UNIVERSITÉ DE SHERBROOKE

Faculté d'éducation

Apprentissage résultant de la modélisation vs. l'instruction régulière dans un cours de physique introductoire au cégep

Learning Outcomes of Modelling vs. Regular Instruction in a CEGEP Introductory Physics Course

par

Stéphan Bourget

Essai présenté à la Faculté d'éducation
en vue de l'obtention du grade de
Maître en éducation (M. Éd.)
Maîtrise en enseignement au collégial

Mai 2020

UNIVERSITÉ DE SHERBROOKE

Faculté d'éducation

Apprentissage résultant de la modélisation vs. l'instruction régulière dans un cours de physique introductoire au cégep

Learning Outcomes of Modelling vs. Regular Instruction in a CEGEP Introductory Physics Course

par

Stéphan Bourget

Essai présenté à la Faculté d'éducation
en vue de l'obtention du grade de
Maître en éducation (M. Éd.)
Maîtrise en enseignement au collégial

Mai 2020

UNIVERSITÉ DE SHERBROOKE

Faculté d'éducation

Apprentissage résultant de la modélisation vs. l'instruction régulière dans un cours de physique introductoire au cégep

Learning Outcomes of Modelling vs. Regular Instruction in a CEGEP Introductory Physics Course

par

Stéphan Bourget

a été évalué par un jury composé des personnes suivantes :

Dr. Stephen G. Taylor, PhD.   Directeur de la recherche
Prénom et nom de famille

Dr. Amir Shoham, PhD.   Membre externe du jury
Prénom et nom de famille

Essai accepté le ________________ 2020

ABSTRACT

Students have many unviable preconceptions in physics. Modelling instruction (MI), which gets students actively modelling physical phenomena, is reported to improve learning and attitudes. This research thus sought to assess the impact of this approach in comparison to regular instruction (RI) and interactive engagement (IE) when applied to an introductory physics course in a Quebec CEGEP. The hypotheses were that modelling students would perform differently and would overall prefer this form of active learning. A mixed-methods design with a strong quantitative strand based on a quasi-experiment was used. Learning outcomes were defined as conceptual understanding as measured by the FCI and RRMCS, and problem-solving skills as measured on the final exam. Based on a qualitative survey, novice MI seemingly produced more dissatisfaction than satisfaction, which is consistent with what is reported for other methods fostering IE. Overall, this research seems to indicate that novice MI compares to both IE and RI as no statistically meaningful difference was found on learning outcomes, with the single exception of FCI post-test scores, for which both MI and IE differed from RI. Yet, there appears to be potential that formal training and further experience with the method would help tap into.

Keywords: CEGEP, modelling instruction, model-based physics, interactive engagement, physics education, introductory physics courses, conceptual learning, Force Concept Inventory, FCI, student beliefs, attitudinal survey, attitudes, Colorado Learning Attitudes about Science Survey, CLASS, Rotational and Rolling Motion Conceptual Survey, RRMCS, pedagogy.

SUMMARY

The purpose of this study was to compare modelling instruction with regular instruction and interactive engagement in terms of students' learning outcomes and attitudes toward physics in an introductory physics course (Mechanics) in an anglophone CEGEP in Montreal, Quebec, Canada. This topic is important because students have many conflicting preconceptions challenging their mastery of physics concepts. They also have an erroneous perception of the scientific process and the nature of physics.

Modelling instruction gets students continuously working collaboratively with teammates and the whole class, actively modelling a physical phenomenon and using whiteboards to exteriorize and communicate their thoughts. The literature shows that modelling instruction improves learning gains and attitudes about physics. Therefore, this study proposed to answer three research questions. First, how does modelling instruction differ from regular instruction or interactive engagement in terms of learning outcomes for CEGEP Mechanics students? Second, how does modelling instruction differ from regular instruction or interactive engagement in terms of attitudes (or beliefs) about physics for those same students? Third, how are CEGEP Mechanics students perceiving (in terms of what they like or don't like) the introduction of modelling instruction? The hypotheses were that students receiving modelling instruction would perform differently on assessments of deep conceptual understanding, on an exam testing problem-solving skills, and on an assessment of expert-like attitudes. It was also thought that they would overall prefer this form of active learning.

A mixed-methods design with a strong quantitative strand based on a quasi-experiment was used. Learning outcomes were defined as students' conceptual understanding as measured by pre/post-test normalized gains (FCI, RRMCS) and students' procedural mastery of the mathematics (tested in a traditional exam with textbook-like problems). Attitudes about physics were defined as students' self-reported beliefs on a pre/post CLASS test. Non-random, intact group samples of Mechanics college students were used. The treatment group was constituted of two of the researcher's sections while the control groups were constituted of ten other sections taught by colleagues. Quantitative data were analyzed through a multivariate analysis of variance (MANOVA) with p-values $\leq .05$ for statistical significance. Then comparisons were made with the control groups and the literature. Responses to a qualitative survey were analyzed to complement the quantitative analysis.

Although interactive engagement surpassed regular instruction on normalized conceptual gains (43% vs. 29%; $p = .002$ on Games-Howell simultaneous test for differences of means), novice modelling instruction couldn't be statistically distinguished from either interactive engagement (41% vs. 43%; $p = .918$) or regular instruction (41% vs. 29%; $p = .229$) for conceptual learning outcomes, except on a slightly different measure. It did statistically distinguish itself from regular instruction ($p = .019$ on Tukey pairwise comparison), on par with interactive engagement ($p = .006$ on a comparison with regular instruction), only on FCI post-test scores that were on average at 66% (mostly like interactive engagement at 64%), whereas they were at 54% for regular instruction. The score of 60% is considered the threshold over which a student barely begins to adopt Newtonian thinking. FCI normalized gains of modelling instruction and interactive engagement were consistent with what other studies have reported. Regular instruction performed better than what has been reported for traditional instruction (29% instead of 22%; $t = 3.00$; $p = .004$); it is believed the limited use of active learning incorporated into class activities might be the cause of it. No statistically meaningful difference was found on procedural learning outcomes ($p > .05$ on all comparisons).

Based on the attitudinal CLASS test results, novice modelling instruction couldn't be distinguished from either interactive engagement or regular instruction ($p > .05$), all methods failing equally at improving attitudes toward physics and its learning. Attitudes didn't appear to worsen either. This observation is surprising for modelling instruction because the literature reports it to be one of the rare cases where attitudes typically improve. This demands to be investigated further.

Based on a qualitative survey, the researcher's novice implementation of modelling instruction seemingly produced more dissatisfaction than satisfaction, which in some way is consistent with what is reported for other methods fostering interactive engagement, despite improved learning. The level of participation in that survey was very low (23%), though, so results might not be representative.

Overall, this research seems to indicate that novice modelling instruction didn't produce better or worse learning gains and attitude shifts than interactive engagement or regular instruction. Yet, there seems to be potential that formal training and further experience with the method would help tap into, thus justifying further attempts at using modelling instruction under a deliberate practice of continuous improvement.

RÉSUMÉ

Le but de cette étude était de comparer l'enseignement par modélisation avec l'enseignement régulier et l'engagement interactif en termes de résultats d'apprentissage et d'attitudes dans un cours d'introduction à la physique (mécanique) dans un cégep anglophone à Montréal, Québec, Canada. Ce sujet est important, car les étudiants ont de nombreuses idées fausses et contradictoires rendant difficile leur maîtrise des concepts. Ils ont également une perception erronée du processus scientifique et de la nature de la physique.

L'enseignement par modélisation permet aux étudiants de travailler continuellement en collaboration avec des coéquipiers et la classe entière, de modéliser activement un phénomène physique et d'utiliser des tableaux blancs pour extérioriser et communiquer leur pensée. La littérature montre que l'enseignement par modélisation améliore les gains d'apprentissage et les attitudes à l'égard de la physique. Par conséquent, cette étude a proposé de répondre à trois questions de recherche. Premièrement, en quoi l'enseignement par modélisation diffère-t-il de l'enseignement régulier ou de l'engagement interactif en termes de résultats d'apprentissage pour les collégiens en mécanique? Deuxièmement, en quoi l'enseignement par modélisation diffère-t-il de l'enseignement régulier ou de l'engagement interactif en termes d'attitudes (ou de croyances) à l'égard de la physique pour ces mêmes collégiens? Troisièmement, comment les collégiens en mécanique perçoivent-ils (en termes de ce qu'ils aiment ou n'aiment pas) l'introduction de l'enseignement par modélisation? Les hypothèses étaient que les étudiants recevant l’enseignement par modélisation performeraient différemment aux évaluations de compréhension conceptuelle, à un examen testant les compétences en résolution de problèmes et à une évaluation des attitudes en termes de ressemblance avec celles des experts. Il était aussi pensé qu'ils préféreraient globalement cette forme d'apprentissage.

Un design quasi expérimental fort, intégré à une approche mixte, a été utilisé pour répondre aux questions. Les résultats d'apprentissage furent définis comme la compréhension conceptuelle des étudiants mesurée par les gains normalisés avant et après intervention (FCI, RRMCS) et leur maîtrise procédurale des mathématiques (testée dans un examen traditionnel avec des problèmes typiques de manuel). Les attitudes à l'égard de la physique furent définies comme les croyances autodéclarées des étudiants (CLASS), avant et après intervention. Des échantillons intacts non aléatoires d'étudiants en mécanique ont été utilisés (groupes classes). Le groupe de traitement était constitué de deux sections enseignées par le chercheur tandis que les groupes témoins étaient constitués de dix autres sections enseignées par des collègues. Les données quantitatives ont été analysées par une analyse multivariée de la variance (MANOVA) avec des valeurs de $p \leq 0{,}05$ pour la signification statistique. Ensuite, des comparaisons ont été faites avec les groupes témoins et la littérature. Les réponses à une enquête qualitative ont été analysées pour compléter l'analyse quantitative.

Bien que l'engagement interactif ait dépassé l'enseignement régulier sur les gains conceptuels normalisés (43% contre 29%; $p = 0{,}002$ sur le test simultané de Games-Howell pour les différences de moyennes), l'enseignement novice par modélisation n'a pu être statistiquement distingué ni de l'engagement interactif (41% vs 43%; $p = 0{,}918$) ni de l'enseignement régulier (41% vs 29%; $p = 0{,}229$) pour les résultats d'apprentissage conceptuels, sauf sur une mesure légèrement différente. Il ne s'est en effet distingué statistiquement de l'enseignement régulier

($p = 0{,}019$ sur la comparaison par paires de Tukey), tout comme l'a fait l'engagement interactif ($p = 0{,}006$ sur une comparaison avec l'enseignement régulier), que sur les scores FCI après intervention : une moyenne de 66% (64% pour l'engagement interactif), alors qu'elle est de 54% pour l'enseignement régulier. Le score de 60% est considéré comme le seuil au-dessus duquel un étudiant commence à peine à adopter une pensée newtonienne. Les gains normalisés du FCI liés à l'enseignement par modélisation et à l'engagement interactif étaient conformes à ce que d'autres études ont rapporté. L'enseignement régulier a donné de meilleurs résultats que ce qui a été rapporté pour l'enseignement traditionnel (29% au lieu de 22%; $t = 3{,}00$; $p = 0{,}004$); on pense que l'utilisation limitée de l'apprentissage actif incorporé dans les activités en classe pourrait en être la cause. Aucune différence statistiquement significative n'a été trouvée sur les résultats d'apprentissage procédural ($p > 0{,}05$ dans toutes les comparaisons).

Sur la base des résultats du test attitudinal CLASS, l'instruction par modélisation novice n'a pas pu être distinguée de l'engagement interactif ou de l'enseignement régulier ($p > 0{,}05$), toutes les méthodes échouant également à améliorer les attitudes envers la physique et son apprentissage. Les attitudes ne semblent pas non plus s'aggraver. Cette observation est surprenante pour l'enseignement par modélisation, car la littérature rapporte qu'il s'agit de l'un des rares cas où les attitudes s'améliorent généralement. Cela nécessite une enquête plus approfondie.

Sur la base d'une enquête qualitative, la mise en œuvre novice de l'enseignement par modélisation a apparemment produit plus d'insatisfaction que de satisfaction, ce qui est en quelque sorte cohérent avec ce qui est rapporté pour d'autres méthodes favorisant l'apprentissage actif. Le niveau de participation à cette enquête était cependant très faible (23%), de sorte que les résultats pourraient ne pas être représentatifs.

En somme, cette recherche semble indiquer que l'enseignement novice par modélisation n'a pas produit de gains d'apprentissage ni de changements d'attitude meilleurs ou pires que l'engagement interactif ou l'enseignement régulier. Cependant, il semble y avoir du potentiel qu’une formation formelle et une expérience plus approfondie de la méthode permettrait d'exploiter, justifiant ainsi de nouvelles tentatives d'enseignement par modélisation dans le cadre d'une pratique délibérée d'amélioration continue.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

ADI	Argument-Driven Inquiry
AMTA	American Modeling Teachers Association
ANOVA	ANalysis Of VAriance
ASU	Arizona State University
CEGEP	Collège d'Enseignement Général Et Professionnel
CLASS	Colorado Learning Attitudes about Science Survey
DV	Dependent Variable
EBAPS	Epistemological Beliefs Assessment for Physical Science
FCI	Force Concept Inventory
FIU	Florida International University
IE	Interactive Engagement
IIR	Interdisciplinary Islands of Rationality
IV	Independent Variable
KR-20	Kuder-Richardson Formula 20
MANOVA	Multivariate ANalysis Of Variance
MAPS	Modelling Applied to Problem Solving
MB	Mechanics Baseline
MD	(Halloun – Hestenes) Mechanics Diagnostic
MI	Modelling Instruction
MPEX	Maryland Physics Expectations Survey

PBL	Problem-Based Learning or Project-Based Learning (depending on context)
PER	Physics Education Research
PET	Physics of Everyday Thinking
PI	Peer Instruction
PSET	Physical Science of Everyday Thinking
RBAI	Research-Based Assessment Instrument
REB	Research Ethics Board
RRMCS	Rotational and Rolling Motion Conceptual Survey
SE	Standard Error
StDev	Standard Deviation
VNOS	Views of the Nature of Science (survey)

To all lovers of science
and passionate educators

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Stephen G. Taylor, for reading all my drafts and giving me feedback and suggestions throughout this project. I would like to thank him for his continued support while I worked on my Master's in Education. I would also like to thank my external evaluator, Dr. Amir Shoham, for reviewing this research report.

I would like to thank Vanier College for their support in completing the research on their premises, along with colleagues and students who participated in it and willingly contributed data.

I further want to thank all my Performa instructors and classmates for the learning experience of the last few years.

Finally, I want to thank my family for their love and support.

INTRODUCTION

The purpose of this study was to compare modelling instruction with regular instruction and interactive engagement in terms of students' learning outcomes and attitudes toward physics in an introductory physics course (Mechanics) in an anglophone CEGEP in Montreal, Quebec, Canada. This topic, part of Physics Education Research (PER), is important because students have many conflicting preconceptions about the way nature operates, hence difficulty mastering physics concepts. They also have an erroneous perception of the scientific process and the nature of physics.

Modelling instruction gets students working collaboratively, actively modelling a physical phenomenon. This should result in possible explanations that can be shared with fellow students for a critical examination leading to improvements, revisions, or paradigm shifts through the guidance of the instructor (Halloun and Hestenes, 1987). Whiteboards mediate classroom discourse and allow students' reasoning processes to be exteriorized and open to scrutiny by their peers (Megowan, 2007). The literature shows that interactive engagement (IE) in general, and modelling instruction (MI) in particular, improve learning gains as measured by the Force Concept Inventory, a standardized test used by most studies. It also shows that modelling instruction tends to produce positive attitude shifts toward science and physics as measured by the Colorado Learning Attitudes about Science Survey (Madsen et al., 2015).

Therefore, this study proposed to answer three research questions. First, how does modelling instruction differ from regular instruction or interactive engagement in terms of learning outcomes for CEGEP Mechanics students? Second, how does modelling instruction differ from regular instruction or interactive engagement in terms of attitudes (or beliefs) about physics for those same students? Third, how are CEGEP Mechanics students perceiving (in terms of what they like or don't like) the introduction of modelling instruction?

The hypotheses were that students receiving modelling instruction would perform differently on the standardized assessments of deep conceptual understanding, on an exam testing problem-solving skills, and on a standardized assessment of expert-like attitudes. It was also thought that they would overall prefer this form of active learning.

To answer research questions, a mixed-methods design, with a strong quantitative strand based on a quasi-experiment using standardized pre- and post-tests comparisons, was used. Learning outcomes were defined as students' conceptual understanding as measured by pre/post-FCI (Force Concept Inventory) and RRMCS (Rotational and Rolling Motion Conceptual Survey) normalized gains, plus students' procedural mastery of the mathematics related to the concepts tested in a traditional exam with textbook-like problems. Attitudes about science were defined as students' self-reported beliefs about physics and their physics course on a pre/post-CLASS (Colorado Learning Attitudes about Science Survey) test.

Non-random, intact group samples of Mechanics college students (first year, second semester) represented the college population of Quebec. The treatment (modelling instruction) group consisted of two of the researcher's sections (35 students, out of whom 20 consented to participate) while the control groups consisted of ten other sections taught by colleagues (103 students in regular instruction out of whom 83 consented to participate; 71 students in interactive engagement out of whom 61 consented to participate). Near the start of the semester, the FCI, RRMCS, and CLASS pre-tests were given, whereas at the end of the semester, the same tests were given again as post-tests and a qualitative survey was added to better understand the perception of students on the new instructional design.

Normalized learning gains of averages, averages of normalized gains, effect sizes, and averages of normalized changes based on the FCI and the RRMCS tests were computed with matched data for both the treatment and control groups. Average group shifts in expert-like attitudes, along with effect sizes, were also computed with matched data based on the CLASS test.

Quantitative data generated by the research were analyzed through a multivariate analysis of variance (MANOVA) with p-values $\leq .05$ for statistical significance. Then comparisons were made with the control groups and the literature. Responses to a qualitative survey were analyzed to complement the quantitative analysis.

As the next few chapters will present in greater detail, it was found that the researcher's novice implementation of modelling instruction couldn't be distinguished from colleagues' implementation of either interactive engagement or regular instruction (which included some interactive teaching on a limited basis) for conceptual or procedural learning outcomes measured by the FCI normalized gains and final exam scores, although interactive engagement surpassed significantly regular instruction on the FCI. Comparisons couldn't be made based on the RRMCS. Novice modelling instruction did distinguish itself from regular instruction, on par with interactive engagement, only on FCI post-test scores that were on average a little over 60%, whereas for regular instruction this average was a little below 60%. Considering there was no distinction on pre-test scores, this has some meaning, that is, that the novice implementation of modelling instruction, just like other forms of interactive engagement, seems to have brought students slightly over the threshold of starting Newtonian thinking, contrary to regular instruction that was nearly there, but not quite. When results were compared with the literature, it was found that the FCI normalized gains of modelling instruction and interactive engagement were consistent with what other studies have reported. Regular instruction performed better than what has been reported in the literature for traditional instruction; it is believed the limited percentage of active learning incorporated into class activities might be the cause of it.

It was also found, based on CLASS results, that the researcher's novice implementation of modelling instruction couldn't be distinguished from colleagues' implementation of either interactive engagement or regular instruction, all methods failing equally at improving attitudes toward physics and its learning. Fortunately, attitudes didn't appear to worsen either. This observation is surprising for modelling instruction because the literature reports it to be one of the rare cases where attitudes typically improve.

It was further found, based on a qualitative survey, that the researcher's novice implementation of modelling instruction produced more dissatisfaction than satisfaction, which, although saddening, is consistent with what is reported for other, non-modelling methods fostering active learning (Wells, 1987; Deslauriers et al., 2019). The level of participation in that survey was very low (23%), though, so results might not be representative.

FIRST CHAPTER. PROBLEM STATEMENT

Students often come to physics courses with many misconceptions, or more properly conflicting and incoherent preconceptions that have been rendered obsolete by the scientific endeavour. As physics teachers can see when they elicit spontaneous explanations of physical phenomena from students, they have constructed naïve personal explanations, based on their everyday life experience. These explanations appear viable to them because they work in certain circumstances, but students don't realize that they don't in others (or they are not bothered by that). The power of scientifically accepted conceptions is that they apply more generally to more phenomena in a consistent manner, and are therefore more powerful. It is therefore desirable to help students evolve and revise their preconceptions toward more viable ones that increase their ability to make sense of the larger world of natural phenomena.

Notably, most of students' preconceptions or mental representations resemble those of Aristotelian or Middle Age physics, before modern experimental science was truly born. An example of the former is when students believe that every motion has a cause and that acceleration requires an increasing force. An example of the latter is when students believe that an internal force (akin to the impetus) proportional to mass and velocity can be imparted to objects by an applied force and wear out, consumed by the motion or dissipated by some resistance. Such preconceptions are difficult to dislodge (just as it was long for historical science to overcome them, keeping in mind that Newtonian mechanics was formulated in *Philosophiæ Naturalis Principia Mathematica* only in 1687) and present obstacles to deep learning and understanding of physics. Furthermore, students even have misconceptions about the nature of the field. For many, physics is just a collection of facts and a matter of finding the right formula to solve a given problem, which is far from the way physics is perceived and done by scientists.

Physicists perceive the field as a set of interconnected and coherent constructs organized in theories, which they use in developing viable models of reality. An instructional method called modelling instruction (MI) has been developed by Halloun and Hestenes (1987) to mirror the real practice of science by experts. As scientists, we develop viable models to explain observations and make sound predictions. That is the way modelling works, including the social component so important to establish accepted scientific theories. In a nutshell, it organizes content around scientific models, engage students in constructing those models collaboratively based on experiments, and pushes them to test conceptual models that have been developed in new contexts leading to a failure that justify a revision and a new cycle of inquiry. This approach appears to have tremendous results in conceptual gains based on the Force Concept Inventory (FCI) test. The results are such that one could easily be enthused by the idea of implementing this method at the core of his or her teaching approach.

Surprisingly, this modelling method appears not to be well known in Quebec, as the researcher hardly found any local reference to it. One can easily find in Quebec complete or partial implementations of other types of non-traditional pedagogies such as flipped classrooms, problem-based learning (PBL), interdisciplinary islands of rationality (IIR), project-based learning (PBL), peer instruction (PI), portfolios, reflective writing, and gamification, but the closest approach to modelling instruction (MI) that could be found was argument-driven inquiry (ADI), used in some secondary schools. However, modelling instruction is becoming more and more popular in

American high schools where it has been widely and successfully used. In 2000, the U.S. Department of Education selected the modelling instruction program at Arizona State University (ASU) as one of seven K-12 educational technology programs designated as exemplary or promising out of 134 (Jackson et al., 2008) and in 2001, modelling instruction was recognized as one of only two exemplary K-12 science programs out of 27 that were evaluated (U.S. Department of Education, 2001). This might explain why, in 2008, Brewe was reporting that approximately ten percent of American high school physics teachers have had some formal training in the method. According to a private communication with Dr. Colleen Megowan-Romanowicz, about 500 new modellers are trained every year through the American Modeling Teachers Association (AMTA), over 90% of whom adopt modelling instruction permanently. That would represent 4000-5000 more teachers since Brewe's paper.

Modelling instruction is also used in some American colleges and universities, although it is less common in higher education (Brewe, 2006, 2008). This rises some questions about why the situation is such. It appears that modelling instruction impact learning outcomes of students positively, as will be shown in the literature review, but that it is more easily implemented in high schools because of the flexibility it requires. Yet, as mentioned, some college and university teachers do use the method.

Like probably many college and university modelling professors in American colleges and universities did in their own context, one could wonder how a whole course of CEGEP physics (like Mechanics, Electricity and Magnetism, or Waves and Modern Physics) could be adapted to use the method. In modelling instruction, theory classes and laboratory experiments are generally fused into studio classrooms as opposed to being separated. This change might be considered a challenge by the current faculty, or else difficult to achieve for lab scheduling reasons. Knowing more about modelling instruction would be beneficial before planning future large-scale experiments in Quebec. Fostering the implementation of the method by innovative teachers could come as a further step in a longer-term project. What we tried to accomplish though this research is a more modest first step, by adapting parts of existing modelling instruction materials and processes to a complete calculus-based, introductory Mechanics course (Physics NYA). It has been chosen to apply the method to a full course because private discussions with practitioners emphasized the necessity of a full semester for potential effects to start manifesting. Furthermore, doing it for only one unit would probably have not given the students the time to adjust and get a grasp of the method. It could also have led to objections regarding the change of method because of the increased emphasis on active learning.

This research is nevertheless important because, as teachers, we want our teaching to be effective, and we certainly want our students to understand, master, and retain the concepts that we teach. To do so, we need to engage students and find ways that produce deep and meaningful learning as opposed to surface or rote learning based on memorization.

It is also important because citizens of tomorrow need to be scientifically literate and to develop higher levels of thinking like the capacity to analyze, evaluate, and create or synthesize. This is necessary to make informed judgments and to think critically. This is also necessary to feel like having a voice worth of hearing and develop a will to participate in techno-scientific debates — like those related to climate change and geoengineering, the use of nanotechnologies, the deployment of the 5G network or the use of nuclear power — while being well equipped to oppose misinformation from lobbies and groups of interest.

In summary, this research has studied the effect of modelling instruction (fostering a form of interactive engagement or active learning) on learning outcomes and attitudes about science when applied to a full calculus-based, introductory mechanics course in CEGEP. This topic, part of what is generally called Physics Education Research (PER), is important because students have difficulty mastering physics concepts due to numerous conflicting preconceptions. Students also have a misunderstanding of the nature of physics. Since physics is a fundamental science, improvements in learning gains and attitudes should translate into a better understanding of the scientific process and its concepts.

SECOND CHAPTER. CONCEPTUAL FRAMEWORK

Methods of teaching physics should be revisited because students have difficulty mastering this discipline, often have conflicting preconceptions that impair their understanding, and develop fragmented knowledge through rote learning instead of a coherent organization of knowledge around core models and principles. This can be observed though the generally low performance of students and the struggle they experience in physics classes. Halloun and Hestenes (1985a, 1985b) report in particular that students' (generally erroneous) common-sense beliefs — such as the ideas that an object requires a force to be in motion, that every motion needs a cause, that motion is produced by the larger of two competing forces, or that an internal force maintains the motion of an object independently from external agents — are very stable and lead to a systematic misinterpretation of introductory physics courses' material. Such stable preconceptions due to extensive personal experience has also been reviewed more recently by Neidorf, Arora, Erberber, Tsokodayi and Mai (2020). To improve learning and provide a more authentic experience fostering the development of scientific competencies, it is worth examining the structure and construction of scientific knowledge, and how this can inform science education. This is the goal of this chapter, which presents the conceptual framework on which this research was designed.

1. CONSTRUCTION OF SCIENTIFIC KNOWLEDGE

Hestenes (1987a, 1987b) has produced interesting work on the role of models in scientific knowledge. Scientific knowledge is comprised of both factual and procedural knowledge. The latter consists of strategies, tactics, and techniques (vaguely called the scientific method) driving a dialectical inference of factual knowledge through an experimental dialogue with nature, which tells us what is viable and what is not. Factual knowledge, which holds despite attempts to invalidate it, is in turn comprised of a set of theories and models serving to interpret and make sense of empirical data. These models are conceptual representations of real things.

Theories are structured around a framework and a semantic base. The former is a core of basic laws relating basic descriptive variables. The latter is made of correspondence rules used to interpret and translate the theory into corresponding phenomena, from the abstract and formal to the concrete and empirical. On this core is built a superstructure of definitions, conventions, and theorems, including derived laws. This superstructure enlarges the applicability and usability of the theory to specific subdomains of application. Subdomains encompassed by the superstructure grow and evolve as science advances (Hestenes, 1987a, 1987b).

Theories are used to build models that are essential to the scientist's understanding of nature. In physics, we use representational and mathematical models of objects, interactions, systems and processes to describe, explain or predict phenomena. All scientific models have their limitations, however, as they are simplifications of reality, constructed as viable representations of some of its aspects of interest at a given time in relation to a given project. Nevertheless, these models have proven to be highly effective and powerful, as we can see by the progress science has enabled humans to achieve.

Those models are developed in four stages: (1) description, (2) formulation, (3) ramification, and (4) validation (Hestenes, 1987a, 1987b). Modelling starts with describing objects under study and agents acting on these objects. This is done by providing names identifying them and by attributing object, state or interaction descriptors (variables) representing diverse properties. These descriptors are then used in state and causal mathematical equations at the second stage of formulating the model. The third stage of ramification consists in solving equations, which then must be validated in the fourth stage through a reasonableness assessment, or, even better, an experimental test. Interpretation is present all through the process, from the start when natural and empirical phenomena must be translated into the abstract and formal concepts of the theory, through the end when mathematical solutions must be reinterpreted in terms of natural phenomena.

As can be seen, models are central to the scientific enterprise. Because models are considered so central, both Hestenes and the researcher could be said to share a semantic view of scientific theories (Develaki, 2007). Scientific theories are viewed as a “family of models” respecting the structure and the assumptions of the theory (Giere, 1988). Good scientific models and theories are testable, revisable, explanatory, conjectural and generative (Windschitl et al., 2008). Lattery (2017) proposes to view a scientific model as “a limited (imperfect), inferential, and external representation of a physical system” (p. 29). He claims that the goal of the scientific modelling process is to build explanatory and empirically accountable conceptual models, or said otherwise, to build conceptual models that synthesize at least elements from empirical (the what extracted from experimental measurements), physical (the why through some material analogy), and prior conceptual models of a physical target system.

One way to illustrate that with a simple example is to look at what could happen when somebody studies the refraction of a monochromatic light beam through a prism of glass. Observations could lead to a physical analogy with a motorized toy truck (analogous to the light) moving at constant velocity and passing from concrete (analogous to air) to mud (analogous to glass). That in turn could inspire some experiments leading to a mathematical expression of the law of refraction describing the phenomenon. Those two models could then be combined in a conceptual model of light rays moving in straight lines at a constant speed till they hit the boundary of a medium where the light moves at a lower speed. At the boundary, the transmitted light rays behave like the law of refraction describes. A new physical analogy could be taught of, like a safeguard (analogous to light) running to save a swimmer in distress in the shortest time possible. The optimal path, from sand (analogous to air) to water (analogous to glass) would correspond to the law of refraction. Combined with the empirical model, this could lead to a further development of the conceptual model by incorporating Fermat's principle of least time. By using such a simple example and thinking of others, it can already be seen how the semantic, model-centred view of scientific knowledge development might be transferred to the classroom, and that's what the next few sections explore.

2. SCIENCE EDUCATION

As science teachers, if we want to facilitate students' learning of science, then we must question ourselves about the essence of science and the goal(s) of science education. It appears that the results of current science education are poor. Surface knowledge is the norm and the failure rate is high (Sadler & Tai, 2001). We already alluded to several obstacles — such as conflicting preconceptions and formula-centred rote learning — facing students who study physics. So, what ought we do?

There is evidence from cognitive science that interactive engagement will foster deeper learning because of the activation of neural pathways leading to long-term memory, hence retention of more usable and transferable knowledge (DeHaan, 2005). This complements the socio-constructivist perspective that argues for in-class learning activities that engage students with one another so they can construct better knowledge through socio-cognitive conflict (Perret-Clermont et al., 2004). Referring to the work of Piaget who theorized that learning was achieved through the processes of assimilation and accommodation (Piaget, 1950), deep learning of viable concepts in physics, which often contradict naïve intuition based on common sense and personal experience, requires teachers to foster accommodation through sound management of cognitive conflicts. This means that teachers should strategically place students in situations where their preconceptions are challenged and guide them on the side through the process of revising and replacing them with more viable conceptions, namely Newtonian concepts if the course is on classical mechanics. McKagan, Perkins, and Wieman (2007) have summarized key principles of reformed physics education as the following:

1. Interactive engagement can lead to higher learning gains (Hake, 1998);
2. Directly addressing common misconceptions can lead to higher learning gains (McDermott, 2001);
3. Unless student beliefs about science are explicitly addressed, these beliefs tend to become more novice-like (Redish et al., 1998; Perkins et al., 2005);
4. People have a limited short-term memory, so the course should focus on important points, have a coherent structure, and eliminate nonessential details to reduce the cognitive load (Mayer, 2003);

5. For students to gain a conceptual understanding, all aspects of the course, including homework and exams, must address conceptual understanding, not just numerical problem-solving.

There are many ways we can imagine applying these ideas in class, for it is the aim of most reformed pedagogies. Modelling instruction is one of them, with the particularity of being centred on the authentic process of scientific modelling.

3. MODELLING INSTRUCTION

Modelling instruction (MI) is a method of teaching sciences, based on a socio-constructivist, interactive pedagogy, meaning that learning is viewed as socially situated and constructed through interactions with others. This method was developed by physicists Halloun and Hestenes (1987) at Arizona State University (ASU), and further adapted to high school physics by Malcolm Wells (Wells et al., 1995). The main idea is to get students working collaboratively, actively modelling physical phenomena in a way that reflects scientific practice more authentically. This should result in possible explanations that can be shared with fellow students for a critical examination leading to improvements, revisions, or paradigm shifts under the guidance of the instructor. The goal is to arrive at a deeper understanding of the why and how of both phenomena and the scientific process itself.

More specifically, modelling instruction is a mode of dialectical teaching by which cognitive conflicts are resolved rationally, mirroring the scientific dialectical inference of experts who seek a rational resolution to conflicts between presently available information and incompatible conceptual frameworks. This process is a reflection of the self-regulation process of human cognition (assimilation and accommodation) presented by Piaget (1950). Dialectical teaching requires that students formulate their common-sense beliefs explicitly, check their consistency with empirical evidence, check their consistency with other beliefs they hold, and compare them with alternative beliefs, in particular relevant scientific beliefs (Hestenes, 1987b). In the case of modelling instruction, this is achieved through cycles of model development and model deployment. The first stage comprises pre-lab discussions, lab investigations, and post-lab discussions, while the second stage may comprise worksheet exercises, quizzes, lab practicums and unit tests to assess learning (Jackson et al., 2008). This is ideally implemented, when available, in studio classrooms where it is easier to merge theory and practice as needed, rather than splitting the two between the classroom and the laboratory on a fixed schedule as it is normally done in traditional courses.

Interestingly, the modelling cycle is consistent with Vygotsky's ideas of social construction (Vygotsky, 1962, 1978) through social interactions and language mediation. Based on his ideas, Adey, Shayer and Yates (2001) have developed five pillars of teaching: concrete preparation, cognitive conflict, construction, metacognition and bridging. These pillars can be used to summarize key aspects of the modelling instruction. The concrete preparation is akin to the very practical paradigm lab. The cognitive conflict happens often as students face failing preconceptions (exposed by carefully planned activities) or observe new phenomena for which prior models fail. Construction happens when students restructure their mental schemata following prompts for cognitive assimilation and accommodation, helped by board meetings and Socrative dialogue. Metacognition is fostered through questions like "How did you solve that?" or "Please explain to the others in your group why you think that." Bridging is a principle that is used through deployment activities, like when the modeller wants to bring students to the realization that a table deforms under an applied force and pushes back as a normal force, even though this deformation is not observable to the naked eye. That is the target concept, often difficult for students, so the modeller takes a few steps back and discuss a case that presents no issue to students: a person pushing on a spring and feeling the spring pushing back. Then a few intermediary cases (bridges) are explored to eventually culminate in the acceptance and deep understanding of the target concept (Camp & Clement, 2010). Bridging can also be seen as making links to the broader curriculum and to real life. Modellers do refer back to previous models or elements and tools thereof as they progress to the study of new phenomena, and they do make connections with real life as they strategically and indirectly confront unviable preconceptions arising from naïve and limited real-life experience. There is bridging between inconsistent, unexamined student models of nature and more viable and robust models developed by scientists.

