A Mixed Methods Study of Student Engagement, Attitude, and Achievement in Grade 10 Mathematics

Jeff Irvine*

To cite this acticle:

Irvine, J. (2020). A mixed methods study of student engagement, attitude, and achievement in grade 10 mathematics. *Journal of Mixed Methods Studies*, Issue 2, 26-46 [Online] www.jomesonline.com DOI: 10.14689/jomes.2020.2.3

Abstract. This paper explores the utility of employing a mixed methods methodology when investigating latent variables that cannot be measured directly. The larger study investigated the efficacy of a classroom intervention in modifying student affective dimensions and student achievement. The study involved 68 students in three classes of Grade 10 Mathematics at one secondary school in Ontario, Canada. Over a one month period, students experienced a classroom intervention that was active, hands-on, connected to real-world situations, made extensive use of student groups, allowed students choice, and utilized manipulatives and technology to enhance student engagement while positively influencing students' attitudes. Statistically significant effect sizes were found for engagement and attitude, but not for achievement. Several notable results were found that would not have been revealed without both the quantitative and qualitative dimensions of mixed methods methodology. The study emphasized the utility of employing a mixed methods methodology, particularly with respect to student affective variables.

Keywords: Mixed methods, engagement, attitude, achievement, mathematics

Introduction

This paper interrogates the value added through the use of a mixed methods methodology in a larger study involving a classroom intervention and its effect on student engagement, attitudes and achievement in secondary school mathematics.

The instructional intervention was designed in conjunction with classroom teachers, who had extensive input into both the activities involved as well as the sequencing of the lessons within the study. The instructional intervention emphasized the active involvement of students in their own learning; extensive use of student groups; opportunities for students to have choice in aspects of their learning; real-world problems and problems related to students' daily lives; technology as a toll where appropriate; manipulatives and model building. The intervention utilized student activities; direct instruction; student opinion student surveys; and metacognitive activities such as goal-setting. The focus of the intervention was to specifically address student motivation and metacognition. Specific cognitive strategies were also used to stimulate student engagement and interest. A wide range of instructional strategies were employed, such as four corners, jigsaw, placemat, and think-pair-share.

Research Questions

This paper utilized the following research questions with respect to the value added in utilizing a mixed methods methodology:

1. What was the additional value to the study of using a quantitative dimension in the methodology?

* Brock University, St. Catharines, Canada, dr.jeffrey.irvine@gmail.com, ORCID: 0000-0002-2151-8699

2. What was the additional value to the study of using a qualitative dimension in the methodology?

3. What, if any, findings would have been downplayed or omitted altogether by using only quantitative or only qualitative methodology in this study?

The main study's goals investigated the impact on student engagement, attitudes towards mathematics and student achievement of instructional strategies, that specifically address student motivation and metacognitive dimensions. The principal hypothesis for the main study was that students in the treatment classes would demonstrate significant increases in engagement, attitudes towards mathematics, and achievement compared to students in the control class. The results of that study with respect to engagement and attitudes are reported elsewhere (Irvine, 2020) although some results are also reported here, to demonstrate the use of a mixed methods methodology and emphasize the new information that was obtained through the use of both quantitative and qualitative methods.

Method

Prior to commencement of the study, volunteer teachers were interviewed to determine attributes such as their attitudes towards the teaching and learning of mathematics, knowledge of assessment and instructional strategies, and details of their teaching careers. An interview guide was developed for this purpose. Teachers were offered up to 2 days of professional development but chose instead to have a series of meetings with the researcher to become familiar with the control and treatment materials. The teachers provided several suggestions with respect to sequencing and strategies, most of which were incorporated into the classroom instructional intervention. This encouraged teacher buy-in and ownership of the intervention. Throughout the intervention, the researcher was available as a resource but did not engage in any classroom teaching. The researcher observed approximately 25% of classes over the duration of the study.

All classes received instruction in the same mathematical content. However, instructional strategies for the two treatment classes involved motivation and metacognition while the control class received lessons using more traditional instruction.

Grade 10 was selected based on the relative homogeneity of prior knowledge, since all students had completed the Grade 9 Academic mathematics course. In addition, confounding factors such as the transition from Grade 8 to Grade 9 and attending a new (and usually larger) school were minimized since the students had attended the same school in the prior academic year. The Grade 10 mathematics curriculum consists of four units linear systems, quadratic relations, analytic geometry, and trigonometry. The instructional intervention involved the quadratic relations unit.

Before the classroom intervention all students completed surveys on attitude and engagement, using technology. Each student was given a unique identifier code, to enable post-treatment comparisons. During the study teachers completed daily reflections evaluating the efficacy of the day's lesson and activities. Students completed weekly reflections, commenting on their affective reactions to the week's lessons. Students also completed written summative assessments, one during the unjt and the other at the end of the unit. The summative assessments consisted of written paper-and-pencil tests, scored with marking schemes. Both assessments were created and evaluated by the teachers involved in the survey, with researcher input. At the end of the unit, students completed a rich assessment task, designed by the researcher but scored with a teacher-created rubric. Students then completed post-intervention surveys on engagement and attitude to determine the effects of the intervention on student affect.

After completion of the treatment, student volunteers were identified to participate in audiotaped interviews. Permission forms were given for parent consent. Five students volunteered, and all were interviewed after receiving completed permission forms. An interview guide for these student interviews was developed. All students were assigned pseudonyms when information was reported in the results

section. At the conclusion of the study, both teachers participating in the research were interviewed again, using a separate targeted interview guide.

In order to make warranted assertions about latent variables such as engagement and attitude (Reeve 2013; Vandecandelaere et al. 2012),, multiple data sources are needed to provide sufficient evidence. Therefore, the study used a mixed methods methodology (Teddlie and Tashakkori, 2009).

Mixed methods methodology is defined as "the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study" (Johnson & Onwuegbuzie, 2004, p. 17). The fundamental principle of mixed research (Johnson, Onwuegbuzie, & Turner, 2007) employs multiple data collection methods using different strategies and approaches to obtain complementary strengths and nonoverlapping weaknesses. Thus, dimensions of the problem that may not be apparent using only one data collection method may be revealed when multiple methods are employed. Using a mixed methods methodology also allows for triangulation, whereby results are supported by independent observations of the same result using different data collection methods (Teddlie and Tashakkori 2009).

On the continuum of quantitative methods to qualitative methods, mixed methods methodology occupies the middle ground. However, quantitative methods and qualitative methods need not be equally represented in a study. The current study can be identified as QUAN+qual (Teddlie and Tashakkori, 2009), meaning that quantitative methods (student surveys) are the dominant method, followed by qualitative methods (student interviews) which support and elaborate on the quantitative findings.

In addition, the study employed content analysis, which Krippendorff (2013) describes as "a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use" (p. 24). Krippendorff's conceptual framework for content analysis encompasses:

- A body of text, the data that a content analyst has available to begin an analytical effort;
- A research question that the analyst seeks to answer by examining the body of text;
- A context of the analyst's choice within which to make sense of the body of text;
- An analytical construct that operationalizes what the analyst knows about the context of the body of text;
- Inferences that are intended to answer the research question, which constitute the basic accomplishment of the content analysis;
- Validating evidence, which is the ultimate justification of the content analysis. (p. 35)

This study also employed constructivist grounded theory (Charmaz, 2014) to identify themes not found through content analysis. Constructivist grounded theory makes no a priori assumptions about whether a relationship exists between variables; instead, the grounded theory methodology (known as the constant comparison method) uses an iterative approach of data collection, data analysis, and additional data collection based on this analysis, until data saturation is reached, and no new insights or new properties are generated by further data collection. Constructivist grounded theory employs multi-level coding, proceeding from initial open coding to more focused coding once themes have emerged, to axial coding, which relates coding categories to subcategories, and finally to theoretical coding that links categories to produce a hypothesis or theory (Noerager Stern & Porr, 2011). Data analysis is continued until no new information or relationships are identified (Sutcliffe, 2016). Grounded theory is a method of seeking an "inductive generalized pathway as opposed to a deductive verificational pathway" (Noerager Stern & Porr, 2011, p. 39). Constructivist grounded theory is a particularly appropriate tool for textual analysis (Charmaz, 2014) as well as policy analysis (Richards & Farrokhnia, 2016).

