Teachers' Expectations Vary: Implications for Canadian Students of Different Ethnic Groups

by

Agnes Flanagan

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Education

In

SCHOOL AND CLINICAL CHILD PSYCHOLOGY

Department of Educational Psychology
University of Alberta

© Agnes Flanagan, 2018

Abstract

In Canada, there is an achievement gap between students of different ethnicities. An important alterable contributor to academic achievement is teacher expectations—the beliefs teachers hold about their students' academic capabilities. Teacher expectations affect students' academic performance—high expectations positively impact academic performance and low expectations negatively impact academic performance. Using a survey, I collected data from 140 teachers at a teacher conference in March 2017. In addition to demographic questions, the survey had two scales that gathered information about teacher expectations and behaviours. Differential item functioning analysis showed that teachers had higher expectations for Asian students and lower expectations for Indigenous students compared with Caucasian students. Explanatory item response model showed teachers with six or more years of teaching experience, and teachers above 35 years old, had higher expectations for their students. Teachers did not report behaving differently towards their students. Ultimately, the results will directly affect students across Canada, whose future educational outcomes and careers are influenced by their teachers' expectations. The first step is to inform teachers and education programs. Next, would be to explore whether these differing expectations affect students' academic achievement outcomes positively or negatively.

Preface

This thesis is an original work by Agnes Flanagan. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name: Examining Teachers' Thoughts, Emotions, and Practices Among Diverse Student Learners, REB No. Pro00069833, January 19, 2017.

Acknowledgments

I would like to start by thanking my supervisor Dr. Damien Cormier. Thank you for your guidance and academic mentorship, and for always finding the time in your busy schedule to share your ideas and feedback with me. I have learned so much from you and I look forward to learning more in the years to come.

Thank you to Dr. Okan Bulut for your guidance with the analyses and teaching me about the R program. Thank you also for your general support and feedback throughout this process. Thank you to Dr. Lia Daniels for giving me the opportunity to join your lab at the teacher conference where I collected my data for this project. I also appreciated your time and expertise while developing the Expectations survey. I would also like to thank Dr. Bulut and Dr. Daniels for being my committee members, for your feedback, and for taking the time to review my work.

Thank you to my family for your endless emotional support from the beginning. Despite being far away from "home," it was always so comforting to chat with you over the phone and to spend some relaxing vacations with you. Special thanks to my Dad for always willing to edit my academic work—I appreciate that more than you know. Last but not least, thank you to Ben for moving to Edmonton so I could pursue my dream career. Thank you also for being so patient and for supporting my (somewhat stressful and) busy lifestyle these past two years.

Thank you to my cohort of six amazing women who have been so supportive these past two years. I am so thankful to have developed such good friendships with each of you and I look forward to staying in touch.

Thanks to Social Sciences and Humanities Research Council of Canada for supporting this project, and for allowing me to devote more time to my research than I could have otherwise.

Table of Contents

ABSTRACT ii
PREFACE iii
ACKNOWLEDGMENTS IV
TABLE OF CONTENTS V
LIST OF TABLES VI
LIST OF FIGURES VI
INTRODUCTION & LITERATURE REVIEW 1
ACADEMIC ACHIEVEMENT GAP 3
FORMING EXPECTATIONS 4
TEACHER EXPECTATIONS AND ACADEMIC ACHIEVEMENT 6
EXPECTATIONS FOR STUDENTS OF DIFFERENT ETHNIC GROUPS 7
THE RELATIONSHIP BETWEEN TEACHER EXPECTATIONS AND BEHAVIOUR 9
MEASURING TEACHER EXPECTATIONS 11
PRESENT STUDY 13
METHODOLOGY 14
PARTICIPANTS 14
PROCEDURE 15
MEASURES 16
DATA ANALYSIS 16
RESULTS 19
MAIN ANALYSES 20
DISCUSSION 24
TEACHER EXPECTATIONS 24
TEACHER BEHAVIOURS 26
DEMOGRAPHICS AND EXPECTATIONS 27
IMPLICATIONS 29
LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH 31
CONCLUSION 33
REFERENCES 34
APPENDIX A 40

List of Tables

Table 1: *Item Factor Loadings from Exploratory Factor Analysis*.....20

Table 2: *IRT Likelihood Ratio Results for Expectations Scale Items*.....21

Table 3: *IRT Likelihood Ratio Results for Behaviour Scale Items*.....23

Table 4: *Results of the EIRM Analyses for the Expectations and Behaviour Scales*.....24

List of Figures

Figure 1. Frequency of participants' age and years of experience.....15

Figure 2. Expectations scale DIF plot of item 2 from model 1 comparing Caucasian (0) to Asian (1) 21

Figure 3. Expectations scale DIF plot of item 1 from model 2 comparing Caucasian (0) to Indigenous (1) 22

Figure 4. Behaviour scale DIF plot of item 1 from model 2 comparing Caucasian (0) to Indigenous (1) 23

Teachers' Expectations Vary: Implications for Canadian Students of Different Ethnic Groups

Canada is one of the world's more multicultural countries; home to people of more than 200 ethnic groups. It is projected that people from visible minority groups will double or triple by the year 2031 (Statistics Canada, 2010). The diversity of our country is reflected in the wide range of languages, cultures, and ethnicities of the students in Canadian classrooms (Guo, 2012). The fact that Canada continues to become more diverse leads to a number of implications for many professionals and public institutions, including teachers. One of the greatest implications for teachers is perhaps their role in reducing the academic achievement gap that exists between students of different ethnicities. For example, Indigenous, African American, and Latino students, on average, perform considerably below their peers and Asian students tend to perform above their peers (Lee, 2015; McKown & Weinstein, 2008; Richards, 2011). Hattie (2003) showed that teachers are the most important variable, aside from individual student characteristics, that accounts for students' academic achievement.

Teachers play an important role in students' success. An important alterable contributor to students' academic achievement is teacher expectations. Teacher expectations are the beliefs teachers hold about their students' academic capabilities and level of achievement (Peterson et al., 2016; Strand, 2013). Teacher expectations are formed based on objective information (e.g., past achievement) and subjective information (e.g., stereotypes and prejudice). This information can help teachers determine whether they will have high expectations or low expectations for their students, even before they meet them. Teacher expectations affect students' academic performance, both positively and negatively—high expectations positively impact academic performance and low expectations negatively impact academic performance (Rubie-Davies et al., 2006). Students' academic achievement is impeded if their teacher has low expectations for them

(Rubie-Davies et al., 2006). In other words, teachers' expectations can bias their evaluations of students and predict the students' achievement level (Jussim & Harber, 2005). These expectations can also be communicated in covert, unintentional ways (Brophy, 1983). For example, teachers are more likely to call on students and teach more material to students for whom they have higher expectations. Teacher expectations are especially dangerous for students from academically stereotyped groups, such as Indigenous students, who consistently fall behind their peers in terms of academic achievement. This stereotype effect was recently shown in a study conducted in New Zealand by Dandy et al. (2015). The gap between academically stereotyped students and their non-stereotyped peers is not reducing in magnitude (Parkin, 2015) and teachers can play an important role minimizing this gap.

There have been many studies that have explored topics related to teacher expectations and ethnic groups and many show stereotypical differences in the expectations teachers have for students from different ethnic groups (e.g., Rubie-Davies, 2015; Weinstein et al., 2004). However, there is a lack of recent Canadian research on this topic. This research is important because the nature of teacher expectations is implicit and usually unintentional. In other words, teachers usually do not even realize they have these differential expectations for their students, and that there can be negative effects for students.

The current study aims to explore whether this is a reality for Canadian teachers. The purpose is to investigate whether teachers have the same high expectations for all students or whether these expectations are lower for students from one ethnic group compared with another. I also investigated whether teachers show differential behaviour for these students. This research may reveal Canadian teacher tendencies to hold unconscious biases in the classroom. If this is the case, then knowing these tendencies exist will be a first step to increasing Canadian teachers'

awareness of their expectations and supporting the reduction of the negative effects poor expectations can have on Canadian students' achievement.

Academic Achievement Gap

There is an identifiable academic achievement gap between ethnic groups in Canada (Statistics Canada, 2013), with Indigenous students performing considerably below their peers and Asian students performing above their peers (Lee, 2015; Richards, 2011). As Canadian classrooms become more diverse, it is critical that this achievement gap be eliminated, to not only ensure *all* students are realizing their full academic potential, but also to reduce the likelihood that they will experience the many negative outcomes associated with poor academic achievement. For example, poor academic achievement is associated with social exclusion, high school drop-out, increased health problems, an increased likelihood of unemployment, and lower income earnings (Brynner, 2008). Based on 2006 Census data, Richard (2011) claims that the most serious education gap in Canada is between Indigenous and non-Indigenous people. Scholars have attributed this to factors such as colonialism, racism, poverty, and inequality of resources outside the school environment (Hoffman-Goetz, Donelle, & Ahmed, 2014).