4. CONCLUDING REMARKS

Considering the previous exposition of our conceptual framework (mapped in figure 1), this research has undertaken to study a small part of the problem stated earlier, which was concerned with the difficulty students experience learning physics due to their conflicting preconceptions on the subject matter and the way science is done by professional scientists. The purpose of this study was therefore to compare modelling instruction (treatment group) with regular instruction and interactive engagement (control groups) in terms of students' learning outcomes and attitudes toward physics (dependent variables) in an anglophone CEGEP in Montreal, Quebec, Canada. This was done by focusing on the following three research questions:

1. How does modelling instruction differ from regular instruction or interactive engagement in terms of learning outcomes for CEGEP Mechanics students?
2. How does modelling instruction differ from regular instruction or interactive engagement in terms of attitudes (or beliefs) about physics for CEGEP Mechanics students?
3. How are CEGEP Mechanics students perceiving (in terms of what they like or don't like) the introduction of modelling instruction?

The first question was useful to assess if modelling instruction would be more efficient than other methods in evolving students' common-sense beliefs about the natural world toward a Newtonian perspective that might in turn help students develop problem-solving skills on more robust foundations. The second question was useful to assess how modelling instruction differed in fostering beliefs or attitudes about physics that would be more in line with those of expert physicists, for it was believed that expert-like attitudes and a deeper understanding of the nature of science lead to more stable, authentic and meaningful learning while also possibly fostering sound civic engagement. The third question was essentially designed to gain insight about the appreciation of students of the modelling approach, as it is also believed that a positive perception would lead to greater participation and motivation, thus enhanced learning for students.

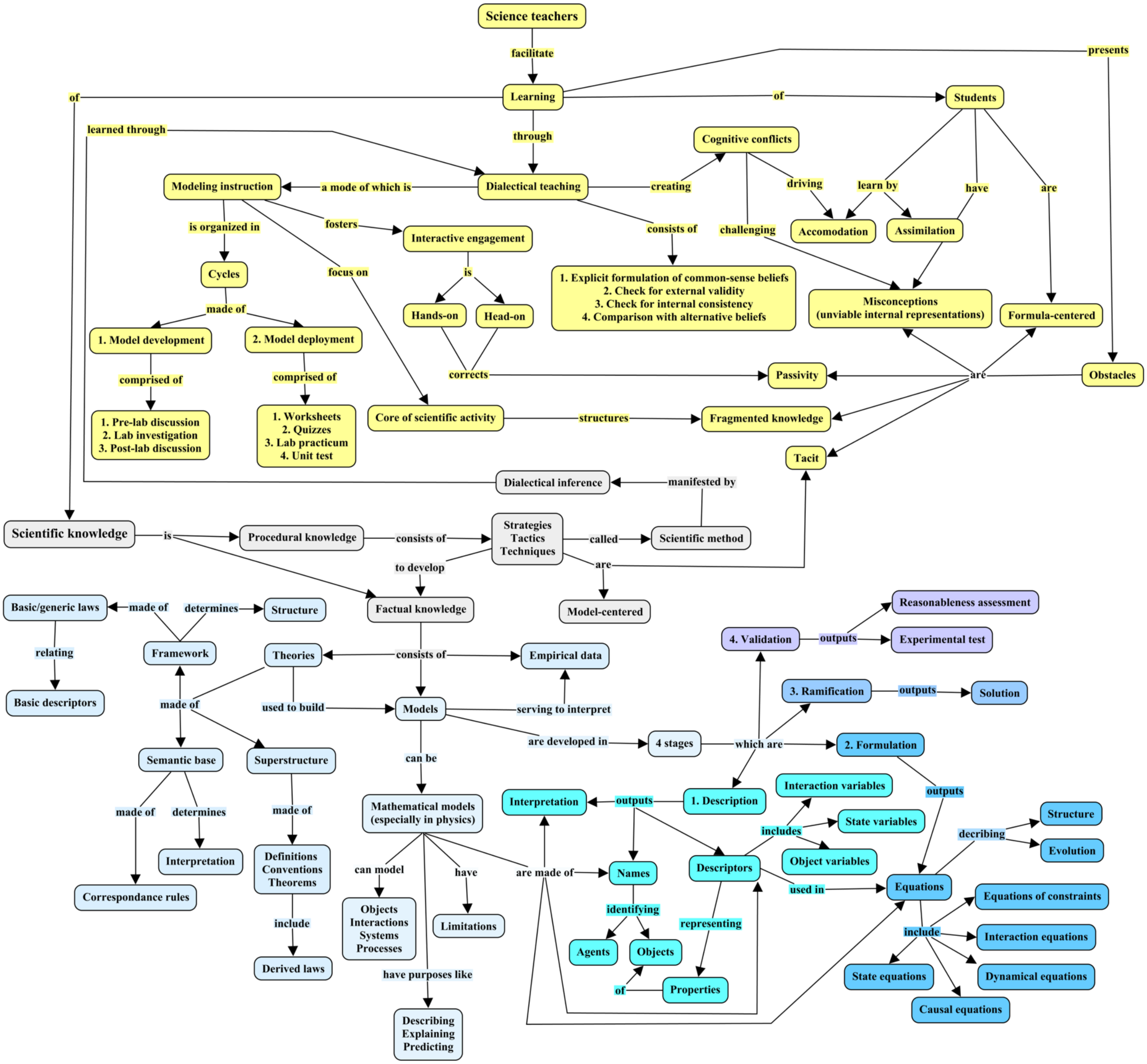


Figure 1. Overview of the conceptual framework.

THIRD CHAPTER. LITERATURE REVIEW

Equipped with the conceptual framework presented in the previous chapter, it is worth researching the literature to further understand novice conceptions and behaviours of students in physics, modelling instruction and the way it is operationalized in the classroom, the reported effectiveness of the method, and the current state of its adoption in higher education. Findings are reported in this chapter. Research questions are then presented and framed against this literature review and the conceptual framework previously discussed.

1. NOVICE CONCEPTIONS AND BEHAVIOURS

Lattery (2017) defines deep leaning in science as being able to explain how and why target scientific concepts succeed while prior ones fail, including processes convincingly bridging from the latter to the former, and how those target concepts apply and transfer to new physical situations. A lack of attention to students' conflicting preconceptions often lead to shallow or rote learning. Actually, two major misconceptions deprive students of deep, meaningful learning in Newtonian physics: the impetus and dominance principles. The former pretends that force is something objects acquire to move, contradicting the first and second laws of Newton. The latter pretends that when two objects interact, the larger or more active object exerts a greater force, contradicting the third law of Newton. Those two common misconceptions seriously limit the understanding of the concept of force at the centre of Newtonian mechanics. Working on them tends to pave the way for eliminating many other misconceptions without direct intervention because most of them rely on a faulty conceptualization of force, which is at the heart of the impetus and dominance principles (Hestenes, 1992).

Furthermore, Malone (2008) reports, based on other research, how novice problem-solving strategies differ from expert ones (table 1) and how different teaching methods, namely modelling versus traditional, can impact the evolution of those strategies. Malone actually completed two studies: one on how modelling instruction influences students' cognitive structures, and another one on how it influences problem-solving and metacognition. In those studies, modelling appeared to foster more expert-like strategies. In particular, modelling students tended to classify problems based on deep structure (e.g., uniformly accelerated motion, Newton's second law, etc.) more than nonmodellers who seemed to privilege surface features (e.g., inclined plane, springs, etc.). This makes sense if one recalls that models are at the centre of the scientific process and, as Malone (2008) reports, that modelling instruction aims at helping students organize knowledge around such basic physics models, with multiple representations. Furthermore, modelling instruction continually asks students to explain how they know what they say and to evaluate their conclusions.

Table 1. Comparison of expert and novice problem-solving strategies as presented by Malone (2008, p. 020107-7).

Expert behaviours	Novice behaviours
• Typically use a working forward strategy except on more difficult problems.	• Typically use a working backward strategy.
• Perform an initial qualitative analysis of the problem situation.	• Usually manipulate equations discovered via equation hunting.
• Construct diagrams during solution process.	• Rarely construct or use diagrams.
• Spend time planning approach sometimes via models of the physical situation.	• Rarely plan approach, simply dive in.
• Use fewer equations to solve problems.	• Use more equations to solve problems.
• Usually solve problems in less time.	• Usually take more time to solve problems.
• Refer to the physical principles underlying the problem.	• Refer to the numeric elements of the problem.
• Concepts more coherent and linked together.	• Concepts not coherent and lack applicability conditions for special cases.
• Fewer errors—concepts usually deployed correctly.	• More errors—concepts usually deployed incorrectly.
• Can use more than one representation to solve problems—which usually allows them to deviate to other solution paths when stuck.	• Usually only utilize a numeric representation to solve problems—once they become stuck rarely can free themselves.
• Check and evaluate solution by a variety of methods (i.e., more flexible).	• Superficially check solution if at all.
• Rarely refer to problem statement or text.	• Frequently refer to problem statement and textbook (especially examples).

2. MODELLING INSTRUCTION

Modelling instruction has been widely used in American high schools, with great success in terms of conceptual gains and attitude shifts compared to traditional methods (Brewe, 2006, 2008). Within this didactical framework, an emphasis is put on the construction and application of models (Jackson et al., 2008). These models are viable constructs built in accordance with the natural world. They are temporal and constantly validated and refined (Brewe, 2008). This is done in a context where the focus is put on inquiry, observation, collaboration, communication, and reasoning, with an instructor acting as a questioner rather than a provider of knowledge (Megowan, 2010).

2.1 Model construction

In high school Newtonian mechanics (USA), Hestenes (1997) identifies five basic particle models which can be subdivided into kinematical and causal models: (1) constant velocity and free particle models; (2) constant acceleration and constant force models; (3) simple harmonic oscillator and linear binding force models; (4) uniform circular motion and central force (with constant radius) models; (5) collision and impulsive force models. On the other hand, Brewe (2002) identifies six major general models in introductory physics: (1) particle model; (2) rigid-body model; (3) constant acceleration model; (4) free-particle model; (5) harmonic oscillator model; (6) field model.

When adapted to specific situations, general models become specified models, which can then be used to solve specific physics problems. Common modelling tools are mainly systemic like the system schema, accounting like energy pie charts, energy bar charts, and the “equation of everything” (the first law of thermodynamics about energy conservation), or functional like interaction energy graphs, potential graphs and equipotential surfaces (Brewe, 2002). Other modelling tools that can be exploited are state diagrams, motion maps, kinematic graphs, force or free-body diagrams, momentum vectors, field lines and field vectors. Models represent the structure in a system or process, and such structure can be systemic (composition, environment, connections), geometric (position, configuration), temporal (descriptive, causal), or interactive (interaction laws) (Hestenes, 1997). Hestenes (2006) later added the object structure (intrinsic properties) as a fifth type, but the researcher would argue that it could also be seen as a detailed part of the composition component of the systemic structure. Figure 2 summarizes the process of constructing a model from a given situation and figure 3 represents the structure in a model for the modified Atwood's machine, as explained by Hestenes (1997).

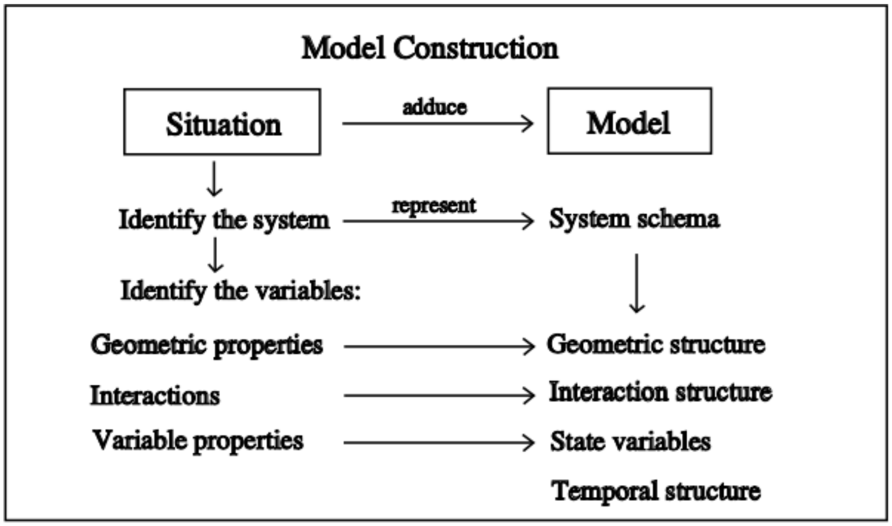


Figure 2. Model construction as presented by Hestenes (1997, p. 948).

A benefit of using such models and representations, besides giving a more authentic experience of the process of science, is that they offer multiple modes of access to a target physical system by which a student can understand and explain it. Thus, they deepen understanding and give meaning as not only the how, but also the why are described. These cognitive tools also bridge between the concrete and the abstract, and most useful for college students who still struggle on their way to formal reasoning (Torkia-Lagacé, 1981), from the visual to the mathematical.

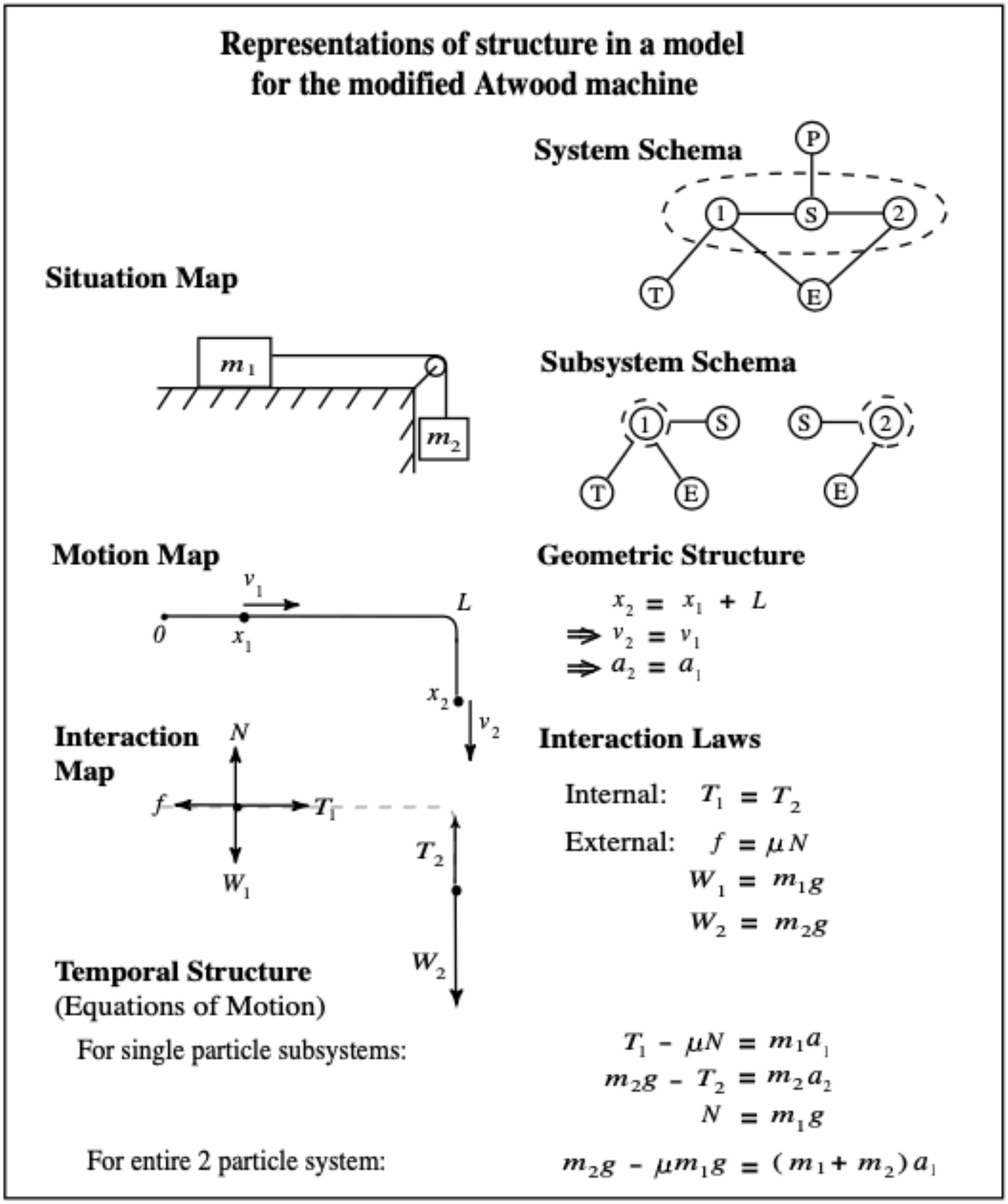


Figure 3. Representation of structure in a model for the modified Atwood's machine as presented by Hestenes (1997, p. 944).

2.2 The modelling cycle

Modelling instruction revolves around the modelling cycle, which students go through multiple times, as in a spiral of repeated practice. Figure 4 presents the method as outlined by Hestenes (1997), with phases of (1) Model Development, (2) Evaluation, and (3) Application in concrete situations. The first part of the figure describes the main objectives of the method whereas the second part summarizes how to apply modelling in the classroom.

Similarly, Colleen Megowan (2010) split the modelling cycle in three phases: (1) Model Construction, (2) Model Testing and Elaboration, and (3) Model Application. In the first stage, a paradigm lab helps students to collaboratively identify a model and its parameters. That paradigm lab generates a series of conversations leading students to organize what they know and what they want to know, generate testable hypotheses, seek evidence and construct an argument (Windschitl et al., 2008). This ends with a post-lab "board meeting" where student-centred discussions occur. In the second phase, the model is collaboratively refined and tested through a variety of tasks and problems that elicit a better mastery of the model initially identified. This is preceded by some homework serving the same purpose. In the third phase, students use the model to solve more complex and contextualized problems, either as homework or in-class whiteboarding exercises. This is completed by a final lab practicum and a final unit assessment testing their ability to solve qualitative and quantitative problems. She recommends groups of three students for collaborative work that occurs before whole-class board meetings essentially led by students themselves.

MODELLING METHOD SYNOPSIS

The Modelling Method aims to correct many weaknesses of the traditional lecture-demonstration method, including the fragmentation of knowledge, student passivity, and the persistence of naïve beliefs about the physical world.

What to teach: Model-centred instructional objectives

- To engage students in understanding the physical world by constructing and using scientific models to describe, to explain, to predict, to design and control physical phenomena.
- To provide students with basic conceptual tools for modelling physical objects and processes, especially mathematical, graphical and diagrammatic representations.
- To familiarize students with a small set of basic models as the content core of physics.
- To develop insight into the structure of scientific knowledge by examining how models fit into theories.
- To show how scientific knowledge is validated by engaging students in evaluating scientific models through comparison with empirical data.
- To develop skill in all aspects of modelling as the procedural core of scientific knowledge.

How to teach: Student-centred instructional design

- Instruction is organized into modelling cycles which engage students in all phases of model development, evaluation and application in concrete situations — thus promoting an integrated understanding of modelling processes and acquisition of coordinated modelling skills.
- The teacher sets the stage for student activities, typically with a demonstration and class discussion to establish common understanding of a question to be asked of nature. Then, in small groups, students collaborate in planning and conducting experiments to answer or clarify the question.
- Students are required to present and justify their conclusions in oral and/or written form, including a formulation of models for the phenomena in question and evaluation of the models by comparison with data.
- Technical terms and representational tools are introduced by the teacher as they are needed to sharpen models, facilitate modelling activities and improve the quality of discourse.
- The teacher is prepared with a definite agenda for student progress and guides student inquiry and discussion in that direction with “Socratic” questioning and remarks.

The teacher is equipped with a taxonomy of typical student misconceptions to be addressed as students are induced to articulate, analyze and justify their personal beliefs.

Figure 4. The modelling method as outlined by Hestenes (1997, p. 941).

The specifics of the modelling cycle have also been detailed in terms of two stages (Hestenes, 1997; Jackson et al., 2008): (1) Model Development and (2) Model Deployment. In Model Development, a demonstration and a pre-lab discussion are followed by a lab investigation (without pre-printed instructions) and a post-lab discussion. Technical terms and representational tools are introduced here, and both Socratic dialogue and whiteboarding (the use of little whiteboards by teams of students to work out their ideas and mediate the conversation) are extensively used. Model Deployment makes use of worksheets, followed by quizzes, a lab practicum, and a unit test.

For Brewe (2008), the modelling cycle rather consists of five stages: (1) Introduction and Representations, (2) Coordination of Representations, (3) Application, (4) Abstraction and Generalization, and (5) Refinement. The stage of Introduction and Representations creates a need for a new model through an inquiry-based laboratory activity and introduces basic concepts and graphs. The Coordination of Representations relates graphs to more common graphical or diagrammatical representations. The Application stage develops equations and applies knowledge and tools to solve problems and draw conclusions. The stage of Abstraction and Generalization identifies common features of special cases where the model is applicable. Refinement improves upon the model based on encounters with new situations.

2.3 Modelling discourse management

Dwain Desbien (2002) has made an important contribution to modelling discourse management in response to some observations that were gathered, starting in 1995, through interviews of honours university physics students that had went through a course using a prior version of modelling instruction. The main points were that many students felt that the representational tools were a burden they would gladly skip if not marked on them, that physics homework and exams should look like those found in a mathematics, and that conceptual questions were unfair because understanding can be shown by the manipulation of equations. In Desbien's words, "the problem solution was not to change the class activities, but rather change how the class was managed. Students needed to develop the models themselves. The epistemology of science needed to be explicit. Use of shared representational tools needed to be developed collaboratively. The class needed to be a community working together like scientists through peer-peer interaction. Modeling discourse management was developed to meet these goals" (p. 49). In his dissertation, he describes a superior, socio-constructivist modelling method of instruction as being based on seven elements:

1. There is a deliberate creation of a cooperative learning community.
2. This community experiences an explicit need for the creation of models in science.
3. Through the process, students create inter-individual meaning.
4. Teachers seed elements in the discussion through well-thought interactions with individual groups.
5. Teachers intentionally don't close the discussion at the end of class; instead, adequate follow-up is provided.
6. They foster inter-student discussion by not controlling or not being involved at all in whole-class discussions.
7. Finally, they actively evaluate students formatively.

Based on his research, Desbien (2002) showed that modelling discourse management seemed to improve both students' understanding and views about science compared to both traditional teaching and traditional modelling classroom management.

2.4 Whiteboarding and board meetings

Whiteboarding is a central feature of modelling instruction (Megowan, 2007) and relates to the theory of inscriptions of Roth and McGinn (1998). It mediates classroom discourse and allows students' reasoning processes to be exteriorized and open to scrutiny by their peers. It is potentially a very effective way to detect misconceptions and to address them, but certain conditions must be met for this to be successful.

1. Instructors must be able to listen to students engaged in productive inscription-mediated discourse leading to model construction.
2. They must also have a clear picture of how the model ought to look and possess a good mastery of metaphors and language linked to students' misunderstanding.
3. Teachers should also develop the skill to orient discourse without giving answers, and they should be aware of how student inscriptions on whiteboards relate to their thinking about models.
4. Finally, instructors should know about the cultural models of schooling students bring to the class and the new models that can be set in their place with skillful management of classroom activities.

Collaborative work can be a waste of time when not properly managed. Besides skillful management from teachers, a curriculum based on a sound choice of tasks fostering inter-student discourse, within their zone of proximal development — the zone where learning is challenging but achievable with guidance and encouragement — (Vygotsky, 1962, 1978), is essential.

2.5 Benefits of modelling instruction

There are many advantages to modelling instruction. From a cognitive perspective, it simplifies the content and reduces the cognitive load on working memory (Gerjets & Scheiter, 2003; Paas et al., 2003) by organizing it around a small number of general models, allowing better and easier learning. Furthermore, it allows students to appreciate physics' coherence as knowledge is integrated and connected to prior models that have been refined; it is more reflective of the way physicists think and work (thus it is more authentic); and it develops transferable problem-solving skills as the ability to develop, test and refine models is one that is shared by other technoscientific disciplines (Brewe, 2006, 2008). It also fosters intelligent engagement in public discourse and debate about technoscientific issues (Jackson et al., 2008). Desbien (2002) noticed in his dissertation that modelling discourse in university physics improved retention rates during first and second semesters at Phoenix institutions part of his research.

3. EFFECTIVENESS OF MODELLING INSTRUCTION

Various studies and meta-analyses have already been performed to assess the effectiveness of modelling instruction in high schools, community colleges and universities (e.g., Hake, 1998; Hestenes, 2000, 2006; Brewe et al., 2010; Madsen et al., 2015). Although none have been found related to the special context of Quebec’s CEGEP system, it is worth exploring and reporting what resulted from those other studies.

3.1 Conceptual understanding of mechanics

Physics education research widely relies on the Force Concept Inventory (FCI) to get insights into the learning outcomes of students in mechanics (Madsen et al., 2017a). It assesses, according to Hestenes (1997), “the effectiveness of mechanics courses in meeting a minimal performance standard: to teach students to reliably discriminate between the applicability of scientific concepts and naïve alternatives in common physical situations” (p. 937). This is a standardized test for which validity and reliability have been demonstrated (Jackson et al., 2008). Details of this multiple-choice questions test are described by Hestenes, Wells, and Swackhamer (1992). Henderson (2002) has estimated that in his practice at the University of Minnesota, at most 2.8% of students don’t take the FCI seriously when it is ungraded, compared to 0.6% when it is graded. If cases in which students refuse to take the test (0.5%) or leave a lot of blanks (1.4%) are removed because they are easily extracted from the sample, at most 0.9% of students might have lower scores on ungraded tests. The percentage appears low enough not to require a grading of the FCI. Furthermore, Henderson has verified that post-test results are not affected in any statistically significant manner by the FCI being given as a pre-test. It should be noted that students are never told that they will have to answer the same questionnaire at the end of the semester when they are given the FCI at the beginning of classes. Hestenes and Halloun (1995) consider FCI scores of 60% as the entry threshold to Newtonian physics and 85% as the Newtonian mastery threshold. This means that at 60% students barely begun to use Newtonian concepts coherently in their reasoning, and below their thinking would displays such characteristics as “(1) *undifferentiated concepts* of velocity and acceleration; lacking a vectorial concept of velocity; (2) lacking a *universal force* concept (i.e., believing that there are other influences on motion besides forces), and unable to reliably identify the agents of forces on an object; (3) *fragmented and incoherent concepts* about force and motion” (p. 505). This also means that above 85%, students could be considered as confirmed Newtonian thinkers.

Seven thousand five hundred (7,500) American high school physics students were involved in the Modelling Instruction Project in 1995–1998. Hestenes (2006) reports 29%/69% pre-test/post-test means for expert modellers compared to 26%/52% for novice modellers and 26%/42% for traditional teachers involved in his sample. The average gain of students taught by expert modellers was more than two standard deviations higher than for traditional teachers. This is consistent with results for 20,000 students (Hestenes, 2000) collected between 1994 and 2000: 26%/69% pre-test/post-test means for expert modellers compared to 26%/53% for novice modellers and 26%/42% for traditional teachers. Vesenka, Beach, Munoz, Judd, and Key (2002) also tested modelling instruction in undergraduate algebra-based physics courses of two universities and found that students achieved over a one-half standard deviation from those receiving traditional lectures, while their FCI normalized gains were two times greater. In 2016, a meta-analysis of 63 papers on FCI gains, representing 31,000 students in 450 classes, was performed. It was found that the mean normalized gain is 22% for traditional lectures and 39% for interactive engagement (Von Korff et al., 2016).

3.2 Attitudes and beliefs about physics

Studies based on the Maryland Physics Expectations Survey (MPEX), the Epistemological Beliefs Assessment for Physical Science (EBAPS), the Colorado Learning Attitudes about Science Survey (CLASS) and the Views of the Nature of Science (VNOS) survey have repeatedly shown that students of all ages have difficulty learning how science knowledge is constructed and in most cases regress in sophistication over a semester-long science course (Otero & Gray, 2008). Recently, Madsen, McKagan, and Sayre (2015) published a meta-analysis of twenty-four North American studies, based on the CLASS and the MPEX, that looked into the attitudes or beliefs of students about physics and how closely they aligned with those of experts. The surveys asked students questions about how they learn physics, how physics is related to their everyday lives, and how they think about the discipline of physics. They report that student attitudes often deteriorate as they go through typical (even reformed) physics classes, but that on the contrary, significant improvements result from an explicit focus on model-building and developing expert-like beliefs. They submit that the teaching method is the main factor, although the class size and the student population also explain part of the variance in shifts. They report that in most large calculus-based courses students often leave with the belief that physics is about memorizing facts and plugging numbers into equations, and not relevant to their life. Modelling instruction would perform better possibly because it makes student work in small groups to mimic the way scientists create knowledge and develop models with multiple representational tools, using labs to refine or revise conceptions and using whiteboard and board meeting discussions to arrive at consensus. By having a more authentic experience of physics, students would develop more expert-like beliefs about the field. The positive impact of modelling instruction is consistent with results from Brewe, Kramer and O'Brien (2008, 2009) at Florida International University (USA) and those of de la Garza and Alarcon (2010) at Tecnológico de Monterrey (Mexico).

4. ADOPTION OF MODELLING INSTRUCTION IN HIGHER EDUCATION

Despite positive results, impediments slow down the widespread adoption of modelling instruction in higher education. Modelling instruction is best implemented using a hands-on approach that fits well with studio-format classes where the separation between theory and experiments is removed, and it requires reduced content coverage in favour of deeper understanding and skills. Furthermore, there is a lack of university-level resources and most textbooks, with rare exceptions like Chabay and Sherwood's *Matter and Interaction* series (2015), are not fostering a model-centred approach by ignoring the role of models and often omitting any extensive use of multiple representations (Brewe, 2006, 2008).

It is, therefore, useful to consider how a partial and novice implementation of some aspects of modelling instruction would impact learning outcomes in the context of CEGEPs where no studies on that type of teaching have been undertaken. If significant learning gains or attitude shifts are detected despite an imperfect implementation, this could encourage early adoption which could be improved over time as expertise is built. On the other hand, if no significant gains are detected, this could inform the community of educators about the necessity of thorough formal training for a successful implementation.

5. RESEARCH QUESTIONS AND HYPOTHESES

This research was guided by two main quantitative questions, each of which are presented with its associated research hypothesis and the statistically null hypothesis.

Q_1: How does modelling instruction differ from regular instruction or interactive engagement in terms of learning outcomes for CEGEP Mechanics students?

> H_1: CEGEP Mechanics students receiving modelling instruction, compared to regular instruction or interactive engagement, will perform differently on pre/post standardized assessments (FCI, RRMCS) of deep conceptual understanding and on more traditional exams testing problem-solving skills.

> H_0: CEGEP Mechanics students receiving modelling instruction, compared to regular instruction or interactive engagement, will perform equivalently on pre/post standardized assessments (FCI, RRMCS) of deep conceptual understanding and on more traditional exams testing problem-solving skills.

Q_2: How does modelling instruction differ from regular instruction or interactive engagement in terms of attitudes (or beliefs) about physics for CEGEP Mechanics students?

> H_2: CEGEP Mechanics students receiving modelling instruction, compared to regular instruction or interactive engagement, will perform differently on a pre/post standardized assessment (CLASS) of attitudes (or beliefs) about physics.

H_0: CEGEP Mechanics students receiving modelling instruction, compared to regular instruction or interactive engagement, will perform equivalently on a pre/post standardized assessment (CLASS) of attitudes (or beliefs) about physics.

A third qualitative question complemented the main quantitative ones. Its purpose was to gain insight into the appreciation and motivation of students experiencing modelling instruction, for it is believed that a positive perception enhance learning.

Q_3: How are CEGEP Mechanics students perceiving (in terms of what they like or don't like) the introduction of modelling instruction?

It is expected that students receiving modelling instruction will overall prefer this form of active learning.

A comparative study of student learning outcomes and attitudes about physics from different instructional formats has been performed to address the first two questions. The mode of instruction (modelling instruction vs. interactive engagement or regular instruction) was the independent variable (IV) whereas the learning outcomes and attitudes or beliefs about physics were the dependent variables (DV1 and DV2). A first learning outcome (DV1.1) was understood and operationalized as students' conceptual understanding as measured by pre/post standardized test score gains in the Mechanics (Physics NYA) course (expressed in different forms later described). Another learning outcome (DV1.2) was the students' procedural mastery of the mathematics related to the concepts that were assessed in a more traditional final exam with textbook-like problems to solve. Attitudes about physics (DV2) was understood and operationalized as students' self-reported beliefs about physics and its learning on pre/post standardized tests validated for this purpose (and again expressed in different forms later described). A qualitative survey, on the other hand, inquired about students' perceptions of the instructional design and thus addressed the third question.

FOURTH CHAPTER. METHODOLOGY

This research used a mixed-methods design with a strong quantitative strand based on a quasi-experiment using pre- and post-test comparisons to answer the first and second research questions about learning outcomes and attitude shifts. The third research question was addressed by a survey with open-ended questions to gain qualitative feedback from students receiving modelling instruction. That second qualitative phase aimed at placing the results of the first quantitative phase in context and exploring participants' views in more depth. The whole design is described in further detail in this chapter.

1. TARGET POPULATION AND SAMPLES

To represent the anglophone college population of first-year science students attending the researcher's college in Montreal, Quebec, non-random, intact group samples of Mechanics students (first year, first semester) were used because they were convenient. Furthermore, it would have been unrealistic to perform a random sampling that would have necessitated to move students around, changing their scheduling, their teacher, and their classmates. Those samples were classes of about 35 students, some of which have been combined for statistical analysis as they received similar instruction. They were typically aged about 17-18 years old. The proportion of males or females was quite variable from one class to another. Females accounted from about 40% up to 72% in individual classes, but merging classes by modes of instruction brought the proportion between 45% and 60%. Correspondingly, the proportion of males would be between 55% and 40%.

The researcher’s college currently receives about 8700 students per year, with slightly less than 25% in Continuing Education. According to its strategic plan (Vanier College, 2015), it is the most multicultural anglophone CEGEP with 94 different countries of origin and one fourth of students born outside Quebec. The student population is mostly female (about 55% in Fall 2013). Most students don’t take a full charge in courses and delay graduation (71% in pre-university studies, 55% in career programs, in Winter 2013).

Students taking Mechanics (Physics NYA) are typically spread between three programs: 200.B1 Health Science, 200.B2 Pure and Applied Science, and 200.C0 Computer Science and Mathematics. If they followed the regular path of studies in Quebec, it is their second all-physics course, their first one being Secondary V Physics taken normally the year before in high school.

To improve external validity, a statistical analysis of the pre-FCI and pre-CLASS data was performed to gain information about the initial physics knowledge and attitude differences between groups.

2. METHOD

Sections of Mechanics studied in this quasi-experiment research were selected based on the willingness of teachers to participate. Two of the researcher’s day sections and ten from colleagues teaching the same course in the same semester were studied.

The treatment group was made of the researcher’s sections to better control conditions of learning/teaching. Although this non-random, intact group sample also introduced possible biases, it made things easier and avoided asking another teacher to become familiar with a teaching method he or she might not know or be reluctant to try. These two sections (35 students) experienced the new instructional design with merged theory and laboratory practice for the whole course.

There were two different kinds of control groups. Three other teachers' classes (Reg1, Reg2, Reg3) received interactive expository instruction with regular labs, closer to, yet not exactly traditional instruction, hence the "regular" qualification. The two remaining classes (Act4, Act5) were highly involved in a different form of interactive engagement or active learning. This classification was based on a qualitative questionnaire sent to all participating teachers (reproduced in appendix H) and further clarified by each with an estimated percentage of active learning based on the following question: "How would you approximately split the classes, in %, between time where students were active driving the discussion and inquiring about the physics compared to the time when you as the teacher were driving the discussion and leading or organizing the learning of the topic? For labs, how would you split the time between investigating, exploratory or instructionless/ill-defined labs versus confirmation labs with well-defined instructions to follow?"

