Teacher Judgment Accuracy of Student Perceptions of Closeness and Conflict in Teacher-Student Relationships

Collin Seastrand
Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Education Commons

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/etd/9688

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Teacher Judgment Accuracy of Student Perceptions of Closeness and Conflict in Teacher-Student Relationships

Collin Seastrand

A thesis submitted to the faculty of Brigham Young University in partial fulfillment of the requirements for the degree of Master of Arts

Richard D. Osguthorpe, Co-chair
Melissa Newberry, Co-chair
Keith Thiede
Sarah Clark

Department of Teacher Education
Brigham Young University

Copyright © 2022 Collin Seastrand
All Rights Reserved
ABSTRACT

Teacher Judgment Accuracy of Student Perceptions of Closeness and Conflict in Teacher-Student Relationships

Collin Seastrand
Department of Teacher Education, BYU
Master of Arts

Teacher decisions are largely influenced by the judgments that they make regarding their students’ states of mind. Thus, it is important for teachers to be accurate in their judgments. The current study investigated the teacher judgment accuracy of student perceptions of the teacher-student relationship. In a Western state of the United States, 40 teachers used a prediction form to predict how their 4th – 6th grade students would rate the closeness and the conflict of their teacher-student relationship via the Student Perception of Affective Relationship with Teacher Scale (SPARTS, Koomen & Jellesma, 2015). Students then took the survey, and teacher predictions and student reports were run through bivariate correlations and t-tests for analysis.

Teacher predictions and student reports had a mean correlation coefficient of .31 for closeness and .39 for conflict. Nine of the 40 teachers had a negative correlation for closeness, conflict, or both, suggesting that a good percentage of teachers are fairly poor at judging how their students perceive their teacher-student relationship. Nonetheless, the t-tests confirmed that on average, teacher judgment accuracy was significantly better than random. The results of this study suggest that there is a need for intervention to help teachers become more accurate judges of how their students perceive the teacher-student relationship. As teacher judgments impact teacher decisions, accurate judgments of the student perceptions of the relationship might allow for teachers to make decisions that would more effectively nurture positive teacher-student relationships.

Keywords: teacher judgment, judgment accuracy, teacher-student relationships
ACKNOWLEDGMENTS

My journey to the end of this project would not have happened if not for the continuous support and encouragement of my committee co-chairs, Dr. Richard Osguthorpe and Dr. Melissa Newberry. Both of them took me under their wing and did everything they could to keep me moving forward. I would also like to thank the other members of my committee for their help and feedback throughout this process, Dr. Sarah Clark, and Dr. Keith Thiede.

I also would like to express my love and gratitude to my wife, Kristen Seastrand, who consistently encouraged me throughout this process, and who spent countless hours tending to our two children, Ryan and Reggie, so that I could spend the necessary time on this project.

Finally, I’d like to acknowledge my father, Dr. Gary Seastrand, for encouraging me to pursue a graduate degree and advance my knowledge in the field of education. His example as a dedicated educator and unceasing learner is an inspiration to me and many others.
TABLE OF CONTENTS

TITLE PAGE ................................................................................................................................... i
ABSTRACT .................................................................................................................................... ii
ACKNOWLEDGMENTS ............................................................................................................. iii
TABLE OF CONTENTS ............................................................................................................... iv
LIST OF TABLES ......................................................................................................................... vi
LIST OF FIGURES ...................................................................................................................... vii
CHAPTER 1: Introduction ............................................................................................................. 1
   Statement of the Problem .......................................................................................................... 4
   Statement of the Purpose .......................................................................................................... 4
   Research Questions ................................................................................................................... 5
CHAPTER 2: Review of Literature ................................................................................................ 6
   Conceptual Framework: Teacher Judgment Accuracy ............................................................. 6
   The Importance of Teacher-Student Relationships ................................................................. 10
   Quantifying the Quality of Teacher-Student Relationships .................................................... 12
CHAPTER 3: Method ................................................................................................................... 14
   Participants .............................................................................................................................. 14
   Setting ..................................................................................................................................... 16
   Instruments .............................................................................................................................. 17
      The Measure of Closeness ................................................................................................ 17
      The Measure of Conflict ................................................................................................... 18
      The Measure of Teacher Predictions ................................................................................ 18
      The Measure of Judgment Accuracy ................................................................................ 18
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Descriptive Statistics for Closeness</td>
<td>25</td>
</tr>
<tr>
<td>Table 2</td>
<td>Accuracy of Closeness by Grade</td>
<td>28</td>
</tr>
<tr>
<td>Table 3</td>
<td>Descriptive Statistics for Conflict</td>
<td>29</td>
</tr>
<tr>
<td>Table 4</td>
<td>Accuracy of Conflict by Grade</td>
<td>31</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1  Histogram of Relative Accuracy of Closeness ........................................................... 27
Figure 2  Histogram of Relative Accuracy of Conflict .............................................................. 30
Figure 3  Scatter Plot of Closeness vs Conflict for Each Teacher ............................................ 32
CHAPTER 1

Introduction

In the course of instruction, teachers constantly judge and formatively assess student learning in an effort to tailor their instruction to meet the individual needs of the learners. It is self-evident in education that teachers are required to make thousands of judgments, often subconsciously, each day. These judgments influence every part of classroom life, especially instructional decisions that involve interactions with students. For example, teachers use their judgment when deciding how to provide feedback, take disciplinary action, intervene with struggling students, create student groups, and plan and pace lessons.

The literature has referred to teaching as an “interplay between teachers’ judgments of student learning and subsequent instruction” (Thiede et al., 2018, p. 106; see also Box et al., 2015; Donovan et al., 2000; Ready & Wright, 2011). This interplay suggests that accurate teacher judgment of student ability and their ongoing learning positions teachers to make more effective instructional decisions (Südkamp et al., 2012; Thiede et al., 2015), which leads to increased student learning (Thiede et al., 2018). The accuracy with which teachers predict student outcomes is referred to in the literature as “teacher judgment accuracy” (Thiede et al., 2018, p. 106).

If instructional decisions are generally being guided by teacher judgments of student learning (Alvidrez & Weinstein, 1999; Ready & Wright, 2011; Shavelson, 1978), then it is imperative for teachers to make accurate judgments (Goodlad et al., 1990). Shavelson (1978) highlighted the notion that teachers need to be accurate in their judgments regarding their students’ states of mind specifically their cognitive, emotional, and motivational states. In other words, teachers need to be able to accurately judge which students are learning the content and
which are not, and also be able to accurately judge the emotional and motivational states of their students.

Since Shavelson’s work in 1978, research in teacher judgment accuracy has expanded greatly. The majority of teacher judgment accuracy research, however, has focused particularly on teacher judgment of student academic achievement (Hoge & Coladarci, 1989; Südkamp et al., 2012). That is, teacher judgment accuracy research has primarily explored the accuracy with which teachers predict a cognitive state of mind (Shavelson, 1978). For example, a meta-analysis by Südkamp et al. (2012) found that teachers can predict student academic outcomes with a moderate level of accuracy. However, they focused exclusively on analyzing teacher judgment accuracy of student achievement, excluding studies that analyzed other states of mind, such as a student’s motivational state (Givvin et al., 2001), and other emotional student characteristics, such as attention, anxiety, and enjoyment (e.g., Karing, 2009; Pohlmann et al., 2004; Spinath, 2005). Research has also explored ways to increase the level of accuracy with which teachers predict academic outcomes (Thiede et al., 2018), relying on the notion that increased judgment accuracy will bring about better instructional decisions, resulting in increased student learning.

This large and growing line of research on teacher judgment accuracy of students’ cognitive abilities stands in contrast to the paucity of research exploring teacher judgment accuracy of students’ emotional and motivational states (see Urhahne & Wijnia, 2021). While some research has explored teacher judgment accuracy of states of mind related to motivation and emotion, such as learning enjoyment, test anxiety, (Stang & Urhahne, 2018), and self-perception (Lohbeck et al., 2015), there does not appear to be any research that explicitly examines teacher judgment accuracy of the student perceptions of the teacher student relationship.
Exploring teacher judgment accuracy of student perceptions of teacher-student relationships is important, as positive teacher-student relationships are widely accepted as a critical component both of student learning and as a moral obligation for teachers (Hattie, 2009; Noddings, 2013). It is unknown how accurately teachers judge their students’ perceptions of their teacher-student relationship. Inasmuch as accurate judgments of student learning allow teachers to effectively tailor instruction to improve cognitive learning, accurate judgments of how the students feel about the teacher-student relationship similarly would allow teachers to effectively build positive individual relationships with their students.