Research has revealed various reasons for this achievement gap, such as socioeconomic status (SES), parents' education, peers, and quality of the school students attend (Richard, 2011; Strand, 2014). However, most of these variables cannot be altered easily to help close this gap. An alterable predictor that appears to contribute to student academic achievement is teachers' academic expectations (Hattie, 2003). Hattie (2003) described that student characteristics, such as gender, ethnicity, and motivation, explain the greatest amount of the variance in their academic achievement (i.e., 50%). He added that teachers account for the second largest variance (i.e., 30%) of students' academic achievement. This means that besides students themselves,

teachers play an important role in students' academic success and they can help reduce this academic achievement gap more easily than we can change students' SES, parents' education and quality of school attended.

Forming Expectations

Teacher expectations are the beliefs teachers hold about their students' academic capabilities and level of achievement (Peterson et al., 2016; Strand, 2013). Teacher expectations are formed based on knowledge teachers have about their students, both objective knowledge such as previous grades, and subjective knowledge such as perceptions of in-class performance and prejudices or stereotypes (Friedrich, Flunger, Nagengast, Jonkmann, & Trautwein, 2015). Rubie-Davies' (2015) book dedicates a whole chapter to describing sources of information, supported by research, that teachers use and integrate as they form expectations for their students.

Prior achievement. One source of information is students' *prior achievement*. There is a bidirectional relationship between teacher expectations and student achievement, meaning that expectations impact achievement, and student achievement or grades impact teacher expectations. Teachers are forming expectations for their students when they receive their records of academic achievement, even before meeting their students. This helps teachers mold their expectations and plan what learning opportunities the individual students will be exposed to. This will be described in greater detail in the next section.

Student effort. A second source of information is *student effort*. Teachers assume that high-achieving students put more effort into their school work and work harder. As a result, they give these students higher grades than they might deserve. In contrast to this assumption, research shows that low-achieving students spend more time on homework (Jussim, Smith,

Madon, & Palumbo, 1998).

Student characteristics. Researchers generally agree that some student characteristics render students more susceptible to teacher expectations than others. These characteristics include having an identified *disability*, children from low-*SES*, belonging to an *ethnic minority group*, and *student gender* (Rubie-Davies, 2015). Gender stereotypes suggest that teachers have higher expectations for boys in mathematics and for girls in reading. Research has shown that having more than one vulnerability may result in additive effects of teachers' expectations (Jussim, Eccles, & Madon, 1996). For example, a student from an ethnic minority group and who is also from a low-SES home would be more susceptible to teacher expectations than a student with just one potential contributing factor.

Stereotypes. Some of the above factors could be considered *stereotypes*, which is another contributor to teacher expectations. Stereotypes are beliefs about the personal attributes of particular groups of people. Besides the characteristics listed above, there are stereotypes about many other groups that Rubie-Davies (2015, p. 29) lists, such as "blondes are dumb, Americans are obese, Asians are good at mathematics, women are moody, African Americans are good at basketball, and Italians are good cooks." Stereotyping is something that humans do, often without being consciously aware of doing it. Stereotypes provide us with cognitive shortcuts that allow us to organize information about certain groups that we can access when we are interacting with someone that we do not know much about. The danger of stereotypes is that they take away the individuality of people within that group. Stereotyping of groups is something that humans do, including teachers. For teachers, stereotyping contributes to the expectations they hold for their students, which can affect many things including student academic achievement.

Teacher Expectations and Academic Achievement

Teacher expectations exist in the classroom and they affect students' academic performance, both positively and negatively (Rubie-Davies et al., 2006). Students' academic achievement will increase when a teacher has high expectations for them. A teacher with low expectations for students will impede those students' academic achievement. There is evidence that teachers' perceptions may bias their evaluations of students and predict the students' achievement level (Jussim & Harber, 2005). This effect is more likely to impact students in the early years and tapers off by fifth grade, where teacher expectancy effects are mediated by children's own expectations (Weinstein, Gregory, & Strambler, 2004). Rubie-Davies (2015) suggested this could be because younger students do not have fixed views of their abilities and they are easily malleable; whereas older students tend to have more established self-views and are therefore harder to influence and change.

There are many theories that have been used to explain this relationship between teacher expectations and student academic achievement. A popular one is the self-fulfilling prophecy whereby erroneous teacher expectations may lead students to perform at levels consistent with those expectations (Brophy & Good, 1974; Rosenthal & Jacobson, 1968). The self-fulfilling prophecy effect was first demonstrated in schools by Rosenthal and Jacobson (1968). In their study, they gave a fake test to students and told teachers the results of the test would indicate which students are likely to "bloom," and show a sudden and dramatic growth in their intellectual development. Researchers randomly selected students who would be likely to "bloom" and gave those results to teachers. The researchers assessed students again at the end of the year. Results showed that students with the most improved intellectual development were the ones whose teachers were told they were expected to "bloom" intellectually. Many empirical

studies were conducted in an attempt to replicate these findings, but only one third of subsequent studies showed similar results (see Brophy & Good, 1974, for a review).

Jussim and Harber (2005) reviewed the past 35 years of empirical research and came to four important conclusions on self-fulfilling prophecies and teacher expectations. First, they confirmed that self-fulfilling prophecies in the classroom do occur, but these effects are typically small. Second, powerful self-fulfilling prophecies may selectively occur among students from stigmatized social groups. For example, they found the general effect of teacher expectations to be mostly small, ranging from $d = .2$ to $d = .41$. However, for some students from ethnic minorities or low-SES communities, expectation effects were much larger between $d = .63$ and $d = .87$. The third conclusion is that it remains unclear whether self-fulfilling prophecies affect intelligence, and whether they in general do more harm than good. Their fourth and final conclusion is that teacher expectations may predict student outcomes because these expectations are accurate, rather than because they are self-fulfilling. This review demonstrates that this topic has been researched by many researchers in the field. The meta-analysis completed by Rosenthal and Rubin, which was published approximately 40 years ago, drew from the 345 studies that were available at that time—and studies continue to be published decades later (e.g., Glock, 2016). Although it is unknown whether this topic is as heavily researched now as it was, this demonstrates that the topic is still of interest today and seemingly important to researchers.

Expectations for Students of Different Ethnic Groups

The finding that teacher expectation effects are more powerful for certain groups of students may be explained by implicit prejudice. Implicit prejudice occurs when an individual has subconscious negative feelings or beliefs about an individual based on the individual's membership in a certain social group (e.g., gender, ethnicity). It occurs with little intention or

control, and is usually based on day-to-day experiences (Rubie-Davies, 2015). Much of the research shows that stigmatized minorities are at an intensified risk of the negative expectancy effects (Weinstein et al., 2004). Although the effects of teacher expectations eventually disappear for most students, they may endure in diluted form for members of at-risk groups for years (Jussim & Harber, 2005).

Most Canadians believe that colonization and racism are issues of the past; however, Battiste and McLean (2005) argue they have become the biggest challenges within the education system today. For example, Whitley (2014) was interested in identifying key influences in supporting educational success for Indigenous students. Whitley collected data by conducting focus groups with teachers. Teachers remarked that the most common response to hearing of an Indigenous student who was academically successful was "one of shock or surprise" (Whitley, 2014, p. 170). Teacher participants in this study felt this response was detrimental to the school experiences of Indigenous students (Whitley, 2014). Knowing teachers are conscious about having these types of reactions and biases certainly helps depict what those teachers might expect academically of their Indigenous students.

Not only do some teachers recognize their biases, but students recognize them as well. McKown and Weinstein (2008) found that students noticed that their teacher treated high- and low-achievers differently. In classrooms where students observed this differential treatment, teachers were likely to expect more from academically non-stereotyped ethnic groups than academically stereotyped students with similar levels of achievement. They also found this teacher bias was intensified when classrooms were more ethnically diverse. Specifically, teachers in these classrooms had higher expectations for European American and Asian American students than they did for African American and Latino students with similar levels of

achievement. This suggests that teachers have more obvious differential expectations in classrooms with greater diversity of students, and this is noticed by their students. Students awareness of this suggests that expectations can be communicated overtly.

The Relationship between Teacher Expectations and Behaviour

Teachers expectations have not only been found to be related to student academic outcomes, but it has also been linked to their behaviour in the classroom. Robert Rosenthal is well-known in this area of study and has contributed many research studies (e.g., Rosenthal, 1994). He proposed a four-factor theory to explain expectancy effects. He argued that teacher expectations are communicated covertly, through very subtle, nonverbal and typically unintended ways. The four key categories of teacher behaviour are: climate, input, output, and feedback. Teachers create warmer climates, teach more material (referred to as input), call more frequently on (referred to as output), and give differential feedback to students for whom they have higher expectations (Rosenthal, 1994).

