the Four Categories 20

Table 3. Correlation Matrix for Four Cue Categories 22

LIST OF FIGURES

Figure 1. Cue-utilization survey....17

Figure 2. Proportion of Teachers Reporting Use of Individual Cues....19

Figure 3. Proportion of Teachers Reporting Use of Cue Categories 21

Figure 4. Relative Accuracy between Teachers with High and Low use of Class Performance Cues....23

Figure 5. Relative Accuracy between Teachers with High and Low use of Personal Attributes Cues....24

Figure 6. Relative Accuracy between Teachers with High and Low use of External Factors Cues....25

Figure 7. Relative Accuracy between Teachers with High and Low use of Class Behavior Cues....26

CHAPTER 1: STATEMENT OF THE PROBLEM

Teachers make a variety of judgments as they teach. How accurately teachers judge student learning is important because these judgments guide subsequent instruction. Thus, more accurate judgments lead to more effective instruction, which leads to greater gains in student achievement (Thiede et al., 2018). In his landmark paper, Koriat (1997) introduced the cue-utilization framework, which identifies factors that affect the accuracy of judgments. The framework suggests judgment accuracy depends on the cues teachers use and the diagnosticity of those cues. Cue utilization refers to what the teachers use to make judgments. For example, they may look for class participation and judge one student to perform well based on high participation and another to perform poorly due to low participation. Cue diagnosticity refers to how related a cue is to an outcome measure. Considering the example above, if class participation is not related to student achievement, the cue is not diagnostic. The primary goals of the present investigation are to explore the cues teachers self-report using for judging student learning, and then examine the degree to which cue use affects judgment accuracy. For this study, religious educators in a Northwest suburb of the United States were asked to predict how well their students would perform on a learning assessment covering the religious content of their course.

This study contributes to the literature in two important ways. First, there are no published articles exploring the cues teachers self-report using to judge student learning; therefore, the present study will provide the first empirical data on cue use. This will be

the first study to examine how self-reported cue use is related to judgment accuracy. Second, this is the first study to examine teacher judgments in a religious studies context.

Teacher Judgments in a Religious Education Context

As noted above, there are no published studies providing data about the cues teachers self report using to judge student learning. Therefore, there is no basis for evaluating whether teachers across different academic domains rely on different cues. This study was set in a religious studies context to examine whether or not religious educators use different cues than teachers in other educational settings. Inasmuch as this setting introduces new cues for judging student learning, it will afford a unique opportunity to examine how cue use affects judgment accuracy.

The present study was set in the Seminaries and Institutes division of the Church of Jesus Christ of Latter-Day Saints. There are some unique aspects of this religious educational setting that are germane to this study. First, the environment (not simply the content) itself is very different than a traditional public school classroom. Religious instruction is provided during the regular school day on a campus adjacent to the public high school, and students are released from public school for one class period to attend. Students walk across the street, leaving school property to get to class. When they enter the class, they see other students they know from school and with whom they participate in other religious observances. These relationships are the foundation for an environment of love, respect, and purpose that leads to open and vulnerable conversations. When the bell rings, students sing a religious hymn together before offering a prayer and listening to a peer share a thought from a religious text. Then a teacher provides instruction, typically leading a discussion focused on religious text. Teachers regularly use questions

to help the students identify and analyze content, think deeply about religious principles, feel the truth and importance of those principles, and apply those truths in their lives. This instruction requires student participation and is often very personal and soul-searching. Finally, class ends with a prayer and an invitation to live according to the religious principles taught that day.

Second, in this particular religious setting, a class is purposefully set up as a student-centered learning environment. There are three stated priorities that focus teachers on the learner. According to Webb (2014), instructors should: (a) develop the ability to see each student's individual needs, strengths, and potential; (b) help students follow principles for having personal spiritual experiences and developing mastery of the religious content of the course; and (c) establish expectations and provide individual opportunities that inspire students to have meaningful experiences with sacred texts and teachings. Laced within each of these priorities is an underlying necessity to understand students. Furthermore, these priorities are emphasized in teacher trainings, and teachers are often asked if they are meeting them (especially after being observed by administrators).

Third, in this religious teaching context, success is measured less by how much students know and more by how much students *change* because of what they know. Similar to how a math teacher doesn't just want their students to know the quadratic formula, but understand how to think quantitatively; these religious educators do not just want their students to know and understand sacred texts and religious content, they also want students to apply what they have learned. The teachers in this study are successful if their students act differently as a result of the lesson. Therefore, if students come to class

and learn nothing about the content of a certain text, but they leave thinking about how they need to be kinder to their siblings, be more honest, or do any other thing that makes them live according to religious principles—mission accomplished! The purpose of learning, in this religious education context, is to help students become better people. In this regard, learning is more of a "means" than it is an "end" of education.

Finally, the emphasis on becoming is illustrated by how the teachers in this study use their preparation time and what a class looks like. These teachers are encouraged to focus on the student's life – family, friends, hobbies, etc. For example, it is common for these teachers to visit students in their homes to get to know them and their families, extend a class assignment, or follow-up on missed classes.

These unique aspects of this religious education setting are important because evidence from the literature suggests that student-centered instruction increases judgment accuracy (Connor et al., 2014; Curby, Rimm-Kaufman, Cameron Ponitz, 2009). Due to the student-centered nature of the classrooms in this study, there is reason to believe that these teachers will have above average judgment accuracy. However, it might also be the case that these teachers have access to cues that are superfluous to predicting student learning and not helpful for making judgments; thus, it is also possible that this added knowledge may focus teachers on cues that do not accurately predict (low cue diagnosticity) student learning.

In sum, working in the context of religious education may introduce a broad range of cues for judging student learning. It is not clear how this will affect the accuracy of teacher judgments of student learning. Thus, in the present investigation, judgment accuracy of these religious educators will be compared to the average level of teachers'

judgment accuracy in the extant literature. More important to advancing teacher judgment research, the present investigation will explore the cues these religious educators self-report using for judging student learning, and then examine the degree to which cue use affects judgment accuracy.

The research questions for the study are the following:

1. How does judgment accuracy for religious educators compare to average accuracy from the teacher judgment literature?
2. What cues do religious educators self-report using to judge student learning?
3. Does self-reported cue use affect judgment accuracy?

CHAPTER 2: LITERATURE REVIEW

There is extensive literature on teacher judgments (for reviews see Hoge & Coladarci, 1989; Südkamp, Kaiser, & Möller, 2012). In this review of literature, I will first describe how teacher judgments are gathered and how judgment accuracy is measured. I will then explain why teacher judgments are important to instruction. Next, I will describe a framework for understanding factors that influence the accuracy of judgments (i.e., the cue-utilization framework, Koriat, 1997), and I will use the framework to explain some of the variations in judgment accuracy found in the literature. Finally, the review of literature will frame the current study within the cue-utilization framework and highlight the contribution this study makes to the literature.

Measuring the Accuracy of Teacher Judgments

The basic method for gathering teachers' judgments is to show or describe a test and ask teachers to predict how their students will perform on it. Students are then given the test. Judgment accuracy is operationalized as the match between teachers' predictions and actual student performance. Although the basic method of gathering teacher judgments has been standardized in the literature (teachers are shown a test and then asked to predict how one or more individuals will score), as noted by Südkamp et al. (2012), differences arise based on (1) whether or not teachers were informed about the nature of the test (i.e. have they seen it or not), (2) judgment specificity (high specificity- the teacher ranks the class high to low. Low specificity- the teacher puts students into groups of low, medium, high), (3) how students were ranked (ranked, categories, Likert

scale, etc.), and (4) domain specificity (are the teachers predicting achievement for a specific subject or overall achievement in school).

Judgment accuracy is measured according to absolute accuracy or relative accuracy. Absolute accuracy describes the match between the magnitude of judgments and actual performance. For example, a teacher's *average* prediction across students on a 10-question test may by 7. If the students *average* score is 7, then the teacher would have perfect absolute accuracy. By contrast, if the students average score was 2, then that teacher would be over-confident by 5 points on average—indicating poor absolute accuracy. Absolute accuracy is important because a teacher who has poor absolute accuracy and is overly confident may move on to a new topics before students have learned materials well enough to proceed.

Relative accuracy describes how well a teacher's predicted order of student performance correlates with actual student performance. Relative accuracy is operationalized as an intra-teacher correlation between predicted and actual performance computed across students. If the teacher has perfect accuracy, the correlation will be 1.0. Relative accuracy is important because a teacher who has poor relative accuracy cannot accurately differentiate students who understand materials from those who do not; thus, the teacher may fail to spend additional time working with struggling students.