An interpersonal connection with the teacher is widely accepted as a crucial component of student development (Davis, 2003; Hamre & Pianta, 2001). Positive teacher-student relationships have been extensively researched and are among the top predictors of student growth (Hattie, 2009). Noddings (2013) posits that teachers have a moral obligation to care for their students and should work to nurture a supportive and warm teacher-student relationship. Sabol and Pianta (2012) claimed “a relationship with at least one caring adult, not necessarily a parent, is perhaps the single most important element in protecting young people…and for many children this adult is a teacher” (p. 213). Thus, a positive teacher-student relationship is a means to an end and also an end in itself, and therefore, the exploration of the accuracy with which teachers judge their students’ perceptions of the relationship is a worthwhile objective.

In an effort to more accurately measure the subjective quality of a teacher-student relationship, research has explored and used several constructs that factor into relationship quality such as closeness, conflict, dependency, and negative expectations. In the current study, student perceptions of the teacher-student relationship were measured only through the constructs of closeness and conflict for two reasons. First, they are perhaps the two most
consistently used and widely accepted constructs (Mason et al., 2017; Pianta, 2001; Zee et al., 2017). Second, the child measure used in this study has the greatest reliability in closeness and conflict (Koomen & Jellesma, 2015). Teachers made predictions about how each of their students would report the closeness and conflict that they perceive in their relationship with the teacher, and students were given a survey to report their perceptions of closeness and conflict. Correlations were run between teacher predictions and students results to find the accuracy of teacher judgments. Thus, this study explores the teacher judgment accuracy of student perceptions of closeness and conflict in the teacher-student relationship.

Statement of the Problem

Teacher judgments impact teacher decisions. Therefore, it is important for teachers to be accurate in their judgments. Teacher judgment accuracy has been traditionally used to explore the accuracy with which a teacher can predict student academic performance. Much less explored is teacher judgment accuracy in emotional and motivational contexts, specifically, the teacher-student relationship. If teachers are poor judges of which students perceive their relationship as close or conflicted, then they cannot effectively work on the improvement of their individual relationships with students. The problem is that research has not explicitly explored whether teachers can accurately predict how their students perceive the teacher-student relationship.

Statement of the Purpose

The purpose of this study is to explore how accurately teachers can judge their students’ perceptions of the teacher-student relationship. In this study, teachers were asked to predict how their students would rate the closeness and the conflict that is present in their relationship, and their students were simultaneously given a survey to report their perceptions of both the
closeness and conflict present in their teacher-student relationship. The correlations between teacher predictions and student results address the research questions below. In other words, the purpose of this descriptive study is to determine the teacher judgment accuracy of student perceptions of the teacher-student relationship.

Research Questions

This study will address the following research questions:

1. How accurately do teachers judge their students’ perceptions of closeness in their teacher-student relationship?

2. How accurately do teachers judge their students’ perceptions of the conflict in their teacher-student relationship?
CHAPTER 2

Review of Literature

The accuracy of teacher judgments on student learning has been explored thoroughly (see Hoge & Coladarei, 1989; Südkamp et al., 2012). One paramount finding from the literature is that accuracy of teacher judgments is positively correlated with gains in student achievement (Thiede et al., 2018). That is, higher teacher judgment accuracy is associated with higher student achievement. An explanation for this finding is that teachers who have a better idea of their students’ state of learning are able to tailor their instruction to better address individual needs, resulting in better outcomes (Thiede et al., 2015). The current study is grounded in the claim that it is important for teachers to be accurate judges of their students’ states of mind (Shavelson, 1978). In particular, one concept with ties to emotional, motivational, and cognitive states mind that has yet to be observed through the conceptual lens of teacher judgment accuracy is the students’ perceptions of the teacher-student relationship.

In this chapter I describe the conceptual framework, encapsulate the growing literature on teacher judgment accuracy, and make a connection to the literature on teacher-student relationships. Later, I suggest the need for the present study, which aims to measure teacher judgment accuracy of student perceptions of teacher-student relationships.

Conceptual Framework: Teacher Judgment Accuracy

The present study is positioned inside the conceptual framework of teacher judgment accuracy. Teacher judgment accuracy is defined as the knowledge that a teacher has about different student characteristics such as cognitive abilities, learning motivation, or work and social behaviors (Urhahne & Wijnia, 2021). This framework centers around the importance of teachers being accurate in their judgments of their students’ states of mind in order to make more
informed instructional decisions (Shavelson, 1978). In other words, it is the notion that having a more accurate knowledge or insight into their students’ states of mind should position teachers to be able to make better adjustments to their pedagogy.

Urhahne and Wijnia (2021) explain the importance of teacher judgment accuracy, "Constant adjustments to inter-individual differences between learners help to make the lesson more effective. Based on person-related judgments, teachers change patterns of social interaction and communication and make instructional decisions that benefit the students” (p. 2). Teachers’ judgments of their students’ states of mind impact their ability to make appropriate adjustments to their instruction and pedagogy in order to effectively address individual needs of their students.

The judgments that teachers make about their students “provide essential information for deciding what and how to teach” (Shavelson, 1978, p. 37). Instructional choices are heavily influenced by teacher judgments. Only when a teacher is accurate in their judgments of how well students are learning the material, can they effectively accommodate for these students. For example, if a teacher’s judgments are inaccurate, they might tailor their instruction to address needs that they perceive, but that do not really exist. It stands to reason that if a teacher cannot make these predictions accurately, their decisions on how to adjust and accommodate for individual learner needs would be insufficient at best. The accuracy of these judgments is of great importance if teachers are to tailor instruction or attend to individual student needs.

Shavelson (1978) highlighted three states of mind, cognitive, motivational, and emotional, to explore through teacher judgment accuracy. The vast majority of research on teacher judgment accuracy has had a fairly exclusive focus on academic and cognitive abilities of students. For example, Thiede et al. (2015) explain that teacher judgment accuracy traditionally
has been used to explore the question, “How well can a teacher evaluate the degree to which students have learned the target materials?” (p. 36). Südkamp et al. (2012) conducted a meta-analysis of research on teacher judgment accuracy of student academic achievement, finding a fairly strong correlation ($r = .63$) between teacher predictions and student achievement.

Although a positive correlation exists, there is certainly room for teachers to improve their ability to predict the academic abilities of their students. Their research also highlights various ways in which teacher judgment accuracy may have an effect on students, such as guiding instruction (Alvidrez & Weinstein, 1999), identifying struggling students (Bailey & Drummond, 2006), influencing teachers’ expectations about students’ ability (Brophy & Good, 1986), shaping feedback to students and parents (Hoge & Coldarci, 1989), and influencing students’ academic self-concept (Möller et al., 2009).

While much is known about teacher judgment accuracy of academic outcomes (Machts et al., 2016), it is imperative that research continues exploring teacher judgment accuracy in the lesser explored emotional (Spinath, 2005) and motivational (Givvin et al., 2001) states of mind as well. The present study uses the lens of teacher judgment accuracy to explore the concept of teacher-student relationships, which arguably involve all three states of mind. However, the present study focuses on how students feel about or perceive the relationship, which perhaps is most in line with an emotional state of mind.

Within the concept of teacher judgment accuracy, there are various schools of thought on methodology. The standard method, traditionally used to gather teachers’ judgments of student academic ability, involves showing teachers a test that their students will take, and asking them to predict how their students will perform on it. Then, students take the test and teacher predictions are compared against the student results through a bivariate correlation. Teacher
judgment accuracy is operationalized as the intra-individual correlation between the teachers’ predictions and the actual student performance computed across students (e.g., Helmke & Schrader, 1987).

In measuring teacher judgment accuracy, it is important to account for a teachers’ ability to distinguish differences in student skill levels (Alvidrez & Weinstein, 1999). Helmke and Schrader (1987) suggested computing a correlation for each individual teacher, rather than stacking all teacher data and all student data before analysis. This intra-individual correlation method is advantageous because it shows the accuracy with which a teacher can judge their students’ relative to each other. The current study aims at measuring an individual teacher’s ability to predict their individual students’ outcomes, or who feels they have a strong relationship and who feels they have a weak relationship, so an intra-individual measurement is appropriate. However, even within this intra-individual correlation, there is an important distinction between absolute and relative accuracy.

Judgment accuracy that looks at the degree to which the average predicted performance is related to the average actual performance is called absolute accuracy (Dunlosky & Rawson, 2012), and is also referred to as confidence bias. Teachers normally overestimate students’ performance, resulting in positive confidence bias (Bates & Nettelbeck, 2001; Rausch et al., 2016). Absolute accuracy is simply calculated by subtracting the average actual performance from the average predicted performance. A score closer to zero represents higher absolute accuracy. A teacher with strong absolute accuracy will know how to pace their lessons, as the score is dependent on the class as a whole.