Employing a mixed methods methodology was appropriate for this study since two of the variables of interest (namely, engagement and attitude) are latent variables that must be imputed from observed or reported behaviours (Reeve, 2013; Vandecandelaere et al., 2012). The third variable of interest, achievement, is usually identified as observable, although some dimensions of achievement, such as

understanding, must also be imputed. Therefore, multiple data sources were needed in order to make warranted assertions about the results of the study. The quantitative phase was exploratory, to determine whether a relationship existed between the classroom interventions and changes in student engagement, student attitude, or student achievement. Because engagement and attitude are latent constructs, the qualitative phase was used to support conclusions drawn in the quantitative phase, through triangulation of data. The qualitative phase was also explanatory, to begin to construct relationships detailing how the classroom intervention affected the student variables, namely engagement, attitude, and achievement.

Theoretical Framework

Mixed methods methodology is grounded in pragmatism, "a philosophy that encourages us to seek out the processes and do the things that work best to help us achieve desirable ends" (Ozmon and Craver 2003, 127). In this definition, we can see elements of the colloquial view of pragmatic, "relating to matters of fact or practical affairs often to the exclusion of intellectual or artistic matters: practical as opposed to idealistic" (Merriam-Webster 2019). On the continuum of idealism to realism, pragmatism certainly lies closer to realism. However, pragmatists like John Dewey would take issue with the aforementioned phrase "often to the exclusion of intellectual or artistic matters"; while pragmatists claim consequences as the final test for thought, these consequences may be social, aesthetic, moral or ethical (Ozmon and Craver 2003).

Johnson and Onwuegbuzie (2004) state that pragmatism as a philosophical foundation for mixed methods methodology is not without significant debates. Quantitative methods are grounded in positivism, the view that there exists an objective truth which can be discerned using scientific methods, and that the observer is separate from the phenomenon that is being observed. In contrast, qualitative methods are based on constructivist or interpretivist views. This position holds that there are multiple versions of reality that are person-centred and that generalisations about phenomena are not possible since they are connected to the viewpoint of the observer. This dichotomy lead to the *incompatibility thesis*, which stated that the two methodologies cannot be part of the same study because of epistemological differences (Teddlie &Tashakkori, 2009). However, subsequent discussion resulted in what Teddlie and Tashakkori (2009) call the *dialectical thesis*, which "assumes that all paradigms have something to offer and that the use of multiple paradigms contributes to greater understanding of the phenomenon under study" (99). This study accepts the dialectical thesis and demonstrates its viability in an instructional intervention environment.

Data Collection

The sample population comprised 68 students across three classes of Grade 10 Mathematics at a secondary school in Ontario, Canada (Table 1).

Table 1.

Classes Involved in the classroom Intervention Study

Class	n	Female	Male	Timetable period(s)	Teacher pseudonym
Control	23	16	7	3	Ms. Alford
Treatment	23	14	9	1,5	Ms. Beckham
Treatment	22	13	9	2,4	Ms. Beckham
Combined treatment classes	45	27	18		Ms. Beckham

The month-long intervention included hands-on activities that were connected to real-world situations, made extensive use of student groups, allowed student choice, and utilized manipulatives and technology to enhance student engagement while positively influencing

students' attitudes; Quantitative data collected were student surveys of engagement and attitude; student written assessments; and student performance on the rich assessment task. Qualitative data were student interviews; student weekly reflections; teacher interviews; teacher daily reflections; researcher classroom observations.

Quantitative Data Collection

Engagement. To measure student engagement quantitatively, pre- and post-intervention surveys using Reeve's (2013) *Dimensions of Student Engagement Survey©* (*DSES*) were employed. This instrument consisted of a total of 39five-point Likert scale questions (*1=strongly disagree, 5=strongly agree*). Written permission to use this survey was obtained from Dr. Johnmarshall Reeve, the copyright holder. This survey had four subscales: behavioural engagement, emotional engagement, cognitive engagement, and agentic engagement. No changes were made to the wording of any question. The survey consisted of blocks of questions on each dimension of engagement (cognitive, emotional, behavioural, agentic). The first five questions in each block asked about engagement; the remaining questions asked about disengagement and were reverse coded when analyzed. The only change made to this survey was to randomize the question order.

The *DSES* had a Cronbach's α of 0.95 on 39 items. All four subscales (cognitive engagement, behavioral engagement, emotional engagement, and agentic engagement) had strong Cronbach's α values as well (Table 2). This is consistent with reliability values found in the literature (Reeve, 2013).

Table 2.

Reliability for Dimensions of Student Engagement Survey and Subscales

Scale	Number of items	Cronbach's α
Engagement (full scale)	39	0.950
Cognitive	12	0.855
Behavioral	10	0.853
Emotional	7	0.898
Agentic	10	0.867

Reproduced with permission from Irvine (2020a).

Attitude. Quantitative analysis for student attitudes employed the *Attitudes Towards Mathematics Inventory©* (*ATMI*; Tapia & Marsh, 2005).This survey used a five-point Likert scale (*1=strongly disagree, 5=strongly agree*). The *ATMI* was a 40-item questionnaire, with subscales value, self confidence, enjoyment, and motivation. Written permission to use this survey was obtained from Dr. Martha Tapia, the copyright holder. This instrument was used verbatim, with no changes to wording or order of questions.

Table 3.

Reliability of Attitudes Towards Mathematics Inventory and Subscales

Scale	Number of items	Cronbach's α
Attitudes (full scale)	40	0.978
Value	10	0.938
Self confidence	11	0.927
Enjoyment	14	0.962
Motivation	5	0.910

The *ATMI* had a Cronbach's α of 0.978. This is consistent with results in the literature (Asante, 2012). All four subscales, value, self confidence, enjoyment, and motivation had strong Cronbach's α values as well (Table 3); again, these values were consistent with reliability values found in the literature (Majeed et al., 2013).

Achievement. The achievement instruments consisted of a rich assessment task and written summative assessments. The control class and the treatment classes received identical assessments. This ensured that all classes involved in the study were evaluated in a consistent manner. Thus, the rich assessment task and the summative assessment were quantitative, and the teacher observations and student interview responses were qualitative. Students' prior- achievement data were collected using self-reported final grade in their last math course taken, as a letter grade. Since this self-report was on a five-point Likert scale, it is possible that students' self-reported grades contained an upward bias. Thus, the variety of data collection methods allowed triangulation of all variables of interest (engagement, attitude, achievement) to increase validity.

Qualitative Data Collection

Qualitative data consisted of student interviews; student reflections; teacher interviews; and researcher classroom observations. Interview protocols (Jacob & Furgerson, 2012; Whiting, 2008) for the student interviews and separate protocols for the teacher interviews were developed. These interview protocols were designed by the researcher based on a review of relevant literature and focused on the three research questions of this study.