Absolute accuracy and relative accuracy inform different kinds of instructional decisions (Thiede, Oswalt, Brendefur, Carney, & Osguthorpe, 2019) and these measures are not necessarily related to one another (Dunlosky & Thiede, 2013). That is, a teacher could be highly accurate in terms of both absolute accuracy and relative accuracy, or a teacher could be highly accurate in terms on one measure and highly inaccurate in terms

of the other. Thiede et al. (2018) recently showed these measures of accuracy were only weakly correlated. Thiede et al. (2018) also showed that teachers' relative accuracy was positively correlated with gains in student achievement; whereas, absolute accuracy was not. Thus, I focused on relative accuracy in the present study.

Why Teacher Judgments Matter

An important aspect of a teacher's professional competence is the ability to accurately assess student understanding (Ready & Wright, 2011; Südkamp et al., 2012), interest, and motivation (Shavelson, 1978). This importance is emphasized by the National Board for Professional Teaching Standards in proposition 3.3 which states that teachers should, "know how to assess the progress of individual students as well as the class as a whole" (Proposition 3.3; National Board for Professional Teaching Standards, 2010).

The importance of teacher judgment is emphasized by the implications it has on different variables that contribute to a student's success. Teachers' judgments of student ability influence expectations of that student (Brophy & Good, 1986). These expectations tend to correlate with student success (De Boer, Bosker, & van der Werf, 2010; Jussim & Eccles, 1992). If a teacher's judgments lead to lower expectations, those expectations could lead to lower future achievement. Teachers' judgments can influence students' academic self-efficacy (e.g., Möller, Pohlmann, Köller, & Marsh, 2009; Trautwein, Ludtke, Köller, & Baumert, 2006), which has been linked to student achievement. If a teacher's judgments lead to lower student academic self-efficacy, this could adversely impact student learning (Marsh, 1990).

More relevant to the day-to-day operations, teachers' judgments of student learning are used to guide instruction (e.g., Alvidrez & Weinstein, 1999; Thiede et al., 2015). That is, judgments guide teachers' choices of classroom activities and materials; they determine the difficulty of classroom assignments, which questions to use, and the grouping of students; and they may signal a teacher to change his/her approach to teaching (Shavelson & Stern, 1981; Südkamp et al., 2012). Teacher judgments are used to identify struggling students who may need additional instruction to master content (e.g., Bailey & Drummond, 2006; Beswick, Willms, & Sloat, 2005; Teisl, Mazzocco, & Myers, 2001). This information is not only important for placement decisions, but teacher judgment is also given heavy weight in decisions regarding intervention (Helwig, Anderson, & Tindal, 2001; Hoge, 1983).

As highly accurate judgments of student learning can lead to more effective instruction, it is not surprising that teachers' judgment accuracy has been linked to student achievement. In particular, greater judgment accuracy has been found to correlate with greater gains in student achievement (Carpenter, Fennema, Peterson, Chiang & Loef, 1989; Peterson, Carpenter, Fennema, & Loef 1989; Thiede et al., 2015; Thiede et al., 2018). Thus, it's no wonder that researchers have sought to find ways to improve the accuracy of teachers' judgments.

The Cue-Utilization Framework of Judgments

The cue-utilization framework (Koriat, 1997) states that judgment accuracy is influenced by two factors: cue utilization and cue diagnosticity. Cue utilization simply refers to the cues people use to make judgments. Cue diagnosticity refers to how predictive a particular cue is to actual performance. If the cues used to make judgments

are highly diagnostic (predictive) of student performance, judgment accuracy will be high. If the cues are not diagnostic of performance, judgment accuracy will be low (e.g., Thiede, Griffin, Wiley, & Anderson, 2010).

For the purposes of this study and in the context of teacher judgments, a cue can be anything said or done that serves as a signal to teachers that a student understands content and will, therefore, perform well on a test. Teachers asked to predict student achievement have hours of student interaction and thousands of "signals" to draw from as the basis for that prediction. In other words, teachers have access to many different cues for making judgments of student learning.

Not every cue is diagnostic of student achievement. Some cues are highly diagnostic of student achievement and others are not. Accuracy is low when judgments are based on cues that are not diagnostic of student achievement (Benjamin, Bjork, & Schwartz, 1998; Hertzog, Dunlosky, Robinson, & Kidder 2003; Robinson, Hertzog, & Dunlosky, 2006; Thiede et al., 2010). Accuracy will also be low when teachers fail to use highly diagnostic cues (van Loon, de Bruin, van Gog, van Merriënboer, & Dunlosky, 2014).

Cues Used to Judge Student Learning

The teacher judgment research has not explicitly examined the cues teachers use to judge student learning. However, research has explored different factors that affect teachers' judgments of student learning. The cue-utilization framework can explain why these factors may affect judgment accuracy.

Gender and Ethnicity

Teacher judgment research has explored whether teacher judgments are informed by students' gender (Helwig, Anderson & Tindal, 2001; Mizala, Martínez, & Martínez, 2015) and ethnicity (Kaiser, Südkamp, & Möller, 2017; Martínez, Stecher, & Borko, 2009). Although it is possible that these demographic variables may be related to certain academic outcomes, the effect of gender and ethnicity on student achievement varies across topics. The weak relation between demographic variables and student achievement suggests that gender and ethnicity are not highly diagnostic cues. Thus, using gender and ethnicity as cues for judging student learning would not support highly accurate judgments. Recent findings by Kaiser et al. (2017) suggest that teachers do not use these cues for judging student learning.

Student Behavior

Student behavior may serve as a cue for judging student learning. In particular, teachers have been shown to use cues such as engagement in class (Kaiser, Retelsdorf, Südkamp, & Möller, 2013; Jenkins & Demaray, 2016), interest (more about teachers' ability to judge interest below) in domain being studied (Kikas, Silinskas, & Soodla, 2015), effort and conduct (Bennett, Gottesman, Rock, & Cerullo, 1993), and verbal assertiveness, compliance, and self-control (Hecht & Greenfield, 2002; Jenkins & Demaray, 2016). Although one can imagine situations in which these behaviors may be related to student achievement, non-academic student behaviors have generally been found to be weakly related to student achievement. Thus, these behaviors are not likely highly diagnostic of student learning, and using student behavior as a cue will not likely support high levels of judgment accuracy.

Past Performance

Teachers use past performance as a cue for judging student learning (Hecht & Greenfield, 2002; Martínez et al., 2009). Past performance is generally diagnostic of student achievement, but the diagnosticity of past performance is determined by the alignment of past performance to the student outcome being predicted by a teacher. When past performance is highly related to the measure being predicted (as with well-designed formative assessments) past performance will be highly diagnostic and using this cue will support highly accurate judgments of student learning (Hecht & Greenfield, 2002; Martínez et al., 2009; Thiede et al., 2019). Teaching practices, like student-centered teaching, that include high levels of dialog between students and the teacher that serve as formative assessments, also lead to improved judgment accuracy (Carpenter et al., 1989; Connor et al., 2014; Curby et al., 2009; Peterson et al., 1989). Past performance as a cue may be less diagnostic when past performance is less aligned to the predicted outcome. This is demonstrated when teachers use past performance in math to predict science performance (Dompnier, Pansu, & Bressoux, 2006), or when teachers use past performance in school in general to predict more specific learning outcomes (Dusek & Joseph, 1983). Generally, more frequent use of past performance to inform teacher judgments will improve judgment accuracy (Martínez et al., 2009; Thiede et al., 2019).

To summarize, the literature has examined a variety of cues that teachers use to judge student learning. The research suggests that cues such as past performance (e.g., formative assessments) are likely highly diagnostic; thus, judgment accuracy should be greater for teachers who use these cues than for those who do not. The research also suggests that cues such as non-academic behavior are not likely diagnostic; thus,

judgment accuracy should be greater for teachers who do not use these cues than for teachers who do.

Teacher Judgments in a Religious Education Context

This is the first study to examine teachers' judgment accuracy in a religious education context. The priorities of the religious organization may influence the process of judging student learning. For example, in the present investigation, the priorities include that instructors should develop the ability to see each student's individual needs, strengths, and potential (Webb, 2014). Thus, within this context, teachers are encouraged to get to know their students on a deeper level. How might this affect the accuracy of teachers' judgments of student learning?

On one hand, the stated priority may encourage teachers to build relationships with students, which could increase judgment accuracy because a strong relationship with students could provide insights into their learning. On the other hand, focusing on aspect could also have a detrimental effect on judgment accuracy because it encourages teachers to consider cues that may not be related to student learning. Given these two possibilities, it is not clear how the religious education context will affect judgment accuracy, which is why it is important to address this issue empirically. Thus, the first research question is: How does judgment accuracy for religious educators compare to average accuracy from the teacher judgment literature?