Relative accuracy, on the other hand, describes “the degree to which judgments discriminate between different levels of performance across students” (Thiede et al., 2018, p.
In other words, this type of accuracy refers to understanding where students are in relation to each other. A teacher with good relative accuracy will have a strong idea of which students could benefit from additional support or attention, and which ones might benefit from extension activities. Relative accuracy allows teachers to place attention on the individual students who actually need it. Thiede et al. (2019) summarized,

*absolute accuracy* describes the match between predicted and actual levels of performance – and could inform the pace at which a teacher proceeds through a lesson.

*Relative accuracy* describes how well a teacher discriminates between students with higher levels of understanding and those with lower levels of understanding – and could inform how a teacher tailors instruction to individual students. (p. 682)

With academic achievement, teacher judgment accuracy varies tremendously. In a meta-analysis of 75 articles on accuracy of teachers’ judgment of student learning, accuracy was reported as high as .80 and as low as -.03 (Methe et al., 2008; Graney, 2008, as cited in Südkamp et al., 2012), with a mean correlation of .63. In a similar fashion, the current study establishes the correlation between student perceptions of the closeness and conflict in the relationship and their teachers’ predictions of such.

**The Importance of Teacher-Student Relationships**

An abundance of educational research suggests that the relational and social aspects of the classroom setting serve as a robust predictor of students’ academic achievement (Ansari et al., 2020; Hamre & Pianta, 2001; Lippard et al., 2018; Martin, 2014; Martin & Dowson, 2009; Pianta & Stuhlman, 2004; Roorda et al., 2011; Sabol & Pianta, 2012). The teacher-student relationship is one large factor of the social and relational classroom setting, and as such, is studied extensively. Ultimately, research shows that children who report positive relationships
with teachers are more likely to succeed academically (Hamre & Pianta, 2001; Maldonado-Carreño & Votruba-Drzal, 2011; Spilt et al., 2012).

Hamre and Pianta (2001) found a correlation between reported teacher-student relationship quality in kindergarten, and achievement outcomes eight years later. Their results clearly accentuate the impact that teacher-student relationships can have on school outcomes, even in the distant future. They emphasized that teacher-student relationships have longitudinal effects on students, academically and socially.

Perhaps one of the most well-known works is Hattie’s (2009) synthesis of over 800 meta-analyses relating to student achievement. In his work, he found the teacher-student relationship to have an effect size of 0.72 on student achievement, where a range of 0.15 to 0.4 is representative of growth in a typical year of schooling. This finding suggests that a positive teacher-student relationship can essentially double the amount of growth in a school year. This point is emphasized further in his influence ranking system, where out of the 138 influences on student achievement that he analyzes, teacher-student relationships rank eleventh. Subsequently, Hattie and Yates (2013) claim that relationships between children and teachers are the bedrock of learning, and that without these relationships, learning will not take place.

As well as academic achievement, teacher-student relationships have been shown to be associated with children’s courage to try new things, attempt difficult tasks, and deal with fears (Dunlap, 2002; Graziano et al., 2007). Teacher-student relationships also have been shown to correlate with children’s ability to develop the social competency and emotional skills necessary for motivation and success in school and life (Buyse et al., 2010; Pianta, 1999; Pianta & Stuhlman, 2004).
In addition to empirical support, there is also a philosophical argument for the importance of teacher-student relationships and why teachers should strive to cultivate them. Velasquez et al. (2013) indicates that education often has a “disproportionate focus on the technical aspects of teaching with little or no focus on its ‘human’ aspects, which attend to students as moral beings and members of larger democratic communities” (p. 162). One vital way these human aspects of teaching get addressed is through the care that teachers give their students (Noddings, 2013). The moral theory known as the ethics of care implies that there is moral significance in the fundamental elements of relationships. This theory has been influential in many fields of research, including philosophy, psychology, social theory, professional ethics, and pedagogy. Noddings (1984) highlights the importance of care ethics in the classroom, emphasizing that moral education is crucial to the cultivation of a caring society. She posits that students must be cared for holistically, which positions education as a moral endeavor lined with moral obligations, one in particular being the development of a positive teacher-student relationship (Goodlad et al., 1990; Noddings, 1984). By nurturing and caring for their students, teachers are fulfilling their moral duties (Goodlad et al., 2004). Because of this, it is important for teachers to strive to build positive relationships with each of their students.

In summary, teacher-student relationships are fundamental for education, and because of this, teacher-student relationships are well worth exploring through the lens of teacher-judgment accuracy. For how can a teacher tailor their pedagogy to build a relationship if they aren’t aware of how their students perceive the relationship?

**Quantifying the Quality of Teacher-Student Relationships**

Research has attempted to measure the quality of teacher-student relationships for decades. Koomen and Jellesma (2015) provide an in-depth summary and critique of the history
of instruments aimed at measuring the teacher-student relationship, including the well-known Student-Teacher Relationship Scale (STRS; Pianta, 2001). Throughout the literature, the teacher-student relationship is divided into different domains, all of which contribute to an overall measurement of the quality of the relationship. Primarily, there are three distinct domains: closeness, conflict, and dependency. These domains have been widely accepted as appropriate measures for quantifying the teacher-student relationship, as they attend to primary strengths and weaknesses that exist in the relationship. However, Koomen and Jellesma (2015) also have highlighted the importance of accurately measuring the student’s experience of the relationship, critiqued instruments that do so inadequately, and proposed a new child measure, the Student Perception of Affective Relationship with Teacher Scale (SPARTS; Koomen & Jellesma, 2015). Because the STRS, along with many other instruments, relies on teachers’ perceptions of the relationship, their goal was to create a student instrument that effectively imitated the STRS. Reliability and validity for SPARTS were tested, comparing results to those from the STRS. They found that SPARTS reliably corresponds to its STRS counterpart in measuring closeness and conflict, but not in dependency.

In summary, research suggests that it is important for teachers to be accurate judges of their students’ states of mind, because those judgments guide instructional decisions, which can lead to better educational outcomes. Research also suggests the importance of positive teacher-student relationships. The current study bridges the gap between these two lines of research, highlighting the importance for teachers to be accurate judges of how their students perceive the closeness and conflict in the teacher-student relationship. The current study is the first that attempts to measure the accuracy with which teachers judge those perceptions.
CHAPTER 3

Method

The data for the present study are part of a data set for a larger study focused on analyzing the specific cues that teachers utilize or pay attention to as they predict student outcomes. That study received Institutional Review Board, university, district, and school approval (see Appendix A), and data were collected in Spring of 2021. Parental permission and child assent forms were distributed and signed prior to data collection.

The present study addresses the following research questions:

1. How accurately do teachers judge their students’ perceptions of closeness in their teacher-student relationship?
2. How accurately do teachers judge their students’ perceptions of the conflict in their teacher-student relationship?

In this chapter, the participants, settings, measures, and procedures for the current study are discussed. The current study uses only a certain portion of the data, in particular, data of teacher predictions of their students’ results from the Student Perception of Affective Relationship with Teacher Scale (SPARTS; Koomen & Jellesma, 2015), and the actual students’ SPARTS results. Because of this, only the procedures necessary for the current study are reported. The additional data that were collected as part of the larger cue-utilization study is not necessary for a replication of the present study, so it is not described here. I also provide information on the statistical analysis that will be used in answering the research questions.

Participants

Data for the current study come from 40 elementary classes in a Western state of the United States. The participants came from seven different schools, located within the same
public school district. The mean student population of these seven schools is 866 and the mean reported student-teacher ratio among these seven schools is 26:1, which is the same as the district. These numbers highlight that class sizes are fairly large throughout the district. These schools are geographically spread throughout the district and represent a relative cross section of other elementary schools in the district.

During the time of data collection, 16 teacher participants were fourth grade teachers, 15 teacher participants were fifth grade teachers, and nine teacher participants taught sixth grade. Teacher participants who taught grades four, five, and six exclusively were used for the present study on account of SPARTS being an instrument that was created and tested for reliability specifically with students in grades four through six (Koomen & Jellesma, 2015). Additionally, as part of the larger study, data were collected from five third grade classes. However, that data are not used as part of the current study.

Because participation was voluntary on both a school and teacher level, it was not a random sampling, and it was difficult to control for equal amounts of teachers in each grade. Nevertheless, a perfect balance of participant grade levels was not deemed fundamental to answering the research questions of this study.