The qualitative data for engagement and attitudes utilized student and teacher interviews. These same interviews were utilized to investigate student perceptions of achievement. Therefore, for all variables of interest, multiple data sources were developed.

The data consisted of student survey data, student achievement measures, student and teacher interview data, student reflections, teacher reflections, and classroom observations by the researcher. With the exception of the researcher classroom observations and student achievement data, all the remaining data were self-reported, which has implications for limitations. The quantitative data consisted of a mix of parametric and nonparametric data. Student achievement data was parametric. Student survey data were nonparametric, and so several comparisons involved a mixture of parametric and nonparametric data.

Quantitative Data Analysis

There is considerable debate concerning the most appropriate statistical methods to employ with Likert scale scores (Mircioiu & Atkinson, 2017). Jamieson (2004) points out that Likert scale scores are ordinal. They do not possess most of the characteristics associated with parametric statistical tests: interval scales, continuous variables, homoscedasticity, or normal distribution; therefore, Jamieson argues that nonparametric tests are the appropriate measures. However, Norman (2010) argues that parametric tests are robust with respect to violation of underlying assumptions, and therefore can be used to generate valid conclusions for Likert scale data. This position was supported by Poncet et al. (2016) using computer simulations of data drawn from various distributions. A third position taken by Peró-Cebollero and Guàrdia-Olmos (2013) specifies that if the sample size is sufficiently large (usually N>30 or N>50) and the distribution is approximately normal (not defined) then parametric tests are appropriate.

A second consideration is statistical power. Parametric tests are frequently cited as having greater statistical power (Zumbo & Zimmerman, 1990); however, using computer simulations, Erceg-Hurn and Mirosevich (2008) and Larson-Hall (2012) found that nonparametric tests had equal or greater power, depending on the shape of the underlying distribution.

This study did not take a position on whether the Likert scale data were parametric or nonparametric. Since the majority of Likert scale data in recent studies (e.g., Argon & Kaya, 2018) were analyzed using parametric statistical tests, parametric methods were the primary statistical analysis method used in this study (Naiman et al.,2000). Paired *t*-tests were used to analyze pre–post comparisons. One-way ANOVA was used for the treatment-control comparisons. Treatment-control comparisons were made for T1-control, T2-control, and T_{Total}-control. Levene's test of equal variances (Derrick, Ruck, Toher, & White, 2018) was used to ensure that results were not affected by differences in variances among the various

classes. Results were confirmed using nonparametric methods (Mann-Whitney U tests, Wilcoxon signed rank tests). The significance level for all tests was set at α=0.05.

Correlational analysis was used to investigate associations among engagement, attitude, achievement, and prior achievement for all students in each class participating in the study, including the control class. Since much of this data was nonparametric, Spearman's rho was the appropriate correlational measure (Naiman et al., 2000). The variables of interest were engagement and attitude (pre- and post), changes in engagement and attitude (pre- and post), incremental changes in engagement and attitude of the treatment classes with respect to the control class, and achievement. Control class and treatment classes were compared and individual students' pre- and post- responses in attitude and engagement, and effect sizes were estimated (Naiman et al., 2000).

Qualitative Data Analysis

Table 4.

A Priori Coding Table

Construct	Look-fors
Engagement	Interesting
	Worked hard
	Slacked off
	Asked questions
	Did homework
	Didn't do homework
	Exciting
	Useful
	Active
	Thought about it
Attitude	Interesting
	Useful
	Valuable
	Enjoyed
	Looked forward to class
	Confident
	Uninteresting
	Boring
	Not useful
	Uncomfortable
	Afraid
	Scared
Achievement	Did well
	Did better
	Didn't do as well
	Understood
	Didn't' understand
	Was difficult
	Did OK
	Higher marks
	Lower marks

All interview data were transcribed by the researcher, and member checking occurred to ensure the accuracy of the transcription. The student and teacher interviews were coded and vetted by the researcher using inductive content analysis (Krippendorff, 2013) to identify themes. An a priori coding table (Saldana, 2014) was developed for this purpose (see Table 1). The entries in the table were based on look-fors cited in the literature for similar studies.

Subsequently, constructivist grounded theory (Charmaz, 2014) was employed to identify unanticipated themes that were not found using primary content analysis. Student interview data and teacher interview data were dealt with separately. While five student interviews are insufficient to independently develop theories using constructivist grounded theory (Charmaz, 2014), the interview data were nonetheless used to support conclusions drawn from other sources through triangulation. The teacher interview data was used to provide supporting evidence of themes identified in the student data analysis.

Ethical Considerations

Any research involving human subjects requires consideration of the ethics involved. Impacts on students from this study were minimized, since (a) all lessons were delivered by their regular classroom teacher; (b) the length of the study was approximately 4 weeks, approximately 20% of the total class time in a semester; and (c) the proportion of the students' final grades directly dependent on this unit of work was commensurate with the proportion of time spent in class on this topic, and the results of the study had no direct impact on students' marks.

Informed consent was obtained from the students as well as their parents or guardians. A letter outlining the intent of the study, its duration, and potential benefits was sent to all participants and their parents, as well as a separate letter to teachers who were considering participating in this study. The letter explained that participation in the study is voluntary, that students may choose to opt out of the study at any time, and that all student data were anonymous and confidential. The letter also contained a mechanism for asking questions. It stressed that the outcome of the study had no bearing on the students' final grades, and that the results of the study could lead to better instruction for students in the future. The researcher personally explained the study to the students and answered any questions. Students who opted out remained in their class, and received instruction from their teacher, but the students' data were excluded from the study.

The mechanism for identifying students who volunteered for follow-up interviews was outlined in a separate letter, again stressing that this participation was voluntary. Additional written consent was obtained for the students who were identified for follow-up interviews, signed by both the students and their parents.

All data collected from the teachers and students were securely stored electronically for the duration of the study, in a password-protected location. Dissemination of results will be anonymized and reported accordingly. At the end of the study, all student and teacher data will be destroyed. This study was approved by the university Research Ethics Board (file #17-096).

Findings

Since this was part of a larger study, some results related to engagement and attitude have already been reported in Irvine (2020a). Tables and figures reproduced with permission. The results described here are included to illustrate both the mixed methods methodology and the results that would have been downplayed or missed completely if only one of quantitative or qualitative methods had been used. Some results from the main study, concerning engagement and attitude, have already been published in Irvine (2020a).

Quantitative Findings

Student pre- and post intervention survey data for both treatment and control classes were analyzed. In addition, treatment-control comparisons were conducted, comparing control class results to each treatment class, and to the entire treatment student population.

Engagement. Quantitative results indicated a positive effect size for the classroom intervention on student engagement.

***Pre–post comparisons*.** 84% of students in the treatment classes had increases in overall engagement(M=0.44, SD=0.816, min=-1.46, max=3.48) as shown in Table 4A statistically significant difference and positive effect size of 0.54 was found (M=0.527, SD=0.694, $t_{(45)}$=5.29, p<0.001) when pre- and post-scores for treatment students was examined. This effect size is considered medium (Cohen, 1992) and suggests that the intervention had a positive impact on student engagement. The overall engagement score and all subscales had statistically significant increases (Table 4). The greatest increase occurred for the agentic engagement subscale, which measures student self-advocacy; this includes behaviours such as asking questions, indicating interest, and telling the teacher which learning activities best fit the student's learning style.

Table 5.