The religious education context was selected because religious educators might be influenced by the religious priorities and content of the educational setting and use additional or different cues to judge student learning. Thus, the second research question is: What cues do religious educators self-report using to judge student learning?

Finally, this is the first study to gather self-reported data on the cues teachers use to judge student learning. These data make it possible to examine how cue use affects judgment accuracy. Thus, the third research question is: Does self-reported cue use affect judgment accuracy?

CHAPTER 3: METHOD

The present study measured the judgment accuracy of 42 seminary teachers in the greater Boise area of Idaho. The purpose was threefold. First, this study will examine whether judgment accuracy for seminary teachers is greater or less than the average accuracy reported in the teacher judgment literature (see Südkamp et al., 2012). Second, this study will gather self-report data to ascertain what cues seminary teachers report using to judge student learning. Third, this study will examine how cue use affects judgment accuracy.

Participants

Forty-two religious educators from a Northwest suburb of the United States participated in this study. Of the 42 teachers that started the study, thirty-six (86%) were male and six (14%) were female. Thirty-eight teachers (90%) were full-time teachers (teaching 4-6 classes); whereas, 4 teachers (10%) were part-time teachers (teaching 2-3 classes). As shown in Table 1, years of experience varied across participating teachers. Judgment accuracy did not differ across the years of teaching experience groups, $F(5, 33) = 1.6$, $MSe = .04$, $p = .19$.

Table 1. Years of Teaching Experience for Teachers

Table 1. Years of Experience					
1-5	6-10	11-15	16-20	21-30	31+
14	7	6	3	7	3

Measures

Knowledge

Knowledge was operationalized as scores on a standard end-of-semester learning assessment that covers the content of a specific religious text and is used throughout the world-wide educational system for the sponsoring religious organization. This test consists of 28 multiple choice questions and four essay questions. The multiple choice each had three options to choose from. The essay questions required students to write about gospel teachings as found in the scriptures. The test has been shown to have good reliability (Cronbach's alpha greater than .80).

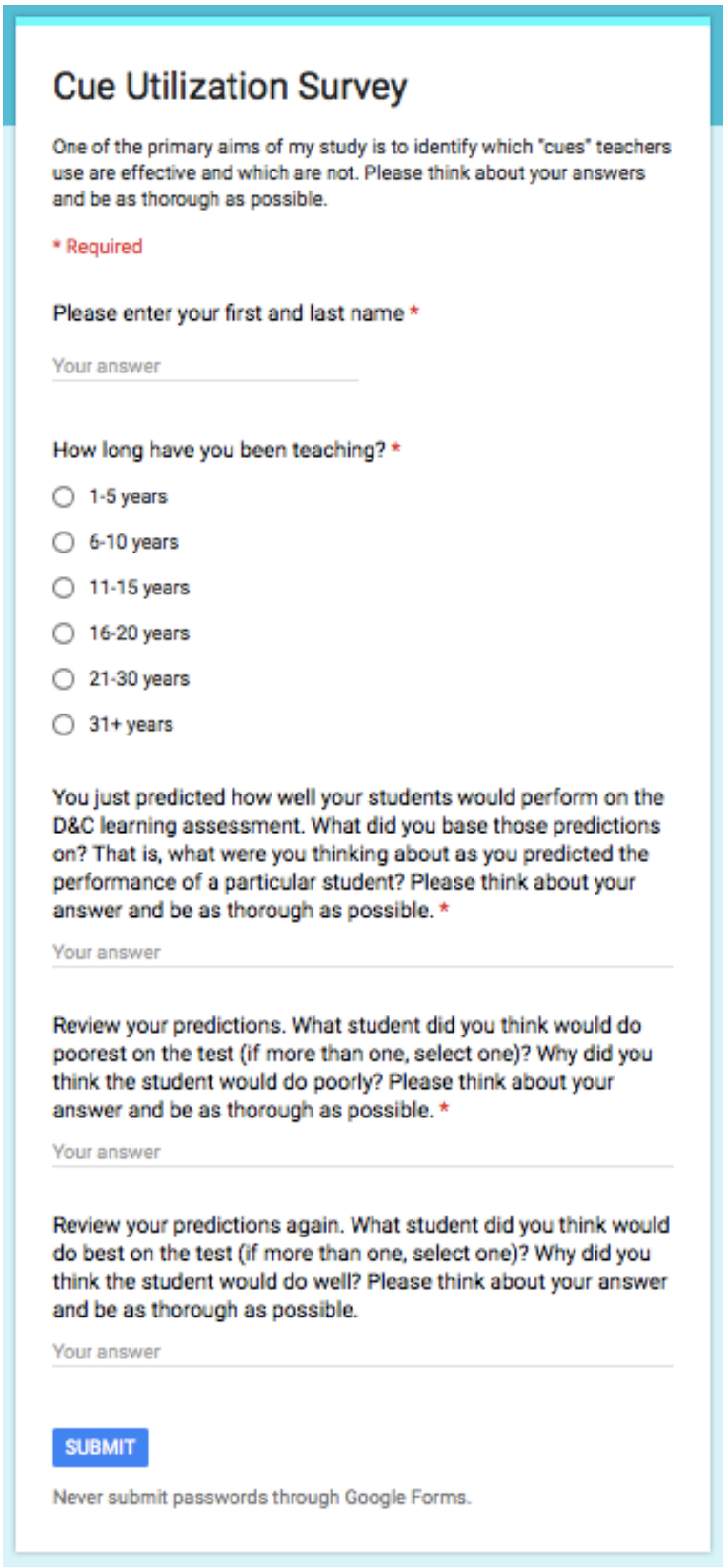
Judgment Accuracy

Judgment accuracy was operationalized as the intra-individual correlation between a teacher's predicted performance and students' actual performance computed across students (as in Helmke & Schrader, 1987).

Procedure

In January of 2019, participating teachers chose one class to predict performance on the end-of-semester learning assessment. After these teachers predicted the performance of all the students in their class, the test was administered mid-January.

Approximately two weeks after collecting the predictions and test results, a survey was sent to teachers asking them to report what cues they used to make their predictions. As shown in Figure 1, teachers were asked to report their cue use in a general sense. They were also asked to identify one student who they predicted would score low and one student they thought would score high and explain the basis for those judgments. The list of cues was created by coding the teachers' responses to all three questions.



Cue Utilization Survey

One of the primary aims of my study is to identify which "cues" teachers use are effective and which are not. Please think about your answers and be as thorough as possible.

* Required

Please enter your first and last name *

Your answer

How long have you been teaching? *

- 1-5 years
- 6-10 years
- 11-15 years
- 16-20 years
- 21-30 years
- 31+ years

You just predicted how well your students would perform on the D&C learning assessment. What did you base those predictions on? That is, what were you thinking about as you predicted the performance of a particular student? Please think about your answer and be as thorough as possible. *

Your answer

Review your predictions. What student did you think would do poorest on the test (if more than one, select one)? Why did you think the student would do poorly? Please think about your answer and be as thorough as possible. *

Your answer

Review your predictions again. What student did you think would do best on the test (if more than one, select one)? Why did you think the student would do well? Please think about your answer and be as thorough as possible.

Your answer

SUBMIT

Never submit passwords through Google Forms.

Figure 1. Cue-utilization survey

CHAPTER 4: RESULTS

This chapter is organized around the three research questions.

Research Question 1

How does judgment accuracy for religious educators compare to average accuracy from the teacher judgment literature?

Südkamp et al. (2012) reported an average accuracy of .63 across 73 studies. Judgment accuracy was operationalized as the intra-individual correlation between predicted and actual performance computed across students. For each participating teacher, judgment accuracy was computed, these intra-individual correlations were then averaged across the participants. Judgment accuracy ranged from 0.01 to 0.94, with an average accuracy of .50 (Standard Deviation = .19). This was significantly lower than the average judgment accuracy computed by Südkamp et al. (2012), $t(35) = 4.01$, $p < .001$.

Research Question 2

What cues do religious educators self-report using to judge student learning?

Teachers self-reported the cues they used to judge student learning. The participating teachers collectively reported using 16 different cues. The number of cues provided ranged from 1 to 10. On average, teachers reported using 4.5 cues. Judgment accuracy was not significantly related to the number of cues used ($r = -.08$, $p = .64$). As shown in Figure 2, participation was the most commonly reported cue used to judge student learning.