Each teacher was pretty much independent in their way to teach Mechanics, except for a common final exam where 78% of the questions were shared (the remaining 22% was to the choice of the teacher). Each section met three times a week: twice with another section for two hours of theory each time (hence a total of four hours, including an optional extra hour to support student success), and once alone for two hours of lab work (or anything else planned by the teacher, like tutorials).

To verify pre-existing knowledge of mechanics concepts and initial attitudes about physics, groups received standardized pre-tests – FCI, RRMCS, CLASS, to be answered on Scantron OpScan sheets (instructions given in appendix E) – on the third week of the Fall 2018 semester (except for Act4 and Act5, who administered the FCI pre-test at the start), after clearance from the Research Ethics Board (REB). It was presumed that didn't affect much the FCI results of the modelling group (Mod) as the researcher had spent the first two weeks on basic lab skills and started kinematics after the pre-tests. The delay may have slightly affected the FCI pre-test scores of Reg1, Reg2 and Reg3, though. The impact on CLASS pre-scores is unknown. At the end of the course, all groups passed the same standardized tests again as post-tests, a few days before the final exam. The RRMCS was given only in the Mod group. The qualitative survey was distributed online to the treatment group on December 4th, two weeks before the end of classes, to obtain feedback on the new instructional design. Responses were received till December 18th, 2018.

The new instructional design was inspired by modelling instruction and materials available from both the American Modeling Teachers Association (AMTA)[1] and the Physics Education Research Group at Florida International University (FIU)[2], the latter being the main source of both curriculum materials and activity plans suggesting how to use them. The design typically consisted of a pre-lab discussion, a lab investigation and a post-lab discussion, followed by worksheet

---

[1] Website: https://www.modellinginstruction.org/

[2] Website: http://univ-modellinginstruction.com/

exercises and sometimes a lab practicum.[3] The course outline can be consulted in appendix I. There were no quizzes nor unit tests to assess learning through each of the modelling cycles, but there were three exams to test kinematics, dynamics, energy and momentum during the semester. Rotational motion was tested on the 78%-common final exam exclusively. The researcher didn't have any formal training nor experience in modelling instruction, so the instructional design should be considered novice modelling instruction at best.

After Scantron OpScan sheets were processed, raw data were exported to Microsoft Excel, then coded and anonymized by the research supervisor. The FCI, RRMCS and CLASS pre- and post-test data were then graded and given scores converted to percentages. If pre- and post-test scores were present, the shift or gain was calculated by subtracting the pre-test score from the post-test score. Only matching results from consenting students were kept.

Considering this group comparison had only one independent categorical variable (mode of instruction), with three levels (modelling instruction, interactive engagement, regular instruction), and three main continuous dependent variables (FCI-based conceptual learning outcomes, final exam-based procedural learning outcomes, CLASS-based attitudes about physics), a multivariate analysis of variance (MANOVA) was performed, looking for *p*-values below .05 for statistical significance. The MANOVA was followed by various ANOVAs to locate significant differences with more precision. After comparing our results with our local control groups, they

---

[3] The specific materials (guiding slides, course pack with lab activities and worksheets) that were used in the modelling instruction group can be requested by contacting the researcher.

were also compared with the literature when published data were available. Extensive research has been done on modelling instruction and interactive engagement versus traditional instruction based on the FCI test. A lot of research has been done on the CLASS test as well. We checked how our results compared to that. Table 2 summarizes the essential elements of the study's timeline.

Table 2. Timeline of the study.

Steps in the methodology	**Dates**	**Goals**
Start of classes	August 22nd, 2018	
Clearance from the Research Ethics Board (REB).	September 10th, 2018	Validating all ethical aspects of the research.
FCI, RRMCS, CLASS pre-tests; consent forms	September 10th to 14th, 2018	Verifying pre-existing knowledge of mechanics concepts and initial attitudes about physics.
FCI, RRMCS, CLASS post-tests	December 3rd to 10th, 2018	Verifying after-instruction knowledge of mechanics concepts and attitudes about physics.
Qualitative survey to modelling students	December 4th to 18th, 2018	Getting insight into the perception of students about modelling instruction.
End of classes	December 10th, 2018	
78%-common final exam	December 12th, 2018	Assessing problem-solving skills.
Survey of participating teachers about their mode of instruction	Winter 2019	Classifying groups according to the main mode of instruction in order to operate proper statistical analyses and answer research questions.
Analysis of data	Winter 2019 to Winter 2020	Answering research questions.

3. INSTRUMENTS

Many dozens of Research-Based Assessment Instruments (RBAIs) have been produced by the physics and astronomy education community based on research into the way students think (Madsen et al., 2017b). The FCI, RRMCS and CLASS (appendix F) were the three that were used in this research. The other two tools for data collection that were used were the common final exam and an online qualitative survey (appendix G). These instruments are described in more detail below.

3.1 Answering the first research question

The FCI, RRMCS and common final exam were used to answer the first research question about how modelling instruction differ from regular instruction or interactive engagement in terms of learning outcomes for CEGEP Mechanics students. Their nature, the type of data generated, and the strategy to analyze it are described here.

3.1.1 Force Concept Inventory (FCI)

The main pre/post-test used to answer the first research question was the last revision (1995) of the Force Concept Inventory (FCI) by Halloun, Hake, Mosca, and Hestenes. This standardized test assesses students' understanding of the basics of Newtonian physics (one-dimensional kinematics, two-dimensional motion with constant acceleration [parabolic motion], Newton’s laws, impulsive forces, vector sums, cancellation of forces, and identification of forces) and provides lots of comparison opportunities from the literature. It is comprised of thirty multiple-choice, conceptual questions asking students to choose between the right Newtonian answer and incorrect common-sense alternatives based on those most often given by students in interviews (see sample questions in appendix F). Students are given thirty minutes to write this test; no marks are deducted for incorrect answers. Brewe (2002) reports a Kuder-Richardson Formula 20 (KR-20) reliability value of .90 for this instrument, consistent with what was found by Lasry, Rosenfield, Dedic, Dahan and Reshef (2011). As previously mentioned, Henderson (2002) found that few students don’t take the FCI seriously when it is ungraded and that post-test results are not affected in any statistically significant manner by the FCI being given as a pre-test.

3.1.2 Rotational and Rolling Motion Conceptual Survey (RRMCS)

A limitation of the FCI is its lack of assessment for rotational kinematics and dynamics, which is a significant part of the introductory Mechanics course in CEGEPs. To alleviate this weakness and be able to compare the conceptual learning outcomes on this topic as well, the Rotational and Rolling Motion Conceptual Survey (RRMCS) was set to be given as a pre- and post-test, along with the FCI. The RRMCS was developed by Rimoldini and Singh (2005) and focuses on rotational kinematics, rotational kinetic energy, moment of inertia, torque, rolling motion, and sliding versus tumbling. They report a reliability index α ranging from .68 to .82. Just like the FCI, this standardized test is comprised of thirty multiple-choice, conceptual questions asking students to choose between the right Newtonian answer and incorrect common-sense alternatives. Contrary to the FCI, though, students are also asked to explain their choices. They are given fifty minutes to write this test; no marks are deducted for incorrect answers. Because of time limitations due to other tests to be conducted in class, students from the treatment group were asked to skip the explanation part of the test and only thirty minutes were allowed to complete it.

Rimoldini and Singh (2005) performed a study with 3000 introductory physics students at the University of Pittsburgh, but the final version of the RRMCS was only given, totally or partially, to 652 of them. They report average test scores, but they make no comparison based on instructional designs, nor do they specify the type of instruction employed (although Dr. Singh told me in a private communication that it was traditional). The present research appears to be the first one to attempt at using the RRMCS in a comparative study and to perform the same statistical analyses as for the FCI. No other teacher besides the researcher accepted to give it, though, so its usefulness is strongly limited. Results will be reported for completion, but little inference will be made.

3.1.3 Final exams

The FCI and RRMCS were complemented by an analysis of final exams composed of problems to solve numerically, akin to those found in the textbook. It has been chosen not to include numerical questions testing problem-solving in the pre-test. Doing so would have helped to treat data uniformly, but the conceptual pre-test is believed to be sufficient to compare groups through descriptive statistics, without getting into the more complex and time-consuming job of duplicating a midterm assessment on a pre-test. Furthermore, that would have been a heavy demand for students to get through such an exam at the start of their first college physics course.

3.1.4 Type of data generated and strategy for analyzing it

With the FCI and the RRMCS, normalized learning gains (in percentages) were computed for everyone based on Hake's (1998) formula, by taking the post-test score minus the pre-test score and dividing it by what could be the maximum possible gain, i.e., the difference between 100% and the pre-test score. Mathematically, the calculation operates as follows:

$$\text{Normalized gain } g \text{ (\%)} = \frac{\text{actual gain } G}{\text{maximum possible gain } G_{max}} = \frac{\text{post-test score} - \text{pre-test score}}{100\% - \text{pre-test score}}$$

For example, if the pre-test score is 20% and the post-test score is 60%, then the actual gain is 40%, the maximum potential gain is 80%, and the normalized learning gain is:

$$\text{Normalized gain } g \text{ (\%)} = \frac{60\% - 20\%}{100\% - 20\%} = \frac{40}{80} = 50\%$$

Normalized learning gains of averages for both treatment and control groups were also computed and compared:

$$\text{Normalized gain of averages } \langle g \rangle \text{ (\%)} = \frac{\langle G \rangle}{\langle G_{max} \rangle} = \frac{\langle \text{post-test score} \rangle - \langle \text{pre-test score} \rangle}{100\% - \langle \text{pre-test score} \rangle}$$

These normalized gains of averages are calculated with the gain of averages and therefore use the average post-test and pre-test scores for matched data, the difference of which (the actual average gain $\langle G \rangle$) is then normalized by dividing it by the maximum possible average gain $\langle G_{max} \rangle$. This is the standard and common way to do this calculation, although many studies rather calculate the average of the individual normalized gains, which can also be argued for. This latter calculation was made to perform statistics while it was compared with the former to check for consistency. Mathematically, it is represented as so:

$$\text{Average of normalized gains } g_{avg} \text{ (\%)} = \langle \frac{G}{G_{max}} \rangle = \langle \frac{\text{post-test score} - \text{pre-test score}}{100\% - \text{pre-test score}} \rangle$$

Hake defines high, medium, and low values of g as $g \geq 0.7, 0.7 > g \geq 0.3,$ and $g < 0.3$, respectively (Hake, 1998).

Although it is most common for physics faculty to compare normalized gains (Madsen et al., 2017a), it might be interesting to also analyze the average raw gain $\langle G \rangle$ and the effect size. The former is less often used, but some studies prefer this measure. Its disadvantage is that comparisons become dependent on pre-test scores (Marx & Cummings, 2007), hence it is not always very useful to use it unless groups are found comparable based on that pre-test score. The latter is more common in social sciences research, but yet provides additional insight. In particular, the effect size accounts for the spread in students' scores and allows comparing classes of different sizes more fairly (Madsen et al., 2017a), without normalizing gain scores on the same scale (Nissen et al., 2018). Furthermore, Nissen, Talbot, Thompson and Van Dusen (2018) state that in contrast to Hake (1998), Coletta and Phillips (2005) found that the normalized gain of averages was correlated with pre-test means. Their comparisons tended to show that normalized gains were biased in favour of high pre-test scores, leading to a recommendation to use Cohen's *d* as an alternative measure. The study of effect sizes though Cohen's *d* may, therefore, be more robust. Mathematically, Cohen's effect size *d* is evaluated the following way:

$$\text{Effect size } d = \frac{\langle \text{post-test score} \rangle - \langle \text{pre-test score} \rangle}{\text{Pooled standard deviation}}$$

$$\text{Pooled std dev. } s_{pooled} = \sqrt{\frac{\left(\left(N_{pre} - 1\right) s_{pre}^2 + \left(N_{post} - 1\right) s_{post}^2\right)}{N_{pre} + N_{post} - 2}} = \sqrt{\frac{\left(s_{pre}^2 + s_{post}^2\right)}{2}}$$

where s represents the standard deviation. Here, $N_{pre} = N_{post}$ because we use matched data. To account for the dependence between pre- and post-tests, a correction can be made (Nissen et al., 2018):

$$\text{Corrected effect size } d_{dep} = \frac{\langle\text{post-test score}\rangle - \langle\text{pre-test score}\rangle}{\sqrt{(s_{pre}^2 + s_{post}^2 - 2rs_{pre}s_{post})/2(1-r)}}$$

where r represents the correlation between the pre-tests and post-tests and serves to deattenuate the effect size. Small, medium and large differences are associated with Cohen's d ~0.2, ~0.5, and ~0.8 respectively (Cohen, 1988).

On the other hand, Marx and Cummings (2007) suggest analyzing the normalized change:

$$\text{Normalized change } c = \begin{cases} \dfrac{\text{post} - \text{pre}}{100\% - \text{pre}} \text{ if post} > \text{pre} \\ \text{drop if post} = \text{pre} = 100 \text{ or } 0 \\ 0 \text{ if post} = \text{pre} \\ \dfrac{\text{post} - \text{pre}}{\text{pre}} \text{ if post} < \text{pre} \end{cases}$$

The rationale behind this proposition is that the normalized gain, which has a low-test score bias, produces a non-symmetric range of scores and doesn't permit the calculation of the average of individual normalized gains when a student has a perfect pre-test score (rather forcing the use of the normalized gain of the groups' averages). The normalized change addresses those limitations. When it comes time to compare average normalized changes, they argue that the average of changes c_{avg} is better than the change of averages $\langle c\rangle$ because it captures the spread more accurately, making the report of the estimated uncertainties calculated from individual scores consistent with the average normalized changes. Furthermore, it removes the ambiguity of whether the data were matched or not.

In all cases, the data were carefully interpreted by keeping in mind that raw scores are typically ordinal (Wright & Linacre, 1989), meaning that our measures should be considered ordinal as well.

With the final exam, raw scores and raw group averages were statistically compared to extract effects on procedural learning outcomes seen as problem-solving skills.

3.2 Answering the second research question

The CLASS was used to answer the second research question about how modelling instruction differ from regular instruction or interactive engagement in terms of attitudes (or beliefs) about physics for CEGEP Mechanics students. Its nature, the type of data generated, and the strategy to analyze it are described here.

3.2.1 Colorado Learning Attitudes about Science Survey (CLASS)

To answer the second research question, the pre/post-test scheme used the popular Colorado Learning Attitudes about Science Survey (CLASS) standardized test, which was developed by Adams, Perkins, Podolefsky, Dubson, Finkelstein, and Wieman (2006). The latest (third) version for physics, released in 2004, was used. This instrument measures students' self-reported beliefs about physics and its learning. The analysis then permits to evaluate how closely these beliefs align with those of experts. The CLASS test is comprised of forty-two questions (see appendix F for a sample) that are answered on a five-point ordinal Likert scale: strongly disagree (1), disagree (2), neutral (3), agree (4), strongly agree (5). Twenty-seven questions fall into one of the following eight empirically determined categories: real-world connections; personal interest; sense-making or effort; conceptual connections; applied conceptual understanding; problem-solving general; problem-solving confidence; and problem-solving sophistication. Nine more are non-categorized, but part of the overall scoring. Finally, another six are excluded from scoring for lack of "expert" response or usefulness in their current form. In particular, the thirty-first statement serves to identify the majority of students who randomly choose answers, and to discard their surveys from the analysis. Students are allowed a total of ten minutes to answer the test. The developers assessed the test-retest reliability (through correlations of answers) to .98-.99 for beliefs in agreement or disagreement with those of experts, and .88 for neutral responses (Adams et al., 2006).

3.2.2 Type of data generated and strategy for analyzing it

With the CLASS, percentages of responses in agreement (percent favourable or percent expert-like response) and in disagreement (percent unfavourable or percent novice-like response) with the experts' view were calculated and averaged. Then, the individual shifts and the average group shifts in attitudes from pre- to post-test for matched data were computed. Mathematically, those calculations are made in the following way:

$$\text{Shift } S\,(\%) = \text{pre-test \% favourable} - \text{post-test \% favourable}$$

$$\text{Average shift } \langle S \rangle\ (\%) = \langle \text{pre-test \% favourable} \rangle - \langle \text{post-test \% favourable} \rangle$$

For example, if the percent favourable scores of a student on the pre-test and post-test are 40% and 60% respectively, then his or her actual shift in attitudes is 20%.

The analysis was again completed by a calculation of Cohen's effect size *d*:

$$\text{Effect size } d = \frac{\langle \text{post-test score} \rangle - \langle \text{pre-test score} \rangle}{\text{Pooled standard deviation}}$$

$$\text{Pooled std dev. } s_{pooled} = \sqrt{\frac{\left(\left(N_{pre} - 1\right) s_{pre}^2 + \left(N_{post} - 1\right) s_{post}^2\right)}{N_{pre} + N_{post} - 2}} = \sqrt{\frac{\left(s_{pre}^2 + s_{post}^2\right)}{2}}$$

where s represents the standard deviation. Here, $N_{pre} = N_{post}$ because we use matched data. To account for the dependence between pre- and post-tests, a correction can be made (Nissen et al., 2018):

$$\text{Corrected effect size } d_{dep} = \frac{\langle \text{post-test score} \rangle - \langle \text{pre-test score} \rangle}{\sqrt{(s_{pre}^2 + s_{post}^2 - 2r s_{pre} s_{post})/2(1 - r)}}$$

where r represents the correlation between the pre-tests and post-tests and serves to deattenuate the effect size. Small, medium and large differences are associated with Cohen's *d* ~0.2, ~0.5, and ~0.8 respectively (Cohen, 1988).

When scoring the CLASS surveys, neutrals are considered neither in agreement nor in disagreement with experts because there are so many reasons that could influence a student's choice for this answer. Agreement and strong agreement are collapsed together, just as disagreement and strong disagreement are. The reason is that students' interpretations of nuances between pairs of terms are not consistent, while at the same time it is important to offer a 5-point rather than a 3-point Likert scale to avoid increasing the number of neutrals (Adams et al., 2006). Unanswered statements are discarded from the calculations, as long as there is not a lot of them, which could justify discarding the survey. Once again, the data were interpreted carefully by keeping in mind that all measures are ordinal.

3.3 Answering the third research question

An online qualitative survey was used to answer the third research question about how CEGEP Mechanics students are perceiving (in terms of what they like or don't like) the introduction of modelling instruction. Its nature, the type of data generated, and the strategy to analyze it are described here.

3.3.1 Qualitative survey

The qualitative survey to answer the third research question was comprised of a few open-ended questions. Likert scales could have been used to make it more quantitative, but this could also have directed comments rather than foster spontaneous reactions from students. The questionnaire was asking about their views on the new instructional design.

3.3.2 Type of data generated and strategy for analyzing it

The qualitative appreciation of the new instructional design was evaluated through a qualitative analysis of comments received. Although not many responses were received, major themes were extracted, and some ideas for successful implementation on a larger scale were gained.

4. ETHICAL CONSIDERATIONS

Some potential ethical issues existed relative to participants, sponsors of the research, and the community of educational researchers, but each of those was addressed.

It is believed that an ethic of respect and freedom is essential to the conduct of any research. The first aspect of that is the necessity of voluntary informed consent. Participation needs to be free and consensual. No student was part of the study without proper consent obtained by a neutral person in the absence of the teacher. The researcher took that role in classes belonging to the regular instruction and interactive engagement groups, whereas a pedagogical counsellor from the college did it for the modelling instruction group that the researcher was teaching. The consent wasn’t revealed to the teacher. Furthermore, no compensation or reward was offered for participation in the research, thus avoiding a potential influence on participants’ decisions.

Consent forms from students were distributed and collected in the third week of the semester due to delays in getting approval from the Research Ethics Board. Technically, some students might have been 17 years old, but it is customary for the administration to consider college students independent from their parents when they enter CEGEP. This research abided by the same standards and the consent from parents was not required.

Proper consent also means that participants understand why their participation is important, how the information will be used and whom it will be reported to. This is closely related to the second aspect of an ethic of respect, namely openness and disclosure. Clear information about the process of the research was thus provided by the researcher or a neutral person before consent. Participants were informed of the research and its goals transparently.

The third aspect is the right to withdraw for any or no reason. Participants were informed of this right and of the process to follow should this happen.

Although whole groups received the treatment, students who didn't (or withdrew) consent didn't have their data compiled and analyzed for the study. Students who didn't want to participate in the treatment (modelling instruction) also had a period at the start of the semester to change sections, according to the official calendar set by the administration. The researcher didn't know who gave consent or not before the end of the semester after final grades had been submitted and the data anonymized.

The fourth aspect is privacy and confidentiality. To address that, the data were anonymous by having the supervisor coding it before any analysis occurred. The storage of data and results needs to be carefully thought through. Data were stored on the researcher's personal computer, which is password-protected and not accessible to unauthorized people. Raw data will be destroyed no later than seven years after the end of the research. That should be sufficient protection of the kind of data that was collected by the researcher.

Finally, the fifth aspect is the disclosure of results. Interested participants were, therefore, offered a summary of the final report without any reference to the standardized tests that were used.

This research was not directly funded by anyone, but it was realized on the premises of the researcher's college. An application was submitted to their Research Ethics Board (RBE) for permission to carry out the project. Copies of the research certification received, of the research recruiting script that has guided the presentation of the project to students, and of the consent form participants had to fill out are included in appendices B to D.

This work was carried out ethically, according to best practices. As such, the community of educational researchers was well represented. To fulfill other responsibilities toward this community, the results of this research will be communicated in conferences and journals after publication, as a way to contribute to the construction of a body of knowledge about instructional methods that improve the learning of students in physics and possibly in other disciplines.

It is believed that this research will benefit physics students by allowing better instruction to occur and that risks associated with it were naught. Indeed, the same instruction would have occurred even if no research had been conducted, except that with this research that novel instructional design was formally evaluated. On the contrary, participants may directly benefit if they take a future physics course where the new instructional design has been adopted (or not) because of results found. More immediately, the new instructional design may have shown them that there are different ways of approaching physics. Furthermore, some students may have gained a more realistic understanding of the scientific process.

FIFTH CHAPTER. COLLECTED DATA AND RESULTS

Having collected raw data from consenting participants, descriptive statistics were performed. We could then proceed to the analysis and interpretation of the data as we sought answers to the research questions. The main results of those statistics for each teaching method category – modelling instruction (Mod), interactive engagement (Act) and regular instruction (Reg) – are presented below, in five parts, the last three separately presenting the collected data and results related to a research question in particular. The first section presents data about the sample and seeks ground for statistics calculated afterwards. The second section provides a general assessment of potential group differences. More detailed analyses supporting results presented herein can be found in appendix A.

1. DATA ABOUT THE SAMPLE

As described in the methodology chapter, the research used a treatment group (Mod: modelling instruction) of 20 consenting students and two control groups (Act: interactive engagement; Reg: regular instruction) of respectively 61 and 83 consenting students. Six classes corresponding to twelve sections (two sections per class, one section per lab) of Mechanics taught by six physics teachers at the college participated in the research. Table 3 summarizes the distribution of sections in the three groups and details the number of consenting students along with participation rates for the various tests that were given. This information was calculated by comparing the number of responses to the total number of students in each group.

Table 3. Research design and participation rates.

<table>
<tr><th>Teacher</th><th>Researcher</th><th>Teacher 4</th><th>Teacher 5</th><th>Teacher 1</th><th>Teacher 2</th><th>Teacher 3</th></tr>
<tr><td rowspan="2">Code [a]</td><td rowspan="2">Mod</td><td>Act4</td><td>Act5</td><td>Reg1</td><td>Reg2</td><td>Reg3</td></tr>
<tr><td colspan="2">Act</td><td colspan="3">Reg</td></tr>
<tr><td>Treatment</td><td>Novice Modelling Instruction</td><td colspan="2">Interactive Engagement (highly interactive)</td><td colspan="3">Regular Instruction (somewhat interactive)</td></tr>
<tr><td>Est. % of active learning [b]</td><td>75%</td><td>70%</td><td>70%</td><td>40%</td><td>20%</td><td>25%</td></tr>
<tr><td rowspan="2">Class sections</td><td rowspan="2">2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td></tr>
<tr><td colspan="2">4</td><td colspan="3">6</td></tr>
<tr><td rowspan="2">Total number of students N</td><td rowspan="2">35</td><td>36</td><td>35</td><td>35</td><td>36</td><td>32</td></tr>
<tr><td colspan="2">71</td><td colspan="3">103</td></tr>
<tr><td rowspan="2">Number of consenting students N_{cons}</td><td rowspan="2">20 (57%)</td><td>31 (86%)</td><td>30 (86%)</td><td>25 (71%)</td><td>32 (88%)</td><td>26 (81%)</td></tr>
<tr><td colspan="2">61 (86%)</td><td colspan="3">83 (81%)</td></tr>
<tr><td>FCI pre/post-tests</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td></tr>
<tr><td rowspan="2">$N_{FCI}^{matched}$</td><td rowspan="2">20 (57%)</td><td>27 (75%)</td><td>24 (69%)</td><td>6 (17%)</td><td>31 (86%)</td><td>24 (75%)</td></tr>
<tr><td colspan="2">51 (72%)</td><td colspan="3">61 (59%)</td></tr>
<tr><td>RRMCS pre/post-tests</td><td>Yes</td><td>No</td><td>No</td><td>No</td><td>No</td><td>No</td></tr>
<tr><td>$N_{RRMCS}^{matched}$</td><td>20 (57%)</td><td>*</td><td>*</td><td>*</td><td>*</td><td>*</td></tr>
<tr><td>78%-common final exam</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td></tr>
<tr><td rowspan="2">N_{final}</td><td rowspan="2">20 (57%)</td><td>30 (83%)</td><td>26 (74%)</td><td>21 (60%)</td><td>32 (89%)</td><td>26 (81%)</td></tr>
<tr><td colspan="2">56 (79%)</td><td colspan="3">79 (77%)</td></tr>
<tr><td>CLASS pre/post-tests</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td></tr>
<tr><td rowspan="2">$N_{CLASS}^{matched}$</td><td rowspan="2">19 (54%)</td><td>21 (58%)</td><td>11 (31%)</td><td>8 (23%)</td><td>29 (81%)</td><td>21 (66%)</td></tr>
<tr><td colspan="2">32 (45%)</td><td colspan="3">58 (56%)</td></tr>
<tr><td>Qualitative survey</td><td>Yes</td><td>No</td><td>No</td><td>No</td><td>No</td><td>No</td></tr>
<tr><td>N_{survey}</td><td>8 (23%)</td><td>*</td><td>*</td><td>*</td><td>*</td><td>*</td></tr>
</table>

[a] Mod: MODelling; Act: InterACTive engagement; Reg: REGular instruction

[b] Based on a rough estimate by the teacher.

Only data from consenting students were kept. That led to a proportion of available data varying from 57% in the modelling instruction group to 86% in the interactive engagement group. The data were further refined by filtering out tests that were less than 80% complete (less than 24 answered questions out of 30 for the FCI and RRMCS, or 36 out of 45 for the CLASS) or that didn't answer the filter question correctly on the CLASS test (there to ensure students were paying attention to the test). We further rejected non-matched sets of scores on the FCI, RRMCS, and CLASS tests to perform statistical analyses of gains, changes, and shifts. In the end, the net response rates varied between 57% and 72% on the FCI, was 57% on the RRMCS, and varied between 45% and 56% on the CLASS. For the final exam, keeping grades of consenting students led to a net response rate varying between 57% and 79%. For the end-of-semester qualitative survey sent to modelling students, the response rate was only 23%. It is on that final and cleaned dataset that we performed further calculations of statistics, presented in the following sections, using the Minitab v.19 software.

It is worth noting that two classes (55 consenting students) received their post-tests under special conditions. Teacher Reg1 (25 consenting students) lacked time at the end of the semester and administered them during the last optional hour of class, which probably explains the very low percentage of matched data (17% for the FCI; 23% for the CLASS) despite a relatively high percentage of consenting students (71%). Relative to the number of consenting students, this corresponds to a drop rate of 76% and 68% for the FCI and CLASS post-tests, respectively. On the other hand, teacher Act5 (30 consenting students) administered the CLASS post-test online, for students to do at home, again because of a lack of time. That may explain the low percentage of matched CLASS data (31%) compared to the FCI (69%). Relative to the number of consenting students, this corresponds to a drop rate of 63% for the CLASS post-test.

2. GENERAL ASSESSMENT OF POTENTIAL GROUP DIFFERENCES

To detect if there was any statistically significant difference between groups, we simultaneously tested the equality of means from favourable and unfavourable attitude shifts S, normalized gains g and final exam grades. To achieve this, we performed a general multivariate analysis of variance (MANOVA) with $\alpha = .05$. The main results are reproduced in table 4.

Table 4. MANOVA tests for groups (modes of instruction).

Criterion	**Test Statistic**	***F***	**DF** **Num**	**Denom**	***P***
Wilks'	0.83035	2.411	8	198	0.0167
Lawley-Hotelling	0.19801	2.426	8	196	0.0161
Pillai's	0.17487	2.395	8	200	0.0173
Roy's	0.15824				

s = 2 m = 0.5 n = 48.5

Wilks' test, just like Lawley-Hotelling's and Pillai's, produce a p-value below $\alpha = .05$, meaning that there is a statistically significant difference between groups. Although we can reject the null hypothesis, that doesn't tell us where that difference is located. A quick overview of ANOVA p-values evaluated through the MANOVA pointed in the direction of the FCI test (the only one with $p < .05$, evaluated at 0.00169), but because there were differing missing values for different dependant variables, we decided to perform separate analyses of variance to investigate relations between groups (novice modelling instruction, active learning, regular instruction) and dependant variables more precisely. Nevertheless, we were able to predict that differences would most probably be detected only on the FCI assessment of learning outcomes, which required an ANOVA to localize between which groups that difference can be observed.

3. FIRST RESEARCH QUESTION: LEARNING OUTCOMES

Our first research question and corresponding hypotheses were the following:

Q_1: How does modelling instruction differ from regular instruction or interactive engagement in terms of learning outcomes for CEGEP Mechanics students?

H_1: CEGEP Mechanics students receiving modelling instruction, compared to regular instruction or interactive engagement, will perform differently on pre/post standardized assessments (FCI, RRMCS) of deep conceptual understanding and on more traditional exams testing problem-solving skills.

H_0: CEGEP Mechanics students receiving modelling instruction, compared to regular instruction or interactive engagement, will perform equivalently on pre/post standardized assessments (FCI, RRMCS) of deep conceptual understanding and on more traditional exams testing problem-solving skills.

To answer this question about learning outcomes, we collected responses to the FCI, RRMCS and 78%-common final exam. We then calculated various statistics like pre- and post-test scores, (normalized) gains and changes, and effect sizes, which all appear in appendix A. Hereafter, FCI raw score will first be examined, then FCI conceptual learning gains, RRMCS conceptual learning gains and final-exam grades will be discussed.

3.1 Conceptual understanding through FCI raw scores

When looking at FCI scores (table 5), it is noteworthy to observe that the average FCI score for the pre-test was below 60% in all groups and that for the post-test it was barely above 60% (modelling instruction and interactive engagement) or slightly below (regular instruction). Based on Hestenes's and Halloun's considerations (1995), this would mean that our students arriving from high schools – who for most completed Secondary V Physics (which covers translational kinematics, Newton's laws and forces, energy, and geometrical optics) a year before – are very deficient conceptually and ill-prepared for our calculus-based mechanics course which resembles the US-equivalent university physics, although many teachers limit the use of calculus that is often learned in parallel, thus giving a course that resembles more the US-equivalent college physics. Considering the initial state of students, perhaps it's not very surprising that FCI post-test scores are not very high, although it's somewhat saddening to observe scores that are close to 60%, meaning that students barely began to use Newtonian reasoning, if at all, toward the end of the course, even with interactive engagement methods.

It is shown in appendix A that differences in FCI pre-test scores were not significant, but we should perform the same kind of analysis, based on a one-way analysis of variance (ANOVA) with $\alpha = .05$, on the post-test scores. Results are in tables 5 and 6, and figures 5 and 6.

Table 5. Group means (%) on FCI post-test scores.

Group	*N*	Mean	StDev	95% CI
Act	51	64.25	19.19	(59.30, 69.20)
Mod	20	66.17	21.20	(58.27, 74.08)
Reg	61	53.55	15.40	(49.03, 58.08)

Pooled StDev = 17.8673

Table 6. Analysis of variance on FCI post-test scores.

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	*F*-Value	*P*-Value
Group	2	4197	9.25%	4197	2098.3	6.57	0.00191
Error	129	41182	90.75%	41182	319.2		
Total	131	45379	100.00%				

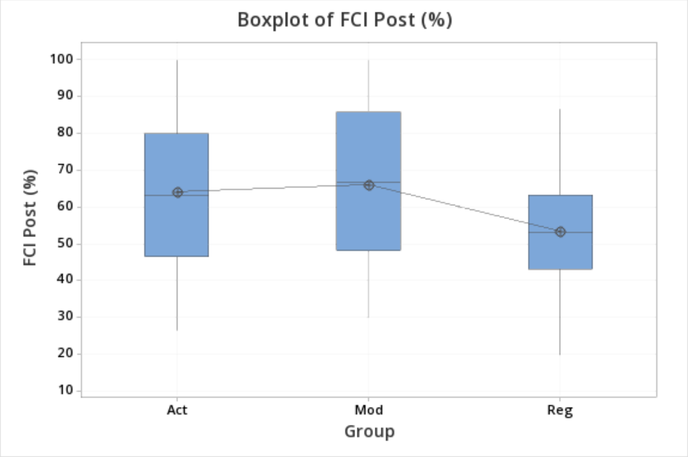
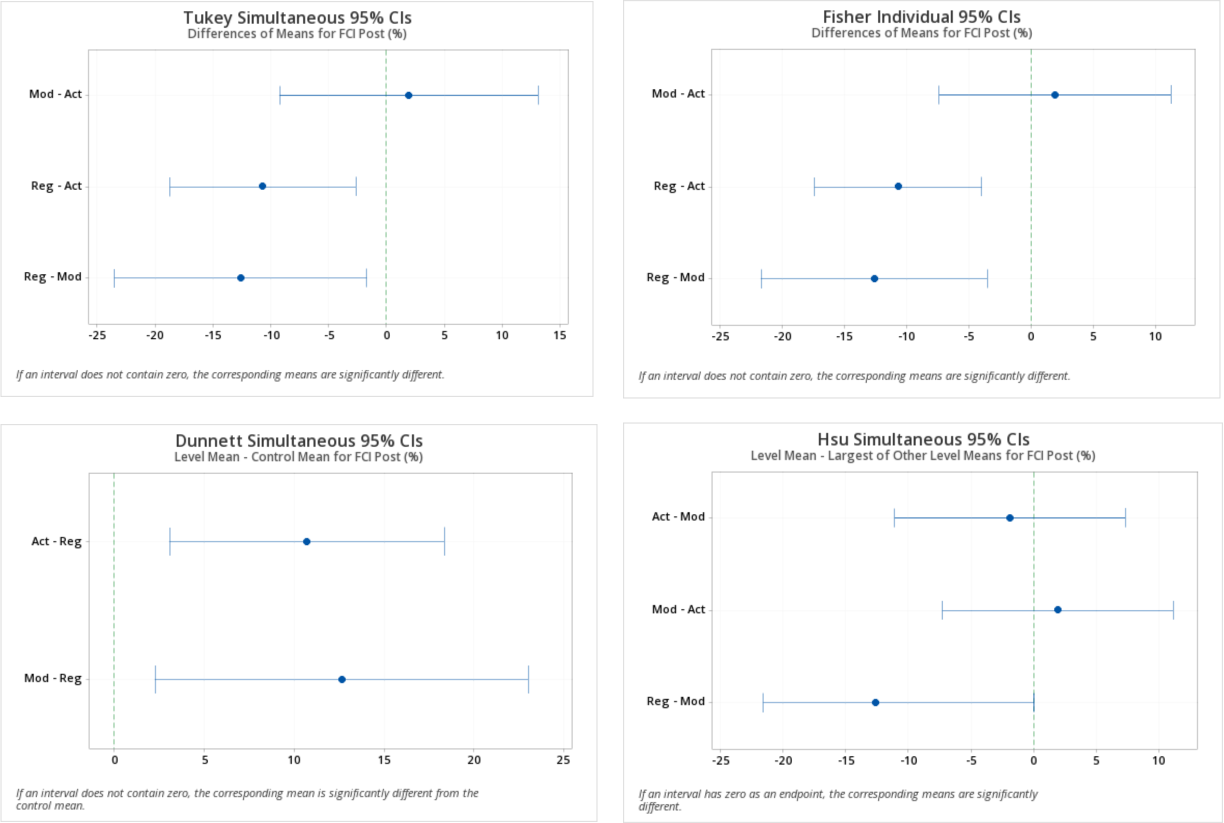


Figure 5. Boxplot of FCI post-test scores.