Furthermore, Brophy (1985) identified 17 teacher behaviours that differed depending on whether it was directed towards a high- or low-achieving student. For example, (a) *wait time* following an answer provided by a low-expectation student is shorter than wait time for a high-expectation student. (b) Teachers' *response following an incorrect answer* from a low-expectation student will typically provide the correct answer or ask another student. For high-expectation students, teachers will typically repeat the question or provide a clue. (c) Teachers will ask high-expectation students to *provide an answer* more often than they will ask low-expectation students. (d) Teachers have less *friendly and warm interactions* with low-expectation students. (e) Teachers make less *eye contact* with low-expectation students. (f) Teachers will praise low-expectation students when they have given an incorrect answer whereas high-

expectation students do not usually get *inappropriate feedback*.

These behaviours have not only been observed in classrooms from an outsider or researchers, but teachers (Whitley, 2014) and students themselves (Weinstein et al., 2004) have also noticed these differential behaviours. One of the key influences identified in Whitley's (2014) study, both from the student and teacher perspective, was academic expectations. Teachers noted that they often had very different life experiences than their Indigenous students, which affected what they expected and how they interacted with their students (Whitley, 2014). In another review, children reported that high achievers are exposed to more challenging material, given more opportunities for autonomy and leadership, and supported more positively. Conversely, the work of low achievers is more structured with drill and repetition being the key focus—it is also more criticized than the work of their peers (Weinstein et al., 2004). Students notice these behaviours and teachers acknowledge it as well.

Brophy and Good (1970) proposed a cycle to explain the relationships between student academic achievement, teacher expectations, and teacher behaviours. The cycle begins with teachers who have their own set of beliefs at the beginning of each school year. These beliefs affect their behaviour in the classroom (e.g., call on certain students more frequently, challenge instruction). Students recognize this differential behaviour, which motivates them to work harder or to work less hard depending on whether the behaviour is positive (i.e., encouraging) or negative (i.e., discouraging). In turn, students' level of motivation or work ethic impacts their academic achievement, positively or negatively. The resulting outcomes with respect to students' academic achievement will feed back into the beginning of the cycle—this will again influence teachers' beliefs and the cycle continues. This cycle will result in different achievement outcomes depending on a number of factors, but the cycle starts with teacher beliefs.

Measuring Teacher Expectations

There are three types of study designs used most often to measure teacher expectations: empirical, naturalistic, and survey studies. Survey studies seem to be the most feasible way of tapping into teacher expectations implicitly. An implicit measure is when participants are not aware of what the measurement reflects (De Houwer, 2006). The benefit of using an implicit or indirect measure is that they avoid response biases, which are especially prevalent when asking questions about sensitive topics. Instead, an indirect measure captures implicit attitudes on these sensitive topics. The anonymity of surveys also helps avoid response biases, such as social desirability.

Surveys measuring teacher expectations or behaviours vary between studies. A more recent study in Germany by Glock (2016) showed differential effects in teacher behaviours. Preservice teachers were presented with a description of a hypothetical student who talked out of turn. The gender and ethnicity of the student described was manipulated systematically but all other parts of the description were the same. The teachers had to read the description and identify what intervention they would give the student. Results showed that male and ethnic minority students received more, and intensive, interventions.

Expectations can also be measured by asking teachers about their current opinions of their students' competencies. McKown and Weinstein (2008) described a collection of three studies conducted in the United States that measured teacher expectations by asking teachers in the fall to rank order their students in order of expected year-end achievement in reading and math. This was completed on a five-point scale: 1 = low; 2 = low average; 3 = average; 4 = high; 5 = outstanding. The researchers measured other variables such as: ethnicity, prior achievement, perceived differential teacher treatment, and classroom diversity. The results of their study

showed the gap in students' academic achievement is attributable to different expectations for equally-achieving children from different ethnic groups.

Dandy and colleagues (2015) conducted a set of studies in New Zealand and used a questionnaire to investigate the perceptions and expectations of minority students among teachers-in-training (Study 1), experienced teachers (Study 2), and students (Study 3). Participants completed a questionnaire containing the expectations measure and demographic questions. The expectations measure was based on questions used in the Michigan Study of Adolescent Life Transitions. Expectation-related questions were completed by participants for hypothetical students from different ethnic groups. The hypothetical students were labeled *Chinese Australian student*, *Anglo-Australian student*, and *Aboriginal student*. For each hypothetical student, five items assessed expectations in two academic subjects, math and reading, regarding academic processes and outcomes. Their results showed that Asian students were expected to perform better in mathematics and expend greater effort than Aboriginal and Anglo-Australian students. There were higher expectations for Anglo-Australians compared with Aboriginal students. The second important finding from this study is that the three groups of respondents (preservice teachers, teachers, and students) answered the questionnaire similarly. These results are stereotypical, and I wondered what the results would be with Canadian participants.

Demographic information. In terms of demographic information, Dandy et al. (2015) found no difference in expectations between teachers based on experience (preservice teacher versus experienced teacher). They also did not find any difference for gender. One might hypothesize that teachers who recently graduated from an education program, or teachers with little experience in the classroom, would have similar expectations for all students. They might

have less experiences that strengthen these biases. In contrast, Rubie-Davies (2010) found that participants with low expectations were those with less teaching experience. This was another area that I wanted to explore.

Present Study

Canada is one of the world's more multicultural countries and the diversity of our country is reflected in the wide range of languages, cultures, and ethnicities of the students in Canadian classrooms (Guo, 2012). Further, there is an academic achievement gap between majority and minority groups within classrooms in Canada. Researchers have discussed different factors that may contribute to this gap such as SES, parent education, peers, and quality of school. However, Hattie (2003) described that, aside from individual student characteristics, teachers contribute the greatest amount of variance in students' academic achievement and they should be the primary focus in reducing this gap. An alterable predictor of academic achievement is teachers' academic expectations. Teacher expectations are formed based on both objective and subjective sources of information. These expectations have been shown to positively and negatively impact students' academic achievement. Students from stigmatized minority groups (e.g., Indigenous students) are especially at-risk of the negative effects. These expectations can also influence teachers' behaviours. Many studies have been completed in this area, but few have been completed in Canada, especially in recent years. Survey studies are useful in providing a snapshot of the attitudes and behaviours of the target population, in this case, Canadian teachers.

The purpose of the study is to examine whether teachers' academic expectations differ for students of different ethnic groups in Canada. I am also interested in examining whether reported teacher behaviours differ for students from different ethnic backgrounds. In addition, researchers had previously considered the influence of teachers' demographic information such

as gender and teaching experience on expectations (Dandy et al., 2015) but the evidence is not consistent. Exploring these key variables may identify a specific group of teachers that require support in changing their expectations for certain students. The current study was intended to replicate the study conducted by Dandy et al. (2015), and also extend it by adding questions about teacher behaviours. The three main research questions for this study are as follows:

1. Do Canadian teachers' expectations differ for students of different ethnic groups?
2. Does teacher behaviour differ in response to students of different ethnic groups?
3. Does a teacher's age, gender, or teaching experience predict their expectations or behaviour?

Methodology

Participants

We recruited participants at a teacher conference in Edmonton, Alberta in March 2017. During the two-day conference, 140 teachers completed the survey. Of these, 34% were secondary school teachers (n = 47), 46% were primary school teachers (n = 64) and 20% (n = 29) did not indicate which level they taught. Participants' age ranged from 20 to 60 years old (M = 36.8). Teachers had, on average, 10.74 years of teaching experience, with a range of 0 to 35 years. Over 80% of participants were female (n = 113), with four participants not indicating their gender. Figure 1 shows the frequencies of age and teaching experience reported by participants.

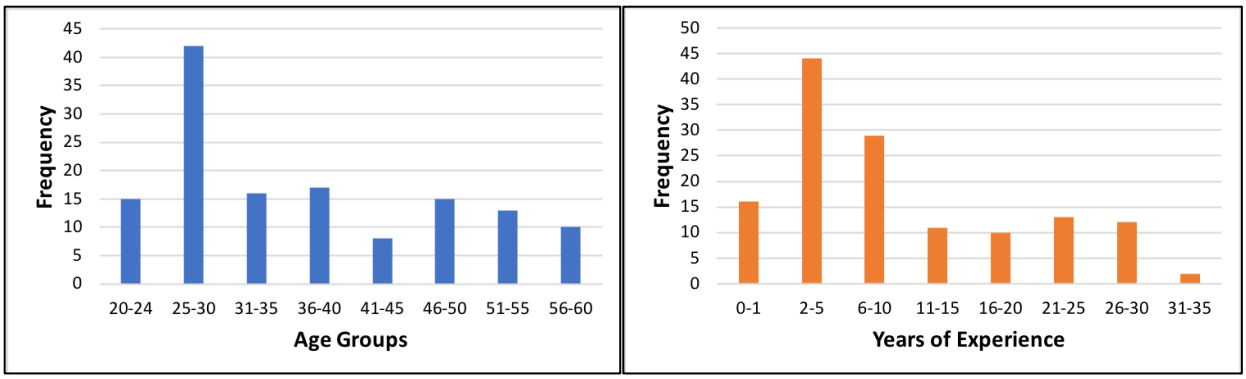

Figure 1. Frequency of participants' age and years of experience

Procedure

This study was carried out using a quantitative, descriptive research design. The data was collected using a survey that was created based on previous research (Dandy et al., 2015). Ethical approval for the study was obtained from the University of Alberta Research Ethics Board (REB2 - Pro00069833). Teachers were eligible for participation if they worked in the greater Edmonton area, could physically access the conference, and were registered as attendees for at least one day of the conference. A booth in the exhibition hall associated with the conference was used to recruit participants. Teachers approached the booth if they were interested in participating in research from the University of Alberta. At this point, they were provided with a brief overview of the survey and asked if they would be interested in completing it. Information letters were available to teachers who initially expressed interest in the study. The information letters provided teachers with more detailed information about the study and the consent process. Teachers who completed a survey were entered to win one of five possible $100 gift certificates from the business of their choice. Teachers provided consent by completing the survey. Survey completion time was approximately 10 minutes for each participating teacher.