Pre- and Post-DSES Scores for Treatment Students (T_{Total})

	n	Pre		Post		*t*	df	Sig.	Cohen d
		Mean	Std. deviation	Mean	Std. deviation				
Engagement	46	3.05	0.079	3.57	0.080	5.209	45	<0.001***	0.54
Emotional	46	2.96	0.756	3.68	0.659	6.216	45	<0.001***	0.65
Behavioral	46	3.39	0.093	3.73	0.618	2.868	45	0.006**	0.38
Agentic	46	2.78	0.087	3.60	0.836	6.991	45	<0.001***	0.73
Cognitive	46	3.12	0.084	3.44	0.947	2.936	45	0.005**	0.31

Note. ** significant at p=0.01; ***significant at p=0.001. Reproduced with permission from Irvine (2020a).

The largest increases (Engagement post minus Engagement pre) were shown by students who had reported the lower scores in the pre-intervention survey (see Figure 1). In Figure 1 each data point gives the change (Engagement post minus Engagement pre) for a particular student in the treatment group. If only qualitative methods had been employed, this important distinction would not have been observed.

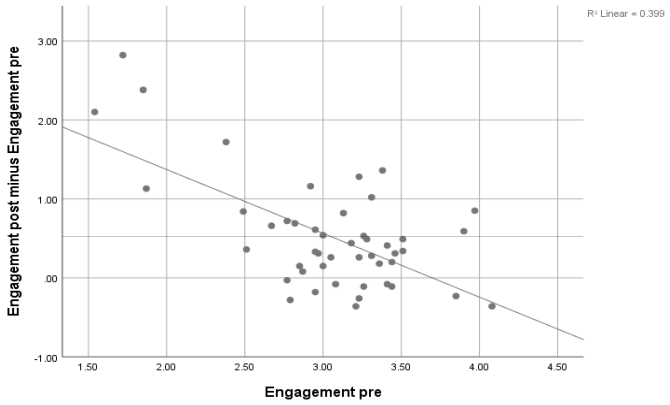

Figure 1. *Comparison of magnitudes (Engagement post minus Engagement pre) for each student, showing that greatest changes occurred for students with initially low engagement pre scores. Trend line shown. Reproduced with permission from Irvine (2020a).*

Treatment-control comparisons. **? Before** the intervention, the control class showed a significant advantage over the treatment classes (M= 0.34, SD= 0.158, $t_{(66)}$=2.140, p=0.036). After the intervention, this difference between the control class and the treatment classes had disappeared.(M=-0.24, SD=1.024, $t_{(21)}$=-1.100, p=0.284). Therefore the engagement levels for the treatment classes had

increased, but not for the control class. Again, this is a distinction that would not have manifested itself if only qualitative data had been collected.

Attitude. Table 5 shows statistics for the *ATMI* and subscales for T_{Total} (all treatment students).

Pre–post comparisons. In the treatment classes, 76% of students showed a positive increase in their attitudes towards mathematics. This represents a significant positive effect size of 0.32 (M=0.270, SD=0.0870, $t_{(45)}$=3.110, p=0.003).While overall attitude was statistically significant, only the self-confidence subscale increase was significant(Table 5).

Table 6.

Pre- and Post ATMI Overall and Subscale Scores for T_{Total}

Category	n	Pre		Post		t	df	Sig.	Cohen d
		Mean	Std. deviation	Mean	Std. deviation				
Attitude	46	3.56	0.772	3.83	0.504	3.110	45	0.003**	0.32
Value	46	3.82	0.088	3.95	0.725	1.556	45	0.123	--
Enjoyment	46	3.67	0.105	3.80	0.090	1.243	45	0.220	--
Motivation	46	3.62	0.103	3.69	0.091	0.928	45	0.358	--
Self Confidence	46	3.56	0.101	3.84	0.087	3.138	45	0.003**	0.33

** *significant at p=0.01*. Reproduced with permission from Irvine (2020a).

Similar to engagement, the students who had the lowest attitude pre scores showed the largest increases in attitude post scores (Figure 2). In Figure 2 each data point gives the change (Attitude post minus Attitude pre) for a particular student in the treatment group. However, the average increases for the attitude scales (Attitude post minus Attitude pre) were smaller (M=0.15, SD=0.617, min=-1.66, max=1.94) than the magnitude of the average increases for the engagement scales (Engagement post minus Engagement pre), M=0.44, SD=0.817, min=-1.46, max=3.48. This difference was statistically significant (M=0.28, SD=0.932, $t_{(68)}$ =2.307, p=0.015).

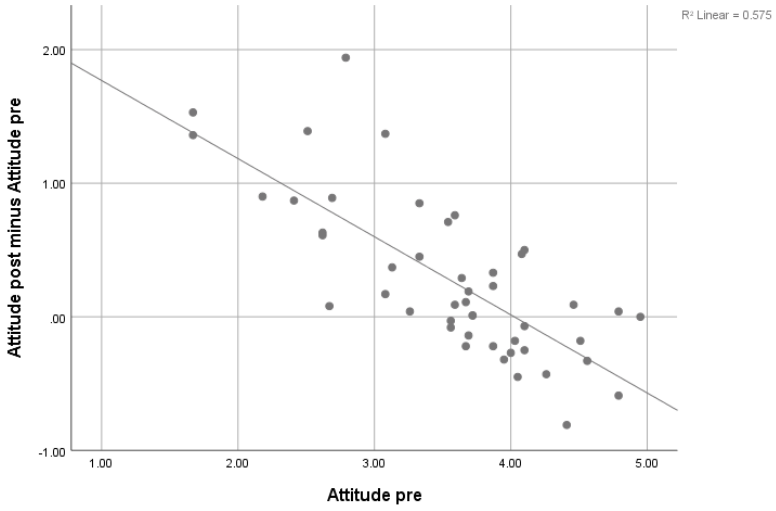

Figure 2. *Magnitudes of changes (Attitude post minus Attitude pre) for each treatment student in T_{Total}. Trend line shown. Reproduced with permission from Irvine (2020a).*

The control class showed a significant negative change in self-confidence (M=3.38, SD=0.660, $t_{(21)}$=-2.608, p=0.016). This somewhat surprising result, which was identified through quantitative analysis,

may reflect that the traditional teaching methods employed with the control class were less effective in sustaining student self-confidence.

Achievement. The student achievement data available for examination included: last year's grade, self-reported on a five-point Likert scale as previously discussed; final term mark, as a percent; the quadratic unit summative mark (denoted Quad tests); the rich assessment task (RAT); and the term mark excluding the quadratics unit (TermXquad). The final term marks contained the quadratic summative mark plus the summative marks for three other units. These data would have had autocorrelation issues if compared to quadratic unit summative marks and therefore were discarded. Since the data, including engagement and attitude scores, were a mixture of parametric and non-parametric data, Spearman's ρ (r_s) was used for correlational analysis (Naiman et al., 2000).

The significant correlations are shown in Table 6. *Grade last year* was positively correlated with engagement scores prior to the intervention (r_s =0.330, p=0.027). This result echoes prior research on the reciprocal relationship between motivation and achievement (e.g., Koller, Baumert, & Schnabel, 2001). A surprising result was finding no significant correlation between the quadratic summative mark and the rich assessment task, given that both assessments are of content from the same unit. At times there was a high level of absenteeism in these classes, with 13 students out of 43 total treatment cases (30%) being absent for the rich assessment task. This may have affected the correlation analysis.

Pre–post comparisons. To compare quadratic summative marks with student self-reported grades for last year, the quadratic mark was converted to a Likert scale using the same scale as the self-reported grade. No significant relationship was found for students in the control class (z=-0.502, p=0.616, ns). However, a strong negative relationship was found for students in the treatment classes (z=-2.824, p=0.005). This implies that students' marks on the unit of study are significantly lower than marks reported for their prior year math course.