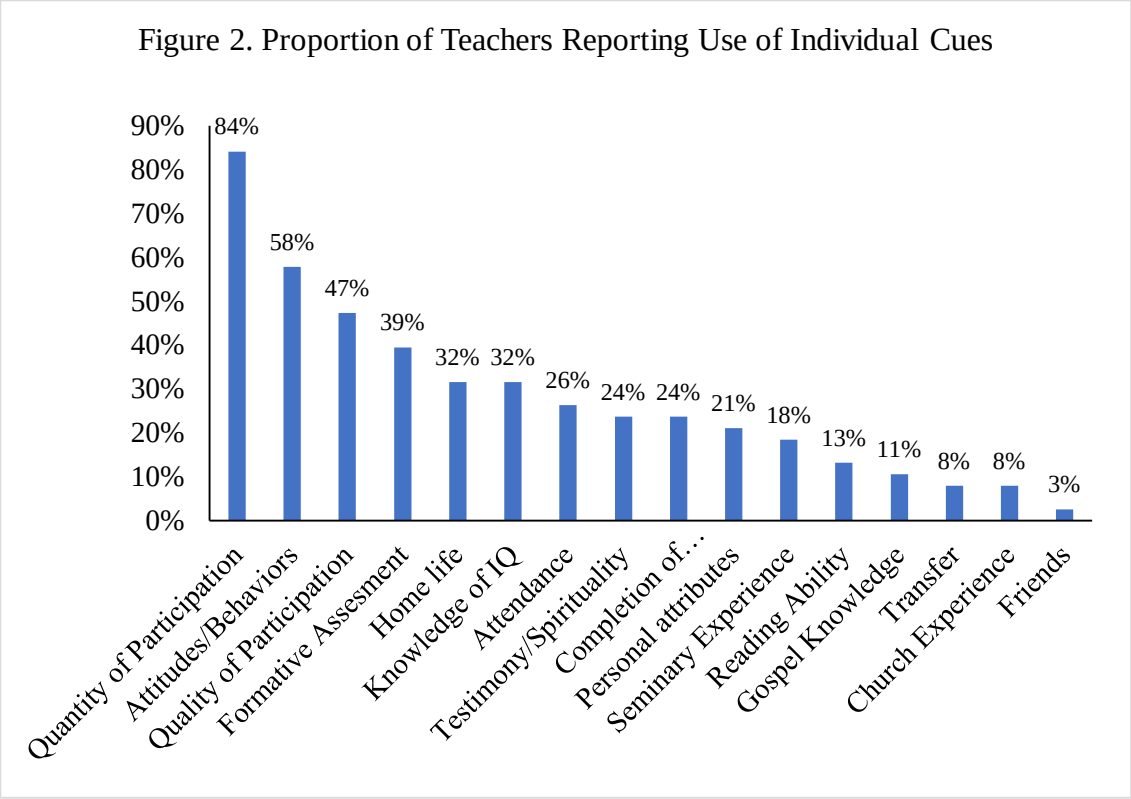


Figure 2. Proportion of Teachers Reporting Use of Individual Cues

These 16 cues were then thematically combined into four broader categories, which align with the literature on teacher judgments. Specifically, the 16 cues were collapsed into ***class performance*** (for more on how this is linked to teacher judgments see Martínez et al., 2009; Thiede et al., 2019), ***personal attributes*** (for more see Hecht & Greenfield, 2002), ***external factors*** (for more see Hauser-Cram, P., Selcuk, R. S., & Stipek, D. 2003), and ***class behavior*** (for more see Kaiser et al., 2013). Table 3 shows the items classified into the four categories.

Table 2. Individual Cues Classified in the Four Categories

Table 2. Individual Cues Classified in the Four Categories		
Individual Cue	**Cue Category**	**# of Teachers Reporting Cue**
Quantity of Participation	Class Behavior	32
Attitudes/Behaviors	Class Behavior	22
Quality of Participation	Class Behavior	18
Attendance	Class Behavior	10
Formative Assessment	Class Performance	15
Completion of Assignments	Class Performance	9
Home life	External Factors	12
Gospel Knowledge	External Factors	4
Transfer	External Factors	3
Church Experience	External Factors	3
Friends	External Factors	1
Knowledge of IQ	Personal attributes	12
Testimony/Spirituality	Personal attributes	9
Personal attributes	Personal attributes	8
Seminary Experience	Personal attributes	7
Reading Ability	Personal attributes	5

Once individual cues were categorized into the four broader categories, each teacher was scored on the number of cues they reported using in each category. Teachers were then classified as high use teachers if they reported using at least two of the cues in the category. By contrast, they were classified as low use teachers if they reported using

fewer than two of the cues in the category. Figure 3 shows the proportion of high use and low use teachers for each category.

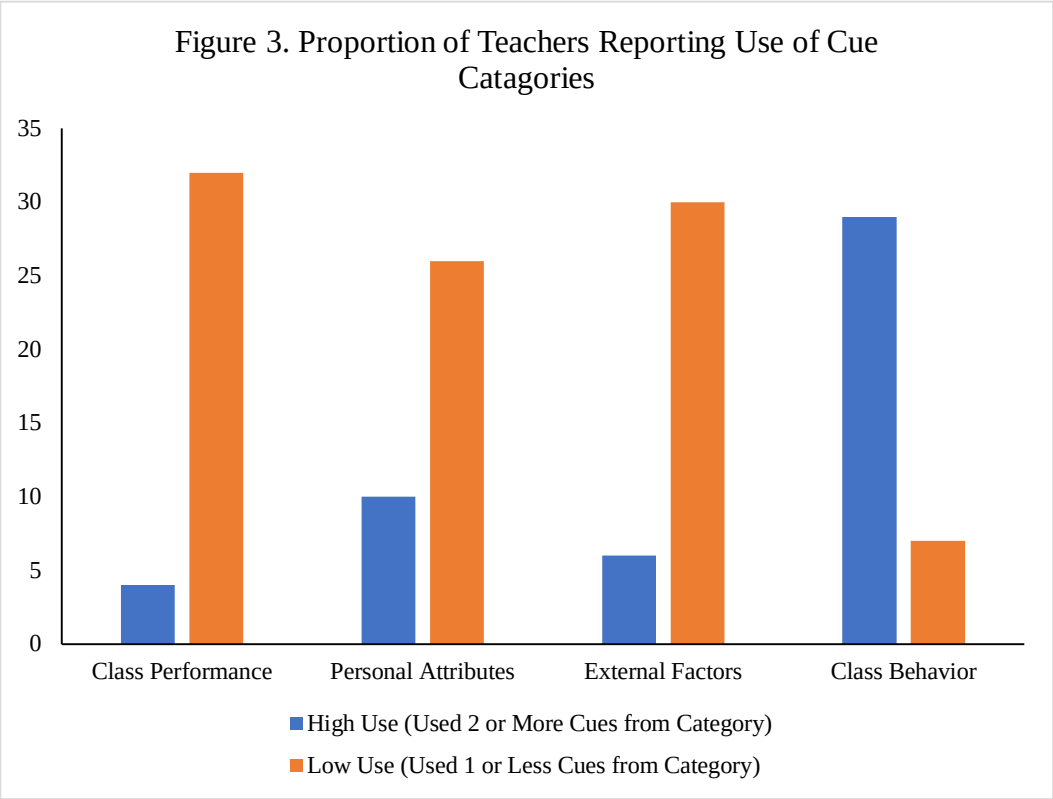


Figure 3. Proportion of Teachers Reporting Use of Cue Categories

It is important to note that these categories are fairly independent of one another. That is, as shown in Table 3, use of one category was not significantly correlated to use of another category. Thus, those teachers who relied heavily on class performance, for instance, were not the same teachers who relied on external factors—else use across categories would have been highly correlated.

Table 3. Correlation Matrix for Four Cue Categories

Table 3. Correlation Matrix for Four Cue Categories				
	Class Performance	**Personal Attributes**	**External Factors**	**Class Behavior**
Class Performance	1.000			
Personal Attributes	-0.022	1.000		
External Factors	0.079	0.222	1.000	
Class Behavior	0.174	-0.165	0.031	1.000

Research Question 3

Does self-reported cue use affect judgment accuracy?

As noted in the introduction, the cue-utilization framework (Koriat, 1997) suggests that judgment accuracy is affected by cue use. To evaluate whether cue use affected judgment accuracy, judgment accuracy was compared for teachers who used cues and teachers who did not use cues. That is, for each of the 16 cues, a t-test was conducted to compare judgment accuracy for those who used the cue versus those who did not. Judgment accuracy was marginally greater for those who reported using *completion of assignments* than for those who did not, $t(36) = 2.0$, $p = .052$. There were no significant differences in accuracy based on use of the other cues, $t(36) < 1.6$, $p > .12$.

Judgment accuracy was also compared for high use and low use teachers across the four broad categories of cues. As shown in Figure 4, class performance affected judgment accuracy, $t(36) = 2.49$, $p = .02$. That is, judgment accuracy was greater for those who reported using cues related to class performance than for those who did not.