Tukey Simultaneous 95% CIs
Differences of Means for FCI Post (%)

If an interval does not contain zero, the corresponding means are significantly different.

Fisher Individual 95% CIs
Differences of Means for FCI Post (%)

If an interval does not contain zero, the corresponding means are significantly different.

Dunnett Simultaneous 95% CIs
Level Mean - Control Mean for FCI Post (%)

If an interval does not contain zero, the corresponding mean is significantly different from the control mean.

Hsu Simultaneous 95% CIs
Level Mean - Largest of Other Level Means for FCI Post (%)

If an interval has zero as an endpoint, the corresponding means are significantly different.

Figure 6. Tukey and Fisher pairwise comparisons, Dunnett multiple comparisons with the control (Reg), and Hsu multiple comparisons with the best (MCB) for FCI post-test scores.

We find a p-value of .002, lower than $\alpha = .05$, leading us to conclude that there is a statistically significant difference between groups. To locate this difference, we performed various tests. Looking at Tukey and Fisher pairwise comparisons, at Dunnet multiple comparisons with the Reg control, and at Hsu multiple comparisons with the best (Mod), all suggest a statistically (and practically) significant difference between the regular instruction group and both interactive engagement and modelling instruction groups. No significant difference is suggested between interactive engagement and modelling instruction groups, however.

FCI post-test scores are not the right variable to infer learning gains as they may have some dependence on FCI pre-test scores, which are not necessarily the same for all groups. Yet, the present results tell us that although all post-test scores were close to 60%, the post-test scores of the regular instruction group were qualitatively different. We indeed infer that students in the Mod and Act groups seemed to barely begin to use Newtonian reasoning at the end of the semester, whereas in the Reg group they seemed to be close to, yet not at this level. The ANOVA allowed us to see this difference as meaningful, and because there was no statistically significant difference in pre-test scores, this appears to be attributable to modes of instruction. That will be further investigated as FCI gains are analyzed.

3.2 Conceptual learning gains on the FCI test

In this section, descriptive statistics will be discussed before analyzing data through a one-way ANOVA and Cohen's *d* (effect size). Then data will be compared with the scientific literature.

3.2.1 Descriptive statistics

Let's first take a look back at some of the descriptive statistics of the FCI test to assess the consistency between various related measures (table 7).

Table 7. Descriptive statistics of FCI gains for different modes of instruction.

Variable	**Group**	***N***	***N******	**Mean**	**SE Mean**	**StDev**
FCI Average Gain <*G*> (%)	Act	51	10	24.90	2.20	15.71
	Mod	20	0	21.84	4.31	19.27
	Reg	61	22	19.07	1.74	13.55
FCI Normalized Gain of Averages <*g*> (%)	Act	51	10	41.06	*	*
	Mod	20	0	39.22	*	*
	Reg	61	22	29.11	*	*
FCI Average of Normalized Gains g_{avg} (%)	Act	51	10	43.39	3.39	24.21
	Mod	20	0	40.56	6.33	28.31
	Reg	61	22	29.13	2.38	18.59
FCI Normalized Change of Averages <*c*> (%)	Act	51	10	41.06	*	*
	Mod	20	0	39.22	*	*
	Reg	61	22	29.11	*	*
FCI Average of Normalized Changes c_{avg} (%)	Act	51	10	43.34	3.41	24.32
	Mod	20	0	40.84	6.20	27.73
	Reg	61	22	29.05	2.40	18.75

If we look at the average raw gain <*G*>, we observe that the highest average is achieved with interactive engagement, followed by modelling instruction, and then by regular instruction. The same order is obtained looking at all other measures, so everything seems consistent. It is however preferable not to work with the average raw gain for it is sensible to differences between groups as expressed through FCI pre-test scores, although a previous analysis (in appendix A) showed that we couldn't detect any statistically significant difference between the groups. The literature rather relies on the normalized gain of averages <*g*> or the average of normalized gains g_{avg}. It is the latter that we can assess through an ANOVA, so we privileged this measure although we calculated both. Some authors have discussed the limitations of normalized gains and proposed to rather use the normalized changes. We calculated them as well, but as can be seen, the differences between the two in our dataset is so minimal that we can safely assume them to be negligible. Therefore, it is sufficient to report only the statistical analysis of the average of normalized gains to evaluate if there is a statistically significant difference between groups.

3.2.2 Analysis of variance

A one-way ANOVA on normalized gains g was thus performed with $\alpha = .05$. Results are presented in tables 8 and 9, and figures 7 and 8.

Table 8. Group means (%) on FCI normalized gains.

Group	***N***	**Mean**	**StDev**	**95% CI**
Act	51	43.39	24.21	(36.59, 50.20)
Mod	20	40.56	28.31	(27.32, 53.81)
Reg	61	29.13	18.59	(24.37, 33.89)

Table 9. Analysis of variance on FCI normalized gains (Welch's test).

Source	**DF Num**	**DF Den**	***F*-Value**	***P*-Value**
Group	2	47.4596	6.35	0.00359

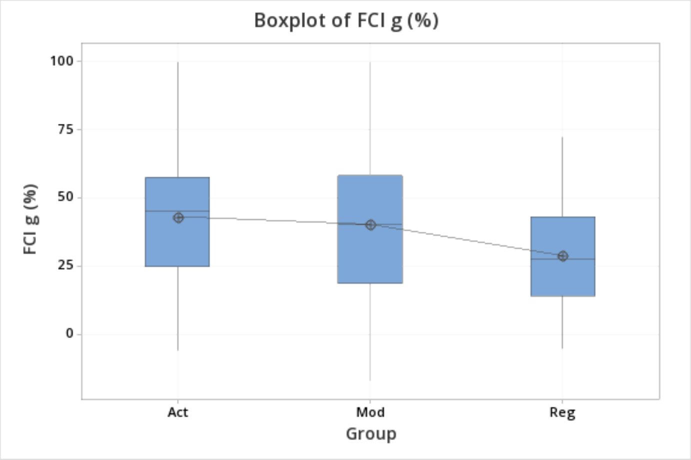


Figure 7. Boxplot of FCI normalized gains.

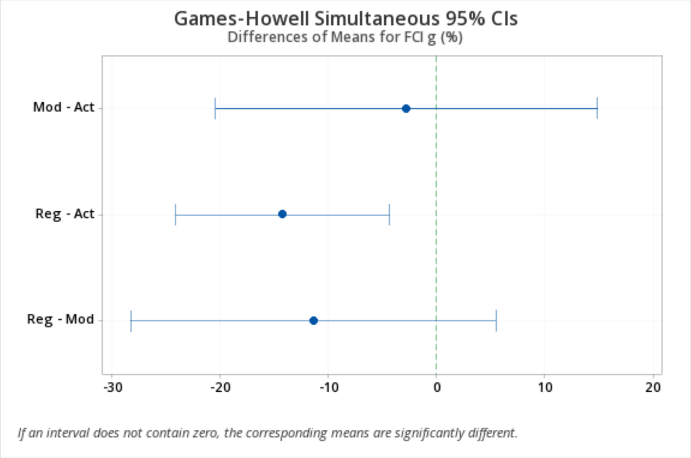


Figure 8. Games-Howell pairwise comparisons for FCI normalized gains.

We find a p-value of .004, lower than $\alpha = .05$, leading us to conclude that there is a statistically significant difference between groups. To locate this difference, we performed Games-Howell pairwise comparisons. This test suggests a statistically (and practically) significant difference between regular instruction and interactive engagement, but it didn't detect any difference between novice modelling instruction and either interactive engagement or regular instruction. The null hypothesis cannot be rejected for novice modelling instruction. However, we do see that the average of normalized gains is higher for interactive engagement (43.4%) compared to regular instruction (29.1%), and we are allowed to wonder, considering that modelling instruction is a special type of interactive engagement, if an expert implementation wouldn't lead to the same kind of results. The observation that the difference of means is smaller between novice modelling instruction and interactive engagement (2.9%) than it is between novice modelling instruction and regular instruction (11.4%) also supports such a hypothesis, which would be worth investigating further.

3.2.3 Effect size

If we leave the ANOVA for a moment and look into the effect size d_G based on the average gain and the pooled standard deviation of FCI pre- and post-test scores, and also into the corrected effect size $d_{G,\ dep}$ taking into account a possible dependence between the two scores modelled by Pearson's coefficient of correlation r, we find results of table 10.

Table 10. Effect size for the FCI gain, for different modes of instruction.

Group	***N***	***N****	**FCI Gain Effect Size d_G**	**FCI Gain Corrected Effect Size $d_{G,\ dep}$**
Act	51	10	1.27	1.26
Mod	20	0	1.08	1.08
Reg	61	22	1.29	1.28

As can be seen in table 10, both Cohen's effect size d_G and the corrected Cohen's effect size $d_{G,\ dep}$ are nearly identical within two decimals. If it is remembered that a negative effect size corresponds to a decrease whereas a positive one corresponds to an increase, and if we consider that ~0.2 is small, ~0.5 is medium, and ~0.8 is large, we find that all teaching methods have a large and positive impact. However, we also observe that the effect sizes of interactive engagement (1.26) and regular instruction (1.28) are nearly the same whereas the effect size of novice modelling instruction (1.08) is lower. It would have been expected otherwise based on the previous ANOVA performed on normalized gains where modelling instruction (40.6%) was situated between interactive engagement (43.4%) and regular instruction (29.1%). Thus, it seems like novice modelling instruction, although producing a large positive effect on FCI gains, falls behind when taking into account the pooled standard deviation. Yet, one should be careful about this, considering limitations to this study. Further research should be performed before arriving at more solid conclusions.

3.2.4 Comparison with the scientific literature

Beside comparing FCI learning gains of a novice modelling instruction implementation with regular instruction and interactive engagement instruction at the researcher's college, it is interesting to do so with results from the literature to put our results into perspective and extract additional insight. Looking for North American meta-analyses, or smaller studies if they are of particular interest, we found a few whose results are summarized in table 11. The main inconvenience of these studies is that the standard error is seldom reported. Yet, the data are still interesting and useful to risk some inferences.

Comparable statistics from our research are presented in table 12. Looking at conceptual learning gains as assessed by FCI normalized gains of averages $<g>$ or averages of normalized gains g_{avg}, we notice that the regular instruction group performed better at 29.1% than the typical 22% reported for traditional instruction in the literature (t-value = 3.00; p-value = .004). We may think that the limited use of active learning strategies, although in moderation, have contributed to this better performance. The interactive engagement group seems to have performed, at 41%-43%, as reported in the literature (39% to 48%). When we look at modelling instruction, the literature reports normalized average gains of about 35-36% for a novice implementation and 56-58% for an expert implementation, keeping in mind that novice modellers are defined as teachers who have completed the first 4-week Modelling Workshop at Arizona State University (ASU) whereas expert modellers are defined as teachers who have completed the full two-summer program of

Modelling Workshops. The researcher's group of modelling instruction performed at 39%-41%, which is also consistent (*t*-value = 0.80; *p*-value = .434). The performance could even be considered a positive surprise as the researcher didn't have any formal training in modelling instruction. Prudence is nevertheless required, as many students (43% of the class) didn't consent to participate in the research and therefore were not taken into account in the data presented.

Table 11. Statistics on FCI gains, based on scientific literature.

Test/Reference	N_{tot}	**Measure (%)**	**Modelling Instruction**	**Interactive Engagement**	**Traditional**
MD, FCI, MB (Hake, 1998)	6,542	<*g*> (StDev)	*	48 (14)	23 (4)
FCI (Hestenes, 2000)	20,000	<*g*>	58 (expert) [a] 36 (novice) [b]	*	22
FCI (Hestenes, 2006)	7,500	<*g*>	56 (expert) [a] 35 (novice) [b]	*	22
FCI (Brewe et al., 2010)	1,016	<*g*>	44.4	*	22.1
		<*G*> (SE)	30.4 (1.1)	*	14.8 (0.5)
FCI (Von Korff et al., 2016)	31,000	<*g*> & g_{avg} (SE) [c]	*	39 (2)	22 (2)

MD: Halloun – Hestenes Mechanics Diagnostic / FCI: Force Concept Inventory / MB: Mechanics Baseline

[a] Expert modellers are defined as teachers who have completed the full two-summer program of Modelling Workshops at Arizona State University (ASU).

[b] Novice modellers are defined as teachers who have completed the first 4-week Modelling Workshop at Arizona State University (ASU).

[c] Estimated graphically from figure 1 in their publication.

Table 12. Comparable statistics on FCI gains for the F2018 research.

Variable	**Group**	***N***	***N****	**Mean**	**SE Mean**	**StDev** [a]
FCI Normalized Gain of Averages <*g*> (%)	Act	51	10	41.06	*	*
	Mod	20	0	39.22	*	*
	Reg	61	22	29.11	*	*
FCI Average of Normalized Gains g_{avg} (%)	Act	51	10	43.39	3.39	24.21
	Mod	20	0	40.56	6.33	28.31
	Reg	61	22	29.13	2.38	18.59

[a] The standard deviation of the average of normalized gains g_{avg} was used for *t*-statistics.

3.3 Conceptual learning gains on the RRMCS test

In this section, descriptive statistics will be discussed before analyzing data through Cohen's *d* (effect size). The data will also be compared with the scientific literature.

3.3.1 Descriptive statistics

Only students from the modelling group were assessed based on the RRMCS standard test to evaluate learning outcomes in rotation. We cannot compare groups, but we can provide descriptive statistics that may become useful for future studies. If we look into table 13, we can first assess the differences between various related measures.

Table 13. Descriptive statistics of RRMCS gains for novice modelling instruction.

Variable	**Group**	***N***	***N****	**Mean**	**SE Mean**	**StDev**
RRMCS Pre-Score (%)	Mod	20	0	35.66	3.26	14.60
RRMCS Post-Score (%)	Mod	20	0	48.49	4.59	20.53
RRMCS Average Gain <G> (%)	Mod	20	0	12.83	2.70	12.05
RRMCS Normalized Gain of Averages <*g*> (%)	Mod	20	0	19.95	*	*
RRMCS Average of Normalized Gains g_{avg} (%)	Mod	20	0	23.41	6.21	27.77
RRMCS Normalized Change of Averages <*c*> (%)	Mod	20	0	19.95	*	*
RRMCS Average of Normalized Changes c_{avg} (%)	Mod	20	0	22.46	6.51	29.11

The first thing we observe is the average pre-test score of 35.7% and the average post-test score of 48.5%. The RRMCS average pre-test score is lower than the one found for the FCI (44.3%). This is understandable as rotation is a topic not covered in high school, contrary to kinematics and forces. Not surprisingly, the RRMCS average post-test score is also lower compared to FCI's (66.2%). We also notice that this post-test score is considerably lower than 60%, leading us to believe that more time and more well-planned activities will be required if one is hoping for a basic conceptual understanding of rotation.

The literature generally relies on the normalized gain of averages $<g>$ or the average of normalized gains g_{avg} to compare learning outcomes of different instructional designs. Both have been calculated. Some authors have discussed the limitations of normalized gains and proposed to rather use the normalized changes. We calculated them as well, but as can be seen, the differences between the two in our dataset is rather minimal. Nevertheless, we report the descriptive statistical analysis of both the average of normalized gains g_{avg} and change c_{avg} in figure 9. It appears that the novice modelling instruction has produced, on average, an increase of 20% to 23% in conceptual understanding of rotation, about half what was found for kinematics and forces with the FCI (39% to 41%).

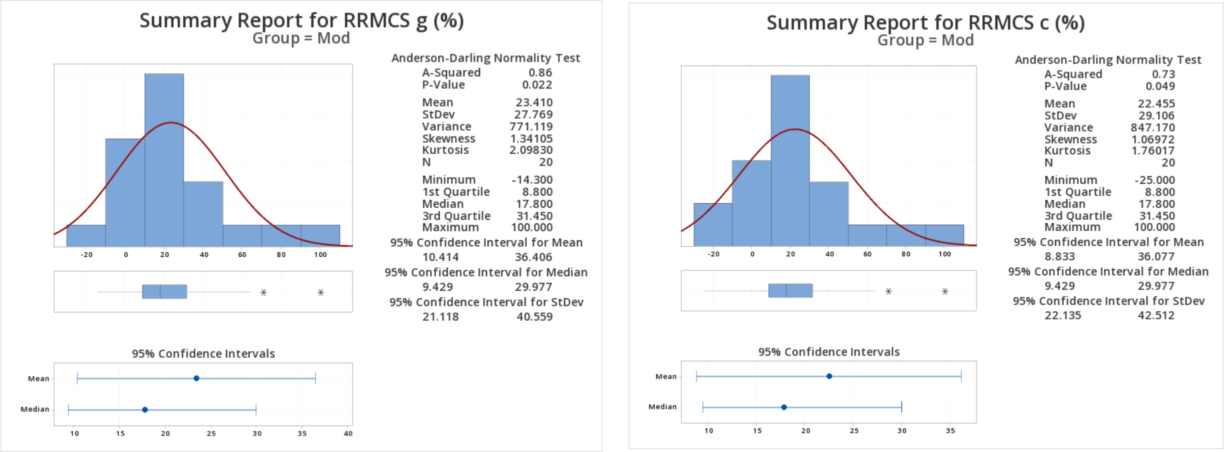


Figure 9. Summary descriptive statistics of RRMCS normalized gains.

3.3.2 Effect size

If we look into the effect size d_G based on the average gain and the pooled standard deviation of RRMCS pre- and post-test scores, and also into the corrected effect size $d_{G,\ dep}$ taking into account a possible dependence between the two scores modelled by Pearson's coefficient of correlation r, we find results of table 14. As can be seen in the table, both Cohen's effect size d_G and the corrected Cohen's effect size $d_{G,\ dep}$ are relatively similar.

Table 14. Effect size for the RRMCS gain, for novice modelling instruction.

Group	***N***	***N****	**RRMCS Gain Effect Size d_G**	**RRMCS Gain Corrected Effect Size $d_{G,\ dep}$**
Mod	20	0	0.72	0.65

If it is remembered that a negative effect size corresponds to a decrease whereas a positive one corresponds to an increase, and if we consider that ~0.2 is small, ~0.5 is medium, and ~0.8 is large, we observe that novice modelling instruction appears to have a moderately high positive impact.

3.3.3 Comparison with the scientific literature

Beside comparing conceptual learning outcomes of a novice modelling instruction implementation with regular instruction and interactive engagement instruction at the researcher's college, it is interesting to do so with results from the literature. Looking for North American meta-analyses, or smaller studies if they are of particular interest, we found one whose results are summarized in table 15. The main inconvenience of that study is that the standard error is not reported. Yet, the data are still interesting and useful to risk some inferences.

Comparable statistics from our research are presented in table 16. When we look at the conceptual mastery of rotation as assessed by the RRMCS scores at the end of a first-semester mechanics course, we observe that the results from the modelling group (48.5%) compare to what Rimoldini and Singh (2005) report for introductory and upper-level traditional courses (44% to 61%). It would be better to have average normalized gains calculated based on an RRMCS pre-test average score, but it is justifiable to skip the pre-test and to focus on the post-test score when students don't have previous knowledge as it is the case of rotation when students take college or university mechanics for the first time. The data from Rimoldini and Singh (2005) mixed introductory and upper-level physics, so it's not perfect. However, only 17 students out of 559 (after removing 93 freshman honour students out of the total of 652 participating in their study) came from those upper-level courses. We can thus hypothesize a relatively low effect of the mix on the results for introductory students. It would seem, therefore, that novice modelling instruction hasn't stood out, although this is based on a single study and both studies have limitations that have to be kept in mind.

Table 15. Statistics on RRMCS scores, based on scientific literature.

Test/Reference	N_{tot}	**Measure (%)**	**Modelling Instruction**	**Interactive Engagement**	**Traditional**
RRMCS (Rimoldini & Singh, 2005)	652	RRMCS scores	*	*	44-61 (introductory and upper level) 75 (freshman honour) [a]

RRMCS: Rotational and Rolling Motion Conceptual Survey

[a] The mode of instruction was not specified, but Dr. Singh is pretty sure it was traditional (private conversation, 2020).

Table 16. Comparable statistics on RRMCS scores for the F2018 research.

Variable	**Group**	***N***	***N******	**Mean**	**SE Mean**	**StDev**
RRMCS Post-Score (%)	Act	0	61	*	*	*
	Mod	20	0	48.49	4.59	20.53
	Reg	0	83	*	*	*

3.4 Problem-solving skills on the 78%-common final exam

Let's now turn to the 78%-common final exam grades as a way to assess problem-solving skills rather than conceptual understanding as it was the case with the FCI and RRMCS. A one-way analysis of variance (ANOVA) on final exam grades was performed with $\alpha = .05$. Results are presented in tables 17 and 18, and figures 10 and 11. The p-values are above $\alpha = .05$ and none of the many tests performed the final exam grades leads us to discard the null hypothesis. Therefore, no statistically significant difference appears between groups.

Table 17. Group means (%) on final exam grades.

Group	***N***	**Mean**	**StDev**	**95% CI**
Act	56	67.61	20.37	(62.58, 72.63)
Mod	20	62.53	15.67	(54.12, 70.94)
Reg	79	68.26	18.79	(64.03, 72.49)

Pooled StDev = 19.0281

Table 18. Analysis of variance on final exam grades.

Source	**DF**	**Seq SS**	**Contribution**	**Adj SS**	**Adj MS**	***F*-Value**	***P*-Value**
Group	2	533.4	0.96%	533.4	266.7	0.74	0.480
Error	152	55034.2	99.04%	55034.2	362.1		
Total	154	55567.6	100.00%				

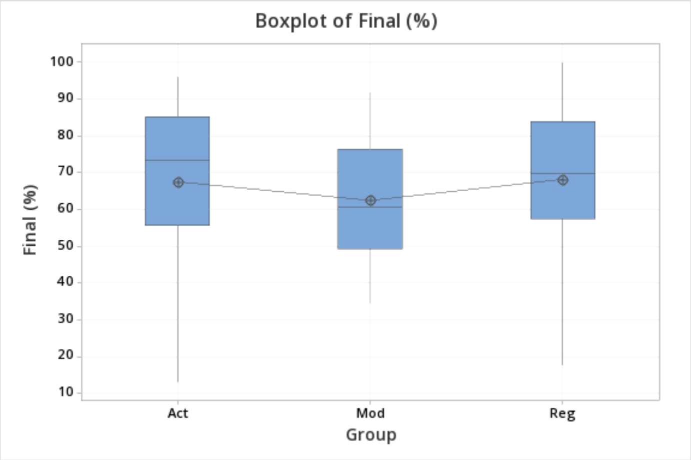


Figure 10. Boxplot of final exam grades.

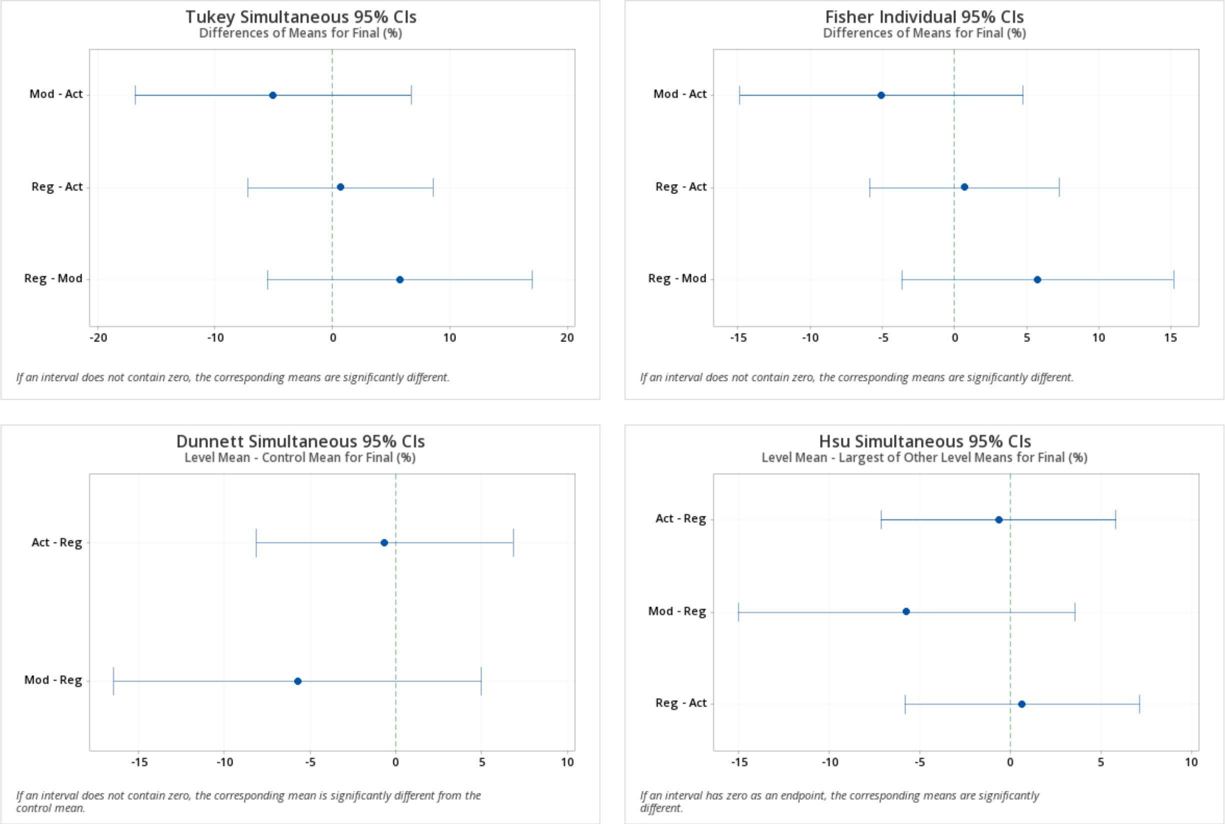


Figure 11. Tukey and Fisher pairwise comparisons, Dunnett multiple comparisons with the control (Reg), and Hsu multiple comparisons with the best (MCB) for final exam grades.

3.5 Discussion

No statistically significant difference was found between novice modelling instruction and either interactive engagement or regular instruction for learning outcomes, be they measured through the FCI for normalized conceptual learning gains or the 78%-common final exam for problem-solving skills. Thus, the null hypothesis couldn't be rejected for modelling instruction. However, a statistically significant difference was found between interactive engagement and regular instruction for FCI normalized conceptual gains (the former apparently surpassing the latter), but not for problem-solving skills.

A surprising thing is that, when looking at effect sizes through Cohen's *d* for FCI conceptual learning gains, interactive engagement and regular instruction appear to have the same effect whereas novice modelling instruction falls behind, although all effects are large and positive. Another thing is that modelling instruction is a particular type of interactive engagement, so it is indeed surprising to see an effect size quite lower, although we also have to remember that it was a novice implementation without formal training in the method. This might be related to the spread of the data affecting the pooled standard deviation and the low participation rate in the Mod group, suggesting further studies with larger sample sizes would be useful to clarify the issue.

Despite not distinguishing itself much, novice modelling instruction overall appears to have at least performed equivalently to other methods. There is still a suspicion that modelling instruction might be closer to interactive engagement, not only from a theoretical perspective, but from an experimental one as well if we recall that we did detect a statistically significant difference between the Mod and Reg groups based on the comparison of FCI post-test raw scores, considering there was no difference on pre-test scores. Comparisons in conceptual learning gains for rotation couldn't be performed for lack of data, but conceptual learning gains for translational motion and forces show that all methods barely succeeded at helping students to transition from a pre-Newtonian to a Newtonian worldview (all FCI average post-scores were close to 60%), interactive engagement (64.3%) and modelling instruction (66.2%) performing significantly better in this regard than regular instruction (53.6%).

4. SECOND RESEARCH QUESTION: STUDENT ATTITUDES ABOUT LEARNING PHYSICS

Our second research question and corresponding hypotheses were the following:

Q_2: How does modelling instruction differ from regular instruction or interactive engagement in terms of attitudes (or beliefs) about physics for CEGEP Mechanics students?

H_2: CEGEP Mechanics students receiving modelling instruction, compared to regular instruction or interactive engagement, will perform differently on a pre/post standardized assessment (CLASS) of attitudes (or beliefs) about physics.

H_0: CEGEP Mechanics students receiving modelling instruction, compared to regular instruction or interactive engagement, will perform equivalently on a pre/post standardized assessment (CLASS) of attitudes (or beliefs) about physics.

To answer this question about student attitudes toward physics and the learning of it, we collected responses to the CLASS. We then calculated various statistics like pre- and post-test scores, shifts, and effect sizes, which all appear in appendix A. Hereafter CLASS favourable and unfavourable attitude shifts will be examined and discussed.

4.1 CLASS favourable attitude shifts

In this section, data will be analysed through a one-way ANOVA and Cohen's d (effect size). The data will also be compared with the scientific literature.

4.1.1 Analysis of variance

A one-way ANOVA on favourable attitude shifts S was performed with $\alpha = .05$. Results are presented in tables 19 and 20, and figures 12 and 13.

Table 19. Group means (%) on CLASS favourable shifts.

Group	***N***	**Mean**	**StDev**	**95% CI**
Act	32	-1.37	17.87	(-7.82, 5.07)
Mod	19	-1.57	12.02	(-7.36, 4.23)
Reg	58	0.39	12.27	(-2.83, 3.62)

Table 20. Analysis of variance on CLASS favourable shifts (Welch's test).

Source	**DF Num**	**DF Den**	***F*-Value**	***P*-Value**
Group	2	44.8163	0.25	0.778

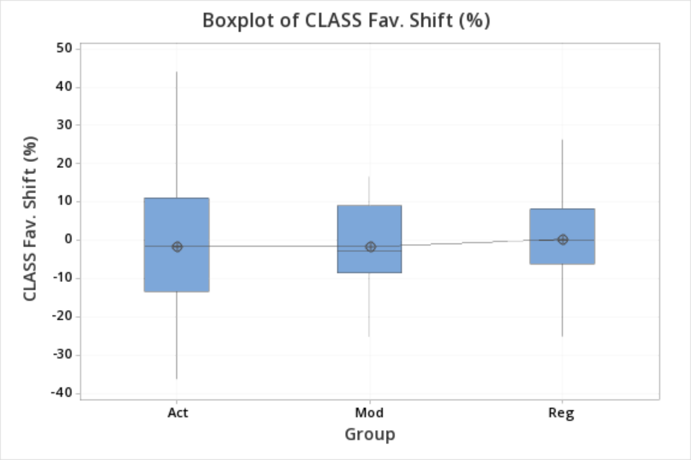


Figure 12. Boxplot of CLASS favourable shifts.

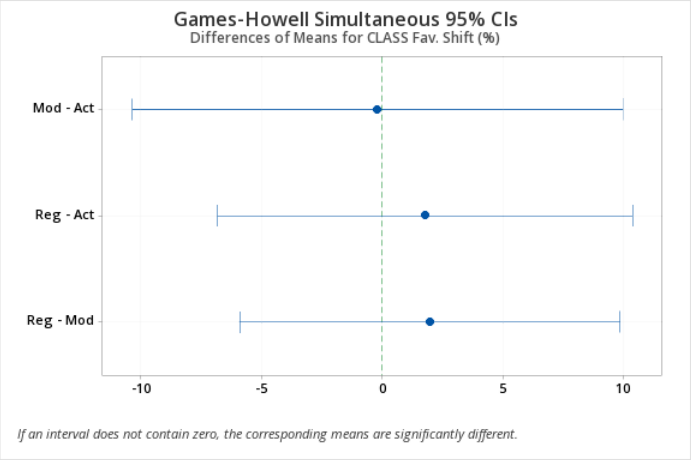


Figure 13. Games-Howell pairwise comparisons for CLASS favourable shifts.

The p-value (.778) is above $\alpha = .05$ and the Games-Howell pairwise comparisons fail to discard the null hypothesis. It appears that all three methods of teaching have produced sensibly the same effects.

4.1.2 Effect size

If we leave the ANOVA for a moment and look into the effect size d_S based on the average shift and the pooled standard deviation of CLASS pre- and post-test favourable scores, and also into the corrected effect size $d_{S,\,dep}$ taking into account a possible dependence between the two scores modelled by Pearson's coefficient of correlation r, we find results of table 21. As can be seen in the table, both Cohen's effect size d_S and the corrected Cohen's effect size $d_{S,\,dep}$ are identical within two decimals.

Table 21. Effect size for the CLASS favourable shift, for different modes of instruction.

Group	***N***	***N****	**CLASS Favourable Shift Effect Size d_S**	**CLASS Favourable Shift Corrected Effect Size $d_{S,\,dep}$**
Act	32	29	-0.08	-0.08
Mod	19	1	-0.10	-0.10
Reg	58	25	0.02	0.02

If it is remembered that a negative effect size corresponds to a decrease whereas a positive one corresponds to an increase, and if we consider that ~0.2 is small, ~0.5 is medium, and ~0.8 is large, we find that novice modelling instruction (-0.10) and other interactive engagement methods (-0.08) appear to have a very small negative impact on students' favourable attitudes toward the field (decreasing them) whereas the effect of regular instruction (0.02) appears nearly inexistent.

4.1.3 Comparison with the scientific literature

Beside comparing attitude shifts of a novice modelling instruction implementation with regular instruction and interactive engagement instruction at the researcher's, it is interesting to do so with results from the literature. Looking for North American meta-analyses, or smaller studies if they are of particular interest, we found a few whose results are summarized in table 22. The main inconvenience of these studies is that the standard error is seldom reported. Yet, the data are still interesting and useful to risk some inferences.

Table 22. Statistics on CLASS attitude shifts, based on scientific literature.

Test/Reference	N_{tot}	Measure (%)	Modelling Instruction	Interactive Engagement	Traditional
CLASS (Brewe et al., 2008)	24	Fav. <*S*>	8.6	*	*
CLASS (Brewe et al., 2009)	22	Fav. <*S*> [a]	12	*	*
CLASS (de la Garza & Alarcon, 2010)	44	Fav. <*S*>	3.05	*	*
CLASS, MPEX (Madsen et al., 2015)	11,131	Fav. <*S*>	9.3	8.5 (explicit focus) 0.7 (some focus)	-3.7

CLASS: Colorado Learning Attitudes about Science Survey

MPEX: Maryland Physics Expectations Survey

[a] Estimated graphically from figure 1 in their publication.

Table 23. Comparable statistics on CLASS favourable attitude shifts for the F2018 research.