Student participants also only came from grades four through six. Not every student in each class participated. Student participation was dependent upon parental consent, student assent, and attendance in class on the day of data collection. There were 788 total participating students with an average of 18.89 participating students per class. Given the average student-teacher ratios, approximately 73% of students in each classroom participated in the study. It should also be noted that some participating teachers had students enrolled in online school, due to the COVID-19 pandemic. However, online students were not included as a part of this study.
Because the study took place in one school district, and in an effort to preserve anonymity, we did not collect individual characteristics for any participants. However, the state publishes district reports of student enrollment for each school, including some of those variables. The sum average student demographics of the seven participating schools are as follows: Students are 48.3% female, and 51.7% male. Student demographics for race are 86.4% White, while 7.9% of students are Hispanic and 5.7% of students are of other or mixed ethnicity. In the whole district, 81% of students are White, 12% are Hispanic, and ~7% are of other or mixed ethnicity. Therefore, the district as a whole is only slightly more diverse than the sample of student participants. Only 15% of the teacher participants were male. The district does not make racial or demographic information of teachers available, so it is not reported here.

**Setting**

There are 61 elementary schools in the participating district, which stretch over a quite large geographical area. The area is predominately White, and also of Christian faith, religiously. The district actively emphasizes a holistic approach to teaching, and encourages teachers to work to build nurturing relationships with their students. For example, the school district is part of a partnership program with a local university. This partnership is dedicated to five commitments: (a) civic preparation and engagement; (b) engaged learning through nurturing pedagogy; (c) equitable access to academic knowledge and achievement; (d) stewardship in school and community; and (e) commitment to renewal. These commitments, nurturing pedagogy in particular, demonstrate the partnership’s emphasis on teacher-student relationships.

The global COVID-19 pandemic also is an important factor in the setting of this study. This district closed schools down in March of 2020 as a safety protocol for the pandemic. In August of 2020, the new school year started with guidelines for face-masks, social distancing,
and sanitation. These guidelines remained in place for the duration of the 2020–2021 school year. This is significant to the study, because data were collected in Spring of 2021, and interpersonal relationships between teachers and students may have been impacted by these guidelines. For instance, close teacher-student relationships might be more difficult to establish when teacher and student cannot see each other’s faces and are required to remain physically distanced from each other.

Instruments

The present study uses the survey items for closeness and conflict from the SPARTS instrument (see Appendix B). This instrument was developed in an effort to reliably assess student perceptions of the teacher-student relationship. In particular, SPARTS was developed with the purpose of exploring whether the characteristics closeness, conflict, and dependency can be used to measure student perceptions, as they are commonly used to measure teacher perceptions, such as in the Student Teacher Relationship Scale (STRS; Koomen & Jellesma, 2015). Koomen and Jellesma (2015) found that SPARTS provided a reliable measure of closeness and conflict. However, dependency was less reliable, and was deemed insufficient for use. Because of this, the current study uses only the SPARTS items that correspond to closeness and conflict, and accepts those two domains as adequate for measuring the quality of the teacher-student relationship from the student’s perspective.

The Measure of Closeness

The SPARTS survey provides statements contributing to an overall closeness score. Each statement is followed by a Likert scale from 1-5 based on agreement with the statement. Each student selects one of the following for each question where 1 = No, that is not true; 2 = That is usually not true; 3 = Sometimes; 4 = That is usually true; 5 = Yes, that is true. Essentially, a 1
indicates low closeness and 5 indicates high closeness. Each student received their own closeness score, calculated as the mean of their reported scores from the eight closeness statements.

**The Measure of Conflict**

The SPARTS survey also provides statements contributing to an overall conflict score. The same format and Likert scale are used here as well. Each student received a conflict score, calculated as the mean of their reported scores from 10 statements. Similar to the closeness scale, a 1 indicates low conflict and a 5 indicates high conflict.

**The Measure of Teacher Predictions**

For reference, teachers were given an exact replica of the survey that students would be taking. They were also given a prediction form with a blank table, with rows designated for individual students, and columns for each teacher prediction, one for closeness and one for conflict (see Appendix B). Each teacher made their predictions according to the same Likert scale that students used. That is, if a teacher thought a certain student to perceive their relationship as having high closeness and low conflict, they might predict that student to score a 5 in closeness and a 1 in conflict.

**The Measure of Judgment Accuracy**

Judgment accuracy is operationalized as the intra-individual correlation between predicted and actual performance computed across the students in a class. Teachers’ predicted scores were compared to actual student mean closeness and conflict scores, analyzing for relative accuracy, and each teacher received an accuracy score for closeness and an accuracy score for conflict. This score is computed with Pearson’s $r$, which is appropriate because it measures a linear correlation between two sets of data. This means the closer the correlation is to +1.0, the more accurate the teacher’s predictions. The closer the correlation is to -1.0, the less accurate the
teacher’s predictions. In other words, a teacher who perfectly predicts the rank order of students reporting high, medium, and low closeness would receive a judgment accuracy score of +1.0. This is also how relative accuracy is calculated in the literature (see Thiede et al., 2019).

For general comparison, teacher judgment accuracy of academic student learning varies widely across studies, but Südkamp et al. (2012) reported the average correlation as .63, which is a high strength correlation according to Cohen (2013), but also shows room for improvement.

 Procedures

As stated previously, a larger study collected teacher predictions and student results for closeness, conflict, math, and reading assessments, along with teacher-reported cue utilization for making their predictions. Reported here are only the procedures relevant to my study, which only includes teacher predictions and student results of closeness and conflict.

 Recruitment

All elementary administrators in the local district were contacted through email in an effort to find schools with willing and able teacher participants in grades four through six. Principals who were interested then gauged teacher interest in their schools, and reported back to the research team. Teachers who were potentially interested (n > 60) attended an online orientation meeting via Zoom to get more details and confirm interest. In this meeting, I, along with other researchers, explained the purpose and goals of the study, and described the procedures of the study to them. Teachers were informed that a researcher would be coming into their classroom, providing them with a copy of the student survey, inviting them to leave the room to make predictions of their students’ SPARTS results. It was explained that SPARTS measures how the child perceives the closeness and the conflict of their teacher-student relationship, and that the researcher would lead the students through the survey. Teachers were
also informed that they would be compensated for their participation with a $100 prepaid VISA gift card, funded by an internal grant at the research team’s university, as approved by the IRB. Potential participating teachers had opportunities to ask questions, and in the end, there were 43 teachers belonging to seven different schools who remained interested and were emailed parental consent forms to send home with their students prior to data collection.

Subsequently, I worked with those teachers and other participating researchers to coordinate schedules and visits to the seven participating schools for data collection. Each individual teacher was scheduled for a specific day and time during an instructional period of their choice that aligned with researcher schedules. In most cases, we were able to coordinate the schedules to allow researchers to dedicate one day per school, where, depending on the number of participating teachers and overlapping schedules, one, two, or three researchers would show up and collect data in one class before moving to the next.

**Data Collection**

Data collection lasted about an hour in each classroom, including introductions, explanations, and administration of the surveys. As mentioned earlier, there were additional surveys administered that pertain to the larger study, but are not utilized in the present study. Only students who had returned the parental consent form participated in the study. All teacher and student participants also filled out an assent form on the day the data was collected. Students who did not participate in the study were simply engaged in other material unrelated to the study, provided by their teacher.

For data collection, a researcher entered a participating teacher’s room, and the teacher introduced the researcher to the students and reminded the students of their elected participation in the study. The researcher spent a few minutes individually with the teacher, assuring the
teacher’s understanding of the overall process and the different forms, including how to record their predictions. The researcher then provided the teacher with an exact copy of the student instrument, along with the teacher prediction form. The teacher then invited non-participating students to begin other student work, turned time over to the researcher, and left the room to record their predictions.

The researcher then further explained the study and the process to the students and passed out child assent forms. Students were allowed to asked questions before filling out their assent forms. They were also informed that they could withdraw from the study at any point. After assent forms were collected, the researcher began the survey with those participating. Each student received a paper consisting of the closeness and conflict SPARTS questions and were asked to put their name on their paper in order to accurately compare the teacher predictions to the student answers. Students were instructed to follow along with the researcher, to not go ahead, and to raise their hand if they had questions or needed more time along the way. The researcher read each question aloud, clearly and slowly, pausing to allow students time to select their answer. This process was repeated with all survey questions.

After all questions were answered, student response forms were collected, and the classroom teacher was invited back into the room. Teacher prediction forms were also collected and kept separate from forms of other teachers and students. That is, each teacher’s prediction form was kept with the response forms of their students. Students and teachers were aware that the method for collecting the forms would guarantee that teachers would never see any of their students’ responses, thus preventing student risk and allowing for honesty in student answers.