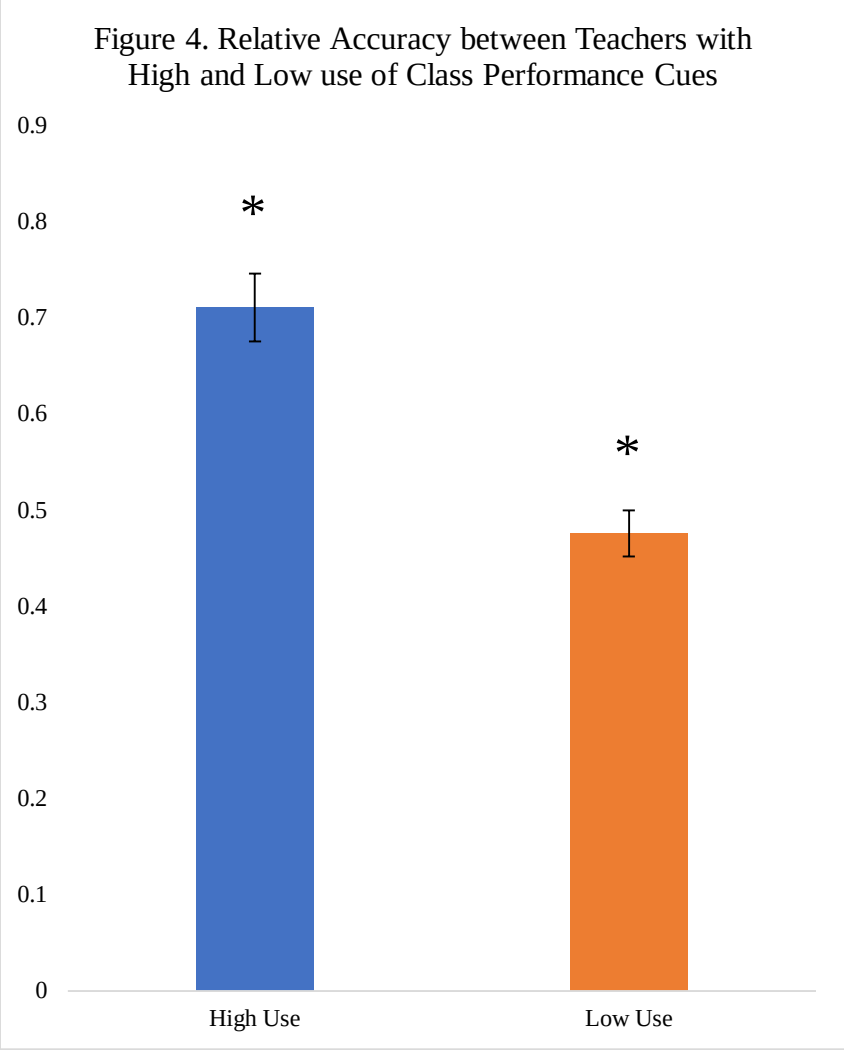


Figure 4. Relative Accuracy between Teachers with High and Low use of Class Performance Cues

As shown in Figure 5, judgment accuracy was also affected by use of personal attributes as a cue. Judgment accuracy was greater for those who did *not* report using this cue than for those who reported using it, $t(36) = 2.27$, $p = .03$.

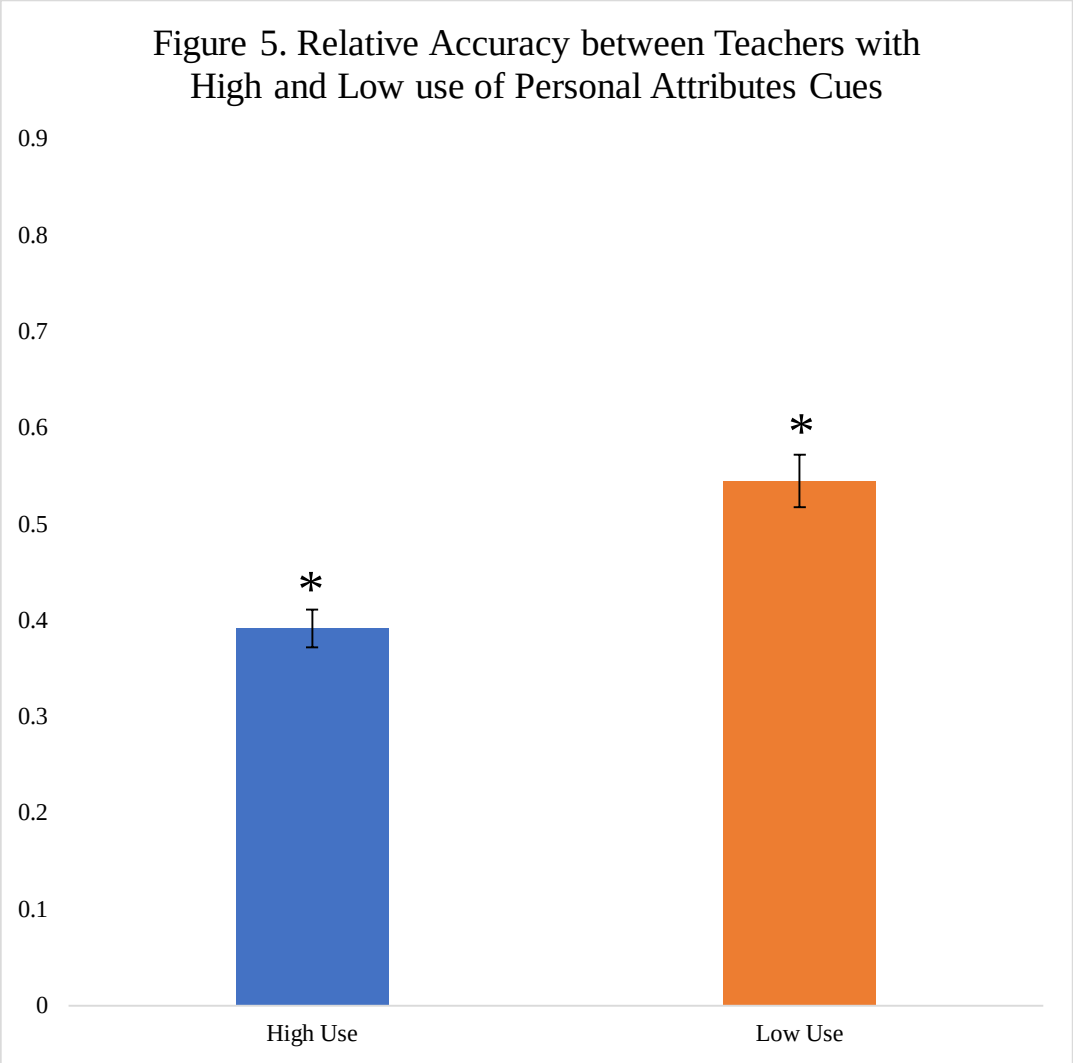


Figure 5. Relative Accuracy between Teachers with High and Low use of Personal Attributes Cues

Judgment accuracy was not affected by use of external factors, $t(36) = .82$, $p = .42$ (see Figure 6).