Variable	Group	*N*	*N**	Mean	SE Mean	StDev
CLASS Average Favourable Attitude Shift <*S*> (%)	Act	32	29	-1.37	3.16	17.87
	Mod	19	1	-1.57	2.76	12.02
	Reg	58	25	0.39	1.61	12.27

Students' views of the nature of physics and of learning physics as assessed by CLASS favourable shifts (table 23) present peculiar disparities. According to the literature, traditional instruction produces negative shifts of the order of -3.7% in favourable attitudes. The regular instruction group performed better at 0.4%, although not by much. The difference is nevertheless statistically significant (*t*-value = 2.54; *p*-value = .014). The interactive engagement group produced a shift of -1.4%, whereas the literature reports positive shifts between 0.7% when there is some focus on developing expert-like beliefs, and 8.5% when there is an explicit focus on that. A one-sample *t*-test shows that the negative shift recorded in this research is small enough to be compatible with nearly inexistent shifts reported by Madsen, McKagan, and Sayre (2015) when there is only some focus on expert-like beliefs (*t*-value = -0.66; *p*-value = .517).

Results from the modelling instruction group are somewhat more difficult to understand. The group performed negatively at -1.6% in favourable attitude shifts whereas modelling instruction is reported to perform between 3.1% in the case of the very small sample of de la Garza and Alarcon (2010) and 12% in the case of the other very small sample of Brewe, Kramer, and O'Brien (2009). Referring with more confidence to the large study of Madsen, McKagan, and Sayre (2015), courses focusing explicitly on model building – like Modelling Instruction (MI) but also including a few more model-focused pedagogies like Physics by Inquiry, Physics of Everyday Thinking (PET), Physical Science of Everyday Thinking (PSET), and Modelling Applied to Problem Solving (MAPS) – show a favourable shift of 9.3%. The difference is quite important (*t*-value = -3.94; *p*-value = .001), and if we believe it reflects reality despite the limitations of our research, it has to be explained. Further research would be beneficial to confirm any attempt at explaining this observation, but it can be hypothesized that the novice implementation of modelling instruction by the researcher didn't optimize discourse management in a way that would have helped shift attitudes and beliefs toward physics more positively. The modelling group essentially performed like the interactive engagement group, which appeared compatible with common results interactive engagement with only some focus on developing expert-like beliefs.

4.2 CLASS unfavourable attitude shifts

In this section, data will be analysed through a one-way ANOVA and Cohen's effect size *d*.

4.2.1 Analysis of variance

A one-way ANOVA on unfavourable attitude shifts *S* was performed with $\alpha = .05$. Results are presented in tables 24 and 25, and figures 14 and 15.

Table 24. Group means (%) on CLASS unfavourable shifts.

Group	***N***	**Mean**	**StDev**	**95% CI**
Act	32	0.03	16.75	(-6.00, 6.07)
Mod	19	2.71	7.19	(-0.76, 6.18)
Reg	58	0.01	9.24	(-2.42, 2.44)

Table 25. Analysis of variance on CLASS unfavourable shifts (Welch's test).

Source	**DF Num**	**DF Den**	***F*-Value**	***P*-Value**
Group	2	48.3104	0.90	0.413

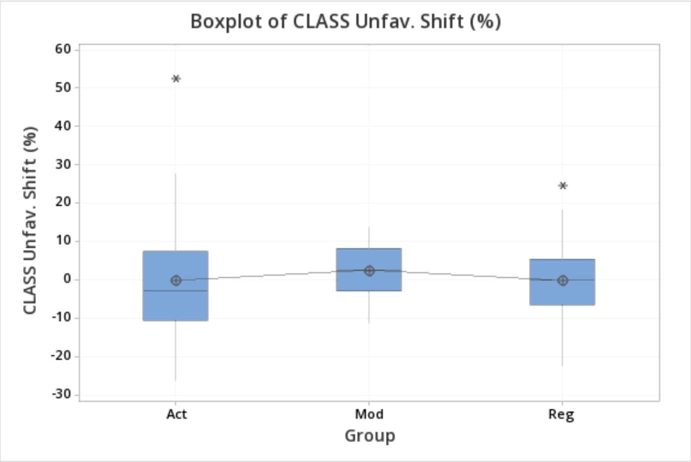


Figure 14. Boxplot of CLASS unfavourable shifts.

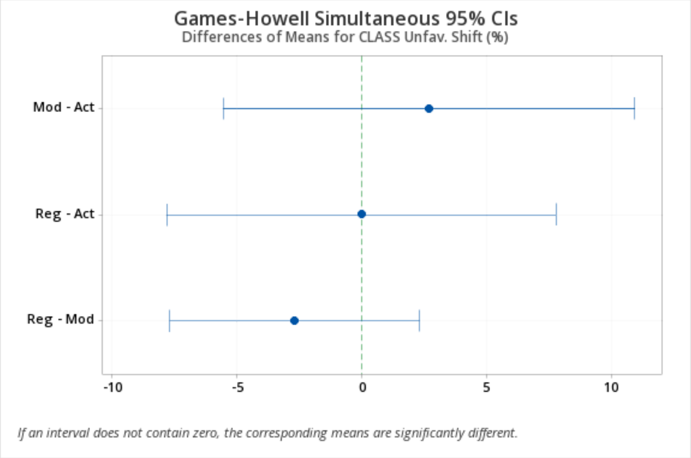


Figure 15. Games-Howell pairwise comparisons for CLASS unfavourable shifts.

The p-value (0.413) is above $\alpha = .05$ and the Games-Howell pairwise comparisons fail to discard the null hypothesis. It appears that all three methods of teaching have produced sensibly the same effects.

4.2.2 Effect size

If we leave the ANOVA for a moment and look into the effect size d_S based on the average shift and the pooled standard deviation of CLASS pre- and post-test unfavourable scores, and also into the corrected effect size $d_{S,\ dep}$ taking into account a possible dependence between the two scores modelled by Pearson's coefficient of correlation r, we find results of table 26. As can be seen in the table, both Cohen's effect size d_S and the corrected Cohen's effect size $d_{S,\ dep}$ are identical within two decimals.

Table 26. Effect size for the CLASS unfavourable shift, for different modes of instruction.

Group	***N***	***N****	**CLASS Unfavourable Shift Effect Size d_S**	**CLASS Unfavourable Shift Corrected Effect Size $d_{S,\,dep}$**
Act	32	29	0.00	0.00
Mod	19	1	0.24	0.24
Reg	58	25	0.00	0.00

If it is remembered that a negative effect size corresponds to a decrease whereas a positive one corresponds to an increase, and if we consider that ~0.2 is small, ~0.5 is medium, and ~0.8 is large, we find that novice modelling instruction (0.24) has a small negative impact on students' unfavourable attitudes toward the field (increasing them) whereas the effect of regular instruction (0.00) and interactive engagement (0.00) appears essentially inexistent.

4.3 Discussion

No difference was found between groups for either favourable or unfavourable attitude shifts, and therefore the null hypothesis couldn't be rejected. One surprise is that the average favourable attitude shift for interactive engagement (-1.4%) and modelling instruction (-1.6%) tends to be negative whereas the average unfavourable attitude shift for the same (Act: 0.03%; Mod: 2.7%) tends to be null or positive, as opposed to a near absence of effect for regular instruction (favourable: 0.39%; unfavourable: 0.01%), as if active learning of college physics, no matter the method, was worsening attitudes toward the field. It is normally reported that students leave typical physics courses believing that physics is less coherent, less logical, and less relevant to their everyday lives than when they started the course (McKagan et al., 2007) unless the course is explicitly focused on model-building and developing expert-like beliefs (Madsen et al., 2015). Yet, looking further at the numbers in our study, the averages (just like the effect sizes) are relatively close to zero and the intervals of confidence go both ways. Perhaps then the average shifts in attitude don't reflect the actual average shifts of the population, considering the intervals of confidence. This would be worth further investigation for confirmation.

5. THIRD RESEARCH QUESTION: STUDENTS' PERCEPTION OF MODELLING INSTRUCTION

Our third research question was the following:

Q_3: How are CEGEP Mechanics students perceiving (in terms of what they like or don't like) the introduction of modelling instruction?

It was expected that students receiving modelling instruction would overall prefer this form of active learning, but we sought to capture the actual perception and reasons for it.

To answer this question, we collected answers to a qualitative and open-ended survey at the end of the semester. Only 8 out of 35 students (23%) answered that anonymous survey online, so it is delicate to generalize to the whole class. Nevertheless, answers are reproduced in their integrality appendix A. A horizontal bar graph summarizing the frequency of main themes identified in modelling students' comments is presented in figure 16. The overall positive or negative counts correspond to a number of students, whereas the various theme counts correspond to a number of times a theme has been expressed by the same or different students in answers to different questions, as a way to assess the weight of the perception.

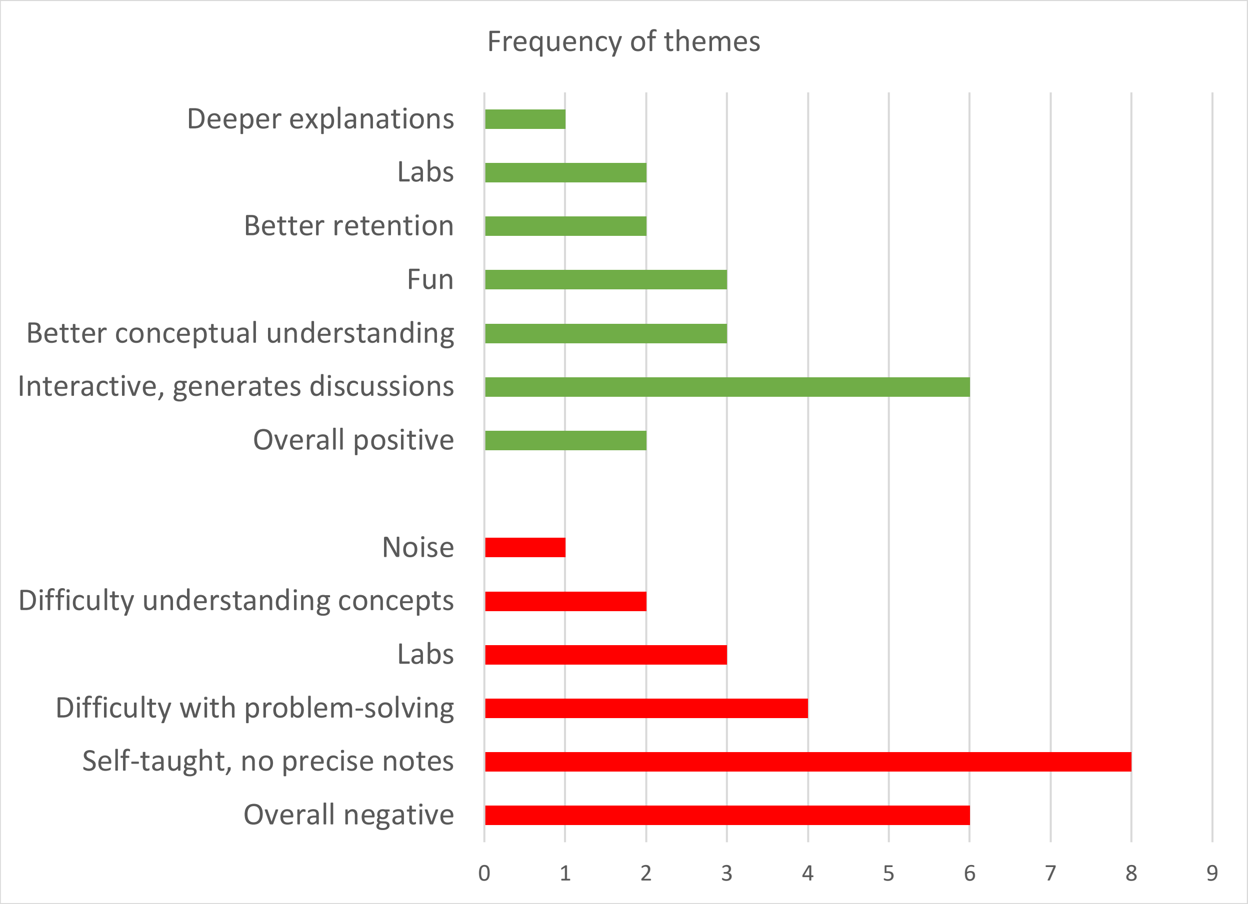


Figure 16. Frequency of themes in modelling students' comments to the end-of-semester appreciation survey.

5.1 Discussion

Of the eight students who answered, two (25%) could be said to have liked the experience, while the other six (75%) didn't much. It's hard to presume the representativeness of such a small sample, but at the same time, we can try to make sense of the experience of those who decided to answer. In a sense, negative feedback is not that different from the one that is reported from other teachers when they move from traditional and more passive classrooms to more active learning classrooms, where the student must be engaged to learn. Wells (1987) reminds us that “[t]he reaction of students to a course in physics that is structured around concept development—nothing to memorize—initially induces high frustration levels, accompanied by frequent requests to be told what they should ‘know’ for the tests” (p. 36).

According to the survey, many students seemed to prefer to be told what to memorize rather than discover by themselves and face the initial uncertainty of building knowledge as real scientists do through experiments, observations, and debates when they try to explain a new phenomenon that previous models cannot fully explain. It’s certainly more cognitively demanding to fully immerse oneself in the process, as student E admits when saying that “[i]t requires more work”, and not all students are willing to take such an active role. We see that in student A complaining about the course being mostly “self-taught” and preferring “some more direct instruction of theory”, or student G feeling “as if it were the students teaching themselves” and preferring “precise class notes to follow along.” Such complaints are quite common in active learning, no matter the method and despite improved performance on average. At the same time, it must be recognized that a novice implementation of modelling instruction was performed. It’s possible that the delicate and optimal balance between autonomous work and guidance was somewhat offset.

Many students seemed to prefer the “sage on the stage” approach, placing in the teacher the role of true knowledge deliverer. Student B thus mentions: “Best way: copying notes and having good explanation from teacher.” Student C also mentions: “The material of the course should be explained by the teacher first then put into application by the students and not the opposite.” Student D reinforces this by explaining that he or she preferred when “the teacher explained all the material and well leaded the class for the final exam" (as in Chemistry). Those complaints could be related to the strong desire for explicit guidance mentioned earlier, or relate to the epistemological figure of the teacher as it is interpreted here. These are also common in active learning. It shows the necessity to develop ways to sustain a transformation of epistemic beliefs of students, or of their relation to experts and scientific knowledge, an aspect that a more expert implementation of modelling instruction might improve more than what appears here based on a relatively limited novice implementation.

Other students seemed to be deficient in their understanding of the experimental foundation of physics, by whishing theory had more importance in the course. It's the case of student C when saying: "I don't like the fact that it relies mostly on experiments rather than theory." But other students did enjoy the stronger emphasis on lab experiments, like students E (with some nuances about the way it was done and their grading) and F, or teamwork like student G (although it's not clear if by teamwork the student was thinking about the labs).

Some students felt they didn't learn as they should how to solve problems. Student C "had difficulty applying [concepts] in problems," adding that "I feel like I didn't learn to apply some of the concepts we saw in this course quite well in problems." Student B mentions: "I don't know how to solve challenging mechanics problems." Student E said the modelling approach "wasn't as useful for exam preparation and getting a good R-score." That being said, this can happen with other approaches as well, especially if solving problems doesn't involve following recipes blindly anymore. On the other hand, some students like C, E and F found modelling fun and helpful in understanding concepts and in retaining the information learned better.

Perhaps most students who participated in the research and answered the survey realized that college physics was not exactly like they imagined from their high-school experience, that simple “plug-and-chug” in equations was not sufficient. This might have affected their attitudes and they may have attributed that to modelling instruction whereas it would have been attributed to any other interactive method of teaching if it had happened to be the one employed to teach the course. It would be interesting to compare attitudes and perceptions when an active class is surrounded by more traditional classes to what they are when a whole department has adopted active learning methods, thus becoming the norm.

In any case, although surprising for modelling instruction (and thus contrary to our expectation), results are consistent with what Deslauriers, McCarty, Miller, Callaghan, and Kestin (2019) report: “When students experience the increased cognitive effort associated with active learning, they initially take that effort to signify poorer learning. […] Compared to students in traditional lectures, students in active classes perceived that they learned less, while in reality they learned more. Students rated the quality of instruction in passive lectures more highly, and they expressed a preference to have ‘all of their physics classes taught this way,’ even though their scores on independent tests of learning were lower than those in actively taught classrooms” (p. 19251).

SIXTH CHAPTER. CONCLUSION

This research measured the impact on learning outcomes and attitude shifts of modelling instruction in a CEGEP introductory physics course, namely Mechanics, compared to regular and interactive engagement instruction. The main objective was to find out if a novice implementation of modelling leads to significant changes or not. A secondary objective was to gain some insights into students' perceptions of the new instructional design. Thus, we had three research questions to investigate.

1. CLOSING STATEMENTS ON RESEARCH QUESTIONS

The first research question was: How does modelling instruction differ from regular instruction or interactive engagement in terms of learning outcomes for CEGEP Mechanics students? According to other studies found in the literature (Hake, 1998; Hestenes, 2000, 2006; Brewe et al., 2010; Von Korff et al., 2016), novice modelling instruction compares with interactive engagement and surpasses traditional instruction on FCI normalized gains, whereas expert modelling instruction surpasses both. It was thus hypothesized that students receiving modelling instruction would perform differently on pre/post standardized assessments of deep conceptual understanding (FCI, RRMCS) and on more traditional exams testing problem-solving skills. No support for this hypothesis was found, overall, through the statistical analysis, except for FCI post-test scores when compared with regular instruction (*p*-value between .007 and .019 depending on the test).

The second research question was: How does modelling instruction differ from regular instruction or interactive engagement in terms of attitudes (or beliefs) about physics for CEGEP Mechanics students? According to the literature (Madsen et al., 2015), modelling instruction compares with interactive engagement that has an explicit focus on developing expert-like beliefs when those attitudes are measured through CLASS or MPEX favourable attitude shifts. On the other hand, it surpasses both interactive engagement with a more limited focus and traditional instruction. It was thus hypothesized that students receiving modelling instruction would perform differently on a pre/post standardized assessment (CLASS) of attitudes (or beliefs) about physics. It was found that the researcher's novice implementation of modelling instruction couldn't be distinguished from colleagues' implementation of either interactive engagement or regular instruction (which included some interactive teaching on a limited basis), all methods failing equally at improving attitudes toward physics and its learning. Fortunately, attitudes didn't appear to worsen either, but our alternative hypothesis couldn't find support.

The third research question was: How are CEGEP Mechanics students perceiving (in terms of what they like or don't like) the introduction of modelling instruction? Because modelling instruction is believed to foster a better understanding of physics, it was expected that students receiving modelling instruction would overall prefer this form of active learning. It was found, based on a qualitative survey and the relative percentage of positive versus negative perceptions, that the researcher's implementation of modelling instruction produced more dissatisfaction than satisfaction, which, although saddening, is consistent with what is reported for other, non-modelling methods fostering active learning (Wells, 1987; Deslauriers et al., 2019). It must be remembered, though, that the level of participation in that survey was very low (23%).

2. LIMITATIONS OF THE STUDY

The primary focus of this study was to find out if modelling instruction produced different learning outcomes and attitudes compared to regular or interactive instruction. Our analysis has not detected any significant difference (be it positive or negative). It therefore appears that novice modelling instruction has been as good as other modes of instructions. Yet, based on the literature, we would have expected modelling instruction to show significant improvements in comparison to regular instruction. We can offer some possible explanations for that.

2.1 Implementation of modelling instruction and regular instruction

First, modelling instruction was applied in a novice manner, in the sense that the researcher investigated the method, but never had formal training to implement it. Hake (1998) mentions, when discussing interactive engagement (IE) courses leading to low normalized gains or averages $\langle g \rangle$, that various implementation problems were apparent in cases he studied: "insufficient training of instructors new to IE methods, failure to communicate to students the nature of science and learning, lack of grade incentives for taking IE activities seriously, a paucity of exam questions which probe the degree of conceptual understanding induced by the IE methods, and use of IE methods in only isolated components of a course" (p. 66). The researcher feels that insufficient training in modelling instruction and lack of grade incentives for taking modelling activities seriously (most of it was formative, not summative) may have affected the situation, although he also feels that the nature of science and learning was properly communicated, that tests included a good balance between problem-solving and conceptual questions, and that the modelling method was systematically applied throughout the whole course. On that latter point, however, it must be mentioned that the rigid structure of two hours of lab time and twice two hours (including an added hour of special support all Mechanics teachers give to help students succeed) of class time once a week made it difficult to adopt a proper pace: most often modelling activities (be they in the development or deployment phases) would be rushed to be ready for the next lab activity which could only be done on a specific day of the week whereas at rarer, less problematic times, the pace would be slowed down and more opportunities for whiteboarding problems would appear. Therefore, it is possible that this novice, not optimized, implementation of modelling instruction, compounded by time issues and some teams not taking whiteboarding sessions as seriously as they should, did reduce effects documented in the literature. An expert, well-trained modeller might more easily deal with constraints and still maintain expected learning gains.

Hestenes (1997) also discusses conditions for effective modelling instruction. He states that "teacher guidance is essential throughout the modelling cycle" (p. 953) and that "[t]he most critical element in successful implementation of the modelling method is the skill of teacher in managing classroom discourse" (p. 954). Again, the balance between too much guidance and not enough is not easy to set, and the novice approach of the researcher, compounded by time constraints whereas a typical modelling cycle is at least two weeks long with about a week for model development and a week for model deployment (Hestenes, 1997), may not have allowed modelling instruction to express its full potential, although it has proven not to be harmful. After all, based on the data that were collected, the researcher's novice modelling instruction couldn't be differentiated from either interactive engagement nor regular instruction, whereas on the FCI we could find a significant difference between interactive engagement and regular instruction.

On the other hand, regular instruction was not purely traditional, lecture-based instruction. Some form of active learning was part of regular instruction, although in a reduced way (on the order of 20-40%). That may have reduced the differences further, although differences were strong enough to be found significant on the FCI averages of normalized gains between the interactive engagement and the regular instruction groups.

2.2 Timing of pre-tests

Second, most pre-tests (FCI, CLASS, RRMCS) were administered a little bit late, three weeks after the start of the semester. Although it is believed the impact on the FCI for the modelling instruction group is rather limited, it might not have been so in other groups if they had already started discussing some topics like kinematics, forces, and/or momentum (the order of the topics depends on the teacher). It was still early in the semester, but not exactly as if questions would have been answered on the very first day of college mechanics. It is possible that some FCI gains were therefore reduced by a larger pre-test score (or at least changed, if we don't assume the direction of that change in case regression precedes learning), or that attitudes assessed on the CLASS had already started to shift. For the RRMCS, an impact would be very surprising as this test is on a topic (rotation) that has never been explored by students in a physics class before. In the only study that we could find using the RRMCS test, the developers used it only after instruction (Rimoldini & Singh, 2005).

2.3 Sampling issues

Third, we must recognize limits to the sampling. This quasi-experiment was based on convenience and intact, non-random samples with relatively small sizes, especially for the treatment, modelling instruction, group (20 consenting students for novice modelling, as opposed to 83 and 61 for regular instruction and active learning, respectively). Furthermore, those consenting students did not represent the totality of the sections (57% for novice modelling, versus 81% and 86% for regular instruction and active learning, respectively), and not all of them completed both the pre- and post-tests (as little as 45% for the CLASS test in the active learning group, and 57% for the FCI and RRMCS tests in the modelling group).

A power analysis can be useful here to gain insights into the possibility of type II errors. Requiring a statistical power $(1-\beta)=.80$ and assuming a pooled standard deviation of 20% (approximately like for the FCI normalized gains, at 22.5%, and final exam grades, at 19.0%) while setting the significance threshold at $\alpha=.05$, we can calculate the maximum difference between means for various sample sizes: 20, 50, 79 (table 27 and figure 17). That maximum difference is the mean difference that can be detected between the factor level that has the smallest mean and the factor level that has the largest mean. Its value should represent the smallest difference that has practical consequences. The novice modelling sample group had a total of 19-20 observations, whereas the regular instruction group had at most 79 observations. Assuming groups of the same size, a size of 20 leads to a maximum difference of 20.2% whereas a size of 79 leads to a value of 9.9%. The range of maximum differences between means remains relatively high, generating doubts about the capacity of the statistical tests to detect small differences in FCI normalized gains and final exam grades, especially between the novice modelling group and other groups.

Doing the same calculations with an assumed pooled standard deviation of 14% (closer to what we have for attitude shifts, at 14.1%), a sample size of 20 leads to a maximum difference of 14.1% whereas a sample size of 79 leads to a value of 7.0% (table 28 and figure 18). The range of maximum differences between means improves and we can have slightly more confidence in the capacity of the statistical tests to avoid type II errors and detect small practical differences in attitude shifts, which we failed to detect.

Table 27. Power analysis for FCI normalized gains and final exam grades.

Sample Size	Power	Maximum Difference
20	0.8	20.1589
50	0.8	12.5435
79	0.8	9.9411

The sample size is for each level.

Table 28. Power analysis for CLASS shifts.

Sample Size	Power	Maximum Difference
20	0.8	14.1112
50	0.8	8.7805
79	0.8	6.9588

The sample size is for each level.

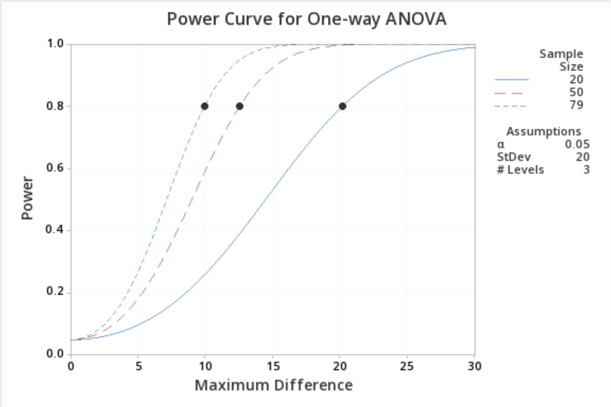


Figure 17. Power analysis for FCI normalized gains and final exam grades.

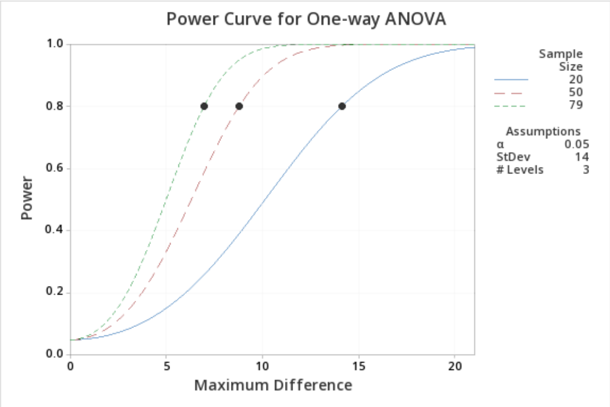


Figure 18. Power analysis for CLASS shifts.

Yet, despite insights gained from power analyses, we must keep in mind that the representativity of results may have been compromised, even for a single class. Indeed, it might happen that those students who didn't consent to participate had a particular profile, and who once removed changed the overall distribution and biased the data. This is a serious limitation that we couldn't assess precisely as we didn't have access to those data for ethical reasons.

2.4 Other methodological issues

Fourth, possible lurking variables could have compromised the study. Differences in students' traits (like autonomy, self-motivation, etc.) and differences in socioeconomic status could have had an impact, yet data about that would have been hard and delicate to gather. One might think that over a semester in a single college with a particular culture and demographics, those elements would average out over a given group and lead to the comparability of other groups from the same institution. It cannot be said so if one would like to generalize to other institutions or locations, though, and thus this should be done very carefully.

Students' skillset differences due to previous knowledge, education, or performance, could also have influenced results, but that is why we had conceived a design where post-test results were compared to pre-test results to extract normalized learning gains and changes, and where effect sizes were calculated to corroborate findings. Our analysis of FCI pre-test scores and CLASS pre-test favourable attitude scores didn't detect any statistically significant difference between groups, so groups appeared comparable in this regard. Nevertheless, intact group sampling was used instead of true random sampling. Participants were not chosen at random, and the distribution of those participants between the treatment group (Mod) and control groups (Reg and Act) were not random either, for we used existing class sections. This limits the generalizability of the research, but it could hardly have been done otherwise.

Finally, because the researcher used his class while also running the research, there was a possibility of bias as behaviours might have been unconsciously driven by the hypothesis he aimed to validate. Considering results, though, this possible bias seems unlikely.

3. IMPLICATIONS FOR FUTURE RESEARCH

Considering limitations, future research is likely required to confirm and expand upon results presented herein. The obvious avenue would be to conduct again the same kind of research, in the same or a different college. It would also be very interesting to do this research with an expert modeller, or at least with a novice modeller after formal training in the method (with one or a few modelling workshops under his or her belt). This was not the case in the present study, although documentation and videos were consulted to gain insights.

Considering that, if confirmed, all methods at best barely succeeded at helping students to reach a stage of Newtonian reasoning, further research on why and how to better achieve this goal through a first Mechanics course after high school would be useful. A broader look at Quebec's educational system and a reflection on it, besides teaching methods, may also be of interest.

A major avenue of further research is the comparison of different methods of instruction in helping students improve their conceptual understanding of rotation. One study using the RRMCS was done on traditional instruction (Rimoldini & Singh, 2005). This research has collected data for novice modelling instruction. We didn’t find any other studies and there is, therefore, a lot of room for further studies on this particular topic.

Another major path of research would be the study of the deterioration or lack of amelioration of attitudes toward physics and its learning as students go through a formal course. Why are such courses typically leading to less expert-like beliefs at worst, no change at best? The literature (Madsen et al., 2015) reports an amelioration for modelling instruction and interactive engagement with an explicit focus on those attitudes and beliefs, but this research hasn’t seen such amelioration. This should be validated as well.

Finally, we reported a significant, although not necessarily representative, dissatisfaction of students with modelling instruction. We know this to be also the case with other interactive engagement methods (Wells, 1987; Deslauriers et al., 2019). It would be interesting to compare attitudes and perceptions when an interactive engagement class is surrounded by more traditional classes to what they are when a whole department has adopted interactive engagement methods, thus becoming the norm.

4. IMPLICATIONS FOR FUTURE TEACHING PRACTICE

This research has shown that novice modelling instruction appears to be as good as other methods of teaching, thus encouraging teachers to give it a try if they wish. However, we suspect it to produce results closer to interactive engagement considering its nature, the mean values calculated and other studies we mentioned previously. The inability to replicate the high conceptual learning gains and favourable attitude shifts reported in the literature seems to indicate that there is great potential yet to uncover, but that to tap into it, formal training in the method given by expert modellers is probably necessary for many to extract all the benefits that have been documented elsewhere. It is thus recommended for curious and potential adopters, including the researcher, to attend workshops (as those developed at ASU or through the AMTA) that provide a living experience of the method as alternatively student and teacher, and develop skills in discourse management. This should be accompanied by a reflection on assessments or other techniques that would further motivate students to change their habits and get actively involved with all modelling activities, and an attempt at securing lab space more often to enhance flexibility in the scheduling of lab or other hands-on modelling activities.

5. CONCLUDING REMARKS

Physics is a discipline in which many students struggle, partly because they are filled with conflicting preconceptions. Therefore, it is important to find ways of teaching that are more effective in dislodging those unviable conceptions while helping students successfully learn both the fundamentals and the culture of science in general, and physics in particular.

Keeping limitations in mind, this research seems to indicate that novice modelling instruction didn’t produce better or worse learning gains and attitude shifts than interactive engagement or regular instruction with limited interactive learning, at least in a statistically significant manner. Thus, it was shown that in the particular implementation and context of the researcher, novice modelling instruction performed equivalently on all grounds. Considering results reported in the literature, it is believed that formal training and further experience with the method would have led to improved results, therefore justifying further attempts at using modelling instruction under a deliberate practice of continuous improvement. Further research should, however, be done to confirm present results and reduce possible biases due to limitations of this study.

REFERENCE LIST

Adams, W. K., Perkins, K. K., Podolefsky, N. S., Dubson, M., Finkelstein, N. D., & Wieman, C. E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics - Physics Education Research*, *2*(1). https://doi.org/10.1103/PhysRevSTPER.2.010101

Adey, P., Shayer, M., & Yates, C. (2001). *Thinking science: student and teachers' materials for the CASE intervention* (3rd ed.). London: Nelson Thornes.

Brewe, E. (2002). *Inclusion of the Energy Thread in the Introductory Physics Curriculum: An Example of Long-Term Conceptual and Thematic Coherence.* Arizona State University. Retrieved March 19, 2020, from https://www.compadre.org/Repository/document/ServeFile.cfm?ID=4761&DocID=205

Brewe, E. (2006). Modeling theory applied; modeling instruction in university physics. *arXiv Preprint Physics/0602086*. Retrieved from https://arxiv.org/abs/physics/0602086

Brewe, E. (2008). Modeling theory applied: Modeling Instruction in introductory physics. *American Journal of Physics*, *76*(12), 1155. https://doi.org/10.1119/1.2983148

Brewe, E., Kramer, L., & O'Brien, G. (2008). CLASS Shifts in Modeling Instruction. *AIP Conference Proceedings*, 79–82. https://doi.org/10.1063/1.3021278

Brewe, E., Kramer, L., & O'Brien, G. (2009). Modeling instruction: Positive attitudinal shifts in introductory physics measured with CLASS. *Physical Review Special Topics - Physics Education Research*, *5*(1). https://doi.org/10.1103/PhysRevSTPER.5.013102

Brewe, E., Sawtelle, V., Kramer, L. H., O'Brien, G. E., Rodriguez, I., & Pamelá, P. (2010). Toward equity through participation in Modeling Instruction in introductory university physics. *Physical Review Special Topics - Physics Education Research*, *6*(1). https://doi.org/10.1103/PhysRevSTPER.6.010106

Camp, C., & Clement, J. J. (2010). *Preconceptions in Mechanics: Lessons Dealing with Students' Conceptual Difficulties* (2nd ed.). American Association of Physics Teachers.

Chabay, R. W., & Sherwood, B. A. (2015). *Matter & interactions* (Fourth). Wiley.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hoboken: Taylor and Francis.

Coletta, V. P., & Phillips, J. A. (2005). Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability. *American Journal of Physics*, *73*(12), 1172–1182. https://doi.org/10.1119/1.2117109

de la Garza, J., & Alarcon, H. (2010). Assessing Students' Attitudes In A College Physics Course In Mexico. *AIP Conference Proceedings, 1289*, 129–132. https://doi.org/10.1063/1.3515178

DeHaan, R. L. (2005). The Impending Revolution in Undergraduate Science Education. *Journal of Science Education and Technology*, *14*(2), 253–269. https://doi.org/10.1007/s10956-005-4425-3

Desbien, D. M. (2002, January 1). *Modeling discourse management compared to other classroom management styles in university physics*. ProQuest Dissertations Publishing.

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences*, *116*(39), 19251–19257. https://doi.org/10.1073/pnas.1821936116

Develaki, M. (2007). The model-based view of scientific theories and the structuring of school science programmes. *Science & Education*, *16*(7-8), 725–749.

Gerjets, P., & Scheiter, K. (2003). Goal Configurations and Processing Strategies as Moderators Between Instructional Design and Cognitive Load: Evidence From Hypertext-Based Instruction. *Educational Psychologist*, *38*(1), 33–41. https://doi.org/10.1207/S15326985EP3801_5

Giere, R. N. (1988). *Explaining Science: A Cognitive Approach* (Ser. Science and its conceptual foundations series). University of Chicago Press.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64–74. https://doi.org/10.1119/1.18809

Halloun, I. A., & Hestenes, D. (1985a). Common sense concepts about motion. *American Journal of Physics*, *53*(11), 1056–1065. https://doi.org/10.1119/1.14031

Halloun, I. A., & Hestenes, D. (1985b). The initial knowledge state of college physics students. *American Journal of Physics*, *53*(11), 1043–1055. https://doi.org/10.1119/1.14030

Halloun, I. A., & Hestenes, D. (1987). Modeling instruction in mechanics. *American Journal of Physics*, *55*, 455–462. http://dx.doi.org/10.1119/1.15130

Henderson, C. (2002). Common Concerns About the Force Concept Inventory. *The Physics Teacher*, *40*(9), 542–547. https://doi.org/10.1119/1.1534822

Hestenes, D. (1987a). Foundations of Mechanics. In *New Foundations for Classical Mechanics* (pp. 574–602). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-94-009-4802-0_9

Hestenes, D. (1987b). Toward a modeling theory of physics instruction. *American Journal of Physics*, *55*(5), 440–454. http://dx.doi.org/10.1119/1.15129

Hestenes, D. (1992). Modeling games in the Newtonian world. *American Journal of Physics*, *60*(8), 732–748. https://doi.org/10.1119/1.17080

Hestenes, D. (1997). Modeling methodology for physics teachers. *AIP Conference Proceedings*, *399*, 935–958. https://doi.org/10.1063/1.53196

Hestenes, D. (2000). *Findings of the modeling workshop project, 1994 - 2000*. One section of an NSF final report.