Measures

Participants completed the expectations survey and answered several demographic questions pertaining to age, gender, teaching level, and years of teaching experience. The expectations survey had two scales, and both scales had five items each. The first scale gathered information about teacher expectations, based on the measure from Dandy et al. (2015), referred to as the *Expectations Scale*. For example, a question from the Expectation Scale was "How well does this student perform on classroom assessments?". The second scale asked questions about teacher behaviours, referred to as the *Behaviour Scale*. This second scale was created using findings from past research. A question from the Behaviour Scale was "I make sure to challenge this student during instruction." The questions from both scales were answered using a five-point rating scale. The response options in the Expectations Scale range from "bottom of the class" to "top of the class", and the response options on the Behaviour Scale range from "never" to "always" (see Appendix A for the full survey).

The Expectations and Behaviour scales were completed for hypothetical students from different ethnic groups. Three hypothetical students were labelled as a typical Asian student, a typical European-American student, and a typical Indigenous student. For each student, five items were used to assess expectations regarding academic processes and outcomes (natural talent, family support, effort, and current performance on tests and homework) and five items were used to measure teacher behaviours (demanding better performance, praise, challenging instruction, feedback, and calling on student).

Data Analysis

Preliminary analysis. I conducted an exploratory factor analysis with principal axis method to examine whether the two scales (Teacher Expectations and Behaviour) are

unidimensional. This is an important step in the data analysis because it determines whether the items selected for inclusion in each of the two scales are in fact measuring a single (unidimensional) latent trait. For example, I want to ensure the items in the Expectations scale are measuring teachers' expectations (the latent trait) as opposed to teacher behaviour. In exploratory factor analysis, the model determines mathematically the number of dimensions underlying the data collected from teachers based on the aforementioned scales. I used the psych package (Revelle, 2016) in R (R Core Team, 2017) for exploratory factor analyses.

Main analyses. One way of explaining the differences in teachers' expectations (research question 1) and behaviours (research question 2) from the collected data is to use differential item functioning (DIF). DIF is a contemporary method for studying item-level bias. The current study utilized a unique application of DIF to answer the study's research questions. Traditionally, an item contains DIF if individuals with the same level of proficiency or latent trait, but belonging to different groups, have different probabilities of a correct response (Penfield, Gattamorta, & Childs, 2009). The application of this analysis in the current study is the same persons (i.e., teachers) answer the same item for three different ethnic groups. When the items are flagged with DIF, it means that teachers responded to the item differently for one group compared with another. It is a unique applicated given that I manipulated an independent variable, so it is not necessarily purely correlational.

Item Response Theory Likelihood Ratio (IRT-LR) test is a method that compares the likelihood ratios of the models based on IRT (Atalay Kabasakal, Arsan, Gök, & Kelecioğlu, 2014). IRT models are nonlinear latent trait models for categorical ordered data. Graded response model (GRM) was the type of IRT model used for this analysis. This model is well suited for items that have two or more ordinal response categories (e.g., a Likert scale). It is

considered a generalization of the two-parameter logistic model (2PL). When used, the scale is dichotomized at each point. The main advantage of using GRM is that it has better discrimination than other models.

DIF detection in the IRT-LR test is carried out with a null hypothesis based on the comparison of item parameters of focal and reference groups. In this analysis, the reference group for Model 1 and 2 is the Caucasian group (assigned value of 0). The focal group for Model 1 is the Asian group (assigned value of 1) and the focal group for Model 2 is the Indigenous group (assigned value of 1).

Compact and augmented IRT models are formed to test this null hypothesis, and the differences in the likelihood of these models are compared. In a compact model, item parameters are considered equal across reference and focal groups. In an augmented model, item parameters of the studied item differ while they remain equal for other items in reference and focal groups. Then, logarithmic transformations are applied to both models, and G^2 value is calculated [$G^2 = -2LL_c-(-2LL_a)$; L_c: log-likelihood of compact model, L_a: log-likelihood of augmented model]. The value of G^2 has a chi-square (χ^2) distribution with the degrees of freedom equal to the difference in the number of estimated item parameters from the augmented and compact models. Each test of item parameters was conducted one by one. If the value of G^2 is statistically significant, sequential tests are applied to determine the type of DIF (i.e., uniform vs. non-uniform DIF) for the items. IRT-LR was used to answer research questions 1 and 2.

We used explanatory IRT modeling (EIRM; De Boeck & Wilson, 2004) to answer the third research question. I used the mirt package (Chalmers, 2012) in R (R Core Team, 2016) for this analysis. When explanatory variables (or covariates) are incorporated into traditional IRT models, the resulting modeling framework is called EIRM (De Boeck & Wilson, 2004).

Explanatory variables can be related to either items (e.g., linguistic complexity) or persons (e.g., gender, age, experience), or both. In this study, I combine three person-related variables with the IRT model presented earlier. I wanted to explore whether teacher responses regarding the Expectations and Behaviour scales are influenced by teachers' demographic characteristics. The three explanatory variables were created as dichotomous variables: gender (0 = Male; 1 = Female), age (0 = 35 years old or younger; 1 = Older than 35), and years of teaching experience (0 = Five years of teaching experience or under; 1 = Six or more years of teaching experience). The cut-points were based on having relatively equal dichotomized groups for the analyses. I used the *mixedmirt* function in the mirt package (Chalmers, 2012) to include the demographic variables in the EIRM analyses. An explanatory IRT model was applied to each scale separately. In terms of significance testing, z scores were used as the significance criterion. The critical z score values when using a 95% confidence level are $z = -1.96$ and $z = +1.96$. If a z score falls above +1.96 or below -1.96, the test is significant at the 95% confidence level (i.e., $\alpha = .05$), suggesting that the explanatory variable significantly affects the way teachers respond to the items on the scale.

Results

We used exploratory factor analysis with principal axis method to see whether the model would identify the two scales. I combined the items from both scales for the exploratory analysis. Item factor loadings from this analysis are shown in Table 1. The factor loadings show that items 1 through 5 are loaded on factor 1 (teacher expectations) and items 6 through 10 are loaded on factor 2 (teacher behaviours). The Tucker Lewis Index (value range 0–1) was .97 and the root mean error of approximation (RMSEA) index (value range 0–1) was .004. A Tucker Lewis Index of .95 or greater is considered to be a good fit. An RMSEA index of .06 or less is indicative of

acceptable model fit (Hooper, Coughlan, & Mullen, 2008). In this study, the two-factor model exceeded these criteria. The cumulative variance was .62, which means that 62% variance is explained by this two-factor solution. The correlation between factor 1 (Expectation scale) and factor 2 (Behaviour scale) is low (r = .15). This small correlation indicates that the two scales are not measuring the same construct. Based on these statistics, the results showed that the model fit well, so a conclusion can be made that the scales are unidimensional.

Table 1

Item Factor Loadings from Exploratory Factor Analysis

Item	Factor 1 (Expectations Scale)	Factor 2 (Behaviour Scale)
1	**.82**	-.02
2	**.86**	.03
3	**.93**	-.02
4	**.92**	-.01
5	**.77**	.03
6	-.07	**.47**
7	-.13	**.79**
8	.11	**.86**
9	-.03	**.77**
10	.13	**.55**

Main Analyses

We conducted a DIF analysis using the IRT-LR test to answer the two main research questions. Items were flagged as having DIF if teachers significantly rated one group higher on the scale than the other. Significance for this DIF analysis is at the alpha level .01.