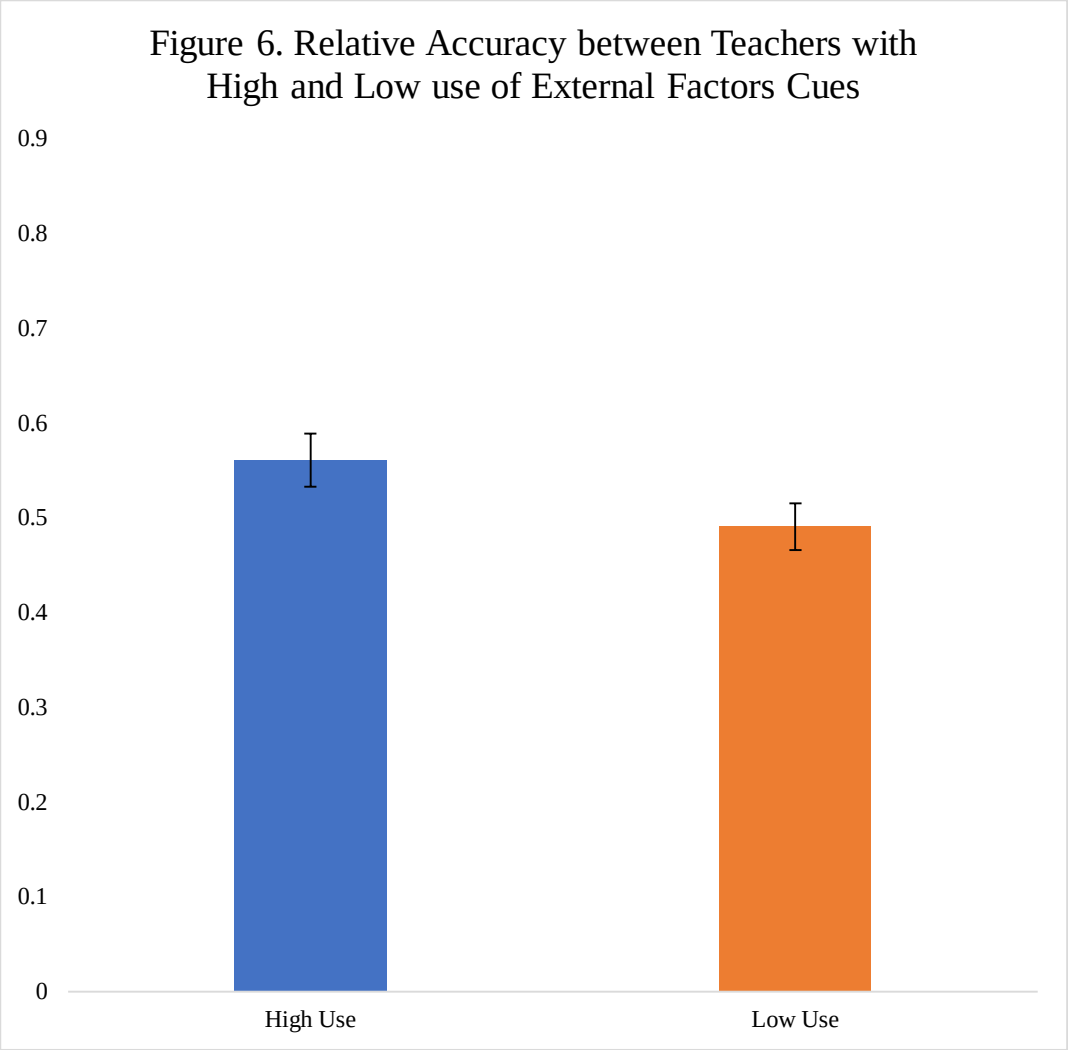


Figure 6. Relative Accuracy between Teachers with High and Low use of External Factors Cues

Judgment accuracy was not affected by use of class behavior, $t(36) = .40$, $p = .50$ (see Figure 7).

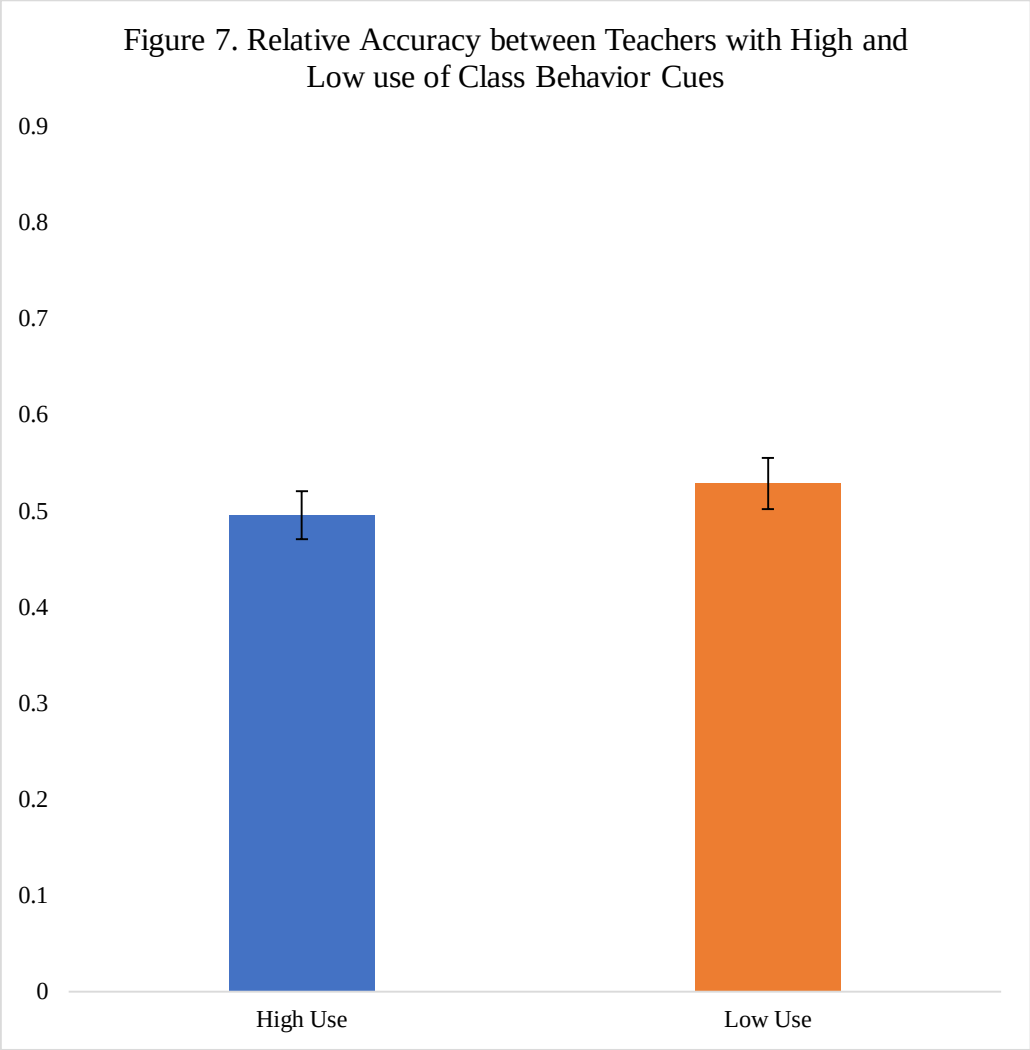


Figure 7. Relative Accuracy between Teachers with High and Low use of Class Behavior Cues

CHAPTER 5: DISCUSSION

Over the past three decades, researchers have explored how accurately teachers judge student achievement (for a recent review see Südkamp et al., 2012) because of the major impact a teacher's opinion has on the futures of his/her students (De Boer et al., 2010; Jussim & Eccles, 1992). Teacher judgments impact the teacher's expectations of students, the student's academic self-efficacy, and the day-to-day decisions teachers make during instruction. All of these have been linked to student achievement.

The accuracy of teacher's judgments of student learning is related to gains in student achievement (e.g., Thiede et al., 2018); therefore, it is important to identify factors that affect the accuracy of teacher judgments. The cue-utilization framework (Koriat, 1997) suggests two key factors that affect judgment accuracy are cue utilization (what teachers use to make judgments) and cue diagnosticity (how related a cue is to an outcome measure).

Thiede et al. (2019) recently suggested the importance of exploring the effect of cue utilization and cue diagnosticity on judgment accuracy. The current study is a first step toward increasing our understanding of cue utilization in teachers' judgments of student learning. This is the first study to ask teachers to self-report on the cues they use to judge student learning.

The findings revealed that teachers participating in this study reported using four broad categories of cues to judge student learning: class performance, personal attributes, external factors, and class behavior. Class performance cues, in this case, had to do with

the teacher's use of formative assessment. Class behavior cues pertained to observable behaviors in class. Personal attribute cues focused on things specific to the student, whereas external factors attended to the environment in which the students live their lives.

Consistent with the cue-utilization framework (Koriat, 1997), the findings showed that judgment accuracy was affected by the cues teachers reported using. In particular, as the literature predicted, teachers who used multiple "class performance" cues judged student learning significantly more accurately than those who did not. This finding validates the importance of consistent formative assessment and student-centered instruction in regard to increasing judgment accuracy (Carpenter et al., 1988; Connor et al., 2014; Curby et al., 2009; Peterson et al., 1989). Although nothing of significance was found in regard to external factors and class behavior, the data found an inverse relationship between teachers who used multiple "personal attributes" cues and judgment accuracy. Teachers who used personal attributes to guide judgment were significantly less accurate than teachers who did not.

The lower than average accuracy by teachers in this study is possibly explained in at least two ways. First, it could be that participating teachers are just bad at predicting student achievement. Despite their personal relationships with students and families, and their stated priority to attend to their individual learning and spiritual needs, it might be that these participating teachers did not have direct access to the cues that predict actual student learning of the religious content of the course.

A second possible explanation is found in the diagnosticity of the cues they used. It might be that participating teachers were more focused on cues that are indicative of

the ultimate purpose of the religious instruction—*transfer* of religious principle to everyday living—as opposed to cues that are indicative of religious content *recall* on an assessment. As a result of their possible focus on transfer (as opposed to recall), it might be that many of the cues teachers used served as distractions. While spirituality may be diagnostic of transfer, for example, it is not necessarily a good cue for student achievement on the learning assessment.