Hestenes, D. (2006). Notes for a modeling theory. In *Proceedings of the 2006 GIREP conference: Modeling in physics and physics education* (Vol. 31, p. 27). Retrieved from https://www.researchgate.net/profile/David_Hestenes/publication/253847244_Notes_for_a_Modeling_Theory_of_Science_Cognition_and_Instruction/links/551ec6f10cf2a2d9e140296c.pdf

Hestenes, D., & Halloun, I. (1995). Interpreting the force concept inventory: A response to March 1995 critique by Huffman and Heller. *The Physics Teacher*, *33*(8), 502–506. Ariane Articles. https://doi.org/10.1119/1.2344278

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, *30*(3), 141–158. https://doi.org/10.1119/1.2343497

Jackson, J., Dukerich, L., & Hestenes, D. (2008). Modeling instruction: An effective model for science education. *Science Educator*, *17*(1), 10.

Lasry, N., Rosenfield, S., Dedic, H., Dahan, A., & Reshef, O. (2011). The puzzling reliability of the Force Concept Inventory. *American Journal of Physics*, *79*(9), 909–912. https://doi.org/10.1119/1.3602073

Lattery, M. J. (2017). *Deep Learning in Introductory Physics: Exploratory Studies of Model-Based Reasoning* (Ser. Science & engineering education sources). IAP, Information Age Publishing.

Levene, H. (1960). Robust tests for equality of variances. In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, I. Olkin et al. eds., Stanford University Press, 278-292.

Madsen, A., McKagan, S. B., & Sayre, E. C. (2015). How physics instruction impacts students' beliefs about learning physics: A meta-analysis of 24 studies. *Physical Review Special Topics - Physics Education Research*, *11*(1). https://doi.org/10.1103/PhysRevSTPER.11.010115

Madsen, A., McKagan, S. B., & Sayre, E. C. (2017a). Best Practices for Administering Concept Inventories. *The Physics Teacher*, *55*(9), 530–536. https://doi.org/10.1119/1.5011826

Madsen, A., McKagan, S. B., & Sayre, E. C. (2017b). Resource Letter RBAI-1: Research-Based Assessment Instruments in Physics and Astronomy. *American Journal of Physics*, *85*(4), 245–264. https://doi.org/10.1119/1.4977416

Malone, K. L. (2008). Correlations among knowledge structures, force concept inventory, and problem-solving behaviors. *Physical Review Special Topics - Physics Education Research*, *4*(2), 020107. https://doi.org/10.1103/PhysRevSTPER.4.020107

Marx, J. D., & Cummings, K. (2007). Normalized change. *American Journal of Physics*, *75*(1), 87–91. https://doi.org/10.1119/1.2372468

McDermott, L. C. (2001). Oersted Medal Lecture 2001: “Physics Education Research—The Key to Student Learning.” *American Journal of Physics*, *69*(11), 1127–1137. https://doi.org/10.1119/1.1389280

McKagan, S. B., Perkins, K. K., & Wieman, C. E. (2007). Reforming a large lecture modern physics course for engineering majors using a PER-based design. *AIP Conference Proceedings*, *883*, 34–37. https://doi.org/10.1063/1.2508685

Megowan, M. C. (2007). *Framing discourse for optimal learning in science and mathematics*. (Doctoral dissertation, Arizona State University).

Megowan, M. C. (2010). The modeling method of instruction in physics: How to do it! *Proceedings of the 2010 Chinese Association of Physics Education and Research Conference.*

Mayer, R. E. (2003). *Learning and Instruction*. Upper Saddle River, NJ: Merrill.

Neidorf, T., Arora, A., Erberber, E., Tsokodayi, Y., & Mai, T. (2020). Review of Research into Misconceptions and Misunderstandings in Physics and Mathematics. In *Student Misconceptions and Errors in Physics and Mathematics: Exploring Data from TIMSS and TIMSS Advanced* (pp. 11–20). Springer International Publishing. https://doi.org/10.1007/978-3-030-30188-0_2

Nissen, J. M., Talbot, R. M., Thompson, A. N., & Van Dusen, B. (2018). A comparison of Hake's g and Cohen's d for analyzing gains on concept inventories. *Physical Review Physics Education Research*, *14*(1), 010115. https://doi.org/10.1103/PhysRevPhysEducRes.14.010115

Otero, V. K., & Gray, K. E. (2008). Attitudinal gains across multiple universities using the Physics and Everyday Thinking curriculum. *Physical Review Special Topics - Physics Education Research*, *4*(2), 020104. https://doi.org/10.1103/PhysRevSTPER.4.020104

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, *38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Perkins, K. K. (2005). Correlating Student Beliefs With Student Learning Using The Colorado Learning Attitudes about Science Survey. *AIP Conference Proceedings*, *790*, 61–64. https://doi.org/10.1063/1.2084701

Perret-Clermont, A. N., Carugati, F., & Oates, J. (2004). A socio-cognitive perspective on learning and cognitive development. In *Cognitive and language development in children* (Vol. 8, pp. 305-332). Oxford: Blackwell.

Piaget, J. (1950). *The Psychology of Intelligence*. London: Routledge.

Redish, E. F., Saul, J. M., & Steinberg, R. N. (1998). Student expectations in introductory physics. *American Journal of Physics*, *66*(3), 212–224. https://doi.org/10.1119/1.18847

Rimoldini, L. G., & Singh, C. (2005). Student understanding of rotational and rolling motion concepts. *Physical Review Special Topics - Physics Education Research*, *1*(1). https://doi.org/10.1103/PhysRevSTPER.1.010102

Roth, W.-M., & McGinn, M. K. (1998). Inscriptions: Toward a Theory of Representing as Social Practice. *Review of Educational Research, 68*(1), 35-59. https://doi.org/10.3102/00346543068001035

Sadler, P. M., & Tai, R. H. (2001). Success in introductory college physics: The role of high school preparation. *Science Education*, *85*(2), 111–136. https://doi.org/10.1002/1098-237X(200103)85:2<111::AID-SCE20>3.0.CO;2-O

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical methods* (8th ed.). Iowa State University Press.

Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*, *69*(347), 730-737.

Torkia-Lagacé, M. (1981). *La Pensée formelle chez les étudiants de collège I : objectif ou réalité?* Research report, Cégep de Limoilou.

U.S. Department of Education, Office of Educational Research and Improvement (2001). *Exemplary and Promising Science Programs, 2001*. Washington, DC. Retrieved from https://eric.ed.gov/?id=ED460863

Vanier College (2015). *Strategic Plan 2015-2020.* Retrieved from https://www.vaniercollege.qc.ca/strategic-plan/strategic-plan-2015-2020/

Vesenka, J., Beach, P., Munoz, G., Judd, F., & Key, R. (2002). A comparison between traditional and "modeling" approaches to undergraduate physics instruction at two universities with implications for improving physics teacher preparation. *J Phys Teach Educ Online 1*(1): 3–7. Retrieved from http://www2.phy.ilstu.edu/~wenning/jpteo/issues/june2002.html

Von Korff, J., Archibeque, B., Gomez, K. A., Heckendorf, T., McKagan, S. B., Sayre, E. C., … Sorell, L. (2016). Secondary Analysis of Teaching Methods in Introductory Physics: a 50k-Student Study. *American Journal of Physics*, *84*(12), 969–974. https://doi.org/10.1119/1.4964354

Vygotsky, L. S. (1962). *Thought and Language*. Cambridge, MA: MIT Press.

Vygotsky, L. S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.

Wells, M. (1987). *Modeling Instruction in High School Physics.* Arizona State University. Retrieved March 26, 2020, from http://modeling.asu.edu/thesis/WellsMalcolm_dissertation.doc

Wells, M., Hestenes, D., & Swackhamer, G. (1995). A modeling method for high school physics instruction. *American Journal of Physics*, *63*(7), 606–619. https://doi.org/10.1119/1.17849

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, *92*(5), 941–967. https://doi.org/10.1002/sce.20259

Wright, B., & Linacre, J. (1989). Observations are always ordinal; measurements, however, must be interval. *Archives Of Physical Medicine & Rehabilitation*, *70*(12), 857-860.

APPENDIX A. DETAILED STATISTICAL DATA ANALYSES

1. DESCRIPTIVE STATISTICS

Descriptive statistics for instruments that were used are presented in tables 29 to 32.

Table 29. Descriptive statistics for the FCI test, for different modes of instruction.

Variable	**Group**	***N***	***N****	**Mean**	**SE Mean**	**StDev**
FCI Pre-Score (%)	Act	51	10	39.34	2.82	20.16
	Mod	20	0	44.34	4.29	19.20
	Reg	61	22	34.48	1.82	14.25
FCI Post-Score (%)	Act	51	10	64.25	2.69	19.19
	Mod	20	0	66.17	4.74	21.20
	Reg	61	22	53.55	1.97	15.40
FCI Average Gain <*G*> (%)	Act	51	10	24.90	2.20	15.71
	Mod	20	0	21.84	4.31	19.27
	Reg	61	22	19.07	1.74	13.55
FCI Normalized Gain of Averages <*g*> (%)	Act	51	10	41.06	*	*
	Mod	20	0	39.22	*	*
	Reg	61	22	29.11	*	*
FCI Average of Normalized Gains g_{avg} (%)	Act	51	10	43.39	3.39	24.21
	Mod	20	0	40.56	6.33	28.31
	Reg	61	22	29.13	2.38	18.59
FCI Normalized Change of Averages <*c*> (%)	Act	51	10	41.06	*	*
	Mod	20	0	39.22	*	*
	Reg	61	22	29.11	*	*
FCI Average of Normalized Changes c_{avg} (%)	Act	51	10	43.34	3.41	24.32
	Mod	20	0	40.84	6.20	27.73
	Reg	61	22	29.05	2.40	18.75
FCI Gain Effect Size d_G	Act	51	10	1.27	*	*
	Mod	20	0	1.08	*	*
	Reg	61	22	1.29	*	*
FCI Corrected Gain Effect Size $d_{G,\ dep}$	Act	51	10	1.26	*	*
	Mod	20	0	1.08	*	*
	Reg	61	22	1.28	*	*

Table 30. Descriptive statistics for the RRMCS test, for different modes of instruction.

Variable	**Group**	***N***	***N****	**Mean**	**SE Mean**	**StDev**
RRMCS Pre-Score (%)	Act	0	61	*	*	*
	Mod	20	0	35.66	3.26	14.60
	Reg	0	83	*	*	*
RRMCS Post-Score (%)	Act	0	61	*	*	*
	Mod	20	0	48.49	4.59	20.53
	Reg	0	83	*	*	*
RRMCS Average Gain <G> (%)	Act	0	61	*	*	*
	Mod	20	0	12.83	2.70	12.05
	Reg	0	83	*	*	*
RRMCS Normalized Gain of Averages <*g*> (%)	Act	0	61	*	*	*
	Mod	20	0	19.95	*	*
	Reg	0	83	*	*	*
RRMCS Average of Normalized Gains g_{avg} (%)	Act	0	61	*	*	*
	Mod	20	0	23.41	6.21	27.77
	Reg	0	83	*	*	*
RRMCS Normalized Change of Averages <*c*> (%)	Act	0	61	*	*	*
	Mod	20	0	19.95	*	*
	Reg	0	83	*	*	*
RRMCS Average of Normalized Changes c_{avg} (%)	Act	0	61	*	*	*
	Mod	20	0	22.46	6.51	29.11
	Reg	0	83	*	*	*
RRMCS Gain Effect Size d_G	Act	0	61	*	*	*
	Mod	20	0	0.72	*	*
	Reg	0	83	*	*	*
RRMCS Corrected Gain Effect Size $d_{G, dep}$	Act	0	61	*	*	*
	Mod	20	0	0.65	*	*
	Reg	0	83	*	*	*

Table 31. Descriptive statistics for the 78%-common final exam, for different modes of instruction.

Variable	**Group**	***N***	***N****	**Mean**	**SE Mean**	**StDev**
78%-Common Final Exam Average Grade (%)	Act	56	5	67.61	2.72	20.37
	Mod	20	0	62.53	3.50	15.67
	Reg	79	4	68.26	2.11	18.79

Table 32. Descriptive statistics for the CLASS test for different modes of instruction.

Variable	**Group**	***N***	***N****	**Mean**	**SE Mean**	**StDev**
CLASS Favourable Pre-Score (%)	Act	32	29	57.71	2.91	16.47
	Mod	19	1	62.26	3.73	16.26
	Reg	58	25	57.34	2.17	16.52
CLASS Favourable Post-Score (%)	Act	32	29	56.35	3.35	18.98
	Mod	19	1	60.67	3.72	16.24
	Reg	58	25	57.73	2.12	16.16
CLASS Average Favourable Attitude Shift $\langle S \rangle$ (%)	Act	32	29	-1.37	3.16	17.87
	Mod	19	1	-1.57	2.76	12.02
	Reg	58	25	0.39	1.61	12.27
CLASS Favourable Shift Effect Size d_S	Act	32	29	-0.08	*	*
	Mod	19	1	-0.10	*	*
	Reg	58	25	0.02	*	*
CLASS Favourable Shift Corrected Effect Size $d_{S,\ dep}$	Act	32	29	-0.08	*	*
	Mod	19	1	-0.10	*	*
	Reg	58	25	0.02	*	*
CLASS Unfavourable Pre-Score (%)	Act	32	29	20.27	1.81	10.26
	Mod	19	1	17.62	2.64	11.52
	Reg	58	25	18.88	1.27	9.70
CLASS Unfavourable Post-Score (%)	Act	32	29	20.32	2.45	13.85
	Mod	19	1	20.31	2.46	10.73
	Reg	58	25	18.88	1.45	11.01
CLASS Average Unfavourable Attitude Shift $\langle S \rangle$ (%)	Act	32	29	0.03	2.96	16.75
	Mod	19	1	2.71	1.65	7.19
	Reg	58	25	0.01	1.21	9.24
CLASS Unfavourable Shift Effect Size d_S	Act	32	29	0.00	*	*
	Mod	19	1	0.24	*	*
	Reg	58	25	0.00	*	*
CLASS Unfavourable Corrected Shift Effect Size $d_{S,\ dep}$	Act	32	29	0.00	*	*
	Mod	19	1	0.24	*	*
	Reg	58	25	0.00	*	*

2. COMPARABILITY OF GROUPS

Before doing any statistical analysis, we must verify that groups are comparable. As a minimal indicator of comparability, a one-way analysis of variance (ANOVA) was performed, with $\alpha = .05$, on both FCI pre-test scores and CLASS pre-test favourable scores.

2.1 FCI pre-test scores

The normality of FCI pre-test scores was first assessed with the Anderson-Darling test (Stephens, 1974) for all three groups: modelling instruction (Mod), interactive engagement (Act), and regular instruction (Reg). Results are presented in figure 19. The absolute skewness varies between 0.12 and 1.00 whereas the absolute kurtosis varies between 0.34 and 0.94. Yet, only one dataset (Act) has a p-value (< 0.005) below $\alpha = .05$ for normality. This is however sufficient to proceed to a test for equal variances without assuming a normal distribution. The test for equal variances fails to reject the null hypothesis with a p-value larger than .05 with both multiple comparisons and Levene's test (Levene, 1960) (see figure 20). Because of that, the ANOVA has been performed with the assumption of equal variances (along with an error rate for comparisons set at 5 in both cases). Results are in tables 33 and 34, and figures 21 and 22. The p-value (.072) for the ANOVA is above $\alpha = .05$ and none of the many tests performed on FCI pre-test scores leads us to discard the null hypothesis (equal means), except for the Fisher pairwise comparison ($p = .031$) and the Hsu comparison with the best in the case of the FCI pre-scores of the Mod versus the Reg groups ($p = .029$). Considering all tests and the ANOVA p-value, it seems reasonable to hypothesize the comparability of groups in terms of initial physics conceptual understanding and to proceed with further analysis of the data.

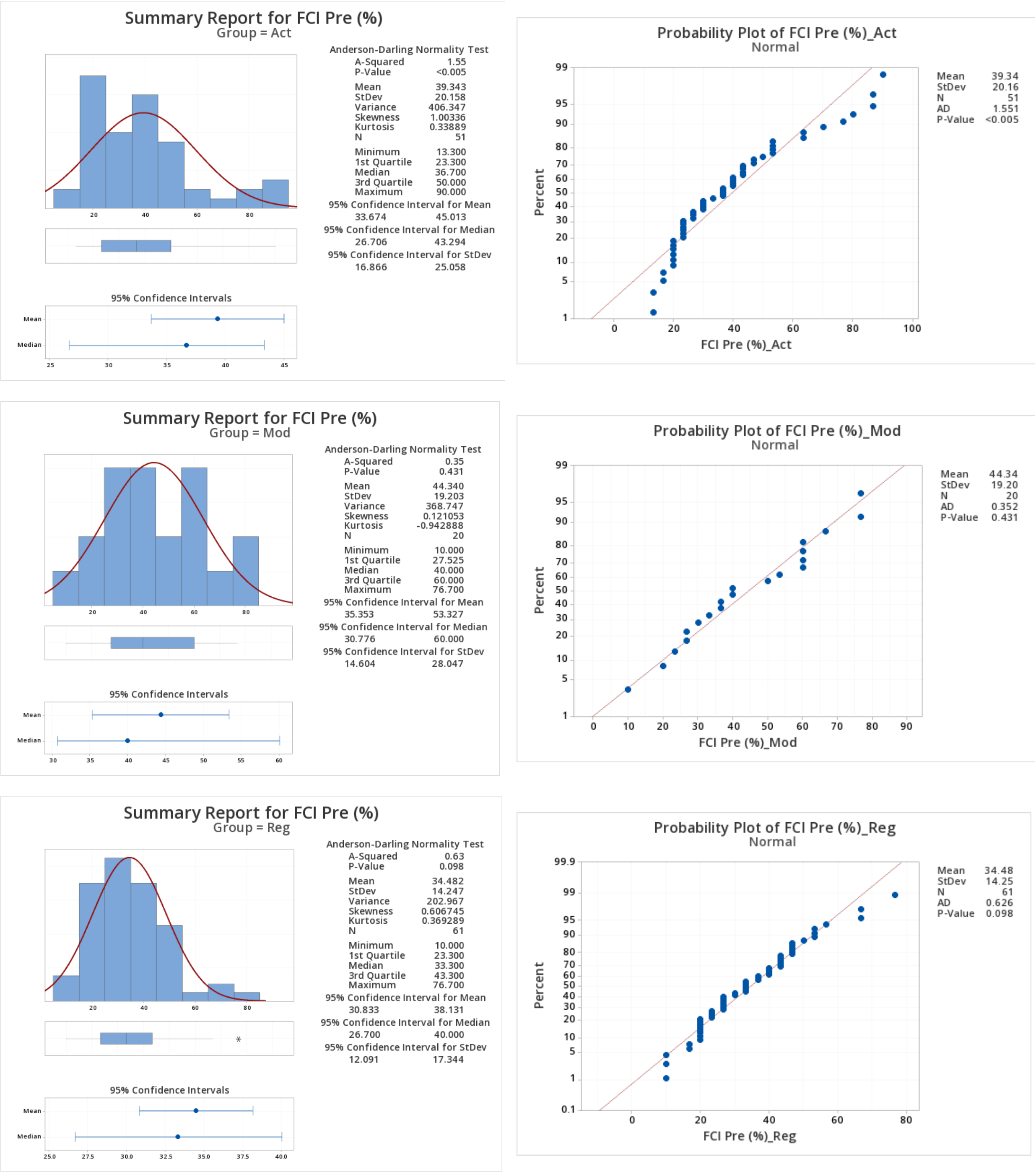


Figure 19. Summary descriptive statistics and probability plots of normality for FCI pre-test scores.

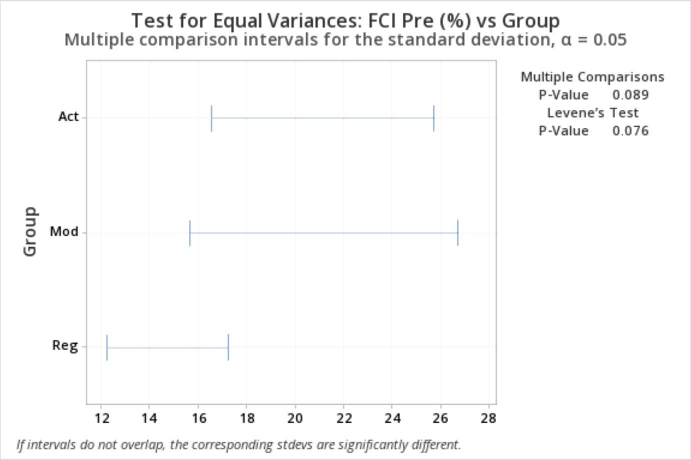


Figure 20. Test for equal variances for FCI pre-test scores.

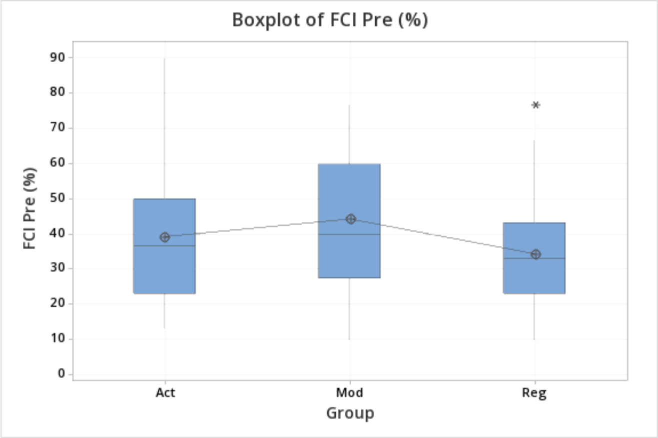
Table 33. Group means (%) on FCI pre-test scores.

Group	***N***	**Mean**	**StDev**	**95% CI**
Act	51	39.34	20.16	(34.50, 44.19)
Mod	20	44.34	19.20	(36.60, 52.08)
Reg	61	34.48	14.25	(30.05, 38.91)

Pooled StDev = 17.4990

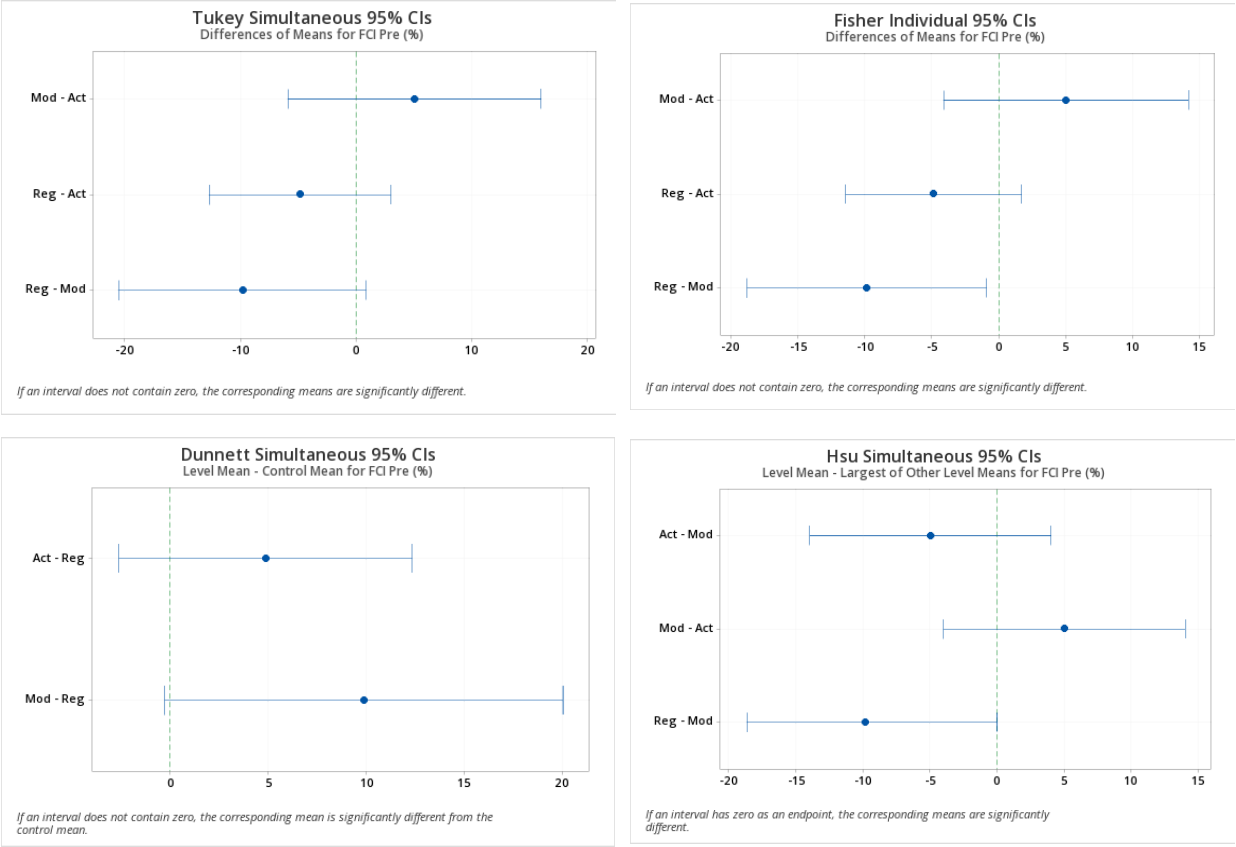
Table 34. Analysis of variance on FCI pre-test scores.

Source	**DF**	**Seq SS**	**Contribution**	**Adj SS**	**Adj MS**	***F*-Value**	***P*-Value**
Group	2	1648	4.01%	1648	824.0	2.69	0.072
Error	129	39502	95.99%	39502	306.2		
Total	131	41150	100.00%				



Boxplot of FCI Pre (%)

Figure 21. Boxplot of FCI pre-test scores.



Tukey Simultaneous 95% CIs
Differences of Means for FCI Pre (%)

If an interval does not contain zero, the corresponding means are significantly different.

Fisher Individual 95% CIs
Differences of Means for FCI Pre (%)

If an interval does not contain zero, the corresponding means are significantly different.

Dunnett Simultaneous 95% CIs
Level Mean - Control Mean for FCI Pre (%)

If an interval does not contain zero, the corresponding mean is significantly different from the control mean.

Hsu Simultaneous 95% CIs
Level Mean - Largest of Other Level Means for FCI Pre (%)

If an interval has zero as an endpoint, the corresponding means are significantly different.

Figure 22. Tukey and Fisher pairwise comparisons, Dunnett multiple comparisons with the control (Reg), and Hsu multiple comparisons with the best (MCB) for FCI pre-test scores.

2.2 CLASS pre-test favourable scores

The normality of CLASS pre-test favourable scores was first assessed with the Anderson-Darling test (Stephens, 1974) for all three groups: modelling instruction (Mod), interactive engagement (Act), and regular instruction (Reg). Results are presented in figure 23. The CLASS pre-test favourable attitude scores present an absolute skewness varying between 0.12 and 0.38 whereas their absolute kurtosis varies between 0.24 and 0.61. All have p-values for normality above $\alpha = .05$ (lowest one is .07), justifying the assumption of a normal distribution to test for equal variances. Bartlett's test (Snedecor & Cochran, 1983) fails to reject the null hypothesis (see figure 24). Because of that, the ANOVA has been performed with the assumption of equal variances (along with an error rate for comparisons set at 5 in both cases). Results are in tables 35 and 36, and figures 25 and 26. The p-value (.514) for the ANOVA is above $\alpha = .05$ and none of the many tests performed on CLASS pre-test scores leads us to discard the null hypothesis (equal means). It thus seems reasonable to hypothesize the comparability of groups in terms of initial attitudes toward physics and to proceed with further analysis of the data.

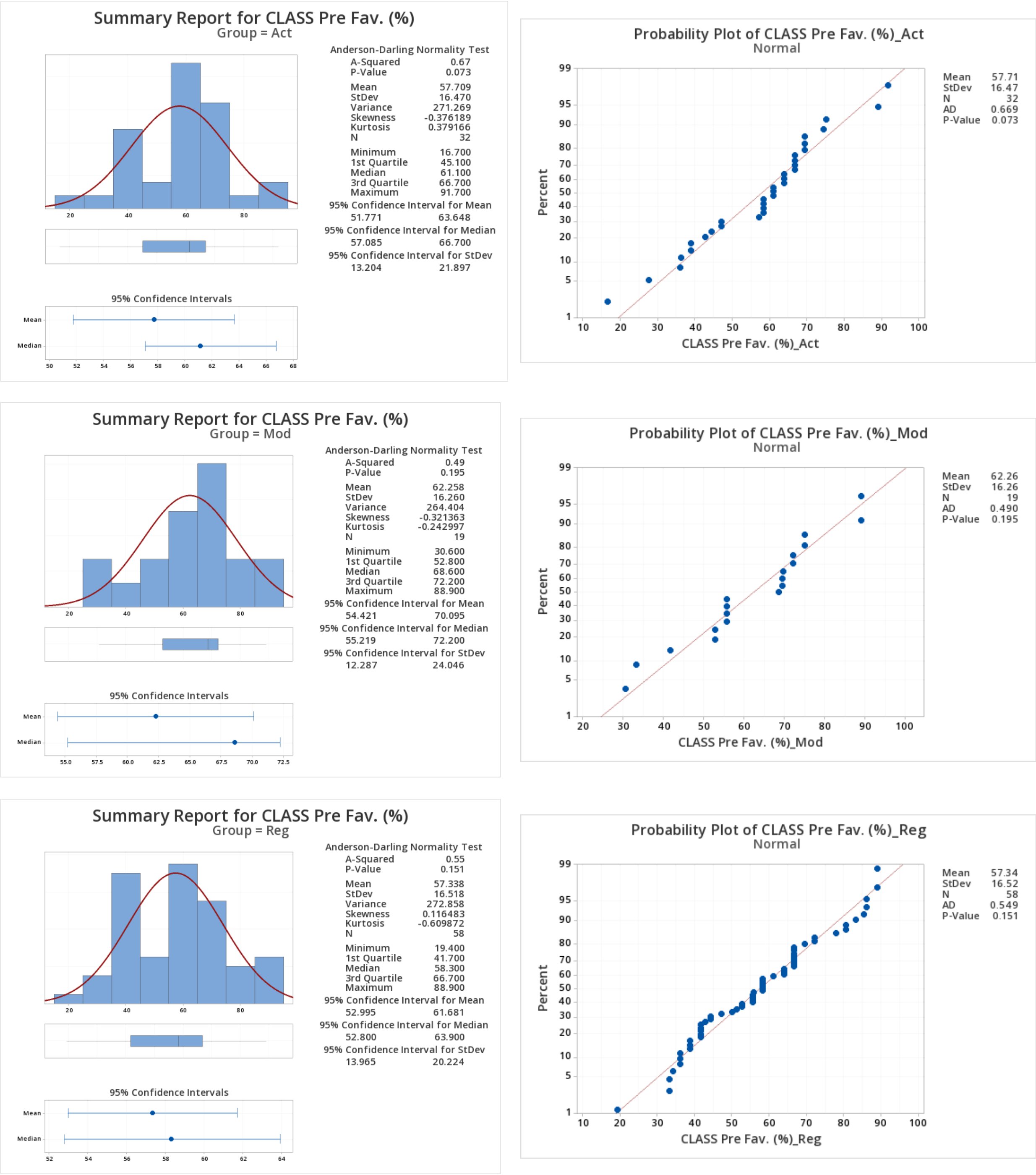


Figure 23. Summary descriptive statistics and probability plots of normality for CLASS pre-test favourable scores.

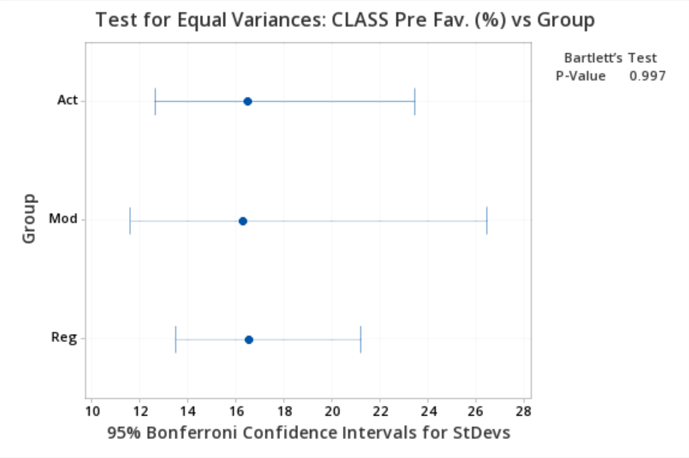


Figure 24. Test for equal variances for CLASS pre-test favourable scores.

Table 35. Group means (%) on CLASS pre-test favourable scores.

Group	***N***	**Mean**	**StDev**	**95% CI**
Act	32	57.71	16.47	(51.94, 63.48)
Mod	19	62.26	16.26	(54.77, 69.74)
Reg	58	57.34	16.52	(53.05, 61.62)

Pooled StDev = 16.4608

Table 36. Analysis of variance on CLASS pre-test favourable scores.

Source	**DF**	**Seq SS**	**Contribution**	**Adj SS**	**Adj MS**	***F*-Value**	***P*-Value**
Group	2	362.5	1.25%	362.5	181.2	0.67	0.514
Error	106	28721.5	98.75%	28721.5	271.0		
Total	108	29084.0	100.00%				

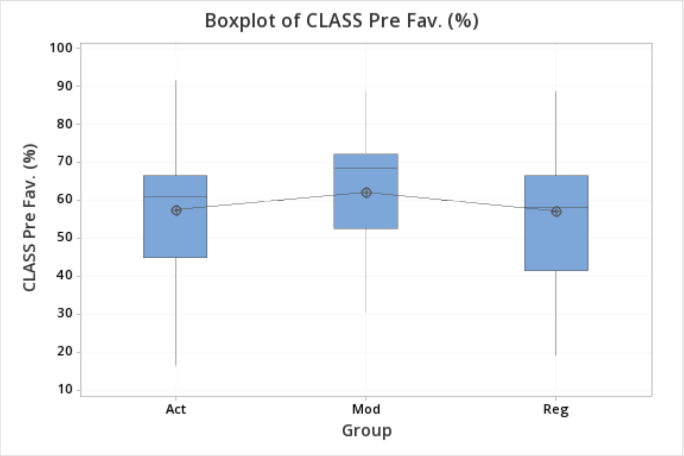


Figure 25. Boxplot of CLASS pre-test favourable scores.