Research question 1. First, I wanted to see whether teachers have differing expectations for students of different ethnic groups. The first model compared the Asian group to the Caucasian group for the first five items (Expectations scale). I found that all items from the Expectations scale showed DIF (see Table 2) and teachers favoured the Asian group for all items (see Figure 2). In other words, teachers had higher expectations for the Asian group compared

with the Caucasian group. The second model compared the Indigenous group to the Caucasian group for the first five items (Expectations scale). Results showed that all items had DIF (see Table 2) and teachers favoured the Caucasian group for all items (see Figure 3). In other words, teachers had higher expectations for the Caucasian group compared with the Indigenous group. It is interesting to note that no teachers expected Euro-American or Asian students to perform at the bottom of the class, while some teachers expected Indigenous students to do so.

Table 2

IRT Likelihood Ratio Results for Expectations Scale Items

Expectation Scale item	χ^2		*df*	
	Model 1	Model 2	Model 1	Model 2
1	75.47*	152.35*	2	2
2	77.32	63.29	2	2
3	70.74	76.52	2	2
4	95.60	83.86	2	2
5	49.78	43.02	2	2

Note: *The difference is significant at α = .01.

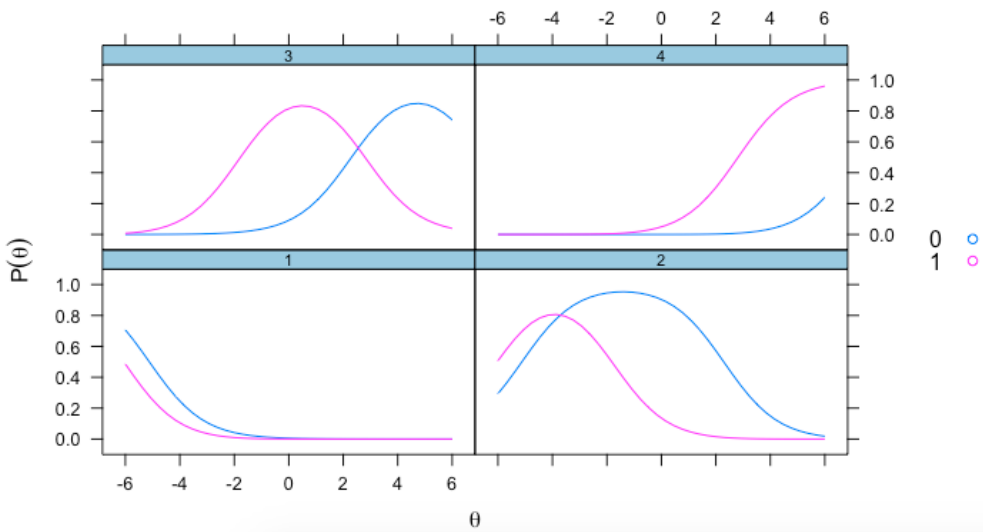

Figure 2. Expectations scale DIF plot of item 2 from model 1 comparing Caucasian (0) to Asian (1) (*Note*: Each plot demonstrates one of the response categories in item 2. Also, the response categories in this comparison were recoded as 1 = Below Average, 2 = Average, 3 = Above Average, 4 = Top of the Class).

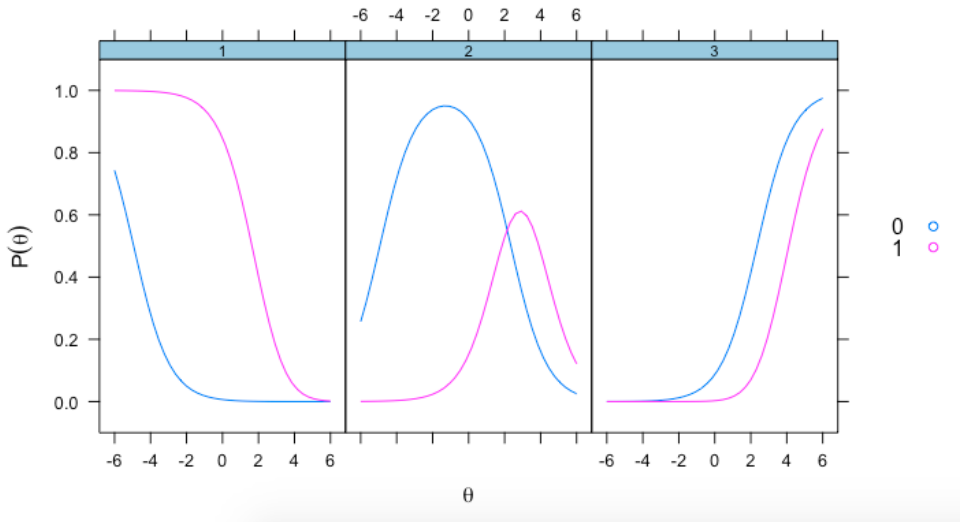

Figure 3. Expectations scale DIF plot of item 1 from model 2 comparing Caucasian (0) to Indigenous (1) (*Note*: Each plot demonstrates one of the response categories in item 1. Also, the response categories in this comparison were recoded as 1 = Bottom of the Class/Below Average, 2 = Average, 3 = Above Average/Top of the Class).

Research question 2. Second, I wanted to see whether teachers behave differently towards students of different ethnic groups. The first model compared the Asian group to the Caucasian group for items from the Behaviour scale. Results showed that teachers behaved favourably for the Asian group on item 1, but all other items were not significant (i.e., the other items did not show DIF) (see Table 3). This means that teachers demand better performance from the Asian group compared with the Caucasian group; however, they behave similarly with regards to praise, challenging instruction, targeted feedback, and calling on students for all three groups. The second model compared the Indigenous group to the Caucasian group for items in the Behaviour scale. Results showed that teachers responded similarly for both groups (see Table 3), which indicates that teachers do not behave differently towards Indigenous and Caucasian students (see Figure 4).

Table 3

IRT Likelihood Ratio Results for Behaviour Scale Items

Behaviour Scale item	χ^2		*df*	
	Model 1	Model 2	Model 1	Model 2
1	25.84	2.73	2	2
2	1.87	2.11	2	2
3	1.12	5.67	2	2
4	2.77	4.08	2	2
5	1.61	6.29	2	2

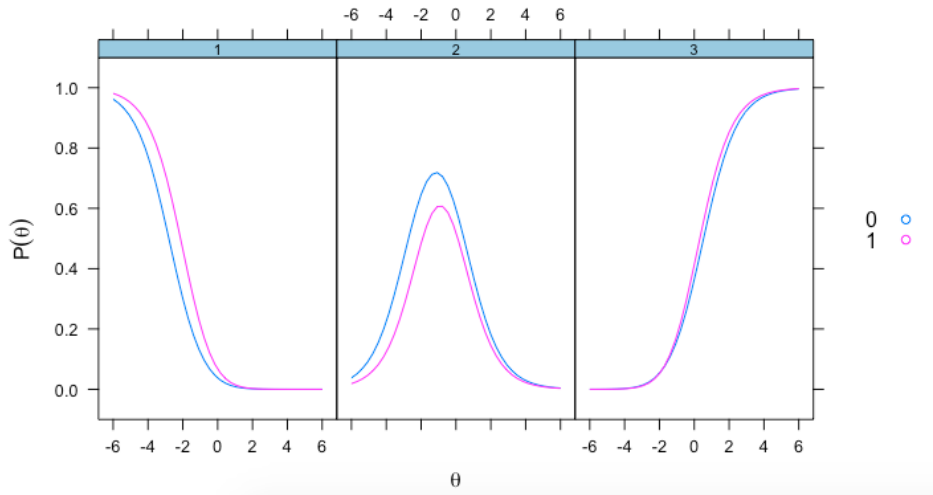

Figure 4. Behaviour scale DIF plot of item 1 from model 2 comparing Caucasian (0) to Indigenous (1) (*Note*: Each plot demonstrates one of the response categories in item 1).

Research question 3. I used EIRM to understand whether teacher characteristics, such as age, gender, and years of experience help predict how teachers will respond to questions on the Expectations and Behaviour scales. Table 4 shows a summary of the results from the EIRM analyses. The explanatory IRT model for the Behaviour scale did not show significant results for any of the variables. This means teachers' age, teaching experience, and gender did not significantly determine teachers' behaviour. The explanatory IRT model for the Expectation scale showed significant results for Age and Years of Experience variables. This means that older teachers (above age 35) had higher expectations for their students and were 1.79 times more likely to choose higher (i.e., positive) response options on the Expectations scale. Teachers with six or more years of teaching experience had higher expectations for their students and were

two times more likely to choose higher response options on the Expectations scale. Teachers' gender did not significantly determine whether teachers had high or low expectations for their students.