Hard copy data are stored in a locked cabinet in a locked office, and the de-identified data were entered into an electronic database that is kept on private computers of the project.
personnel that are password protected. The data will be kept for seven years following the completion of the study.

**Research Design**

Overall, this is a descriptive study. The goal is to describe the accuracy with which teachers can predict student perceptions of closeness and conflict in the teacher-student relationship. In the context of data collection, a survey research design was utilized with the student participants. They were asked to answer a series of questions derived from the closeness and conflict scales of the SPARTS assessment, which assesses student perceptions of the teacher-student relationship. This is an appropriate design because it gathers large amounts of data on students’ perceptions of the teacher-student relationship. In order to answer the research questions, it is necessary to administer a reliable survey that collects student perceptions of closeness and conflict in the teacher-student relationship in such a way that teachers can attempt to predict their results.

In contrast, teachers were not asked to answer any survey questions on themselves, but rather to predict the results of the student survey for each individual participating student. This design has been used before in the teacher-judgment accuracy literature (see Thiede et al., 2015). The correlation between their predictions and the actual student responses is what provides the grounds for the descriptive nature of the overall study.

**Data Analysis**

Being a descriptive study, descriptive statistics are provided on teachers’ predicted closeness scores, students’ actual closeness scores, and relative accuracy of closeness scores. The same are reported for conflict. Relative accuracy scores are a result of running a bivariate correlation between predictions and actual results of both closeness and conflict. This is
appropriate because it provides a ratio of covariance of the two variables and the product of their standard deviations. The result is a value between -1 and 1. This analysis is appropriate because it finds the linear correlation, or in other words, finds how correlated teacher predictions are with student outcomes. If a teacher predicts the student responses correctly, $r$ is closer to 1. If there is no correlation, or in other words, no better than random, $r$ is closer to 0. Backwards predictions result in a negative score, where the closer it is to -1, the more wrong the prediction. This is the best available analysis to measure relative teacher judgment accuracy (Thiede et al., 2018).

Additionally, t-tests were run to evaluate whether teachers mean accuracy scores were significantly greater than 0, or in other words, better than random guesses. Histograms were created to show the spread of teacher judgment accuracy for both closeness and conflict.

It is necessary to note two abnormalities related to the participants in the data. The first is that one teacher had incomplete predictions in their predicted conflict scores, and so their conflict data as a whole were omitted from the results. For this reason, there appears to be one less data point in the results for conflict than for closeness. The second abnormality is due to one of the teacher participants teaching two homeroom classes, which were reported as separate classes. Because of this, there appears to be one extra data point for both closeness and conflict. To summarize, though there were 40 total teacher participants, the closeness data are reported as if there were 41 teachers, and the conflict data are reported as if there were 40 teachers.
CHAPTER 4

Results

The present study aims to describe the accuracy with which teachers can predict their students’ perceptions of the closeness and conflict in their teacher-student relationships. In order to answer the two research questions, teachers made predictions about how their individual students would respond to survey items contributing to both a closeness score and a conflict score representative of the quality of their teacher-student relationship while students responded to the survey items. Bivariate correlations were run to see the correlation between predicted and actual results. That correlation is presented as the relative accuracy of the teachers. T-tests were also conducted to evaluate whether or not their mean accuracy scores were significantly related to each other. Additionally, histograms are included in the results to show the spread and range of teacher judgment accuracy of both closeness and conflict.

As stated in the participants section, it is worth noting that one of the 40 teachers had incomplete predictions for conflict, so for conflict data, their data is omitted and there is one less participant. Additionally, one teacher reported scores for two separate homeroom classes. Thus, although there were 40 total teacher participants, the data is reported here as if there were 41 teachers contributing to closeness data, and 40 teachers contributing for conflict data.

Teacher Judgment Accuracy of Closeness

The first research question was: How accurately do teachers judge their students’ perceptions of closeness in their teacher-student relationship? In order to answer this question, each teacher predicted how each of their students would rate the closeness of their relationship with their teacher, through the nine-item Closeness Scale on the Student Perception of Affective Relationship with Teacher Scale (SPARTS; Koomen & Jellesma, 2015). Their predictions were
according to a 5-point scale, where 1 indicated low closeness and 5 indicated high closeness. For each teacher, a mean predicted value of closeness was computed across students in the class, and is reported in the top row of Table 1.

The actual closeness scores are the scores provided by the students in each class. For each teacher, mean actual closeness was also computed across the students in their class. That is, the mean item score was computed across the nine items of the closeness scale for each student. The mean actual score was then computed across the nine items of the closeness scale for each student and the mean actual score was then computed across the students in the class for each teacher. The mean actual closeness was then computed across all teachers and is reported in the second row of Table 1.

**Table 1**

*Descriptive Statistics for Closeness*

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Closeness</td>
<td>3.89</td>
<td>.52</td>
</tr>
<tr>
<td>Actual Closeness</td>
<td>3.59</td>
<td>.27</td>
</tr>
<tr>
<td>Relative Accuracy Closeness</td>
<td>.31</td>
<td>.27</td>
</tr>
</tbody>
</table>

Relative Accuracy was operationalized as the intra-individual correlation between predicted and actual ratings of closeness. An accuracy score was computed for each teacher. The mean accuracy for closeness was then computed across all teachers and is reported in the bottom row of Table 1.

To evaluate whether teachers can accurately predict students’ perceptions of closeness in the teacher-student relationship, a t-test was conducted to compare mean accuracy for closeness ($r = .31$) to zero. This test showed that the mean accuracy for closeness was significantly greater
than zero, $t(40) = 7.42, p < .001$. Thus, teachers, on average, were able to predict students’ perceptions of closeness better than chance.

Although it is important that on average, teachers had significant accuracy for closeness, it is also important to examine the range of accuracy, as it helps to understand how the mean relates to the spread. Figure 1 shows the spread of the accuracy of closeness for all teachers. Accuracy for closeness ranged from .77 to -.40, with five teachers having negative accuracy scores. A negative accuracy score is the result of a teacher predicting higher closeness scores from students who actually gave lower scores, and predicting lower closeness scores from students who gave higher scores. For reference, Cohen (2013) established benchmarks for the magnitude or strength of the correlations. Correlation coefficients between .00 and .10 are considered very small or insubstantial, between .10 and .30 are considered low, between .30 and .50 are considered moderate, between .50 and .70 are high or strong, between .70 and .90 are very high, and above .90 is considered almost perfect.
In an attempt to more thoroughly analyze the spread of the data, exploratory analyses were conducted to examine accuracy specifically by grade level taught, which was the only demographic data collected from the teachers. Mean Accuracy for closeness by grade is reported in Table 2. The comparison of accuracy for closeness across grades reveals that accuracy did not differ significantly across grades, $F(2, 38) = 1.15, p = .33$. 
Table 2

*Accuracy of Closeness by Grade*

<table>
<thead>
<tr>
<th>Grade</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourth</td>
<td>17</td>
<td>.30</td>
<td>.19</td>
</tr>
<tr>
<td>Fifth</td>
<td>15</td>
<td>.25</td>
<td>.36</td>
</tr>
<tr>
<td>Sixth</td>
<td>9</td>
<td>.42</td>
<td>.20</td>
</tr>
</tbody>
</table>

**Teacher Judgment Accuracy of Conflict**

The second research question was: How accurately do teachers judge their students’ perceptions of conflict in their teacher-student relationship? In order to answer this question, each teacher predicted how each of their students would rate the conflict of their relationship with their teacher, through the ten-item Conflict Scale on the SPARTS. Their predictions were according to a 5-point scale, where 1 indicated low conflict and 5 indicated high conflict. For each teacher, a mean predicted value of conflict was computed across the students in the class, and is reported in the top row of Table 3.

The actual conflict scores come from the scores provided by the students in each class. For each teacher, mean actual conflict was computed across the students in their class. That is, the mean item score was computed across the 10 items of the conflict scale for each student, and then the mean actual score was then computed across all students in the class for each teacher. The mean actual conflict was then computed across the teachers and is reported in the second row of Table 3.
Relative accuracy was operationalized as the intra-individual correlation between predicted and actual ratings of conflict. An accuracy score was computed for each teacher. The mean accuracy for conflict was computed across all teachers and is reported in the bottom row of Table 3.

To evaluate whether teachers can accurately predict their students’ perceptions of conflict in the teacher-student relationship, a t-test was conducted to compare mean accuracy for conflict ($r = .39$) to zero. This test showed that the mean accuracy for conflict was significantly greater than zero, $t(39) = 8.79, p < .001$. Thus, on average, teachers were able to predict students’ perceptions of the conflict in the teacher-student relationship better than chance.