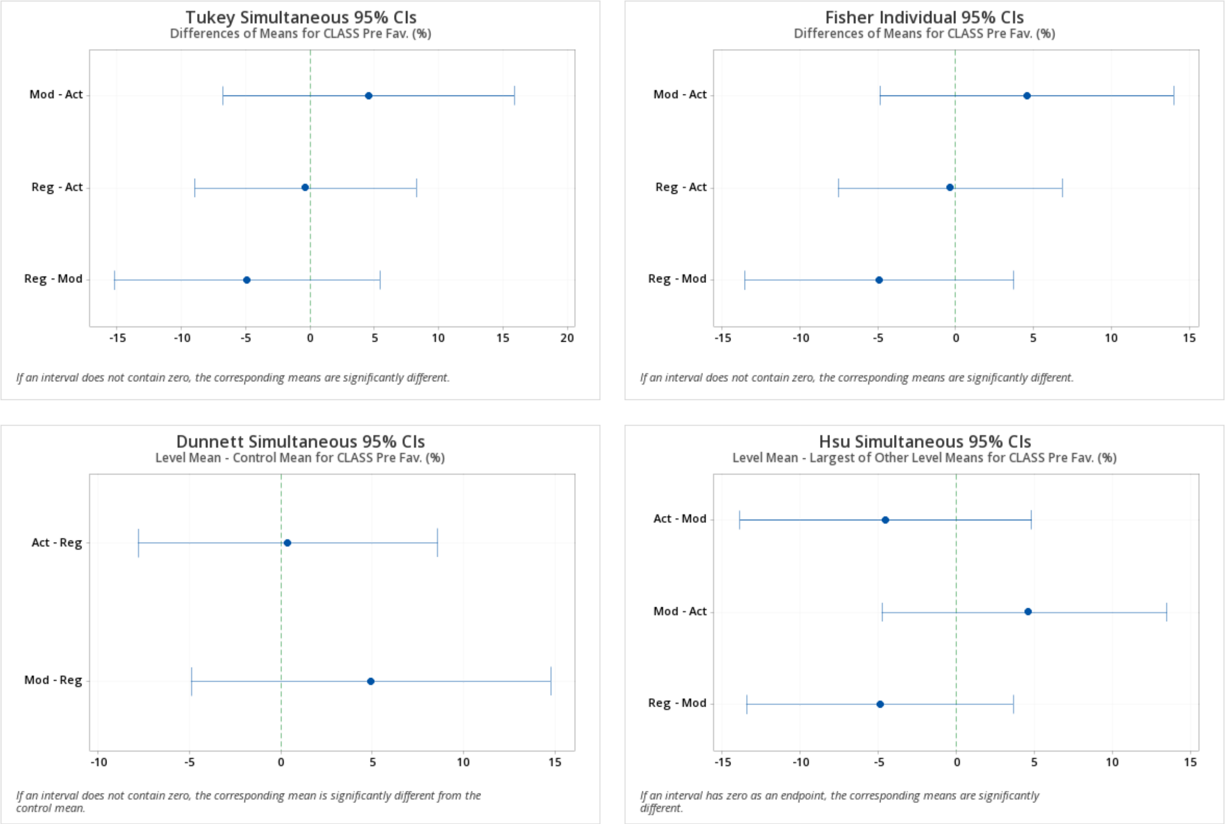
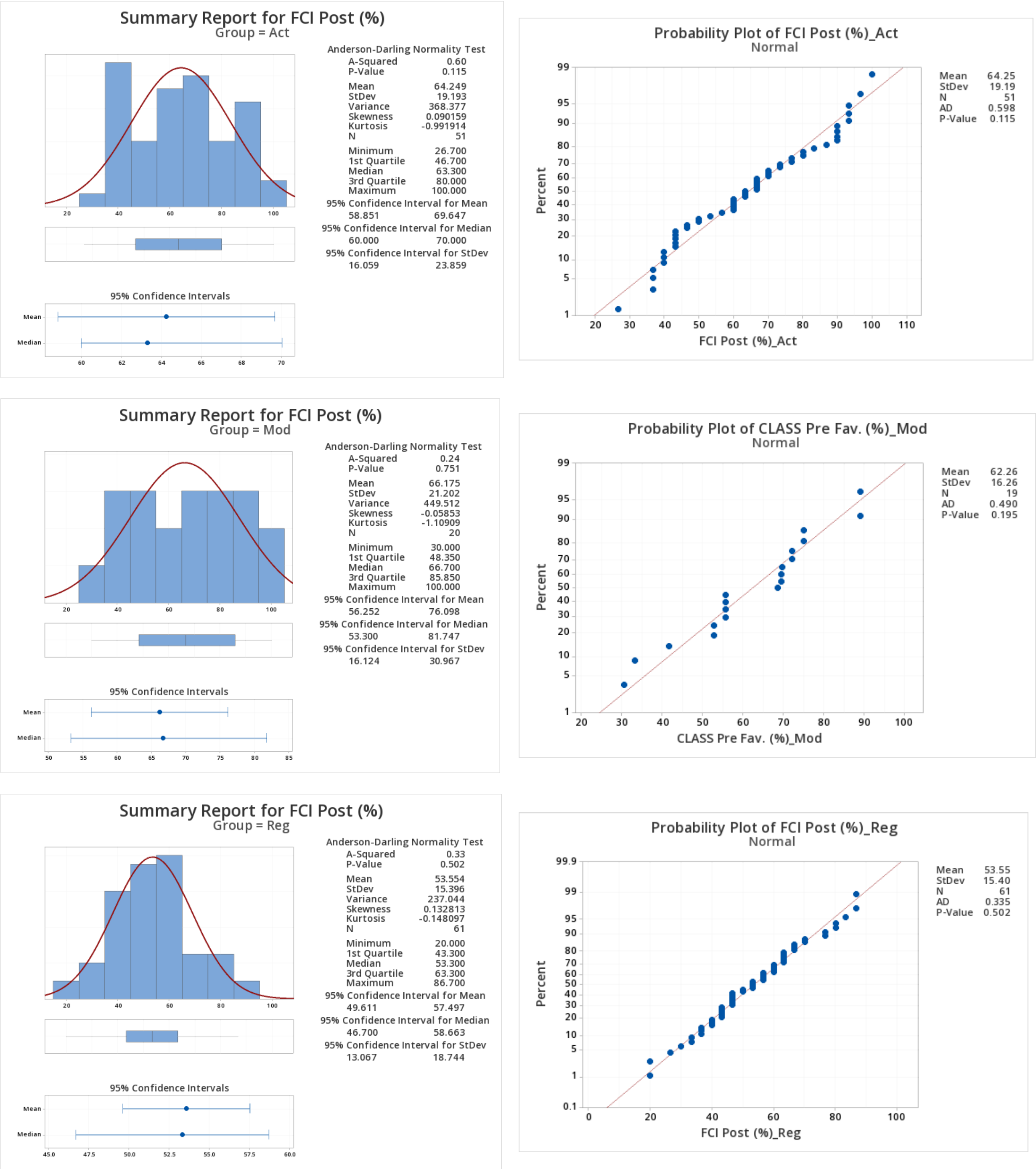


Figure 26. Tukey and Fisher pairwise comparisons, Dunnett multiple comparisons with the control (Reg), and Hsu multiple comparisons with the best (MCB) for CLASS pre-test favourable scores.

3. FIRST RESEARCH QUESTION: LEARNING OUTCOMES

3.1 Conceptual understanding through FCI raw scores

A one-way analysis of variance (ANOVA) with $\alpha = .05$ was performed on the FCI post-test scores. The normality was first assessed with the Anderson-Darling test (Stephens, 1974) for all three groups: modelling instruction (Mod), interactive engagement (Act), and regular instruction (Reg). Results are presented in figure 27. The absolute skewness varies between 0.06 and 0.13 whereas the absolute kurtosis varies between 0.15 and 1.11. All datasets have p-values above $\alpha = .05$ for normality, the lowest one being .115. We could, therefore, proceed to a test for equal variances with the assumption of a normal distribution. The Bartlett's test (Snedecor & Cochran, 1983) for equal variances finds a p-value of .126, which cannot reject the null hypothesis (see figure 28). Because of that, the ANOVA on FCI post-test scores was performed assuming equal variances (along with an error rate for comparisons set at 5). Results are in tables 37 and 38, and figures 29 and 30. We find a p-value of .002, lower than $\alpha = .05$, leading us to conclude that there is a statistically significant difference between groups. To locate this difference, we performed various tests. Looking at Tukey and Fisher pairwise comparisons, at Dunnet multiple comparisons with the Reg control, and at Hsu multiple comparisons with the best (Mod), all suggest a statistically (and practically) significant difference between the regular instruction group and both interactive engagement and modelling instruction groups. No significant difference is suggested between interactive engagement and modelling instruction groups, however.



Summary Report for FCI Post (%)
Group = Act

```
Anderson-Darling Normality Test
  A-Squared            0.60
  P-Value             0.115

  Mean              64.249
  StDev             19.193
  Variance         368.377
  Skewness        0.090159
  Kurtosis       -0.991914
  N                     51

  Minimum           26.700
  1st Quartile      46.700
  Median            63.300
  3rd Quartile      80.000
  Maximum          100.000
95% Confidence Interval for Mean
  58.851            69.647
95% Confidence Interval for Median
  60.000            70.000
95% Confidence Interval for StDev
  16.059            23.859
```

Probability Plot of FCI Post (%)_Act
Normal

```
Mean      64.25
StDev     19.19
N            51
AD        0.598
P-Value   0.115
```

Summary Report for FCI Post (%)
Group = Mod

```
Anderson-Darling Normality Test
  A-Squared            0.24
  P-Value             0.751

  Mean              66.175
  StDev             21.202
  Variance         449.512
  Skewness        -0.05853
  Kurtosis        -1.10909
  N                     20

  Minimum           30.000
  1st Quartile      48.350
  Median            66.700
  3rd Quartile      85.850
  Maximum          100.000
95% Confidence Interval for Mean
  56.252            76.098
95% Confidence Interval for Median
  53.300            81.747
95% Confidence Interval for StDev
  16.124            30.967
```

Probability Plot of CLASS Pre Fav. (%)_Mod
Normal

```
Mean      62.26
StDev     16.26
N            19
AD        0.490
P-Value   0.195
```

Summary Report for FCI Post (%)
Group = Reg

```
Anderson-Darling Normality Test
  A-Squared            0.33
  P-Value             0.502

  Mean              53.554
  StDev             15.396
  Variance         237.044
  Skewness        0.132813
  Kurtosis       -0.148097
  N                     61

  Minimum           20.000
  1st Quartile      43.300
  Median            53.300
  3rd Quartile      63.300
  Maximum           86.700
95% Confidence Interval for Mean
  49.611            57.497
95% Confidence Interval for Median
  46.700            58.663
95% Confidence Interval for StDev
  13.067            18.744
```

Probability Plot of FCI Post (%)_Reg
Normal

```
Mean      53.55
StDev     15.40
N            61
AD        0.335
P-Value   0.502
```

Figure 27. Summary descriptive statistics and probability plots of normality for FCI post-test scores.

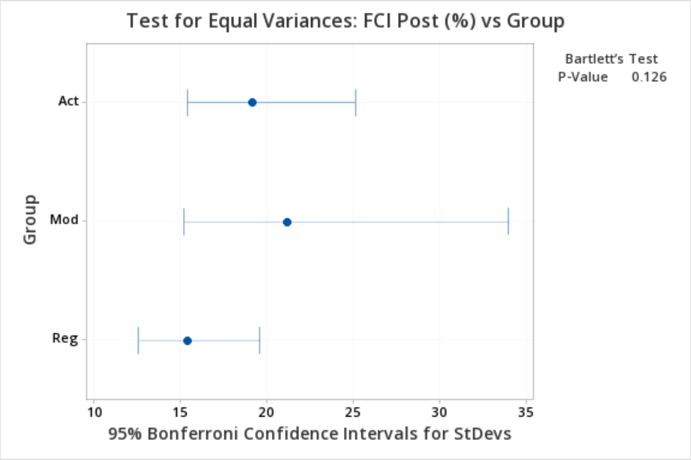


Figure 28. Test for equal variances for FCI post-test scores.

Table 37. Group means (%) on FCI post-test scores.

Group	***N***	**Mean**	**StDev**	**95% CI**
Act	51	64.25	19.19	(59.30, 69.20)
Mod	20	66.17	21.20	(58.27, 74.08)
Reg	61	53.55	15.40	(49.03, 58.08)

Pooled StDev = 17.8673

Table 38. Analysis of variance on FCI post-test scores.

Source	**DF**	**Seq SS**	**Contribution**	**Adj SS**	**Adj MS**	***F*-Value**	***P*-Value**
Group	2	4197	9.25%	4197	2098.3	6.57	0.00191
Error	129	41182	90.75%	41182	319.2		
Total	131	45379	100.00%				

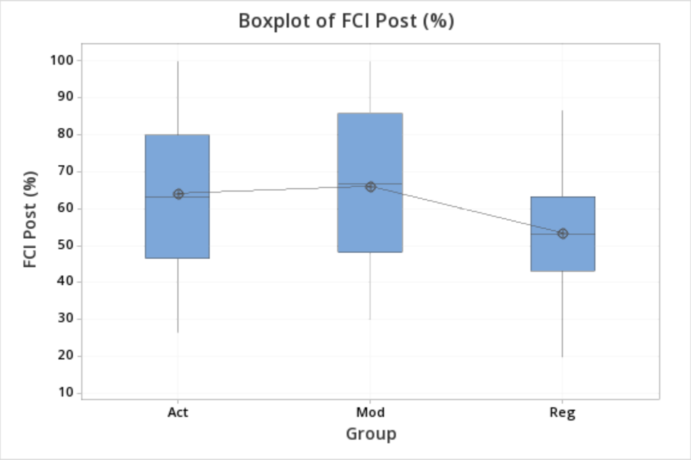


Figure 29. Boxplot of FCI post-test scores.

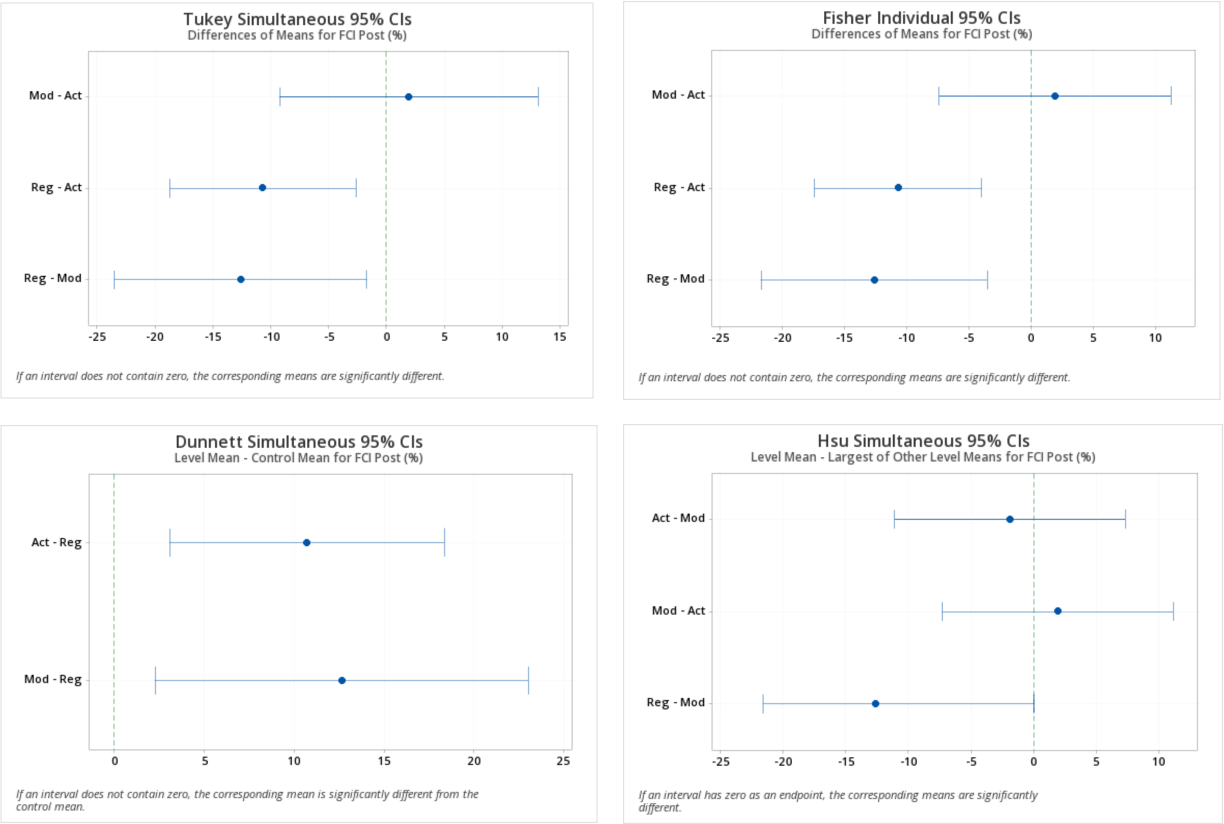


Figure 30. Tukey and Fisher pairwise comparisons, Dunnett multiple comparisons with the control (Reg), and Hsu multiple comparisons with the best (MCB) for FCI post-test scores.

3.2 Conceptual learning gains on the FCI test

3.2.1 Analysis of variance

A one-way ANOVA on normalized gains g was performed with $\alpha = .05$. The normality was first assessed with the Anderson-Darling test (Stephens, 1974) for all three groups: modelling instruction (Mod), interactive engagement (Act), and regular instruction (Reg). Results are presented in figure 31. The absolute skewness varies between 0.12 and 0.32 whereas the absolute kurtosis varies between 0.28 and 0.68. Furthermore, p-values for normality are all above $\alpha = .05$, the lowest being .341. The normality test thus fails to reject the null hypothesis. We, therefore, proceeded to a test for equal variances, assuming a normal distribution. This leads to a rejection of the null hypothesis with Bartlett's test (Snedecor & Cochran, 1983) giving a p-value equal to .035 (figure 32), lower than $\alpha = .05$, although not by much. Because of that, the ANOVA on normalized gains was performed without assuming equal variances (error rate for comparison set at 5). Results are presented in tables 39 and 40, and figures 33 and 34. We find a p-value of .004, lower than $\alpha = .05$, leading us to conclude that there is a statistically significant difference between groups. To locate this difference, we performed Games-Howell pairwise comparisons. This test suggests a statistically (and practically) significant difference between regular instruction and interactive engagement, but it didn't detect any difference between novice modelling instruction and either interactive engagement or regular instruction. The null hypothesis (equal means) cannot be rejected for novice modelling instruction.

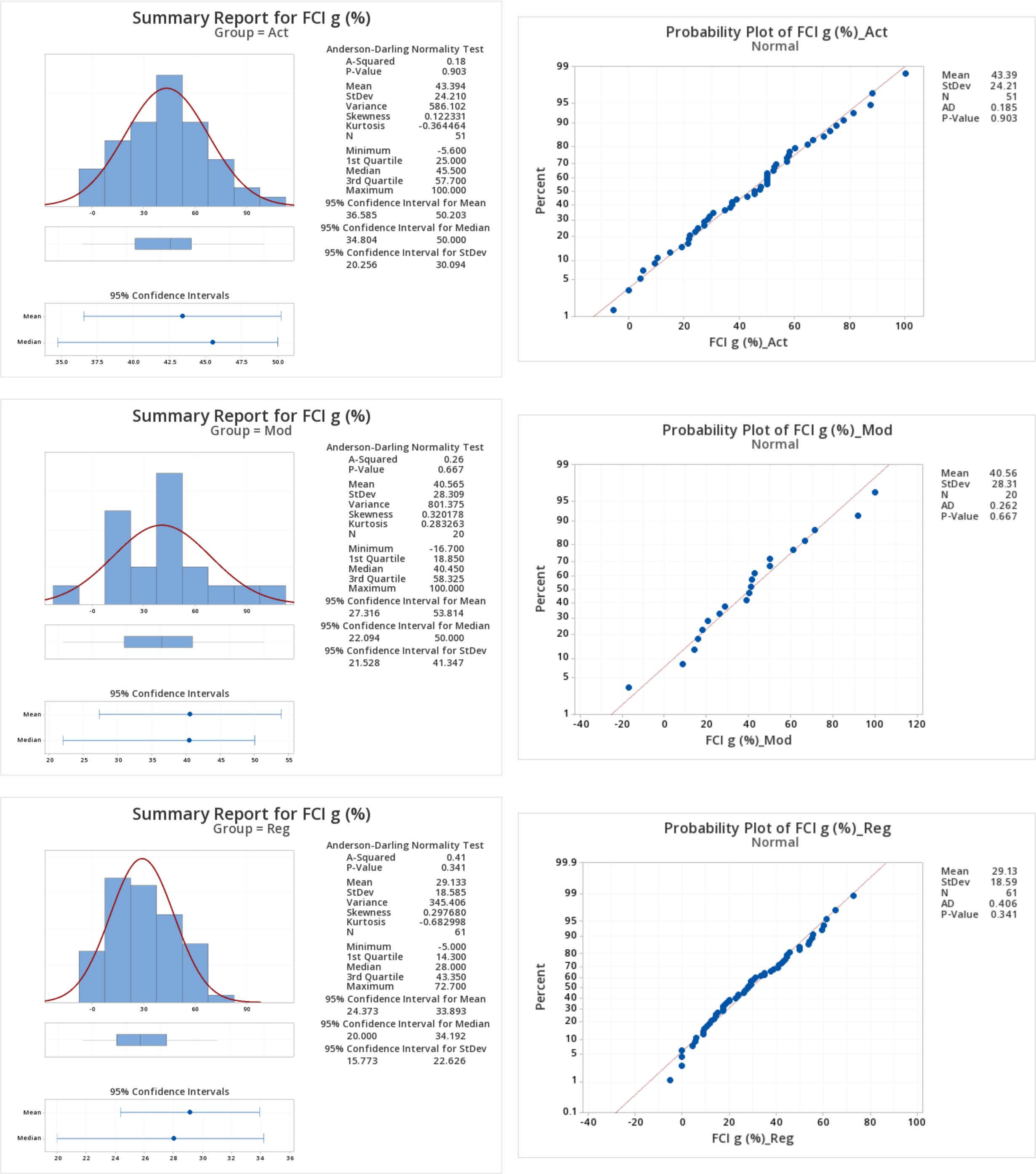


Figure 31. Summary descriptive statistics and probability plots of normality for FCI normalized gains.

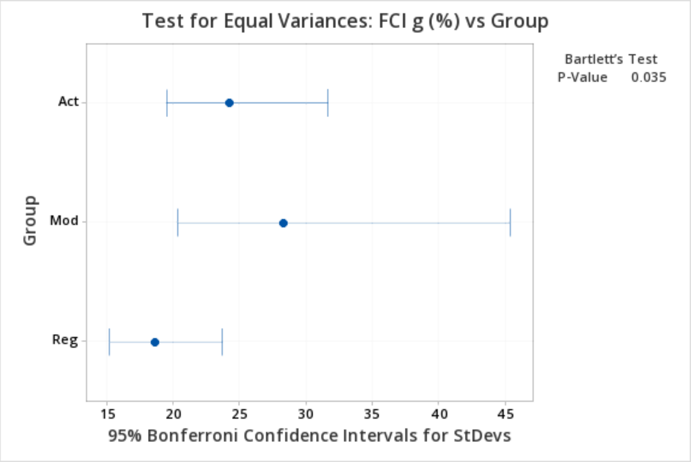


Figure 32. Test for equal variances for FCI normalized gains.

Table 39. Group means (%) on FCI normalized gains.

Group	***N***	**Mean**	**StDev**	**95% CI**
Act	51	43.39	24.21	(36.59, 50.20)
Mod	20	40.56	28.31	(27.32, 53.81)
Reg	61	29.13	18.59	(24.37, 33.89)

Table 40. Analysis of variance on FCI normalized gains (Welch's test).

Source	**DF Num**	**DF Den**	***F*-Value**	***P*-Value**
Group	2	47.4596	6.35	0.00359

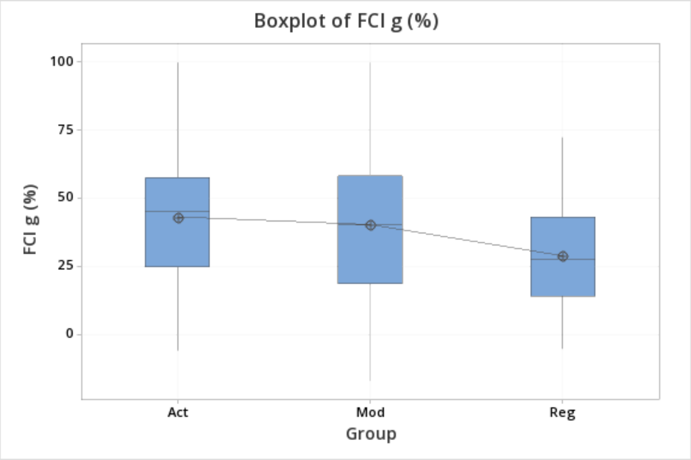


Figure 33. Boxplot of FCI normalized gains.

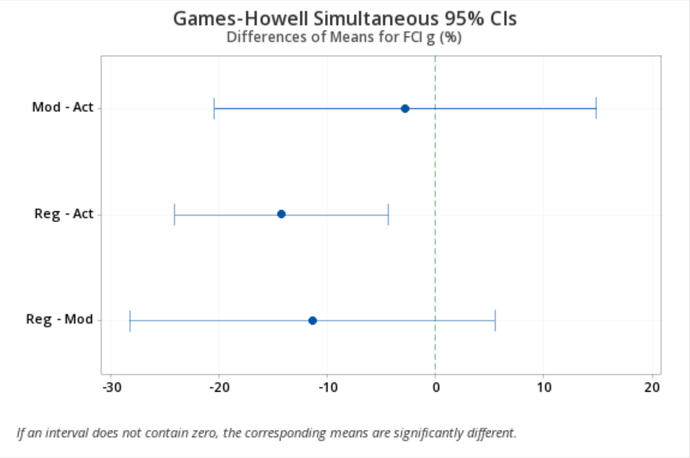


Figure 34. Games-Howell pairwise comparisons for FCI normalized gains.

3.2.2 Effect size

If we look into the effect size d_G based on the average gain and the pooled standard deviation of FCI pre- and post-test scores, and also into the corrected effect size $d_{G,\ dep}$ taking into account a possible dependence between the two scores modelled by Pearson's coefficient of correlation r, we find results of table 41. Figure 35 shows the matrix plots of FCI post-test scores versus pre-test scores leading to an evaluation of Pearson's r. As can be seen in table 41, both Cohen's effect size d_G and the corrected Cohen's effect size $d_{G,\ dep}$ are nearly identical within two decimals.

Table 41. Effect size for the FCI gain, for different modes of instruction.

Group	N	N^*	FCI Gain Effect Size d_G	FCI Gain Corrected Effect Size $d_{G,\ dep}$
Act	51	10	1.27	1.26
Mod	20	0	1.08	1.08
Reg	61	22	1.29	1.28

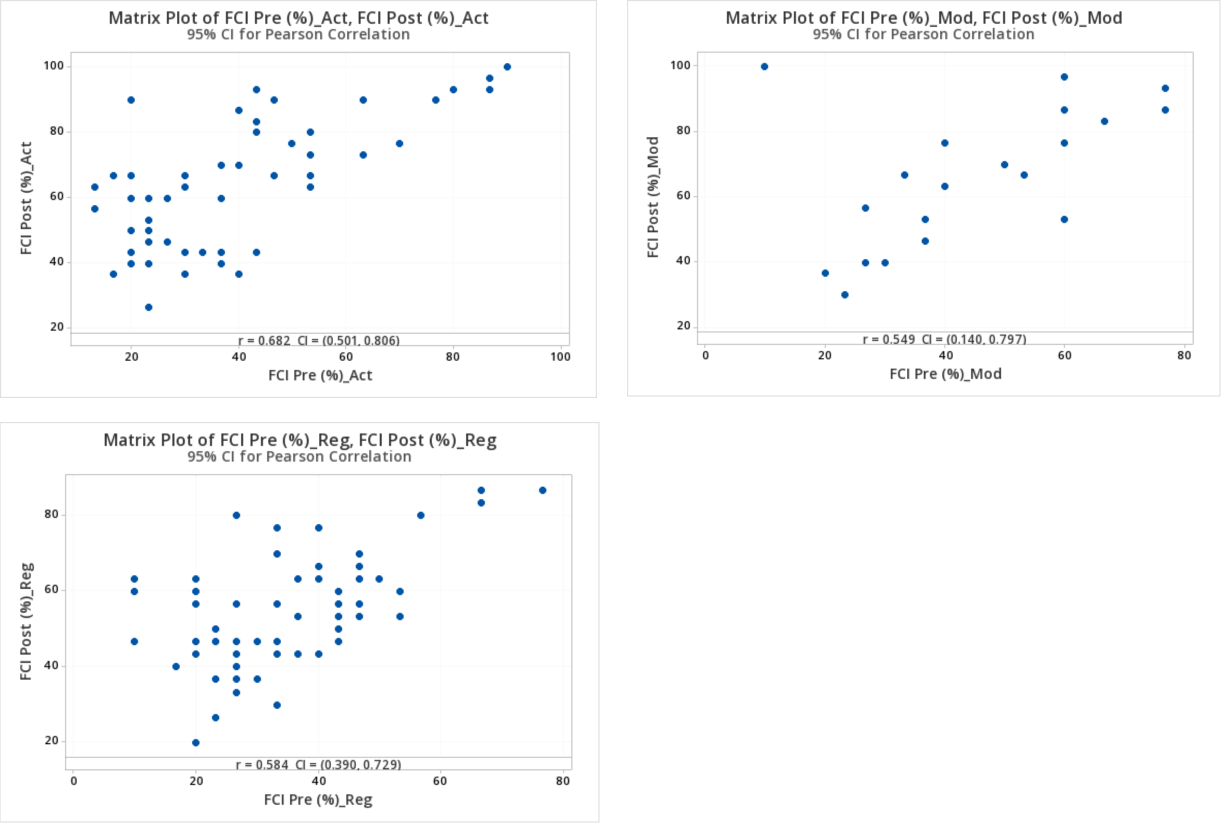


Figure 35. Matrix plots of FCI post-test scores versus pre-test scores.

3.3 Conceptual learning gains on the RRMCS test

3.3.1 Effect size

If we look into the effect size d_G based on the average gain and the pooled standard deviation of RRMCS pre- and post-test scores, and also into the corrected effect size $d_{G,\,dep}$ taking into account a possible dependence between the two scores modelled by Pearson's coefficient of correlation *r*, we find results of table 42. Figure 36 shows the matrix plot of RRMCS post-test scores versus pre-test scores leading to an evaluation of Pearson's *r*. As can be seen in table 42, both Cohen's effect size d_G and the corrected Cohen's effect size $d_{G,\,dep}$ are relatively similar.

Table 42. Effect size for the RRMCS gain, for novice modelling instruction.

Group	***N***	***N****	**RRMCS Gain Effect Size d_G**	**RRMCS Gain Corrected Effect Size $d_{G,\,dep}$**
Mod	20	0	0.72	0.65

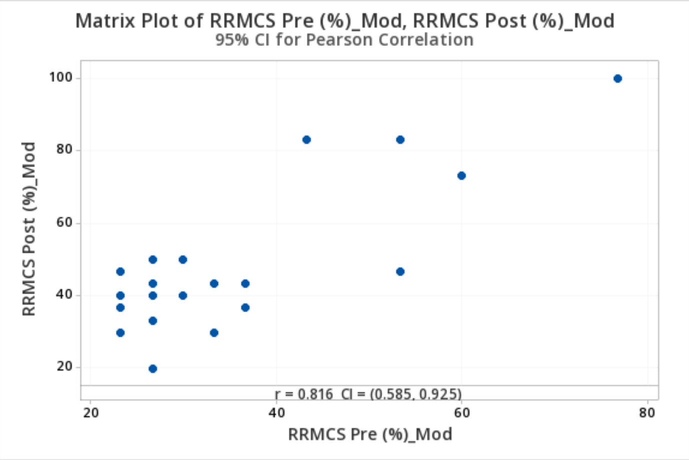
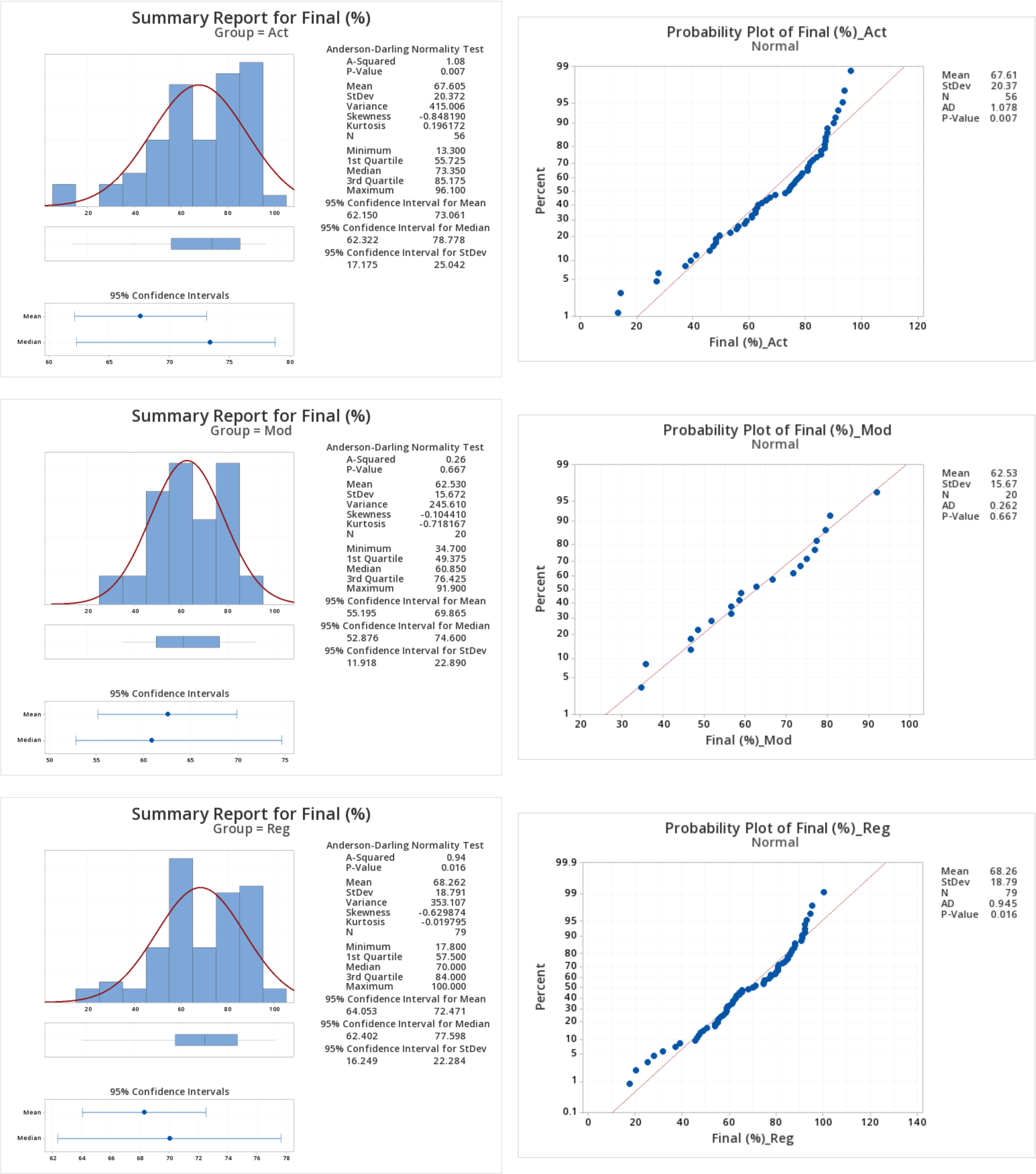


Figure 36. Matrix plot of RRMCS post-test scores versus pre-test scores for the Mod group.