Table 4

Results of the EIRM Analyses for the Expectations and Behaviour Scales

Variables	Expectations scale			Behaviour scale		
	Estimate	Standard Error	z value	Estimate	Standard Error	z value
Age	.58	.29	1.98*	-.18	.18	-.99
Gender	.17	.34	.49	.06	.25	.25
Years of Experience	.72	.26	2.75*	-.15	.19	-.79

Note: *p < .05

Discussion

In this study, I began by verifying and confirming that both scales (Expectations and Behaviour scales) were unidimensional. Next, results from the DIF analyses showed that teachers had significantly different expectations for certain ethnic groups when compared to others (research question 1). Subsequent DIF analyses showed that reported teachers' behaviour did not differ across the three ethnic groups with the exception of one item in one model (research question 2). Finally, results showed that older teachers (over 35 years of age) and teachers with six or more years of teaching experience had higher expectations for their students (research question 3).

Teacher Expectations

The first research question focused on teachers' expectations. Teachers showed higher expectations for the Asian group compared with the Caucasian group. Teachers also showed lower expectations for the Indigenous group compared with the Caucasian group. I hypothesized this result based on past research showing these stereotypes exist and that negative expectancy

effects are more likely to occur for stigmatized groups (Jussim & Harber, 2005).

One explanation for results from the first model (Asian and Caucasian groups) is that positive stereotypes can be advantageous, which is consistent with previous research with Asian American populations (Lee, 2015). Asian immigrants are perceived as smart, high-achieving and successful, which Lee (2015) describes as being largely due to the influence of some highly-educated immigrant Asian groups. In Lee and Zhou's (2015) study, the Chinese and Vietnamese participants reported that teachers and guidance counsellors perceived them as smart and expected them to excel. The Mexican students in their study reported they were perceived as low achievers who did not value education. Lee also found that expectations can enhance the academic performance of some mediocre Asian-American students. I described the benefits of positive stereotypes and expectations; however, there are risks that positive stereotypes may result in teachers failing to detect special needs. For example, Hui-Michael and García (2009) found that in American schools, teachers were less likely to attend to struggling Asian students unless they also manifested behavioural problems. This suggests that there are both risks and benefits of these positive stereotypes, which we should be equally aware of.

One explanation for the second model (Caucasian and Indigenous groups) is the implicit prejudice theory. Whitley (2014) described that teachers in their focus groups noted that their Indigenous students have very different experiences than them, which affected what they expected and how they interacted with those students. Teachers biases may also stem from the recurring statistics that Indigenous students perform below non-Indigenous students. This may be an explanation for my results showing that teachers had higher expectations for Caucasian students compared with Indigenous students.

One important question to consider is whether these results are the reality and truly what

teachers experience in their classrooms or if it is their implicit biases. Hahn, Judd, Hirsh, and Blair (2014) discussed how teachers' expectations are usually quite consistent with what actually happens. Unfortunately, in this study, I cannot determine the basis for these expectations; as such, they only provide limited insight about teachers' expectations.

Reported Teacher Behaviours

The second research question focused on reported teachers' behaviours. Teachers answered only one item from model 1 that was flagged with DIF from one model (item flagged: "I demand better performance from this student"). Teachers responded that they demand better performance from Asian students compared with Caucasian students. No other questions were flagged for DIF, meaning teachers reported that they behave similarly towards all three groups of students with regard to praise, challenging instruction, targeted feedback, and calling on students. Teachers reported demanding better performance for Caucasian and Indigenous students equally. I did not expect to find that teachers behave similarly towards all three groups of students, considering that past research has shown that teachers' expectations influence their behaviours.

One plausible reason for this unexpected result could be related to the fact that the questions from the Behaviour scale could have been interpreted differently by teachers. For example, the first question from this scale was, "I demand better performance from this student." This could have been interpreted as demanding better performance from a student who is struggling in academics, and therefore high achieving students do not need to be asked to perform better. Alternatively, it could be interpreted as demanding better performance or challenging students who are doing well. The fact that teachers could have interpreted this question in different ways could explain the result that teachers did not respond homogeneously.

Another plausible reason for the unexpected result of teachers behaving similarly towards

all three groups of students is the fact that the questions from this scale are asking teachers directly about their actions, which are more likely to be conscious decisions, whereas the Expectations scale is asking about how they perceive the students from these different groups to perform. The questions from the Behaviour scale may be better answered by direct observations of teacher behaviours in the classroom. That being said, Glock's (2016) survey research has shown teachers have reported differential behaviour for certain groups. Glock's (2016) survey differed from the one used in the current study because they had descriptions of students, which varied based solely on the gender and ethnicity of the student. Glock's results demonstrate that it is possible to get teachers to report behaving differently towards students of different ethnic groups using a survey, despite this being a more conscious decision. Future research targeting teacher behaviours may want to either observe teachers in classrooms or use surveys that include a description of a hypothetical student before answering questions.

Demographics and Expectations

The third research question is related to teachers' demographic information and whether certain characteristics influence teachers' expectations. The analyses showed that older teachers (above age 35) had higher expectations for their students and were 1.8 times more likely to choose higher (i.e., positive) response options on the Expectations scale. Teachers with six or more years of teaching experience were twice as likely to choose higher response options on the Expectations scale. This aligns with results from Rubie-Davies (2010) who reported that teachers with less teaching experience tended to have lower expectations for their students. Contrary to this, results from a study conducted by Dandy and colleagues (2015) showed that preservice teachers and experienced teachers had similar expectations for student achievement. Much of the other expectation research has not considered teacher experience so it might be the case that

more research is needed to get more consistent and accurate results.

Results from the third research question could be explained by the fact that teachers with more experience are exposed to students who contradict the stereotypes. Guo (2012) reviewed research showing that most preservice teachers reported that they had attended predominantly white, middle-class schools and had limited prior experiences with diversity, leading to fear, anxiety and uncertainty about their preparedness to work effectively with diverse student populations. Preservice teachers reported that they learned about diversity in their university education classes, but that they did not know how that would translate to practice. Although the current study was not able to gather information about participants' ethnicity, Guo's paper demonstrates that some preservice teachers may be lacking the experience of working with students of different backgrounds. As a result, teachers may have to use stereotypes because stereotypes provide cognitive shortcuts that allow us to organize information about certain groups. These cognitive shortcuts can be used when interacting with students that teachers do not know much about, such as new students at the beginning of the year. Again, the danger of doing this is that it takes away the individuality of people within that group. Teachers with more experience might not have to use these cognitive shortcuts (i.e., stereotypes) as much because they have more experience that suggests that there are unique individual differences for students within these groups.

Dei (1996) wrote there is a tendency to regard "difference as deficit," and that this acts as a barrier to teachers' learning about diversity. Rather than drawing on different cultural groups as sources of alternative strengths, experiences, knowledge, and perspectives, teachers may ignore diversity or perceive it as an obstacle to the learning process. This deficient thinking includes overgeneralizations about family background and having low expectations (Dei, 1996). As a

result, such generalizations suggest that these students are "deficient and in need of remediation" (Guo, 2012, p. 6). Again, these could be beliefs and stereotypes of teachers when they are in the early years of their career—once they get more experience with these students, they may realize that it is not true for everyone. Ultimately, more research is needed to gather more consistent results.

The results from the present study showed that males and females did not differ significantly in terms of their expectations, also found in other research (Dandy et al., 2015). Only 20% of participants from the present study were male so perhaps with a larger sample of males, there would be a different result. There were no significant results for any of the items in the Behaviour scale, which is likely because the majority of teachers responded that they behaved the same for all students (except for one item for Asian students where teachers reported demanding better performance compared with Caucasian students).

Implications

The results of this study will provide information and have strong implications for teachers and other professionals working with children from diverse backgrounds, education programs, and students themselves. This research has shown that teachers have lower expectations for Indigenous students and higher expectations for Asian students. Another result showed that less experienced teachers have lower expectations for students compared with more experienced teacher. These results have implications for education programs in order that new teachers may avoid having to acquire six or more years of teaching experience before developing higher expectations for students. Preservice teachers should be getting explicit instruction on working with students from different backgrounds, and instruction about the differential impact that expectations can have on students. Inadequate preparation of new teachers has been

identified as a contributing factor in the over-representation of minority children in remedial special education (Irvine, 2012). Teacher education programs could do more to promote culturally sensitive and reflective practices.

It is also essential, whether within teacher education programs or professional development events, that teachers and broader groups of school staff engage in deep reflection about the perceptions they hold of Indigenous students (Whitley, 2014) and other stigmatized groups. Anecdotally, one of the participants who completed this survey said she really liked the survey and thought it would be a great exercise to get teachers to complete this survey at the beginning of the year, even if it is used simply to check their biases and stereotypes. This type of activity or intervention can be compared to therapeutic assessment, which is a short-term intervention in which tests are used collaboratively with clients to help them understand themselves better and find solutions to their persistent problems. The main goal is to facilitate positive change in clients by providing them with feedback, which happens to be the assessment results. In the present case, the clients would be teachers or other professionals using this survey, and the "problem" is the lower expectations teachers have for some students of different ethnic groups. Poston and Hanson (2010) used meta-analytic techniques to calculate and analyze effect sizes across 17 published studies of psychological assessment as a therapeutic intervention. Their results showed a significant overall Cohen's *d* effect size of 0.42, which is considered a moderate effect. These results suggest that therapeutic assessment can have positive, clinically meaningful effects on treatment. This suggests that simply the act of bringing these results to respondents' awareness—when done with collaboration, respect, humility and compassion—can promote positive change, which in this case would be to have higher expectations for all students, but

particularly for Indigenous students and other groups of students who come from negatively stereotyped groups.