Although it is important that on average teachers had significant accuracy for conflict, it is also important to examine the range of accuracy, as it paints a better picture of how the sample compares to the mean. Figure 2 shows the spread of the accuracy for all teachers. Accuracy for conflict ranged from .90 to -.26, with five teachers having negative accuracy scores. Negative accuracy scores result from a teacher predicting higher conflict scores for students who actually gave lower scores, and predicting lower conflict scores for students who gave higher scores. Cohen’s (2013) benchmarks for magnitude of correlation are described in the previous section.
In an attempt to more thoroughly analyze the spread of the data, exploratory analyses were conducted to examine accuracy specifically by grade level taught, which was the only demographic data collected from the teachers. Mean accuracy for conflict is reported in Table 4. The comparison of accuracy for conflict across grades reveals that accuracy did not differ significantly across grades, $F(2, 37) = 0.10, p = .91$. 

Figure 2

*Histogram of Relative Accuracy of Conflict*
Table 4

Accuracy of Conflict by Grade

<table>
<thead>
<tr>
<th>Grade</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourth</td>
<td>16</td>
<td>.42</td>
<td>.33</td>
</tr>
<tr>
<td>Fifth</td>
<td>15</td>
<td>.38</td>
<td>.24</td>
</tr>
<tr>
<td>Sixth</td>
<td>9</td>
<td>.37</td>
<td>.29</td>
</tr>
</tbody>
</table>

Combined Teacher Judgment Accuracy

Teacher judgment accuracy of student perceptions of closeness and conflict varied greatly across teachers. Overall, accuracy scores ranged from .90 to -.40. Although the research questions present closeness and conflict as separate constructs, further exploratory analysis revealed variability between closeness and conflict at the teacher level. That is, for the average teacher, high accuracy for closeness didn’t necessarily predict high accuracy for conflict, and vice versa. Figure 3 shows this relationship between closeness accuracy and conflict accuracy at the teacher level, as each point on the plot is representative of one teacher.
There were only two teachers who had a negative relative accuracy in both closeness and in conflict ($r < 0$). On the other extreme, there was only one teacher whose predictions for closeness and conflict were both strongly correlated with student reports. Contrarily, there were also teachers with high correlations in one construct and low correlations in the other. For example, one teacher’s closeness predictions were strongly correlated with student survey results ($r = .74$), but their conflict predictions were negatively correlated ($r = .04$). All in all, teacher accuracy for closeness and teacher accuracy for conflict have a low strength correlation with each other ($r = .28$).

In conclusion, there is a positive correlation between teacher predictions and student perceptions of both closeness ($r = .31$) and conflict ($r = .39$). The correlations for both constructs are only of moderate strength, yet they are statistically significant. Lastly, there is only a low correlation between accuracy of conflict and accuracy of closeness among teachers ($r = .28$).
CHAPTER 5

Discussion

Teacher decisions are heavily influenced by the judgments that they make regarding their students’ states of mind. Accordingly, accurate judgments about students’ states of mind should allow for more effective teacher decisions such as pedagogical or interactional adjustments (Thiede et al., 2018). In order to provide students with the specific support they may need, teachers should first have an accurate understanding of which students need what type of help. The concept of teacher judgment accuracy underscores the importance for teachers to be accurate in their judgments about their students’ cognitive, emotional, and motivational states of mind, and looks to quantify the accuracy of such judgments. Through the lens of teacher judgment accuracy, a relatively unexplored emotional state of mind is that of the teacher-student relationship. The present study addresses this gap in the literature by exploring the following research questions:

1. How accurately do teachers judge their students’ perceptions of closeness in their teacher-student relationship?
2. How accurately do teachers judge their students’ perceptions of the conflict in their teacher-student relationship?

It is worth noting that this study relies on a clear distinction between teachers’ perceptions of the relationship and teachers’ predictions of student perceptions of the relationship. While a teacher’s perception of a relationship may be a factor in their predictions, it is not necessarily proxy for their predictions of student perceptions. For example, a teacher may perceive a student to be overly needy or bothersome, thus elevating the level of conflict in the teacher’s perception of the relationship. However, that teacher may very well still predict that
same student to perceive absolutely no conflict at all. This distinction between teachers’ own perceptions and teachers’ predictions of student perceptions is crucial in understanding how this study differentiates from existing research. In summary, this is a descriptive study focused on determining the accuracy of the judgments teachers make about how their students perceive their teacher-student relationship.

**Findings**

The range of accuracy for both closeness and conflict were relatively large. This considerable variability across teachers suggests that some teachers are quite accurate in their judgments, while others are not accurate at all. The majority of teachers’ accuracy scores, however, were above 0. In other words, most teachers were able to predict student perceptions better than chance. Combined with the *t*-tests, this finding suggests that on average, teachers are able to judge with at least some accuracy how their students perceive the closeness and conflict in their teacher-student relationship.

Another noteworthy finding came about from the comparison between the means of teacher prediction scores and student perception scores presented in Tables 1 and 3. Teacher prediction scores as whole were relatively close to student results for closeness, and even closer in conflict. This may suggest that teachers have a decent idea of how their students view their relationship with the teacher as a whole class, but perhaps tend to struggle more on the intra-individual level. In other words, teachers might be better at judging how their class views the relationship than how each of their students view the relationship.

Also noteworthy is the fact that for both closeness and conflict, the mean of teacher prediction scores was higher than the mean of actual student scores. This was especially true for closeness, suggesting that on average, teachers were more likely to overestimate the perceived
closeness of their students than the perceived conflict. This may have ties to self-serving bias theories (Shepperd et al., 2008), which highlights the tendency for people to hyperfocus on what they do right, often ignoring or justifying what they do wrong. This would be understandable within this context, as it may be difficult for teachers to admit or even accept that certain children view their relationship as conflicted. It also may be common for teachers to justify the conflict by assuming the relationship is closer or stronger than it really is.

It was unanticipated that there would be as many negative accuracy scores among teachers. Overall, there were two teachers with negative accuracy for both closeness and conflict. Additionally, six separate teachers had negative accuracy in one category or the other. That is, the students they predicted high closeness or conflict scores from actually reported low closeness or conflict scores and vice versa. These data points stand out as something that could certainly be seen as problematic and worrisome. A teacher who mistakenly thinks that a certain student views their relationship as close is not likely to make interactional adjustments to strengthen that relationship, when in fact that relationship is one in particular that needs strengthening.

Furthermore, although most teachers had positive correlation scores, it is worth noticing that the majority of those teachers only had a moderate or weaker accuracy score for both conflict and closeness. Interestingly, teachers were generally better at predicting student perceptions of conflict than closeness. Almost twice as many teachers (n = 17) had at least a strong correlation for conflict ($r > .50$), when compared to the strong correlations for closeness (n = 9). The mean accuracy score was higher for conflict, along with the minimum and maximum accuracy scores. This all suggests that teachers are more accurate when judging student perceptions of conflict than closeness. This could possibly be due to conflict potentially standing out more to teachers as they consider the relationships with all their students.
Finally, it is important to discuss the mean relative accuracy scores which directly answer the research questions. Teacher predictions of closeness were moderately correlated with actual results for closeness \( r = .31 \). A moderate correlation was also found between teacher predictions and actual results for conflict \( r = .39 \). T-tests revealed statistical significance between prediction scores and actual scores in both closeness and conflict. While there is significance, the moderate correlations suggest that on average, teachers are only somewhat accurate in their judgments of student perceptions of the teacher-student relationship. As mentioned in previous chapters, and for comparison, Südkamp et al. (2012) established .63 as the average correlation between teachers’ predictions and student results for academic achievement. This fits the existing literature which suggests that teachers are quite better at predicting academic outcomes than states of mind that are more emotional or motivational (see Urhahne & Wijnia, 2021).

**Limitations**

One limitation of the study is the age of the students who took the survey. There was a chance that some participating students might not be on adequate reading levels to fully understand the survey questions. In an effort to minimize the effects of this limitation, time was spent in each classroom to read each survey question out loud, pausing to assure understanding from the student participants before allowing them time to select each answer. This study is also limited in its scope due to the use of self-reported data, which has been critiqued as posing potential threats to internal and external validity. Additionally, the sample was not particularly diverse. All participants were from one localized public school district. Another limitation is that the sample was not truly random. Because principals and teachers had to opt in, there is a chance that participation implies interest in teacher-student relationships, which perhaps could influence
results. Another limitation is found in that only closeness and conflict were used as measures of the teacher-student relationship, when the literature also recognizes additional constructs. Another limitation is that the entire study was presented and conducted only in English, including the orientation meeting, consent and assent forms, and the Student Perception of Affective Relationship with Teacher Scale (SPARTS; Koomen & Jellesma, 2015). Any participants without English fluency may have misinterpreted certain aspects of the study.