In fact, there is evidence of this type of distraction in the literature. In addition to student learning, teachers make judgements about motivation, interest, and attitudes. The teacher judgment literature has shown that judgment accuracy is fairly poor when judging student motivation (Givvin, Stipek, Salmon, MacGyvers, 2001), interest (Swanson, 1985; Middleton, 1995; Givvin et al., 2001), and attitudes (Lewis, 1979; Swanson, 1985; Givvin et al., 2001). This difficulty most likely arises from teachers' inaccessibility to highly diagnostic cues, which would explain the poor judgment accuracy. Even though religious content competency is important in the classrooms in this study, participating teachers may have conflated religious content recall and transfer of religious principles—thus focusing on cues that were diagnostic of the ways that students applied the religious instruction, but not of student achievement on the learning assessment. For example, in a case where a teacher considered formative assessment but also considered spirituality, then it is possible that this later cue might have swayed the teacher to predict lower (or higher), despite the formative assessment.

The teachers that focused on performance cues the most had the highest judgment accuracy of student achievement. This finding confirms what has been shown in the literature. For example, one study found that there is relevant (in the case of this study,

oral and written achievement in mathematics, i.e. formative assessment) and irrelevant (e.g. family background) information that is available as cues for teacher judgment (Kaiser, Möller, Helm, & Kunter, 2015). Another study found that, "teachers who were only provided with information on students' oral and written achievement in mathematics, made more accurate judgments of fictional students' mathematics grades than teachers who were additionally provided with student characteristics (i.e., students' engagement, minority status, gender, intelligence, and German dictation exercise grade)" (Oudman, van de Pol, Bakker, Moerbeek, & van Gog, 2018, p.216). Based on these findings, and the findings of this study, it is recommended that teachers focus exclusively on performance cues when judging student achievement.

That said, if the religious educators were less accurate due to their focus on transfer, then if they were asked to judge transfer (instead of recall), would they do so more accurately? A partial answer to this question is found in the literature that examines judgment accuracy of states of mind, including motivation. In addition to achievement, motivation plays a major role in academic achievement (Ladd & Dinella, 2009; Skinner, Kindermann, & Furrer, 2009). In one study that looked at 17 fourth- through sixth-grade teachers and 100 of their students, however, the researchers found that teacher perception of motivation was significantly inaccurate (Givvin et al., 2001). This is likely for the same reasons mentioned above, namely, judging a state of mind is harder than judging a student's current knowledge because it is difficult to identify diagnostic cues. Nonetheless, with the unique cues available to the participating teachers in this study and the unique focus of their work, it would be interesting to conduct future research that

examines the accuracy with which religious educators can judge transfer of religious principles to everyday living.

As in any study, this research had limitations. The most glaring limitation was the relatively small sample size. A future study with a larger sample size would enhance the contribution to literature. Second, the timing of the survey may have elicited more generic responses for teachers who made their predictions and then failed to give them a second thought.

In conclusion, this study contributed to the current teacher judgment literature in three ways. First, it identified that religious educators performed below-average in terms of judging student achievement and speculated this is most likely due to the cues they used. Second, it provided the first data regarding the cues teachers self-report using to judge student learning. Third, this study validated the theory that cue use affects judgment accuracy.

REFERENCES

Alvidrez, J., & Weinstein, R. S. (1999). Early teacher perceptions and later student academic achievement. Journal of Educational Psychology, 91(4), 731.

Bailey, A. L., & Drummond, K. V. (2006). Who Is at Risk and Why?: Teachers' Reasons for Concern and Their Understanding and Assessment of Early Literacy. Educational Assessment, 11(3&4), 149-178.

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. Journal of Experimental Psychology: General, 127, 55-68.

Bennett, R. E., Gottesman, R. L., Rock, D. A., & Cerullo, F. (1993). Influence of behavior perceptions and gender on teachers' judgments of students' academic skill. Journal of Educational Psychology, 85, 347 – 356.

Beswick, J. F., Willms, J. D., & Sloat, E. A. (2005). A comparative study of teacher ratings of emergent literacy skills and student performance on a standardized measure. Education, 126, 116–137.

Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), Handbook of Research on Teaching (pp. 328-375). New York: Macmillan.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. American educational research journal, 26(4), 499-531.

Connor, C. M., Spencer, M., Day, S. L., Giuliani, S., Ingebrand, S. W., McLean, L., & Morrison, F. J. (2014). Capturing the complexity: Content, type, and amount of instruction and quality of the classroom learning environment synergistically

predict third graders' vocabulary and reading comprehension outcomes. Journal of Educational Psychology, 106, 762 – 778.

Curby, T W., Rimm-Kaufman, S. E., &Cameron Ponitz, C., (2009). Teacher-child interactions and children's achievement trajectories across kindergarten and first grade. Journal of Educational Psychology, 101, 912 – 925.

De Boer, H., Bosker, R. J., & van der Werf, M. P. (2010). Sustainability of teacher expectation bias effects on long-term student performance. Journal of Educational Psychology, 102(1), 168.

Dompnier, B., Pansu, P., & Bressoux, P. (2006). An integrative model of scholastic judgments: Pupils' characteristics, class context, halo effect and internal attributions. European Journal of Psychology of Education, 21(2), 119-133.

Dunlosky, J., & Thiede, K. W. (2013). Four cornerstones of calibration research: Why understanding students' judgments can improve their achievement. Learning and Instruction, 24, 58-61.

Dusek, J. D. & Joseph, G. (1983). The bases of teacher expectancies: A meta-analysis. Journal of Educational Psychology, 75, 327 – 346.

Givvin, K. B., Stipek, D. J., Salmon, J. M., MacGyvers, V. L. (2001). In the eyes of the beholder: Students' and teachers' judgments of students motivation. Teaching and Teacher Education, 17, 321-331.

Hauser-Cram, P., Selcuk, R. S., & Stipek, D. (2003). When teachers' and parents' values differ: Teachers' ratings of academic competence in children from low-income families. Journal of Educational Psychology, 95, 813 – 820.

Hecht, S. A. & Greenfield, D. B. (2002). Explaining the predictive accuracy of teacher judgments of their students' reading achievement: The role of gender, classroom behavior, and emergent literacy skills in longitudinal sample of children exposed to poverty. Reading and Writing, 15, 789 – 809.

Helmke, A., & Schrader, F. W. (1987). Interactional effects of instructional quality and teacher judgement accuracy on achievement. Teaching and Teacher Education, 3(2), 91-98.

Helwig, R., Anderson, L., & Tindal, G. (2001). Influence of elementary student gender on teachers' perceptions of mathematics achievement. Journal of Educational Research, 95, 93–102. doi:10.1080/ 0022067010959657

Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue utilized for judgments about learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 29, 22-34.

Hoge, R. D. (1983). Psychometric properties of teacher-judgment measures of pupil aptitudes, classroom behaviors, and achievement levels. The Journal of Special Education, 17, 401–429. doi:10.1177/ 002246698301700404

Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. Review of Educational Research, 59(3), 297-313.

Jenkins, L. N. & Demaray, M. K. (2016). Teachers' judgments of academic achievement of children with and without characteristics of inattention, impulsivity, and hyperactivity. Contemporary School Psychology, 20, 183 – 191.

Jussim, L., & Eccles, J. S. (1992). Teacher expectations: II. Construction and reflection of student achievement. Journal of personality and social psychology, 63(6), 947.

Kaiser, J., Mo€ller, J., Helm, F., & Kunter, M. (2015). Das schülerinventar: Welche Schülermerkmale die Leistungsurteile von Lehrkra€ften beeinflussen. Zeitschrift Fur Erziehungswissenschaft, 18(2), 279e302. https://doi.org/10.1007/s11618-015-0619-5.

Kaiser, J., Retelsdorf, J., Südkamp, A., & Möller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. Learning and Instruction, 28, 73 – 84.

Kaiser, J., Südkamp, A., & Möller, J. (2017). The effects of student characteristics on teachers' judgment accuracy: Disentangling ethnicity, minority status, and achievement. Journal of Educational Psychology, 109(6), 871-888.

Kikas, E., Silinskas, G., & Soodla, P. (2015). The effect of children's reading skill and interest on teacher perceptions of children's skill and individualized support. International Journal of Behavioral Development, 39, 402 – 412.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. Journal of Experimental Psychology: General, 126, 349-370.

Ladd, G.W., & Dinella, L.M. (2009). Continuity and change in early school engagement: Predictive of children's achievement trajectories from first to eighth grade? Journal of Educational Psychology, 101, 190–206.

Lewis, J. (1979). A reading attitude inventory for elementary school pupils. Educational and Psychological Measurement, 39, 511}513.