Table 7.

Correlations Among Achievement-Related Variables for Treatment Students

	RAT	Quad tests	Term Xquad	Mark last year	Attitude Post	Attitude pre	Engagement post	Engagement pre
RAT								
Quad tests	0.287							
TermXquad	0.379*	0.894**						
Mark last year	-0.132	-0.037	-0.011					
Attitude post	-0.098	-0.225	-0.153	0.141				
Attitude pre	-0.309	-0.134	-0.137	0.120	0.614**			
Engagement post	0.344	0.214	0.211	-0.095	0.048	0.050		
Engagement pre	0.301	0.424**	0.395**	0.330*	0.185	0.150	0.349*	

**Correlation significant at the 0.05 level; **Correlation significant at the 0.01 level*

Treatment-control comparisons. No significant differences were found between T_{Total} and the control class for either grade last year or final term marks excluding the quadratic unit (TermXquad) for this year. However, a difference in grades on the quadratic unit tests was identified, with the control class (M=79.83, SD=11.664) higher than the total in the treatment classes (M=65.58, SD=20.945), significant at p=0.01 ($t_{(67)=}$3.024, p=0.004). A similar result was found for the rich assessment task (RAT), ($t_{(51)}$=3.216, p=0.002). So, for both the quadratic unit test and the RAT, the control class scored significantly higher than T_{Total}.

Qualitative Findings

Qualitative data was used to enhance the quantitative findings. This served to put a more human face on the data as well as to allow the researcher to probe concerning the underlying reasons for the increases

seen in the quantitative findings. In addition to teacher pre- and post interviews, a voluntary sample of five students from the treatment classes were interviewed after the classroom intervention. All were assigned pseudonyms for purposes of analysis. Table 7 shows some data on the five students, which were used as background information to inform the semi-structured student interviews.

Table 8.

Convenience Sample of Student Volunteers Who Were Interviewed

Pseudonym	Sex	Grade last year	Quad mark %	RAT %	Term XQuad %	Attitude		Engagement	
						Pre	Post	Pre	Post
Wendy	Female	80%+	75	68	87	2.67	3.33	4.10	4.60
Lina	Female	80%+	44	78	62	3.44	3.64	2.18	3.08
Dani	Female	80%+	88	68	90	3.13	3.95	2.62	3.23
Rob	Male	70%–79%	56	68	62	3.31	4.33	2.79	4.73
Shelly	Female	70%–79%	63	78	85	3.00	3.15	3.87	4.10

In addition, teacher information and pseudonyms are shown in Table 8. This information, which was obtained from the pre-intervention interviews and classroom observations, was used to inform the post-intervention semi-structured interviews with the teachers.

Table 9.

Teacher Characteristics

Category	Sex	Pseudonym	Qualifications	Years of experience	Years in current school	Dominant teaching styles*
Treatment	Female	Ms. Beckham	Honours Specialist in Mathematics	22	<1	Command; Practice; Guided discovery
Control	Female	Ms. Alford	Honours Specialist in Mathematics	10	<1	Command; Practice; Guided discovery

Note. *Based on teaching styles classifications in Fernandez-Rivas & Espada-Mateos (2019).

Engagement. Grounded theory analysis of student interview comments identified several themes: increased fun during math classes; more social classes due to use of student groups; more hands-on activities involving manipulatives and technology, which increased engagement; a sense of classrooms being more energetic. These themes are explained below and supported with student comments.

Fun was a theme commonly cited in the interviews. Classes during the intervention were considered significantly more fun than the students' usual math classes.

Increased use of student groups was identified by all five interviewees as promoting the social and collaborative culture in class. Again, this was contrasted with the usual math classes, which tended to be teacher-directed and utilized students sitting in rows and working individually. Specific references were made to the increased social atmosphere in the class due to numerous group activities. When asked to identify one activity that was very engaging, Wendy stated

> When we did a bunch of activities in little groups and we passed them around the class. Because we got to work on the questions with other people. So, you got to put in your input but also have them. So, if you

> didn't know something, they might know it and if you complete the question the class you felt you were really pleased and tired. (Post-intervention interview, May 7, 2018)

The use of hands-on activities, manipulatives, and technology was a third theme identified by the students. While the school was in a bring your own device (BYOD) school district and technology was routinely employed, activities using manipulatives such as algebra tiles and other hands-on materials was seen as a positive change by the student interviewees. Three students identified the active learning stance of the unit as a better fit for their personal learning modality compared to the normal classroom routine which often employed a document camera and teacher-lead lessons; that their level of engagement was content and topic specific, in that they felt more engaged during some activities than others; and that there was an increased energy in the classroom compared to the normal classroom routine. Shelly commented on how this differed from the usual atmosphere in math class.

Another theme was energy, since the interviewees felt that the classroom exhibited more energy and activity levels than normal. This energy was attributed to increased social activities and variety of learning opportunities. This trend was identified by Rob, who commented on an increased energy in class when he referred to his surprise that other students became more engaged:

> I participate in class almost all the time. But what I've seen is more kids getting engaged. Kids nround me that I don't expect to get engaged getting engaged. So that was interesting (Post-intervention interview, May 8, 2018).

This student's behaviour appeared to be contagious in that three students described other students in the class appearing to be more engaged than usual and that this behaviour seemed to influence additional students to become more engaged as well.

Student comments identified the increased variety of activities, and classroom organization changes as an additional theme that influenced engagement. This contrasted with the usual classroom routines which involved teacher-led question and answer for most of the classes.

Finally, students identified positive changes in work habits based on interesting and relevant relationships between the mathematical content and the real world. Three of the students interviewed indicated that prior to the intervention, their approach to classwork and homework was somewhat sporadic since they found the mathematical content to be uninteresting and sometimes boring.

Both the students and the teachers emphasized that engagement was task and context specific:

> I liked the activities where you had to work with your peers. Obviously, it's a very different model than you usually do, come with the whole equation together so that was obviously [sic] you're forced to work with everybody else. So, think about it. I like working with others. It's interesting to see what they say. So, it kind of breaks [sic] my thoughts. Usually math class is not like social. (Dani, post-intervention interview, May 8, 2018)

The quantitative and qualitative results are consistent with comments made by Ms. Beckham in her post-intervention interview: she indicated that she observed a noticeable difference in student engagement during the classroom intervention. She attributed this increased engagement to the active nature of the lessons; increased social interactions due to more group work; the use of manipulatives; and technology. She also noted that when she reverted to more traditional, teacher-lead lessons, the levels of engagement decreased.

Ms. Beckham's comments were consistent with comments of the five students who were interviewed. All students interviewed specifically mentioned the rich assessment task as very engaging, since it was active, hands-on, employed manipulatives and technology, and student groups for the data collection stage. Another strategy specifically mentioned by the interviewees was the teacher's use of individual student whiteboards. Comments on this instructional strategy identified uniqueness, immediate feedback,

and an active learning stance in the classroom. Ms. Beckham reinforced the context and task-specific stance:

> So, it's hard but I just found by doing different things you're hitting the different kids all is okay. This kid may not be engaged in this but now they're engaged in that area so that was pleasing [sic]. But I think the couple of conversation ones that we did worked really well, and there were some other ones. (Post-intervention interview, May 9, 2018)

The increased engagement in the treatment classes contrasted with the comments on Ms. Alford's control class:

> My students were sometimes engaged in the work, but I didn't see any dramatic differences compared to the earlier unit. Students talk to each other, and my students asked why we weren't doing the neat things that Ms. Beckham's classes were doing. So, I sometimes integrated some of the activities from the other classes into my class, and that seemed to get my students more engaged. (Post-intervention interview, May 9, 2018)

Based on teacher interviews and classroom observations, I noted that the dominant teaching styles of both Ms. Beckham and Ms. Alford were teacher-directed, command and practice, which would be classified as traditional (Fernandez-Rivas & Espada-Mateos, 2019). However, both teachers also indicated that they sometimes engaged in guided discovery. Ms. Beckham stated "I do a lot of investigations" (Pre-intervention interview, February 21, 2018). During researcher-observed classes that were rated as *low* on fidelity of implementation, Ms. Beckham utilized command and practice styles exclusively. This may have influenced the overall findings of this study, since these more traditional teaching styles are not compatible with the reform-based, active lesson structure of the intervention.