Controls were made for the effects of the Hawthorne effect in this study by asking each teacher to leave the room while the students took the surveys. It was clarified for the students that their teacher would not be seeing their results, and that there was no specific desired result other than their honest answers.

It is also worth noting that this analysis ignores different class sizes. Not every student in each class participated, and each class had a different number of participating students. This is significant because these factors may have influence on the accuracy score of the teachers. For example, it could be relatively easier to accurately predict the rank order of a smaller number of students in the class than a larger number. For example, a teacher predicting 12 student responses might more easily show higher judgment accuracy than a teacher predicting 30 student responses, simply due to less permutations.

**Implications**

As stated before, there was great variability in the judgment accuracy scores for both closeness and conflict. There could be many potential factors of this variability, some of which may incite further research. For example, it would be interesting to explore the effects of more specific demographic information. It is possible that accuracy scores are affected by years of experience, race, gender, and how such characteristics compare to student characteristics. By
exploring more specific demographic variables, perhaps patterns might emerge providing insight into what might lead to increased judgment accuracy of the teacher-student relationship. Although teachers are better than random at predicting student perceptions of the relationship, .39 and .31 are only moderate correlations. That means the field is ripe for exploring ways to increase such teacher judgment accuracy in this area. For example, Thiede et al. (2018) found that academic judgment accuracy could be increased through purposeful professional development programs. Perhaps there are similar or different ways to increase teacher judgment accuracy of student perceptions of the teacher-student relationship. This topic might also be worth studying in a longitudinal setting. It is possible that teachers’ judgment accuracy increases with experience, as they become more familiar with the nuances involved in the teacher-student relationship.

A related implication for future research would be to explore the cues or characteristics that teachers pay attention to as they make their judgments or predictions. It is possible that there may be noticeable trends for the teachers with high accuracy scores and for the teachers with low accuracy scores. For example, it is feasible that teachers who pay specific attention to body language of students or nuances of conversation during lessons are more aware of how their students feel about their teacher-student relationship.

During data analysis, a curiosity arose surrounding the confidence with which teachers predict student outcomes. Were the teachers with low accuracy or high accuracy confident in their judgments? How would that factor into their actual accuracy? This could be further explored not only in the context of closeness and conflict of relationships but in all domains of teacher judgment accuracy research.
There are implications for practicing teachers and policy makers as well. For example, perhaps there should be more focus or training for educators to not only build positive teacher-student relationships, but also to learn how to point out the relationships that may need additional support. As stated earlier, it seems teachers are better at judging the relationship quality of the class as a whole than on an intra-individual level. It is possible that explicit training or professional development could help teachers learn how to pay better attention to cues that are truly diagnostic of the student perceptions. This study could also encourage a discussion among educators about their relationships with their students, especially as perceived by the students. Conversations might be had regarding the importance of accurate teacher judgments, especially when judging the closeness and conflict that the students perceive in the relationship. Such dialogue could inspire and motivate teachers to explore strategies that might increase their teacher judgment accuracy, understanding that more accurate judgments might allow them to make more effective decisions and adjustments.

Conclusions

This study explored the teacher judgment accuracy of how students perceive the closeness and conflict in the teacher-student relationship. This study found that on average, teachers can predict student perceptions of closeness and conflict in the teacher-student relationship with moderate accuracy, and that they are statistically better than random. However, the mean correlations, the substantial variability, and the fact that some teachers had negative correlations all suggest that there is considerable room for improvement in this regard.

This study did not consider specific demographic factors, and it is suggested that future research does so. It is also suggested that future research explores ways to increase the judgment accuracy of teachers in regards to judging how their students perceive the teacher-student
relationship. Educators are encouraged to place additional emphasis on paying attention to how their students might be perceiving the relationship. This understanding should allow teachers to more effectively adjust interactional and pedagogical decisions, which may lead to increased relationship quality.
REFERENCES


Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 328-375). Macmillan. https://hdl.handle.net/123456789/8681


https://doi.org/10.4324/9781315885025

https://doi.org/10.1016/0742-051x(87)90010-2

https://doi.org/10.3102/00346543059003297


APPENDIX A

Institutional Review Board Approval Letter/Consent Forms

Memorandum

To: Richard Oguthone
Department: BYU - EDUC - Center for Improvement of Teacher Ed. & Schooling
From: Sandie Ana, MPA, HRPP Associate Director
       Wayne Larsen, MAcc, IRB Administrator
       Bob Ridge, Ph.D., IRB Chair
Date: December 09, 2020
IRB#: IRB2020-157
Title: Teacher Judgment Accuracy of Student Perceptions of Nurturing Pedagogy

Brigham Young University’s IRB has approved the research study referenced in the subject heading as expedited level, category 7.

The approval period is from 12/09/2020 to 12/08/2021. Please reference your assigned IRB identification number in any correspondence with the IRB. Continued approval is conditional upon your compliance with the following requirements:

1. A copy of the approved informed consent statement and associated recruiting documents (if applicable) can be accessed in IRIS. No other consent statement should be used. Each research subject must be provided with a copy or a way to access the consent statement.
2. Any modifications to the approved protocol must be submitted, reviewed, and approved by the IRB before modifications are incorporated in the study.
3. All recruiting tools must be submitted and approved by the IRB prior to use.
4. In addition, serious adverse events must be reported to the IRB immediately, with a written report by the PI within 24 hours of the PI’s becoming aware of the event. Serious adverse events are (1) death of a research participant; or (2) serious injury to a research participant.
5. All other non-serious unanticipated problems should be reported to the IRB within 2 weeks of the first awareness of the problem by the PI. Prompt reporting is important, as unanticipated problems often require some modification of study procedures, protocols, and/or informed consent processes. Such modifications require the review and approval of the IRB.
6. A few months before the expiration date, you will receive a prompt from IRIS to renew this protocol. There will be two reminders. Please complete the form in a timely manner to ensure that there is no lapse in the study approval. Please refer to the IRB website for more information.

Instructions to access approved documents, submit modifications, report complaints, and adverse events can be found on the IRB website under IRIS guidance: https://irb.byu.edu/iris/training-resources.
Memorandum

To: Richard Osguthorpe  
Department: BYU - EDUC - Center for Improvement of Teacher Ed. & Schooling  
From: Sandee Aina, MPA, HRPP Associate Director  
Wayne Larsen, MAcc, IRB Administrator

Date: February 19, 2021  
IRB#: IRB2020-157  
Title: Teacher Judgment Accuracy of Student Perceptions of Nurturing Pedagogy

Brigham Young University’s IRB has reviewed the amendment submitted to change study personnel. The IRB determined that the amendment does not increase risks to the research subject and the aims of the study remain as originally approved. The amendment has been approved. The revised consent statement and recruiting script have been approved and stamped for your files.

The approval of this protocol expires on 02/18/2022. All conditions for the continued approval period remain in effect. Any modifications to the approved protocol must be submitted, reviewed, and approved by the IRB before modifications are incorporated in the study.
Child Assent  Teacher Judgment of Nurturing Pedagogy Study

What is this research about?
My name is Rich Osoguthrope, and I am a professor at Brigham Young University. I want to tell you about a research study I am doing. A research study is a special way to find the answers to questions. We are trying to learn more about the relationship between you and your teacher and how that relationship connects to your academic achievement. You are being asked to join the study because your teacher volunteered to participate.

If you decide you want to be in this study, this is what will happen.
1. You will answer 7 very brief questions on a paper about your teacher
2. You will read a short passage and answer 5 questions about that passage;
3. You will respond to 10 short math problems

Can anything bad happen to me?
No, your teacher will not see your answers, and your answers will not be a part of your grade. You will lose instructional time, but your teacher will account for this loss in future instruction.

Can anything good happen to me?
We don’t know if being in this study will help you. But we hope to learn something that will help other people some day.

Do I have other choices?
You can choose not to be in this study.

Will anyone know I am in the study?
We won’t tell anyone you took part in this study. When we are done with the study, we will write a report about what we learned. We won’t use your name in the report.

What happens if I get hurt?
Your parents/guardians have been notified that you will be participating in the study, and your teacher will be there to help you, if needed.