Marsh, H. W. (1990). Causal ordering of academic self-concept on academic achievement: A multiwave, longitudinal panel analysis. Jour- nal of Educational Psychology, 82, 646–656. doi:10.1037/0022- 0663.82.4.646

Martínez, J. F., Stecher, B., & Borko, H. (2009). Classroom assessment practices, teacher judgments, and student achievement in mathematics: Evidence from the ECLS. Educational Assessment, 14, 78 – 102.

Middleton, J. A. (1995). A study of intrinsic motivation in the mathematics classroom: A personal constructs approach. Journal for Research in Mathematics Education, 26(3), 254}279.

Mizala, A., Martínez, F., Martínez, S. (2015). Pre-service elementaryschool teachers' expectations about student performance: How their beliefs are affecteds by their mathematics anxiety and student's gender. Teaching and Teacher Education, 50, 70-78.

Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. Review of Educational Research, 79(3), 1129-1167.

Oudman, S., van de Pol, J., Bakker, A., Moerbeek, M., & van Gog, T. (2018). Effects of different cue types on the accuracy of primary school teachers' judgments of students' mathematical understanding. Teaching and Teacher Education, 76, 214-226.

Peterson, P. L., Fennema, E., Carpenter, T. P., & Loef, M. (1989). Teacher's pedagogical content beliefs in mathematics. Cognition and instruction, 6(1), 1-40.

Ready, D. D., & Wright, D. L. (2011). Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities. American Educational Research Journal, 48, 335–360. doi:10.3102/ 0002831210374874

Robinson, E. A., Hertzog, C & Dunlosky, J. (2006). Aging, Encoding Fluency, and Metacognitive Monitoring. Aging, Neuropsychology, and Cognition, 13, 458-478.

Shavelson, R. J. (1978). Teachers' estimates of student' states of mind' and behavior. Journal of Teacher Education, 29(5), 37-40.

Shavelson, R. J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. Review of Educational Research, 51, 455–498. doi:10.2307/1170362

Skinner, E.A., Kindermann, T.A., & Furrer, C.J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. Educational and Psychological Measurement, 69, 493–525.

"Standards for NBPTS: The Five Core Propositions." NBPTS, 2019, www.nbpts.org/standards-five-core-propositions/.

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. Journal Of Educational Psychology, 104(3), 743-762. doi:10.1037/a0027627

Swanson, B. B. (1985). Teacher judgments of "first-graders' reading enthusiasm. Reading Research and Instruction, 25(1), 41}46.

Teisl, J. T., Mazzocco, M. M. M., & Myers, G. F. (2001). The utility of kindergarten teacher ratings for predicting low academic achievement in first grade. Journal of Learning Disabilities, 34, 286–293. doi:10.1177/ 002221940103400308

Thiede, K. W., Brendefur, J. L., Carney, M. B., Champion, J., Turner, L., Stewart, R., & Osguthorpe, R. D. (2018). Improving the accuracy of teachers' judgments of student learning. Teaching and Teacher Education, 76, 106-115.

Thiede, K. W., Brendefur, J. L., Osguthorpe, R. D., Carney, M. B., Bremner, A., Strother, S., . . . Jesse, D. (2015). Can teachers accurately predict student performance? Teaching and Teacher Education, 49, 36–44. doi:10.1016/j.tate.2015.01.012

Thiede, K. W., Griffin, T. D., Wiley, J., & Anderson, M.C.M. (2010). Poor metacomprehension accuracy as a result of inappropriate cue use. Discourse Processes, 47, 331–362.

Thiede, K., Oswalt, S., Brendefur, J. L., Carney, M. B., Osguthorpe, R. D. (2019). Teachers' judgments of student learning of mathematics. In J. Dunlosky & K. A. Rawson (Eds.). Cambridge Handbook of Cognition and Education (pp. 678 – 695). New York, NY: Cambridge University Press.

Trautwein, U., Lu¨dtke, O., Köller, O., & Baumert, J. (2006). Self-esteem, academic self-concept, and achievement: How the learning environment moderates the dynamics of self-concept. Journal of Personality and Social Psychology, 90, 334–349. doi:10.1037/0022-3514.90.2.334

Van Loon, M. H., de Bruin, A. B. H., van Gog, T., van Merriënboer, J. J. G., & Dunlosky, J. (2014). Can students evaluate their understanding of cause-and-effect relations? The effect of diagram completion on monitoring accuracy. Acta Psychologica, 151, 143 – 154.

Webb, Chad. "The Objective and Priorities." The Objective and Priorities of Seminaries and Institutes of Religion, 18 Mar. 2014, www.lds.org/si/objective?lang=eng&site=seminary.

APPENDIX A

Description of Cues

Cue 1- Formative assessment- This cue represents when a teacher used previous tests and/or quizzes that were scored, and the teacher used the score as a basis for his prediction. When this cue was used, the teacher would say something like, “we had previously done class quizzes, and I felt like I had a pretty good idea on how students would do.”

Cue 2- Testimony/spirituality- The word testimony means, “witness,” but in this group, the word is used to denote the strength of belief. For example, if a student has a “strong testimony,” it means that they believe what is taught. It also means that they live the principles outside of school. One teacher wrote, “I knew [this student] struggles on his own testimony.” Another responded to why he rated a certain student highly, “this student is one that strives to live those things that are being assessed. His learning isn’t just centered in book study but applied outside the classroom.”

Cue 3- Home life- One of the major teachings/mottos of The Church of Jesus Christ of Latter-day Saints is that it is a, “home-centered, church-supported” church. Thus, most of what students know should come from study they do in the home and in church on Sundays. This cue looks at whether or not the student’s parents attend church, if they are studying the scriptures as a family, etc. One teacher said, “I feel like I have a general sense of who has support in gospel learning at home.” This teacher used this knowledge to judge how well the student would do.

Cue 4- Other Personal attributes- Are they outgoing, energetic, extraverted, introverted, a good friend, self-disciplined, self-efficacious, and/or eidetic?

Cue 5- Quantity of participation- This cue, along with the cue, "quality of participation" is self-explanatory. The question here is how often does the student share? A common way to express the use of this cue was for a teacher to mention a high or low level of participation.

Cue 6- Attitudes/behaviors- Do they want to be there? Are they excited about seminary? Are they easy or difficult to engage? Etc.

Cue 7- Knowledge of IQ- In some cases, teachers had a general sense of how smart a student is was, typically because the teacher knew his/her grades. The students who were more intelligent, generally speaking, were rated higher because doing well is just what they do. Here are some statements from teachers regarding knowledge of IQ, "… is a young man who is very intelligent." "I also have an impression or attitude about each student as to their overall intellectual abilities." "I know of some students who get good grades in other classes." "I picked three that I thought would get a perfect score because they are 4.0 students at the high school…"

Cue 8- Seminary experience/age- With this cue, teachers considered the students age (freshman, sophomore, etc) and seminary experience (e.g. if a student was a senior,

but it was his first year taking seminary). A student with less experience was judged to do worse than if that student would have had more experience.

Cue 9- Quality of participation- Just because a student shares a lot does not mean he is sharing something of substance. For example, a student who needs to feel validated may share long, drawn out stories frequently that don't necessarily pertain to the scriptures being taught. Of course, these students are treated with love, but many teachers noted the difference between a student who shares things that are meaningful and insightful, and a student who simply shares. For example, teachers would write, "… has very good comments and meaningful participation," "… his response to questions was thoughtful," or "this student also clearly understands the materials based on the comments that they share in classroom discussions."

Cue 10- Reading ability- Teachers judged that a student who struggles reading and writing would inherently struggle taking any test. "I also considered how well they read when they read orally in class."

Cue 11- Attendance- A student who comes less will likely do worse than if he/she attended more. Speaking of a student the teacher thought would do poorly, the teacher said, "She has attended seminary less than half the time during her high school years…"

Cue 12- Transfer- "He can relate scriptures to real life"

Cue 13- Church experience- This centers on how long the student has been a member of the church. Some teachers judged students to perform worse because they had only recently been baptized (joined the church).

Cue 14- Completion of Assignments- Some teachers considered how often students complete assignments as a cue. Those who consistently turn in assignments were judged to do better. The teachers said things like, “I considered whether or not they are doing their readings.”

Cue 15- Gospel Knowledge- Statistically speaking, most of the students have been raised in the church their whole lives. Consequently, there are some basic gospel principles that are understood by most of the class. Some teachers could tell when a student didn’t meet this threshold of understanding and assumed that this would lead to greater difficulty on the test. The teachers reported things such as, “She lacks knowledge of many of the basics of the gospel,” and “He often has a hard time understanding basic doctrines in class.”

Cue 16- Friends. This focuses on whether or not the teacher deemed the students friends to be good or bad influences.