Attitude. Students often mentioned that classes in this unit were more "fun," which may reflect a conflation of engagement and attitudes. The student responses reflected that attitude is a complex construct; interactions of attitude with other factors—such as prior and current achievement, peer and family influences, future plans, and other activities such as sports or the arts—all influenced the students' attitudes toward mathematics. The following attitude-related themes were identified in student interviews: self-confidence, teacher style, interest, persistence, and motivation to continue taking mathematics. These themes are discussed below, supported by student comments.

Increased self-confidence was cited by four students, based on changes in classroom organization from teacher-centred to the use of small groups and activities that addressed different learning modalities. Students shared that they felt confident to contribute in group settings, whereas they might not do the same in whole-class discussions.

All five students identified the variety of activities and the linking of mathematical content to real-world situations as resulting in more interesting classes. The students indicated that the increased number of activities was different than the norm and that the uniqueness of this unit positively impacted their attitudes and resulted in classes being more fun.

Another theme was persistence. Students stated that they tended to persist in solving problems related to the classroom intervention, and that they worked to achieve understanding of the mathematical content.

There was a very positive response to asking students' opinions. Four of the five students indicated that they had never been asked their opinions before, especially in mathematics classes. In particular, being asked about their feelings around engagement and attitudes were new and appreciated.

> I liked a lot of your activities. I also liked the fact that you're interested in what the students think. I feel like sometimes our opinion gets overlooked and we kind of just get passed along to Grade 12. (Lina, post-intervention interview, May 8, 2018)

The students also liked specific personal goal-setting activities during the unit. Again, no one had ever asked them to set personal goals before this unit.

Other themes identified in student interviews included student comfort levels (related to teacher or classroom culture) and frustration based on inadequate understanding of material. Perceived teacher care was also identified.

> [Ms. Beckham] is a great teacher. You know that she cares about you. I feel like she teaches just the right way. A high school teacher like she doesn't hold your hand. But she doesn't push you off a cliff, you know. (Shelly, post-intervention interview, May 7, 2018)

An additional theme was related to teacher style, with the teacher becoming more participatory with students in the learning process. Increased interest was a major theme identified. Students also stated that their more positive attitudes towards mathematics had resulted in increased interest and motivation to take additional mathematics courses beyond what was required, both in the senior grades of high school and potentially at university.

Ms. Beckham indicated that she felt that the attitudes toward mathematics of some students in her classes had increased, although she provided only anecdotal evidence. All students who were interviewed indicated that the increased social culture of the classroom positively affected their attitude toward mathematics, even though two of the interviewees came into the unit with decidedly negative attitudes towards the subject.

Achievement. Themes identified in student interviews related to the distinction between grades and understanding; links between work habits and achievement; and no perceived relationship between increased engagement and increased achievement. These themes are discussed below, with student comments.

Student interview comments indicated that achievement was divided sharply between understanding material and demonstrating competence as indicated by student grades. Several students felt that their personal level of understanding was not adequately reflected in their marks on tests and evaluations:

> I did decently well but I felt like I know more than what I get on my tests and quiz marks. I feel kinda like when I see my grades especially when I work really hard. I see my grades not accurately representing what I know and it really hurts mentally, my confidence. (Rob, post-intervention interview, May 8, 2018)

All students indicated they felt that a strong link existed between achievement and work habits and homework completion:

> I struggle but obviously I know if I do all my homework every night or my questions are complete it does make me remember a lot better so I don't struggle, but it depends on the unit. (Dani, post-intervention interview, May 7, 2018)

Interestingly, while all those interviewed indicated increases in engagement, none of the students related this increased engagement to their personal achievement in mathematics. The students indicated that there was no perceived relationship between the students' increased engagement and attitude scores and their achievement on the unit as represented by the quadratics summative assessment marks reported by the students.

One aspect not identified in student interviews was the relative difficulty of the mathematics content in this unit. Both teachers involved in this study felt that the quadratics unit was the most difficult unit in the course, and noticeably more difficult than the linear systems unit that had preceded the classroom intervention. This was echoed by student interview comments. All students agreed that the current quadratics unit was more difficult than Grade 9 mathematics and also more difficult than the previous unit in this course, linear systems. However, Ms. Alford (control class) stated that her students were

generally strong academically and she did not perceive that they struggled significantly with the content of the unit:

> My students right now are awesome. They are very, very good for the most part. They're determined and they're very organized and they are very driven. I don't really have to keep on top of them, so I don't have to micromanage them for the most part. (Post-intervention interview May 9, 2018)

Discussion

The discussion below focuses on the value of utilizing a mixed methods methodology.

For all three variables of interest, engagement, attitude, and achievement, the quantitative and qualitative data were congruent and reflected the value of a mixed methods methodology. In addition, utilizing both methodologies identified several questions that would not have been exposed without both dimensions of the mixed methods methodology. One of the key benefits of utilizing a mixed methods methodology is that each of quantitative and qualitative methods informs the other. For example, in this study, quantitative information about the students who were interviewed helped shape the semi-structured interview questions for the qualitative portion of the study. The information may also deepen understanding of the phenomena under investigation. In this study, the qualitative data helped increase understanding of the quantitative results. Identification of themes such as fun, persistence, activity, energy, social interactions helped explain the increases in engagement and attitudes found in the quantitative analysis. One methodology may also identify anomalies that might otherwise have been missed. An example of this was the differential impact of the intervention, which was identified through quantitative methods, that would have been missed if only qualitative methods had been employed. Another example is the novelty effect. Conversely, qualitative data concerning the students' distinctions around understanding versus achievement may have pointed out possible issues with the achievement measures, which provided quantitative results but may not have accurately reflected student achievement in the broader sense. This may also have pointed to the need for a more refined pre-measure of student achievement should this study be replicated by other researchers.

Qualitative methods also provided information on the reasons why the classroom intervention in this study was successful. While success was measured quantitatively, qualitative data concerning the benefits of utilizing lessons and activities that were hands-on, active , social and that provided student choice and increased student agency increased understanding of the reasons why the intervention was successful. Some of these issues are discussed in more detail below.

Questions Arising From a Mixed Methods Methodology and Potential Answers

A number of questions arose when the study results were analysed. These included: a decoupling of increased student engagement but decreased student achievement; the relationship of the instructional strategies employed by the classroom intervention and reform mathematics principles; a differential impact of the intervention with respect to pre-intervention student survey results; and the novelty effect of students being asked their opinions.

Increased Engagement but Decreased Achievement. The hypothesis that students in the treatment classes would demonstrate significant increases in student achievement compared to students in the control class was not supported. This decoupling of achievement and engagement was supported by both the quantitative and qualitative data analysis. Student achievement in this unit of study, measured by grades, was lower in the treatment classes than in the control class. This led to a question concerning the link between engagement and achievement. While medium positive effect sizes of 0.54 were found for engagement, and 0.32 for attitude, no positive effect size was found for achievement. This contradicts a number of research studies claiming a positive association between these two constructs (e.g.,

Bodovski & Farkas, 2007), although the claim of association between engagement and achievement is not unanimous in the literature (Dotterer & Lowe, 2011; Marks, 2000).