This study also has implications for other professionals, such as school psychologists and counsellors, working with children from a variety of backgrounds. It is not only teachers who have this stereotyped, "deficient" thinking; it is widespread, and becomes a greater problem when it negatively affects individuals' outcomes, such as academic achievement or therapy outcomes. Given that research has shown that therapists' hope in their clients is significantly related to client outcomes (e.g., Coppock, Owen, Zagarskas, & Schmidt, 2010), this study could even be replicated in psychotherapy, to see if psychologists' or counsellors' expectations or hope differ for clients of different ethnic groups.

Ultimately, the results of this study indicate direct effects for students across Canada and elsewhere, whose future educational outcomes and careers are influenced by their teachers' expectations. Teacher expectations, when negative, are especially dangerous for students from academically stereotyped groups, such as Indigenous students, who consistently fall behind their peers in terms of academic achievement.

Limitations and Directions for Future Research

One of the limitations of this study is the possibility that one question (or more) from the Behaviour scale could be interpreted in various ways depending on the teacher. This is a concern regarding the quality of the items. For future research using this survey, I would recommend getting feedback from teachers on the items from the Behaviour scale to see how they interpret the questions and I would work on improving those items. Artino and colleagues (2014) described two techniques used to get participants' thoughts while answering the survey items—these are often used when developing surveys. The first technique is *think-aloud* where

respondents are asked to verbalize every thought they have while answering each item. The examiner asks the respondent to keep talking and to record what is said for subsequent analysis. Although this technique provides valuable information, Artino et al. (2014) state that it tends to be "unnatural and difficult for most respondents." The second technique they described was *verbal probing*, and they specifically suggest a sub-technique of this called *immediate retrospective probing*. This form of data collection involves the examiner asking a series of probe questions designed to elicit specific information (e.g., "I noticed that you hesitated. Tell me what you were thinking."; "Can you restate the question in your own words?"). This approach limits the interruptions and decreases the artificiality of the process. In reality, many researchers use a mixture of these two techniques to better identify errors. In an effort to get feedback on the items from the Behaviour scale, I recommend using one, or a mixture of both techniques. Another method of measuring teacher behaviours is to videotape or observe teachers' behaviour in the classroom, which has been done successfully in Pianta's research (Spiegel, 2012).

A second limitation of this research is the sample size and scope of the sample. Ideally when doing these types of analyses, there should be more participants. Also, this research only gives a snapshot of the teachers living in the greater Edmonton area. It would be interesting to get responses from teachers in other Canadian provinces, which would also provide a larger sample size.

A third limitation and recommendation, related to our study design, is that I do not know whether teachers' expectations are predicting or causing student outcomes, and if they are more accurate than they are self-fulfilling. To answer those questions, which I think are important,

future research would need to measure students' achievement in addition to our measure. This would involve conducting a longitudinal investigation including standardized achievement tests.

Another recommendation for future research is to include other demographic information from teachers such as their ethnicity and information about the school(s) where they have worked (i.e., ethnic makeup of the school, SES). This may influence teachers' implicit biases.

Finally, this study could be modified and replicated to include students of different ethnic groups. For example, given that there are many new students in Canada that have immigrated from Syria recently, it would be relevant and interesting to examine teachers' expectations and beliefs held about those students. Furthermore, it could be worthwhile exploring teachers' expectations for other groups of students that have negative academic stereotypes, such as African Americans (McKown & Weinstein, 2008).

Conclusion

Hattie (2003) concluded that we need to focus on the source of variance of academic achievement that can make the greatest difference — teachers. Teachers' implicit biases have been under investigation for decades. The results from this study show that teachers have differing expectations for students of different ethnic groups, and specifically, teachers reported having higher expectations for Asian students and lower expectations for Indigenous students. Previous research demonstrated that expectations can have both a positive and negative impact on students' lives. In particular, it can have a negative effect on stigmatized students. Could this be the key to reducing the academic achievement gap? Future research is needed to answer the question of whether these teachers' expectations impact or influence students' academic achievement and consequently the academic achievement gap found between students of different ethnic groups.

References

Abada, T., Hou, F., & Ram, B. (2009). Ethnic differences in educational attainment among the children of Canadian immigrants. *Canadian Journal of Sociology, 34*(1), 1-28.

Artino, A. R., La Rochelle, J. S., Dezee, K. J., & Gehlbach, H. (2014). Developing questionnaires for educational research: AMEE Guide No. 87. *Medical Teacher*, *36*(6), 463-474. doi:10.3109/0142159X.2014.889814

Atalay Kabasakal, K., Arsan, N., Gök, B., & Kelecioğlu, H. (2014). Comparing performances (Type I error and power) of IRT likelihood ratio SIBTEST and Mantel-Haenszel methods in the determination of differential item functioning. *Kuram Ve Uygulamada Eğitim Bilimleri*, *14*(6), 2186-2193.

Battiste, M. & McLean, S. (2005). State of First Nations Learning. Retrieved from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.601.8068&rep=rep1&type=pdf

Brophy, J. E. (1985). Teacher–student interaction. In J. B. Dusek (Ed.), *Teacher Expectancies* (pp. 303–328). Hillsdale, NJ: Lawrence Erlbaum.

Brophy, J., & Good, T. (1974). *Teacher-student relationships: Causes and consequences*. New York: Holt, Rinehart and Winston.

Brynner, J. (2008). Illuminating disadvantage. *Education Journal*, *110*, 34–35.

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software, 48*(6), 1-29.

Coppock, T. E., Owen, J. J., Zagarskas, E., & Schmidt, M. (2010). The relationship between therapist and client hope with therapy outcomes. *Psychotherapy Research, 20*(6), 619-626.

Dandy, J., Durkin, K., Barber, B. L., & Houghton, S. (2015). Academic expectations of Australian students from Aboriginal, Asian and Anglo backgrounds: Perspectives of teachers, trainee-teachers and students. *International Journal of Disability, Development and Education, 62*(1), 60-82. doi:10.1080/1034912X.2014.984591

De Boeck, P., & Wilson, M. (2004). *Explanatory item response models: A generalized linear and nonlinear approach*. New York: Springer-Verlag.

De Houwer, J. (2006). What are Implicit Measures and Why are We Using Them? In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of implicit cognition and addiction* (pp. 11–29). Thousand Oaks, CA: SAGE Publications, Inc.

Dei, G. (1996). *Anti-racism education: Theory and practice*. Halifax: Fernwood Publishing.

Friedrich, A., Flunger, B., Nagengast, B., Jonkmann, K., & Trautwein, U. (2015). Pygmalion effects in the classroom: Teacher expectancy effects on students' math achievement. *Contemporary Educational Psychology*, *41,* 1-12. doi:10.1016/j.cedpsych.2014.10.006

Glock, S. (2016). Does ethnicity matter? The impact of stereotypical expectations on in-service teachers' judgments of students. *Social Psychology Of Education*, *19*(3), 493-509. doi:10.1007/s11218-016-9349-7

Guo, Y. (2012). Exploring linguistic, cultural, and religious diversity in Canadian schools: Preservice teachers' learning from immigrant parents. *Journal of Contemporary Issues in Education, 7*(1), 4-23.

Hahn, A., Judd, C. M., Hirsh, H. K., & Blair, I. V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General*, *143*(3), 1369-1392.

Hattie, J. (2003). Teachers Make a Difference, What is the Research Evidence? Retrieved from:

http://research.acer.edu.au/cgi/viewcontent.cgi?article=1003&context=research_conference_2003

Hoffman-Goetz, L., Donelle, L., & Ahmed, R. (2014). Health literacy in Canada: A primer for students. Toronto, ON: Canadian Scholars' Press Inc.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. Retrieved from: http://arrow.dit.ie/cgi/viewcontent.cgi?article=1001&context=buschmanart

Hui-Michael, Y., & García, S. B. (2009). General educators' perceptions and attributions about Asian American students: Implications for special education referral. *Multiple Voices for Ethnically Diverse Exceptional Learners, 12*, 21–37.

Irvine, J. J. (2012). Complex relationships between multicultural education and special education: An African American perspective. *Journal of Teacher Education, 63*, 268–274.

Jussim, L., Eccles, J.S., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 281–388). San Diego, CA: Academic Press.

Jussim, L., & Harber, K. D. (2005). Teacher Expectations and Self-Fulfilling Prophecies: Knowns and Unknowns, Resolved and Unresolved Controversies. *Personality And Social Psychology Review*, *9*(2), 131-155. doi:10.1207/s15327957pspr0902_3

Jussim, L., Smith, A., Madon, S., & Palumbo, P. (1998). Teacher expectations. In J. E. Brophy (Ed.), *Advances in research on teaching: Expectations in the classroom* (pp. 1–48). Greenwich, CT: JAI Press.