What if I do not want to do this?
You don’t have to be in this study. It’s up to you. If you say yes now, but change your mind later, that’s okay too. All you have to do is tell us. Before you say yes to be in this study, be sure to ask Rich Osoguthrope to tell you more about anything that you don’t understand.

If you want to be in this study, please sign and print your name.

Name (Printed): __________________________ Signature __________________________ Date: __________________
Parental Permission for a Minor

Introduction
My name is Rich Osuttorpe. I am a professor from Brigham Young University. I am conducting a research study about teacher judgment accuracy of nurturing pedagogy. I am inviting your child to take part in the research because (he/she) is in a school that volunteered to participate.

Procedures
If you agree to let your child participate in this research study, your child will be asked to do the following tasks:

- Complete a 10-item math test (15 minutes)
- Complete a 5-item reading test (15 minutes)
- Complete a teacher relationship survey (20 minutes)
- This will take place in their regular classroom as part of the teacher’s scheduled curriculum.

Risks
There is minimal risk associated with participating in this study. Your child is taking tests that are commonly used in school. However, your child may be disappointed that he or she didn’t do as well as expected. Performance on the tests does not affect your child’s grade and your child’s performance will be kept confidential. We will combine your child’s scores with others when we analyze performance, so no one will know how your child did as an individual.

There is a risk of loss of privacy, which the researcher will reduce by not using any real names or other identifiers in the written report. The researcher will also keep all data in a locked file cabinet in a secure location and/or in a secure location in the cloud. Only the researcher will have access to the data. At the end of the study, data will be retained for three years.

There will also be a brief loss of instructional time, which the teacher will try to minimize through additional instruction on subsequent days.

There may be some discomfort caused by being asked some of the questions. You child may answer only those questions that your child wants to, or you child may stop the entire process at any time without affecting his/her standing in school or grades in class.

Confidentiality
The research data will be kept in a secure location (or password protected and encrypted) and only the researcher will have access to the data. At the conclusion of the study, all identifying information will be removed and the data will be kept in a locked cabinet or office for period of three years.

Benefits
There are no direct benefits for your child’s participation in this project.
Compensation
There will be no compensation for student participation in this project.

Questions about the Research
Please direct any further questions about the study to Rich Osguthorpe at 801-422-6064.

Questions about your child's rights as a study participant or to submit comment or complaints about the study should be directed to the IRB Administrator, Brigham Young University, A-285 ASB, Provo, UT 84602. Call (801) 422-1461 or send emails to irb@byu.edu.

You have been given a copy of this consent form to keep.

Participation
Participation in this research study is voluntary. You are free to decline to have your child participate in this research study. You may withdraw your child's participation at any point without affecting your child's grade/standing in school, treatment, or benefits, etc.

Child's Name: ________________________________

Parent Name: ____________________ Signature: ____________________ Date: __________
Consent to be a Research Subject

Introduction
This research study is being conducted by Rich Osguthorpe at Brigham Young University and Keith Thiade at Boise State University to determine teacher judgment accuracy of nurturing pedagogy. You were invited to participate because your school principal gave permission to extend this invitation.

Procedures
If you agree to participate in this research study, the following will occur at your school and in your classroom:

- review the purpose of the study (15 minutes);
- meet with the researchers for a brief orientation (20 minutes);
- provide the researchers with the opportunity to administer the care assessment to each student (20 minutes of class time);
- administer the brief assessment of math and reading (30 minutes of class time);
- predict students’ perceptions of nurturing pedagogy (15 minutes);
- predict students’ scores on the math and reading assessments (15 minutes);
- respond to one open-ended question related to cue-utilization (15 minutes).

Risks/Discomforts:
There is a possibility you will experience some discomfort. You might worry that a person in position of authority will see your judgment scores and consider them in an evaluation. These potential concerns will be addressed by not providing any identifiable data to anyone in a position of authority. All teacher and student data will be de-identified and the researcher will ensure that no student data is identifiable and linked in an identifiable way to you.

Benefits
There are no direct benefits from participation in this study.

Confidentiality
The research data will be kept in a secure location/on password protected computer, as well as in a secure location in the cloud, and only the researcher will have access to the data. At the conclusion of the study, all identifying information will be removed and the data will be kept in the researcher’s locked cabinet/office for three years.

Compensation
You will receive a $150 Visa gift card for participating and completing all of the tasks in this study. The gift card will be sent to you via mail, or delivered in person, once the data collection at your school is complete.

Participation
Participation in this research study is voluntary. You have the right to withdraw at any time or refuse to participate entirely without jeopardy to your employment or standing at the school.

Questions about the Research
If you have questions regarding this study, you may contact Rich Osguthorpe at 801-422-6064 for further information.

IRB NUMBER: IRB2020-157
IRB APPROVAL DATE: 12/09/2020
IRB EXPIRATION DATE: 12/08/2021
Questions about Your Rights as Research Participants

If you have questions regarding your rights as a research participant contact IRB Administrator at (801) 422-1461; A-285 ASB, Brigham Young University, Provo, UT 84602; irb@byu.edu.

Statement of Consent

I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study.

Name (Printed): ___________________ Signature ___________________ Date: ______________
APPENDIX B

Instruments

Teacher-Student Relationship Survey

Name: ___________________________ School: ___________________________ Teacher: ___________________________

Read each statement below and circle the response that best describes your relationship with your teacher.

I feel relaxed with my teacher.
1=No, that is not true  2=That is usually not true  3=Sometimes  4=That is usually true  5=Yes, that is true

Other children are less punished.
1=No, that is not true  2=That is usually not true  3=Sometimes  4=That is usually true  5=Yes, that is true

When I feel uncomfortable, I go to my teacher for help and comfort.
1=No, that is not true  2=That is usually not true  3=Sometimes  4=That is usually true  5=Yes, that is true

I easily have quarrels with my teacher.
1=No, that is not true  2=That is usually not true  3=Sometimes  4=That is usually true  5=Yes, that is true

I try to do things just as my teacher does them.
1=No, that is not true  2=That is usually not true  3=Sometimes  4=That is usually true  5=Yes, that is true

My teacher particularly tells me what I do wrong and not what I do right.
1=No, that is not true  2=That is usually not true  3=Sometimes  4=That is usually true  5=Yes, that is true

I tell my teacher things that are important to me.
1=No, that is not true  2=That is usually not true  3=Sometimes  4=That is usually true  5=Yes, that is true

My teacher treats me unfairly.
1=No, that is not true  2=That is usually not true  3=Sometimes  4=That is usually true  5=Yes, that is true

I feel most at ease when my teacher is near.
1=No, that is not true  2=That is usually not true  3=Sometimes  4=That is usually true  5=Yes, that is true

I guess my teacher thinks I whine a lot.
1=No, that is not true  2=That is usually not true  3=Sometimes  4=That is usually true  5=Yes, that is true

TURN PAGE OVER TO CONTINUE
I think I have a good relationship with my teacher.
1=No, that is not true 2=That is usually not true 3=Sometimes 4=That is usually true 5=Yes, that is true

My teacher thinks I do things sneaky.
1=No, that is not true 2=That is usually not true 3=Sometimes 4=That is usually true 5=Yes, that is true

I ask my teacher for help, also when I know it isn’t really necessary.
1=No, that is not true 2=That is usually not true 3=Sometimes 4=That is usually true 5=Yes, that is true

I don’t like it when my teacher comes near me.
1=No, that is not true 2=That is usually not true 3=Sometimes 4=That is usually true 5=Yes, that is true

I guess my teacher gets tired of me in class.
1=No, that is not true 2=That is usually not true 3=Sometimes 4=That is usually true 5=Yes, that is true

If I have a problem, I can share it with my teacher.
1=No, that is not true 2=That is usually not true 3=Sometimes 4=That is usually true 5=Yes, that is true

I feel my teacher doesn’t trust me.
1=No, that is not true 2=That is usually not true 3=Sometimes 4=That is usually true 5=Yes, that is true

I can be very angry with my teacher.
1=No, that is not true 2=That is usually not true 3=Sometimes 4=That is usually true 5=Yes, that is true

If my teacher says something is not allowed, I often do it anyway.
1=No, that is not true 2=That is usually not true 3=Sometimes 4=That is usually true 5=Yes, that is true
## Nurturing Pedagogy Study
### Teacher Prediction Table

<table>
<thead>
<tr>
<th>Last Name</th>
<th>First Name</th>
<th>Closeness (1-5)</th>
<th>Conflict (1-5)</th>
<th>Math (#1-5) (1-5)</th>
<th>Math (#6-10) (1-5)</th>
<th>Reading (1-5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example</td>
<td>One</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Example</td>
<td>Two</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>