There are a number of possible explanations for the lack of a positive effect size for achievement. One possibility is that the pre-intervention measure of achievement may have been flawed. Students were asked to record their marks from their last mathematics course taken, as a letter grade. This may have resulted in a "rose-coloured glasses" effect, with students overstating their previous marks. This was confirmed by Ms. Beckham for at least one of the student interviewees. A second possibility is that the treatment classes were less academically prepared than the control class. This possibility has some statistical support, as documented earlier. A third possibility is that the mathematical content of the unit under study (quadratics) is more difficult than other units in the course. This position also has some anecdotal support from student and teacher interviews.

Marks (2000) found that the linkage between engagement and achievement was weak, but that engagement stands as a goal of education, separate from any potential relationship to student achievement. This agrees with Collie and Martin (2017) who cite engagement as a goal of education irrespective of its possible impact on other educational variables.

Attribution of achievement to engagement/effort did not occur for the students interviewed in this study, except for homework completion. These students focused more on the perceived dichotomy between achievement as measured by more traditional instruments such as written tests, and achievement as measured by understanding of concepts. None of the students interviewed felt that their achievement as measured by grades accurately reflected their understanding. This decoupling of engagement and achievement has important implications for life-long learning, as an espoused goal of education (Fredricks et al., 2004).

Reform Mathematics and the Classroom Intervention. The classroom intervention employed in this study resulted in significant effect sizes related to student engagement and student attitudes towards mathematics. The classroom intervention utilized strategies that are active, involve students working in groups, employ problem-based learning and ask students to explain their thinking (collectively referred to as reform mathematics curricula). Instructional strategies have a major impact on student engagement (Smith & Star 2007).Reform mathematics curricula were found by Moyer, Robison and Cai (2018) to have long-lasting effects on student engagement. An additional instructional strategy that encourages engagement is student choice (Irvine, 2018). Self-determination theory identifies three basic needs, autonomy, competence, and relatedness together with needs for competence and relatedness (Deci& Ryan 2008). Student choice addresses students' need for autonomy, providing some level of student control over their learning. By providing students with some level of choice, such as a choice of solution methods, choice of problems to be investigated or choice of product to demonstrate their learning – teachers support student autonomy, which has been found to foster increased engagement (Deci, Vallerand, et al.,1991). Teacher-directed learning in mathematics classes, without student choice, has been found to have a significant negative effect on engagement and an increase in students' use of avoidance strategies (Turner et al.,2002). Both quantitative and qualitative data supported this active stance for instruction. These relationships to self-determination theory would not have surfaced without qualitative methodologies being part of this study.

Differential Impact of the Classroom Intervention. Both student engagement and student attitudes showed a clear trend, whereby students who initially reported low engagement or low attitude scores showed markedly greater improvements in the post-intervention responses (see Figures 1 & 2).

The pattern observed for the magnitudes of attitude changes was similar to the pattern for engagement (Figure 1); that is, students who had the lowest attitude pre scores showed the largest increases in attitude post scores (see Figure 2). In Figure 2 each data point gives the change (Attitude pre minus Attitude post) for a particular student in the treatment group. This observation demonstrates the value of mixed

methods methodology. If only qualitative methodologies had been used, this differential impact of the classroom intervention would have been missed.

The intervention was most effective for students who initially reported low levels of engagement and for students who reported more negative attitudes towards mathematics. This effect may be explained by the more active, reform mathematics instructional strategies employed by the classroom intervention. These instructional strategies involve real-world connections, hands-on activities, students working in groups and the use of manipulatives to increase student involvement. Such strategies were found to increase and sustain student engagement (Moyer, Robison, & Cai, 2018; Smith & Star 2007). In the classroom intervention, these strategies had the greatest impact on the most disengaged students and students with negative attitudes. It may be that these students' learning styles were a much better fit for the active instructional strategies than the dominant instructional strategies employed by Ms. Alford and Ms. Beckham, who employed teacher-centred strategies to a large extent. IT may also be argued that increases in student attitudes and engagement resulted in more active involvement of students in their own learning. This result provides supporting evidence of the efficacy of reform mathematics strategies and emphasises that such strategies may have the greatest impact on students who most need to become more reengaged, as well as students with more negative attitudes towards mathematics.

Novelty Effect. The classroom intervention made extensive use of brief student surveys. Students were frequently asked: Was today's lesson interesting?; was today's lesson useful to you?; was today's lesson fun?.. Four of the five students indicated that they enjoyed being asked about their levels of engagement and attitudes, and that no one had ever asked them about these attributes before (see Lina's comment, noted previously). Here again, we see evidence of the usefulness of mixed methods methodologies. If only quantitative methodologies had been employed, this dimension would have been missed. One conclusion that could be drawn is that it is beneficial to students to have teachers solicit their opinions concerning what activities occur in class. This use of student surveys has potential to be a viable topic for future research (see Irvine, J. (2020b).

Conclusion

The appropriateness of employing a mixed methods methodology, especially when examining latent variables that are not readily observable, was strongly reinforced. Mixed methods methodology identifies a methodology after research questions have been formulated, rather than specifying a methodology a priori (Johnson and Onwuegbuzie 2004).

It is informative to consider what information would have been missed if only one methodology had been employed in this study. For instance, if only quantitative methodology had been used, since statistically significant effect sizes were found, the results would be generalizable, but no data on the reasons why the instructional intervention was effective could be discussed. Thus, the "what happened" could be found, but without any information on the "why it happened."

Adding the qualitative dimension to the methodology allowed for a deeper understanding of the phenomena; this allowed for more probing of the complexities of the attitude construct as well as consideration of dimensions such as the novelty effect of students being asked their opinions and feelings. This dimension is apparently rarely considered in mathematics classes, as no mention was found in the literature.

However, if only a qualitative methodology was employed, while the in-depth phenomena of student behaviour would have been accessible, it would not be possible to compute aggregate outcomes such as effect sizes, and the issue concerning the linkage between engagement and achievement may have been minimised or even missed altogether. This has important implications for policy development across broad jurisdictions such as Ontario and Canada. Education policy is influenced by aggregate measures that point the way forward to positively impact student learning for large numbers of students.

Substantial changes in education policy can only be achieved through providing a body of evidence of the efficacy of the changes.

Recommendations

By utilizing a mixed methods methodology that includes both quantitative and qualitative methodologies, researchers are able to reinforce findings from each methodology; surface issues that would not be revealed if only one methodology was used; and provide directions for future research. In this study future research directions included investigation of the novelty effect on student success; refinement of the concept of achievement to reflect not only student grades but also student understanding; longitudinal studies of imp0lementing reform mathematics principles and their impact on student affective variables; and differential effects of reform mathematics principles that revealed greater impact on students with lower pre-study scores.

Therefore, educational researchers should consider a pragmatic approach to selecting methodologies. By first identifying the research questions for a study, and allowing the research questions to drive the choice of methodology, educational researchers will potentially find richer and more multi-dimensional answers to research questions than are accessible if the methodology drives the formulation of the research questions. This does not mean that all studies should employ mixed methods methodology; but once the research questions are specified, the appropriate methodology should be selected in order to provide fulsome answers to the research questions of the study.