Lee, J. (2015). For Asian-American students, stereotypes help boost achievement. Retrieved from: https://theconversation.com/for-asian-american-students-stereotypes-help-boost-achievement-46052

Lee, J., & Zhou, M. (2015). *The Asian American achievement paradox*. New York, NY, US: Russell Sage Foundation.

McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology, 46*(3)*,* 235-261. doi: 10.1016/j.jsp.2007.05.001

Parkin, A. (2015). International Report Card on Public Education: Key Facts on Canadian Achievement and Equity. Retrieved from: http://www.environicsinstitute.org/uploads/institute-projects/environics%20institute%20-%20parkin%20-%20international%20report%20on%20education%20-%20final%20report.pdf

Penfield, R. D., Gattamorta, K., & Childs, R. A. (2009). An NCME instructional module on using differential step functioning to refine the analysis of DIF in polytomous items. *Educational Measurement: Issues And Practice*, *28*(1), 38-49. doi:10.1111/j.1745-3992.2009.01135.x

Peterson, E. R., Rubie-Davies, C., Osborne, D., & Sibley, C. (2016). Teachers' explicit expectations and implicit prejudiced attitudes to educational achievement: Relations with student achievement and the ethnic achievement gap. *Learning and Instruction, 42*, 123-140. doi:10.1016/j.learninstruc.2016.01.010

Poston, J. M., & Hanson, W. E. (2010). Meta-Analysis of Psychological Assessment as a Therapeutic Intervention. *Psychological Assessment*, *22*(2), 203-212.

R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Revelle, W. (2016). psych: Procedures for personality and psychological research. Retrieved from: https://CRAN.R-project.org/package=psych Version = 1.6.12.

Richards, J. (2011). *School dropouts: Who are they and what can be done?* Retrieved from: https://www.cdhowe.org/sites/default/files/attachments/research_papers/mixed//ebrief_109.pdf

Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions In Psychological Science*, *3*(6), 176-179. doi:10.1111/1467-8721.ep10770698

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectations and student intellectual development*. New York: Holt.

Rubie-Davies, C. M. (2015). *Becoming a high expectation teacher: Raising the bar*. London, UK: Routledge.

Rubie-Davies, C., Hattie, J., & Hamilton, R. (2006). Expecting the best for students: Teacher expectations and academic outcomes *British Journal Of Educational Psychology*, 76(3), 429-444. doi:10.1348/000709905X53589

Spiegel, A. (2012). Teachers' Expectations Can Influence How Students Perform. Retrieved from: http://www.npr.org/sections/health-shots/2012/09/18/161159263/teachers-expectations-can-influence-how-students-perform

Statistics Canada. (2013). Skills in Canada: First results from the Programme for the International Assessment of Adult Competencies (PIAAC). Retrieved from

http://cmec.ca/Publications/Lists/Publications/Attachments/315/Canadian-PIAAC-Report.EN.pdf

Statistics Canada. (2010). *Projections of the diversity of the Canadian population, 2006-2031*. Retrieved from: http://www.statcan.gc.ca/pub/91-551-x/91-551-x2010001-eng.pdf

Statistics Canada. (2006). Aboriginal children's survey. Ottawa, ON: Statistics Canada.

Strand, S. (2013). *Insights: What accounts for ethnic achievement gap in secondary schools in England*. Retrieved from http://www.education.ox.ac.uk/wordpress/wp-content/uploads/2013/12/Insights4_Ethnic-Achievement_Strand.pdf

Strand, S. (2014). Ethnicity, gender, social class and achievement gaps at age 16: Intersectionality and 'Getting it' for the white working class. Retrieved from: https://www.researchgate.net/publication/264412254_Ethnicity_gender_social_class_and_achievement_gaps_at_age_16_Intersectionality_and_'Getting_it'_for_the_white_working_class

Weinstein, R. S., Gregory, A., & Strambler, M. J. (2004). Intractable Self-Fulfilling Prophecies Fifty Years After Brown v. Board of Education. *American Psychologist*, *59*(6), 511-520. doi:10.1037/0003-066X.59.6.511

Whitley, J. (2014). Supporting educational success for Aboriginal students: Identifying key influences. *McGill Journal of Education*, *49*(1), 155–181. doi:10.7202/1025776ar

Appendix A

Survey

PLEASE READ ALL TEXT ON THIS PAGE CAREFULLY
As a teacher, you have to take into account many factors when considering a student's performance. Many times, you have to proceed partly on the basis of your intuitions, although these are informed by your experience. You might have recognized patterns that can provide you with at least an approximate starting point for evaluating a specific student. You will be asked to estimate likely educational outcomes for a typical student in a specific cultural group, highlighted in yellow. Of course, you will be aware that there are always exceptions to general rules. But, for the purposes of this exercise, please give your best estimate based on your general expectations of the typical student in the identified group.

*Please think of a typical **Canadian student from an Asian cultural background** (i.e., a student who was born and raised in Canada but whose parents are Asian).*

From your experience, indicate how a typical Asian student tends to compare to other students regarding:	**Bottom of the Class**	**Below Average**	**Average**	**Above Average**	**Top of the Class**
1. the level of support this student receives from family that contributes to success in school.	☐	☐	☐	☐	☐
2. the amount of effort you observe in the classroom.	☐	☐	☐	☐	☐
3. how well this student performs on classroom assessments (e.g., activities, tests).	☐	☐	☐	☐	☐
4. how well this student performs on assigned homework.	☐	☐	☐	☐	☐
5. this student's natural academic abilities.	☐	☐	☐	☐	☐

Now, think about your behaviour in response to a typical Asian student. How often do you:	**Never**	**Rarely**	**Sometimes**	**Often**	**Always**
1. demand better performance from this student?	☐	☐	☐	☐	☐
2. praise this student for good performance?	☐	☐	☐	☐	☐
3. make sure to challenge this student during instruction?	☐	☐	☐	☐	☐
4. provide this student with targeted feedback?	☐	☐	☐	☐	☐
5. call on this student to provide the answer?	☐	☐	☐	☐	☐

*Please think of a typical **Canadian student from a European-American cultural background** (i.e., Caucasian student).*

From your experience, indicate how a typical Caucasian student tends to compare to other students regarding:	**Bottom of the Class**	**Below Average**	**Average**	**Above Average**	**Top of the Class**
1. the level of support this student receives from family that contributes to success in school.	☐	☐	☐	☐	☐
2. the amount of effort you observe in the classroom.	☐	☐	☐	☐	☐
3. how well this student performs on classroom assessments (e.g., activities, tests).	☐	☐	☐	☐	☐
4. how well this student performs on assigned homework.	☐	☐	☐	☐	☐
5. this student's natural academic abilities.	☐	☐	☐	☐	☐

Now, think about your behaviour in response to a typical Caucasian student. How often do you:	**Never**	**Rarely**	**Sometimes**	**Often**	**Always**
1. demand better performance from this student?	☐	☐	☐	☐	☐
2. praise this student for good performance?	☐	☐	☐	☐	☐
3. make sure to challenge this student during instruction?	☐	☐	☐	☐	☐
4. provide this student with targeted feedback?	☐	☐	☐	☐	☐
5. call on this student to provide the answer?	☐	☐	☐	☐	☐

Please think of a typical ***Canadian student from an Indigenous cultural background*** *(i.e., First Nations, Métis, Inuit).*

From your experience, indicate how a typical Indigenous student tends to compare to other students regarding:	**Bottom of the Class**	**Below Average**	**Average**	**Above Average**	**Top of the Class**
1. the level of support this student receives from family that contributes to success in school.	☐	☐	☐	☐	☐
2. the amount of effort you observe in the classroom.	☐	☐	☐	☐	☐
3. how well this student performs on classroom assessments (e.g., activities, tests).	☐	☐	☐	☐	☐
4. how well this student performs on assigned homework.	☐	☐	☐	☐	☐
5. this student's natural academic abilities.	☐	☐	☐	☐	☐

Now, think about your behaviour in response to a typical Aboriginal student. How often do you:	**Never**	**Rarely**	**Sometimes**	**Often**	**Always**
1. demand better performance from this student?	☐	☐	☐	☐	☐
2. praise this student for good performance?	☐	☐	☐	☐	☐
3. make sure to challenge this student during instruction?	☐	☐	☐	☐	☐
4. provide this student with targeted feedback?	☐	☐	☐	☐	☐
5. call on this student to provide the answer?	☐	☐	☐	☐	☐

**Did you have a specific student(s) in mind when asked to think of a typical student from these cultural backgrounds (check all that apply)?	☐ Asian	☐ Caucasian	☐ Indigenous