References

Argon, T., & Kaya, A. (2018). Examination of grit levels of teachers according to personal variables. *Journal of Education and Training Studies,6*(3a), 45–53. https://doi.org/10.11114/jets.v6i3a.3157

Asante, K. O. (2012). Secondary students' attitudes towards mathematics. *IFE PsychologIA*, *20*(1), 121–133.https://www.ajol.info/index.php/ifep/article/view/74705

Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: Beginning knowledge, student engagement, and instruction. *The Elementary School Journal*, *108*(2), 115–130. https://doi.org/10.1086/525550

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Sage.

Deci, E., & Ryan, R. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology*, *49*(3), 182–185. https://doi.org/10.1037/a0012801

Deci, E., Vallerand, R., Pelletier, L., & Ryan, R. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist*, *26*(3–4), 325–346.https://doi.org/10.1080/00461520.1991.9653137

Derrick, B, Ruck, A., Toher, D., & White, P. (2018). Tests for equality of variances between two samples which contain both paired observations and independent observations. *Journal of Applied Quantitative Methods*, *13*(2), 36–47.http://www.jaqm.ro/issues/volume-13,issue-2/3_BEANDEPA.PHP

Dotterer, A., & Lowe, K. (2011). Classroom context, school engagement, and academic achievement in early adolescence. *Journal of Youth and Adolescence*, *40*, 1649–1660. https://doi.org/10.1007/s10964-011-9647-5

Erceg-Hurn, D. M., & Mirosevich, V. M. (2008). Modern robust statistical methods: An easy way to maximize the accuracy and power of your research. *The American Psychologist*, *63*(7), 591–601. https://doi.org/10.1037/0003-066X.63.7.591

Fernandez-Rivas, M., & Espada-Mateos, M. (2019). The knowledge, continuing education and use of teaching styles in Physical Education teachers. *Journal of Human Sport and Exercise*, *14*(1), 99–111. https://doi.org/10.14198/jhse.2019.141.08

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Fredricks, J., Blumenfeld, P., & Paris, A. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109.https://doi.org/10.3102%2F00346543074001059

Irvine, J. (2018). A framework for comparing theories related to motivation in education. *Research in Higher Education Journal*, *35*,Art. 5, 1–30.https://eric.ed.gov/?id=EJ1194268

Irvine, J. (2020a). Positively influencing student engagement and attitude in mathematics through an instructional intervention using reform mathematics principles. *Journal of Education and Learning*, *9*(2), 48–75. https://doi.org/10.5539/jel.v9n2p48

Irvine, J. (2020b). Using surveys to assess student attitudes towards mathematics. *Ontario Association for Mathematics Education Gazette*, *58*(4), *27-30*

.Jacob, S. A., & Furgerson, S. P. (2012). Writing interview protocols and conducting interviews: Tips for students new to the field of qualitative research. *The Qualitative Report*, *17*(42), 1–10. https://nsuworks.nova.edu/tqr/vol17/iss42/3

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, *38*(12), 1217–1218. https://doi.org/10.1111/j.1365-2929.2004.02012.x

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14–26. https://doi.org/10.3102%2F0013189X033007014

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, *1*(2), 112–133. https://doi.org/10.1177%2F1558689806298224

Koller, O., Baumert, J., & Schnabel, K. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. *Journal for Research in Mathematics Education*, *32*(5), 448–470. https://doi.org/10.2307/749801

Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd ed.). Sage.

Larson-Hall, J. (2012). Our statistical intuitions may be misleading us: Why we need robust statistics. *Language Teaching*, *45*(4), 460–474. https://doi.org/10.1017/S0261444811000127

Majeed, A., Darmawan, G., & Lynch, P. (2013). A confirmatory factor analysis of attitude toward mathematics inventory (ATMI). *The Mathematics Educator*, *15*(1), 121–135. http://math.nie.edu.sg/ame/matheduc/journal/v15_1/v15_7.aspx

Marks, H. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, *37*(1), 153–184. https://doi.org/10.3102/00028312037001153

Mircioiu, C., & Atkinson, J. (2017). A comparison of parametric and non- parametric methods applied to a Likert scale. *Pharmacy: Journal of Pharmacy, Education and Practice*, *5*(2).https://doi.org/10.3390/pharmacy5020026

Moyer, J., Robison, V. & Cai, J. (2018). Attitudes of high-school students taught using traditional and reform mathematics curricula in middle school: A retrospective analysis. *Educational Studies in Mathematics*, *98*(1), 115–134. https://doi.org/10.1007/s10649-018-9809-4

Naiman, A., Rosenfeld, R., & Zirkel, G. (2000). *Understanding statistics* (2nd rev. ed.). McGraw-Hill.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education: Theory and Practice*, *15*(5), 625–632.https://doi.org/10.1007/s10459-010-9222-y

Ozmon, H.A., & Craver, S.M. (2003). *Philosophical foundations of education* (7th ed.). Merrill Prentice Hall.

Peró-Cebollero, M., & Guàrdia-Olmos, J. (2013). The adequacy of different robust statistical tests in comparing two independent groups. *Psycológica*, *34*(2), 407–424.https://files.eric.ed.gov/fulltext/EJ1018822.pdf

Poncet, A., Courvoisier, D. S., Combescure, C., & Perneger, T. V. (2016). Normality and sample size do not matter for the selection of an appropriate statistical test for two-group comparisons. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *12*(2), 61–71. https://doi.org/10.1027/1614-2241/a000110

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology*, *105*(3), 579–595.https://doi.org/10.1037/a0032690

Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, *98*(1), 209–218. https://doi.org/10.1037/0022-0663.98.1.209

Reeve, J., & Lee, W. (2013). Students' classroom engagement produces longitudinal changes in classroom motivation. *Journal of Educational Psychology*, *106*(2), 527–540. https://doi.org/10.1037/a0034934

Reeve, J., Lee, W., & Won, S. (2015). Interest as emotion, as affect, and as schema. In K. A. Renninger, M. Nieswandt, & S. Hidi (Eds.), *Interest in mathematics and science learning* (pp. 79–93). AERA.

Saldana, J. (2014). *The coding manual for qualitative researchers*. Sage.

Smith, J., III, & Star, J. (2007). Expanding the notion of impact of K-12 standards-based mathematics and reform calculus programs. *Journal for Research in Mathematics Education*, *38*(1), 3–34. https://doi.org/10.2307/30034926

Tapia, M., & Marsh, G. (2005). Attitudes towards mathematics inventory redux. *Academic Exchange*, *9*(3), 272–276.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research*. Sage.

Turner, J., Midgley, C., Meyer, D., Gheen, M., Anderman, E., Kang, Y., & Patrick, H. (2002). The classroom environment and students' reports of avoidance strategies in mathematics: A multimethod study. *Journal of Educational Psychology*, *94*(1), 88–106. https://doi.org/10.1037/0022-0663.94.1.88

Vandecandelaere, M., Speybroeck, S., Vanlaar, G., Fraine, B., & Van Damme, J. (2012). Learning environment and students' mathematics attitude. *Studies in Educational Evaluation*, *38*(3–4), 107-120. https://doi.org/10.1016/j.stueduc.2012.09.001

Whiting, L. S. (2008). Semi-structured interviews: Guidance for novice researchers. *Nursing Standard*, *22*(23), 35–40. https://doi.org/10.7748/ns2008.02.22.23.35.c6420

Zumbo, B. D., & Zimmerman, D. W. (1993). Is the selection of statistical methods governed by level of measurement? *Canadian Psychology/Psychologie canadienne*, *34*(4), 390–400. https://doi.org/10.1037/h0078865