3.4 Problem-solving skills on the 78%-common final exam

A one-way analysis of variance (ANOVA) on final exam grades was performed with $\alpha = .05$. The normality was first assessed with the Anderson-Darling test (Stephens, 1974) for all three groups: modelling instruction (Mod), interactive engagement (Act), and regular instruction (Reg). Results are presented in figure 37. The absolute skewness of final exam grades varies between 0.10 and 0.85 whereas their absolute kurtosis varies between 0.02 and 0.72. Furthermore, some p-values for normality are below $\alpha = .05$, the lowest being .007. The normality test thus leads us to reject the null hypothesis as there seems to be a significant difference for some groups. We, therefore, proceeded to a test for equal variances, not assuming a normal distribution. This fails to reject the null hypothesis, for both multiple comparisons and Levene's test (Levene, 1960) give p-values well above $\alpha = .05$ (figure 38). Because of that, the ANOVA on final exam grades was performed by assuming equal variances (error rate for comparison set at 5). Results are presented in tables 43 and 44, and figures 39 and 40. The p-values are above $\alpha = .05$ and none of the many tests performed on the final exam grades leads us to discard the null hypothesis (equal means). Therefore, no statistically significant difference appears between groups.



```
Anderson-Darling Normality Test
  A-Squared              1.08
  P-Value               0.007
  Mean                 67.605
  StDev                20.372
  Variance            415.006
  Skewness          -0.848190
  Kurtosis           0.196172
  N                        56
  Minimum              13.300
  1st Quartile         55.725
  Median               73.350
  3rd Quartile         85.175
  Maximum              96.100
95% Confidence Interval for Mean
  62.150               73.061
95% Confidence Interval for Median
  62.322               78.778
95% Confidence Interval for StDev
  17.175               25.042
```

```
Mean      67.61
StDev     20.37
N            56
AD        1.078
P-Value   0.007
```

```
Anderson-Darling Normality Test
  A-Squared              0.26
  P-Value               0.667
  Mean                 62.530
  StDev                15.672
  Variance            245.610
  Skewness          -0.104410
  Kurtosis          -0.718167
  N                        20
  Minimum              34.700
  1st Quartile         49.375
  Median               60.850
  3rd Quartile         76.425
  Maximum              91.900
95% Confidence Interval for Mean
  55.195               69.865
95% Confidence Interval for Median
  52.876               74.600
95% Confidence Interval for StDev
  11.918               22.890
```

```
Mean      62.53
StDev     15.67
N            20
AD        0.262
P-Value   0.667
```

```
Anderson-Darling Normality Test
  A-Squared              0.94
  P-Value               0.016
  Mean                 68.262
  StDev                18.791
  Variance            353.107
  Skewness          -0.629874
  Kurtosis          -0.019795
  N                        79
  Minimum              17.800
  1st Quartile         57.500
  Median               70.000
  3rd Quartile         84.000
  Maximum             100.000
95% Confidence Interval for Mean
  64.053               72.471
95% Confidence Interval for Median
  62.402               77.598
95% Confidence Interval for StDev
  16.249               22.284
```

```
Mean      68.26
StDev     18.79
N            79
AD        0.945
P-Value   0.016
```

Figure 37. Summary descriptive statistics and probability plots of normality for final exam grades.

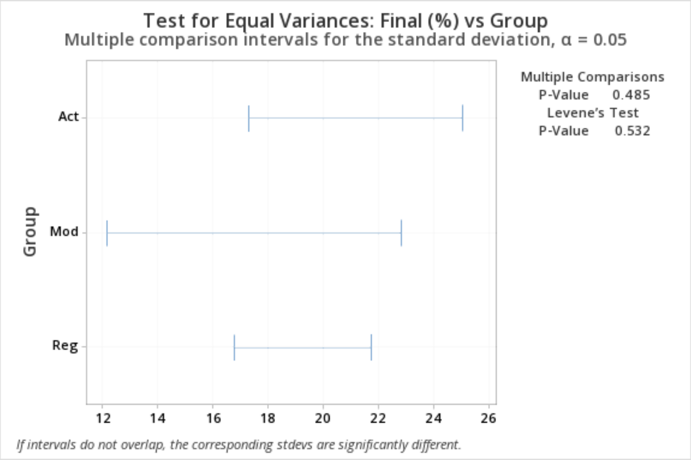


Figure 38. Test for equal variances for final exam grades.

Table 43. Group means (%) on final exam grades.

Group	***N***	**Mean**	**StDev**	**95% CI**
Act	56	67.61	20.37	(62.58, 72.63)
Mod	20	62.53	15.67	(54.12, 70.94)
Reg	79	68.26	18.79	(64.03, 72.49)

Pooled StDev = 19.0281

Table 44. Analysis of variance on final exam grades.

Source	**DF**	**Seq SS**	**Contribution**	**Adj SS**	**Adj MS**	***F*-Value**	***P*-Value**
Group	2	533.4	0.96%	533.4	266.7	0.74	0.480
Error	152	55034.2	99.04%	55034.2	362.1		
Total	154	55567.6	100.00%				

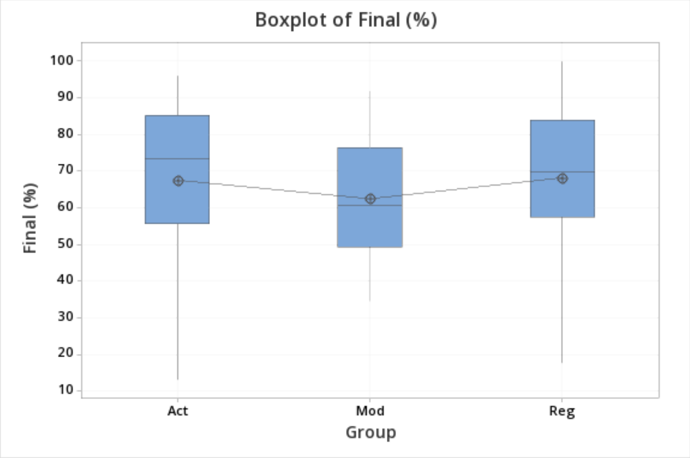


Figure 39. Boxplot of final exam grades.

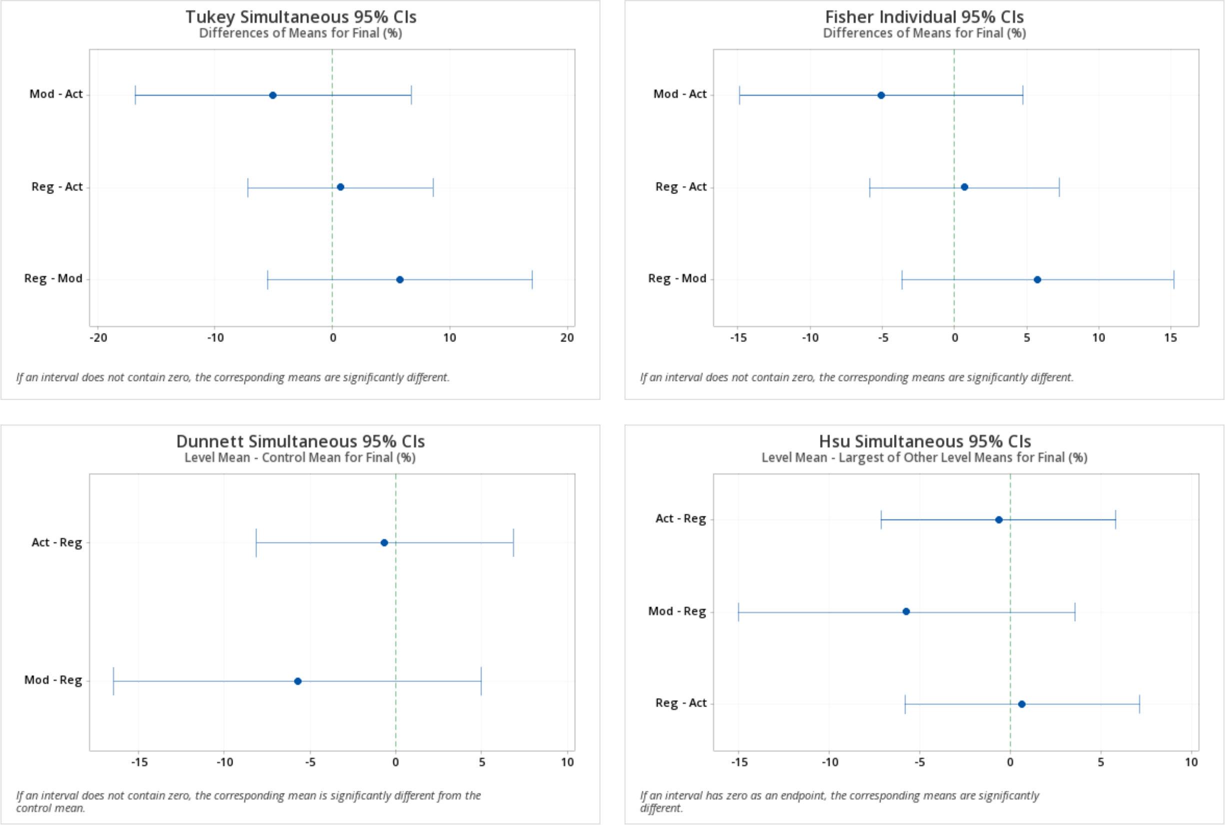


Figure 40. Tukey and Fisher pairwise comparisons, Dunnett multiple comparisons with the control (Reg), and Hsu multiple comparisons with the best (MCB) for final exam grades.

4. SECOND RESEARCH QUESTION: STUDENT ATTITUDES ABOUT LEARNING PHYSICS

4.1 CLASS favourable attitude shifts

4.1.1 Analysis of variance

A one-way ANOVA on favourable attitude shifts S was performed with $\alpha = .05$. The normality was first assessed with the Anderson-Darling test (Stephens, 1974) for all three groups: modelling instruction (Mod), interactive engagement (Act), and regular instruction (Reg). Results are presented in figure 41. The absolute skewness varies between 0.05 and 0.18 whereas the absolute kurtosis varies between 0.17 and 0.38. Furthermore, p-values for normality are all above $\alpha = .05$, the lowest one being .617. This justifies proceeding to a test for equal variances with the assumption of a normal distribution. This test leads to a rejection of the null hypothesis, with a p-value lower than $\alpha = .05$ (although not by much). Bartlett's test (Snedecor & Cochran, 1983) gives a p-value of .033 (figure 42). Because of that, the ANOVA on favourable attitude shifts has been performed without assuming equal variances (error rate for comparison set at 5). Results are presented in tables 45 and 46, and figures 43 and 44. The p-value (.778) is above $\alpha = .05$ and the Games-Howell pairwise comparisons fail to discard the null hypothesis (equal means). It appears that all three methods of teaching have produced sensibly the same effects.

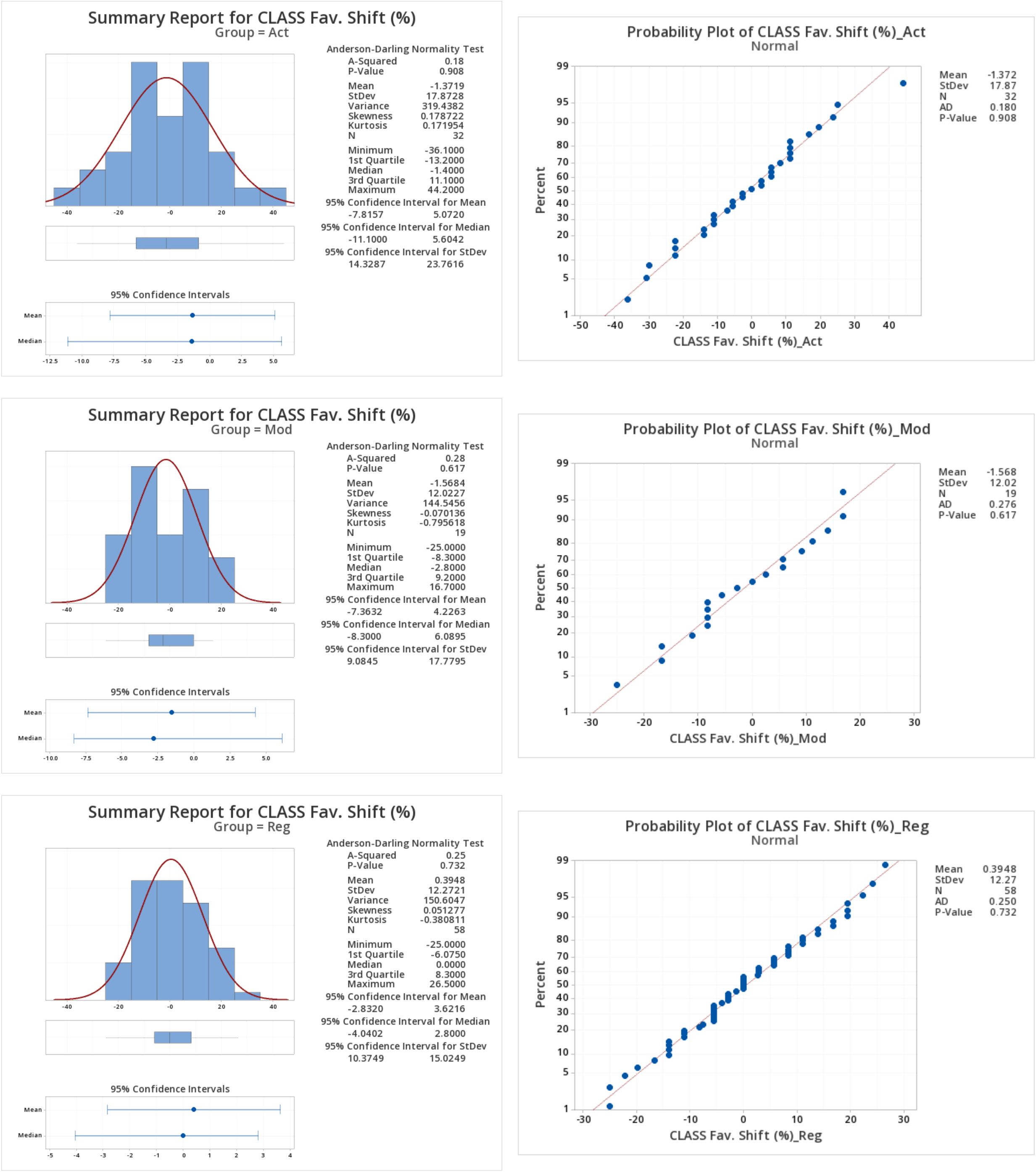


Figure 41. Summary descriptive statistics and probability plots of normality for CLASS favourable shifts.

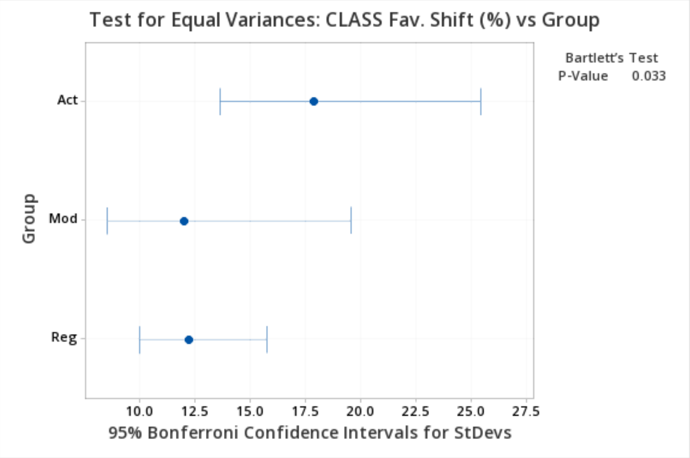


Figure 42. Test for equal variances for CLASS favourable shifts.

Table 45. Group means (%) on CLASS favourable shifts.

Group	***N***	**Mean**	**StDev**	**95% CI**
Act	32	-1.37	17.87	(-7.82, 5.07)
Mod	19	-1.57	12.02	(-7.36, 4.23)
Reg	58	0.39	12.27	(-2.83, 3.62)

Table 46. Analysis of variance on CLASS favourable shifts (Welch's test).

Source	**DF Num**	**DF Den**	***F*-Value**	***P*-Value**
Group	2	44.8163	0.25	0.778

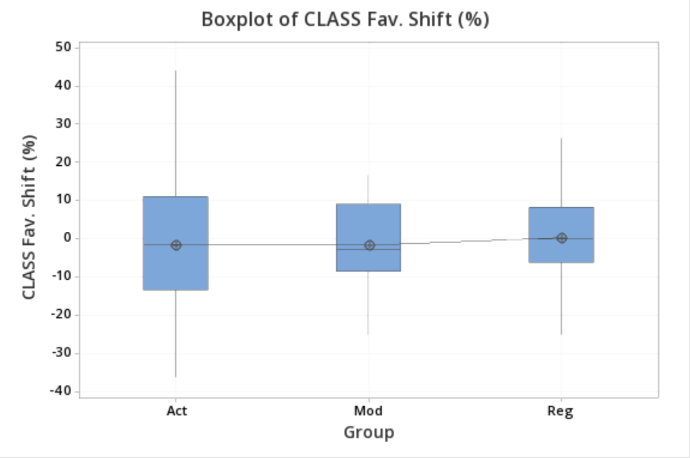


Figure 43. Boxplot of CLASS favourable shifts.

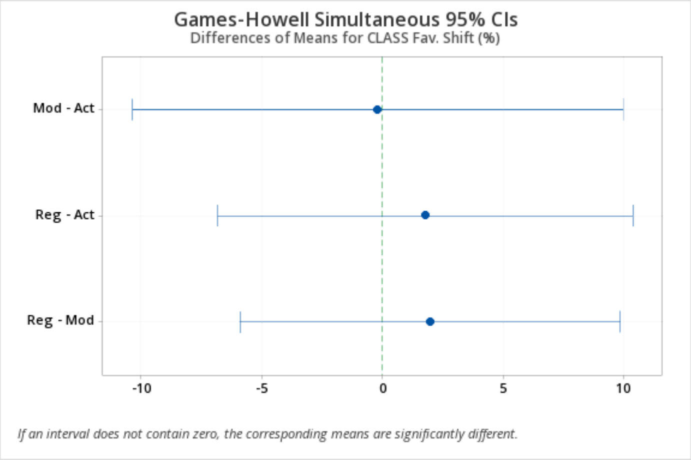


Figure 44. Games-Howell pairwise comparisons for CLASS favourable shifts.

4.1.2 *Effect size*

If we look into the effect size d_S based on the average shift and the pooled standard deviation of CLASS pre- and post-test favourable scores, and also into the corrected effect size $d_{S, dep}$ taking into account a possible dependence between the two scores modelled by Pearson's coefficient of correlation r, we find results of table 47. Figure 45 shows the matrix plots of CLASS post-test scores versus pre-test scores leading to an evaluation of Pearson's r for favourable shifts. As can be seen in table 47, both Cohen's effect size d_S and the corrected Cohen's effect size $d_{S, dep}$ are identical within two decimals.

Table 47. Effect size for the CLASS favourable shift, for different modes of instruction.

Group	***N***	***N****	**CLASS Favourable Shift Effect Size d_S**	**CLASS Favourable Shift Corrected Effect Size $d_{S,\,dep}$**
Act	32	29	-0.08	-0.08
Mod	19	1	-0.10	-0.10
Reg	58	25	0.02	0.02

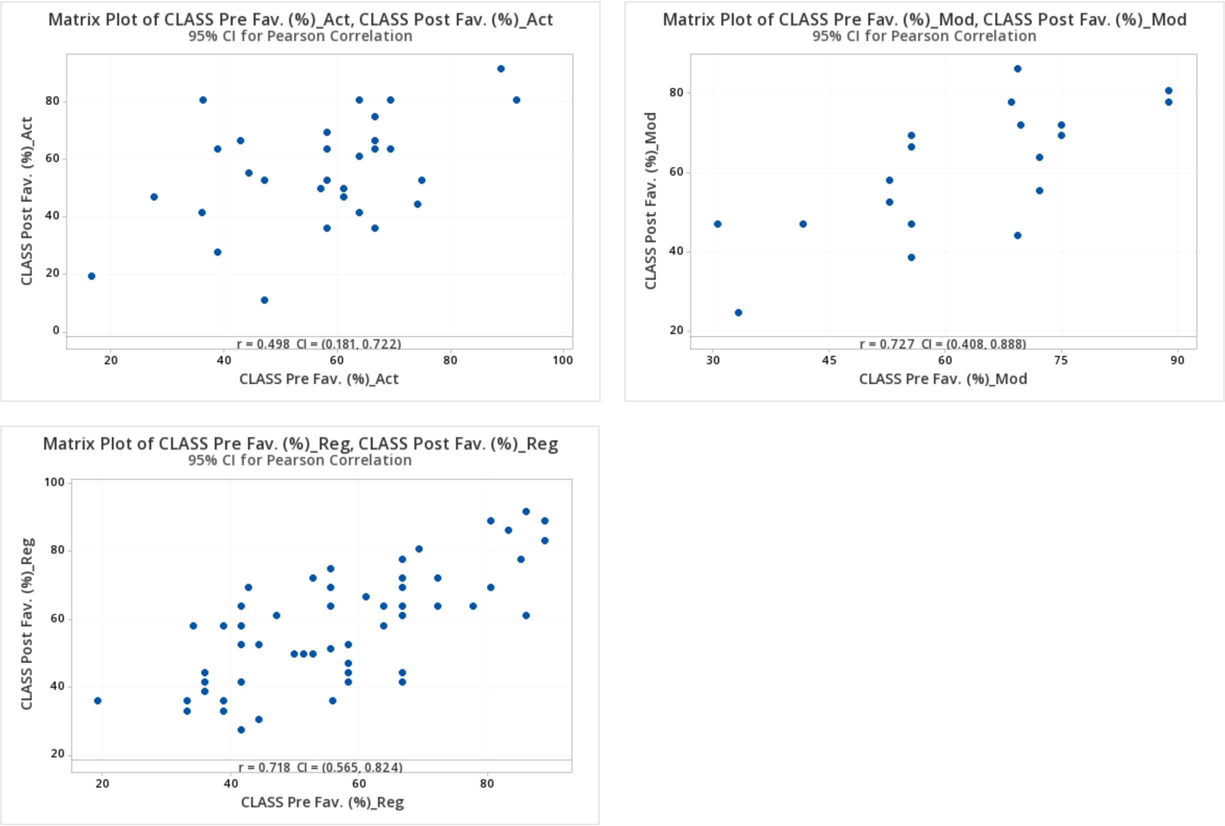


Figure 45. Matrix plots of CLASS favourable post-test scores versus pre-test scores.

4.2 CLASS unfavourable attitude shifts

4.2.1 Analysis of variance

A one-way ANOVA on unfavourable attitude shifts S was performed with $\alpha = .05$. The normality was first assessed with the Anderson-Darling test (Stephens, 1974) for all three groups: modelling instruction (Mod), interactive engagement (Act), and regular instruction (Reg). Results are presented in figure 46. The absolute skewness varies between 0.02 and 0.99 whereas the absolute kurtosis varies between 0.55 and 2.06. Despite some large numbers, all p-values for normality are above $\alpha = .05$ (although the lowest one is close at .06), justifying the assumption of a normal distribution to test for equal variances. This test leads to a rejection of the null hypothesis with a p-value lower than $\alpha = .05$. Bartlett's test (Snedecor & Cochran, 1983) gives a p-value .000(0195) (figure 47). Because of that, the ANOVA on unfavourable attitude shifts has been performed without assuming equal variances (error rate for comparison set at 5). Results are presented in tables 48 and 49, and figures 48 and 49. The p-value (0.413) is above $\alpha = .05$ and the Games-Howell pairwise comparisons fail to discard the null hypothesis (equal means). It appears that all three methods of teaching have produced sensibly the same effects.

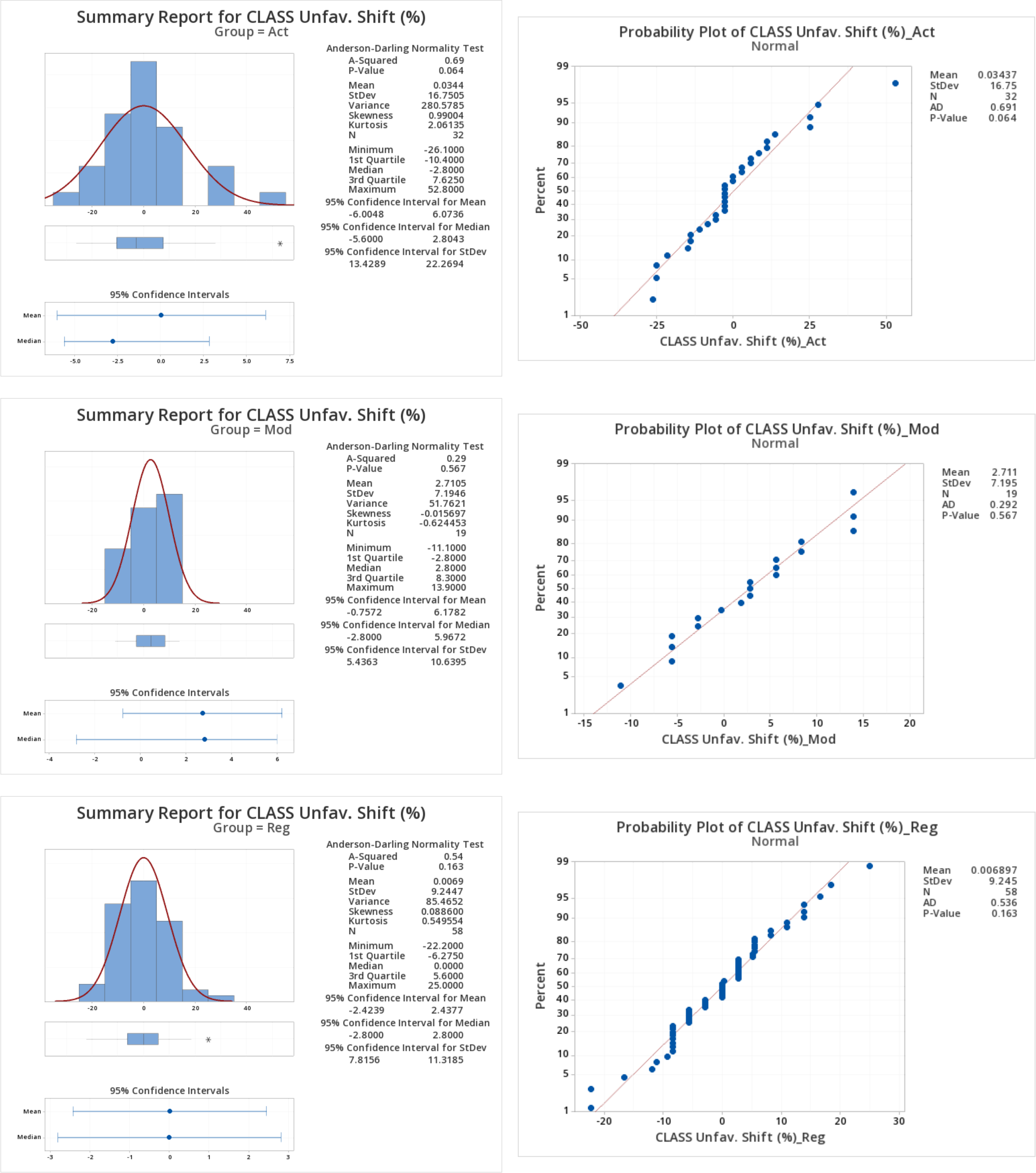


Figure 46. Summary descriptive statistics and probability plots of normality for CLASS unfavourable shifts.

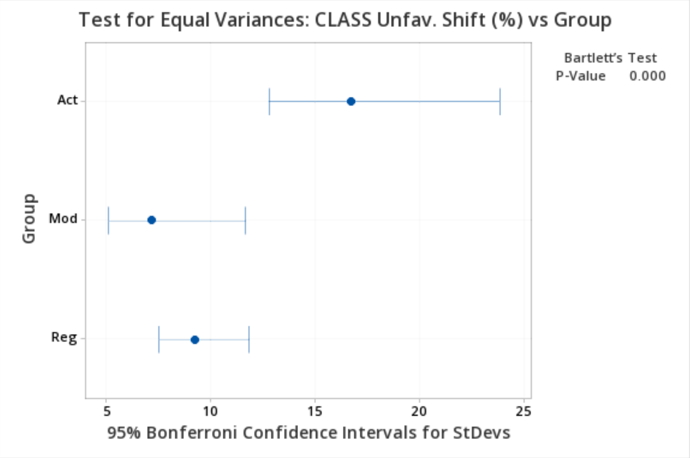


Figure 47. Test for equal variances for CLASS unfavourable shifts.

Table 48. Group means (%) on CLASS unfavourable shifts.

Group	***N***	**Mean**	**StDev**	**95% CI**
Act	32	0.03	16.75	(-6.00, 6.07)
Mod	19	2.71	7.19	(-0.76, 6.18)
Reg	58	0.01	9.24	(-2.42, 2.44)

Table 49. Analysis of variance on CLASS unfavourable shifts (Welch's test).

Source	**DF Num**	**DF Den**	***F*-Value**	***P*-Value**
Group	2	48.3104	0.90	0.413

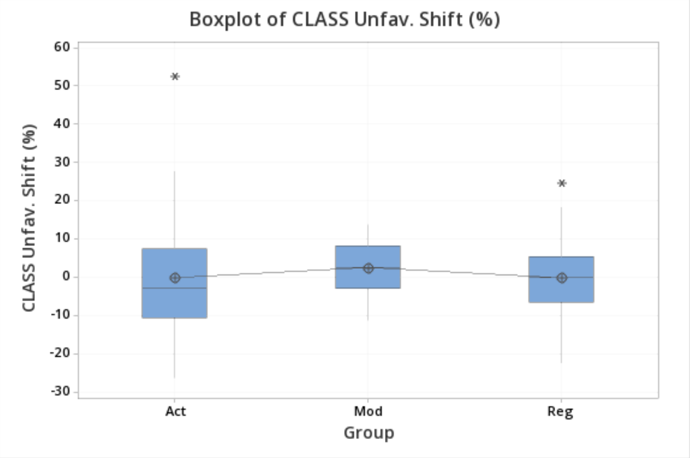


Figure 48. Boxplot of CLASS unfavourable shifts.

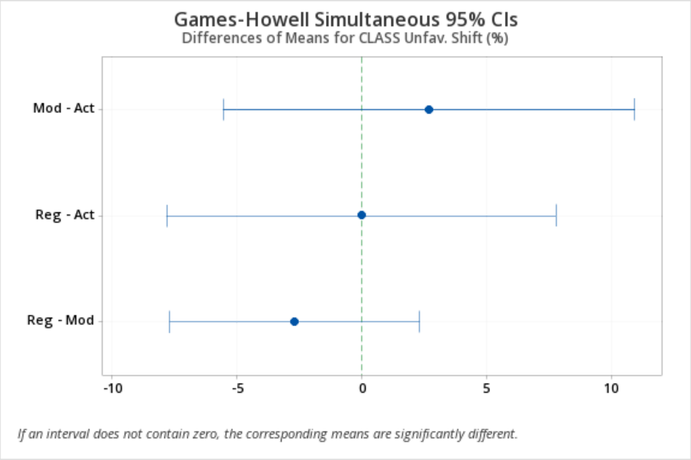


Figure 49. Games-Howell pairwise comparisons for CLASS unfavourable shifts.

4.2.2 *Effect size*

If we look into the effect size d_S based on the average shift and the pooled standard deviation of CLASS pre- and post-test unfavourable scores, and also into the corrected effect size $d_{S,\ dep}$ taking into account a possible dependence between the two scores modelled by Pearson's coefficient of correlation r, we find results of table 50. Figure 50 shows the matrix plots of CLASS post-test scores versus pre-test scores leading to an evaluation of Pearson's r for unfavourable shifts. As can be seen in table 50, both Cohen's effect size d_S and the corrected Cohen's effect size $d_{S,\ dep}$ are identical within two decimals.

Table 50. Effect size for the CLASS unfavourable shift, for different modes of instruction.

Group	***N***	***N****	**CLASS Unfavourable Shift Effect Size d_S**	**CLASS Unfavourable Shift Corrected Effect Size $d_{S,\,dep}$**
Act	32	29	0.00	0.00
Mod	19	1	0.24	0.24
Reg	58	25	0.00	0.00

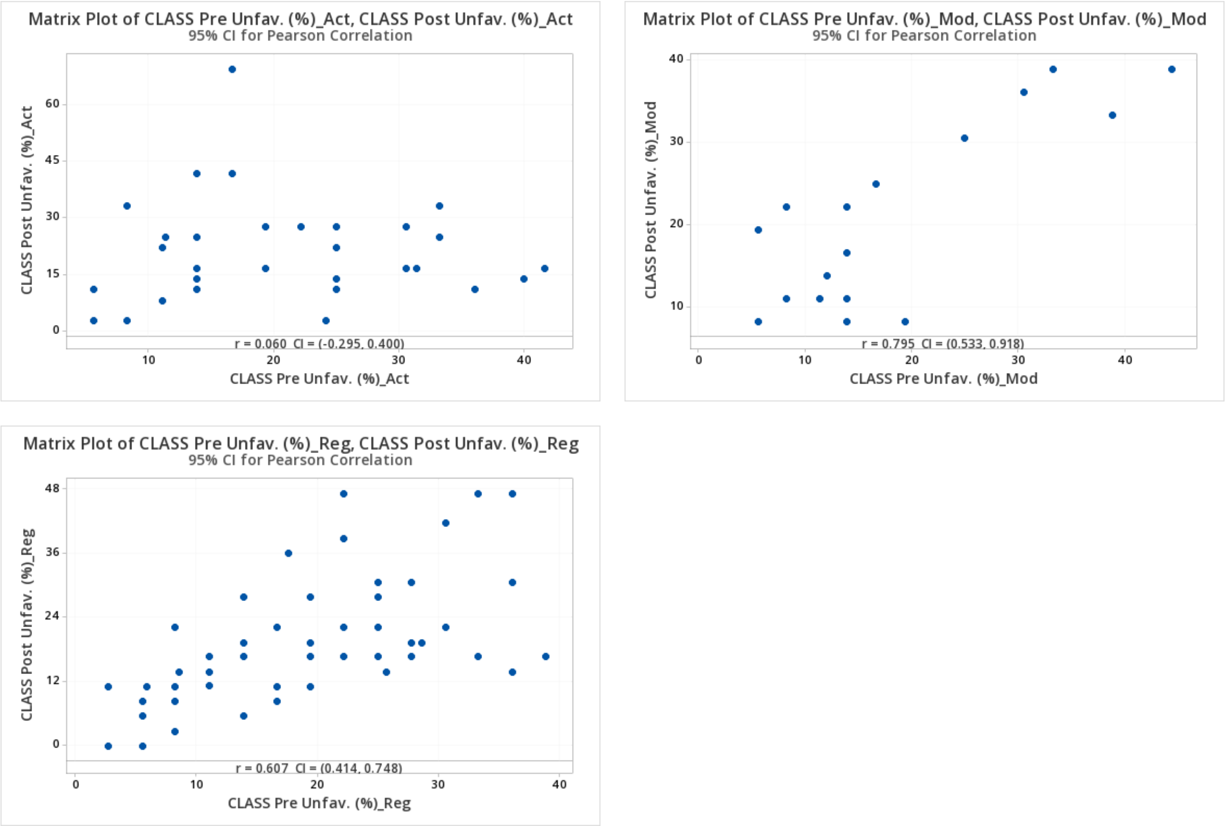


Figure 50. Matrix plots of CLASS unfavourable post-test scores versus pre-test scores.

5. THIRD RESEARCH QUESTION: STUDENTS' PERCEPTION OF MODELLING INSTRUCTION

To answer the third research question about students' perception of modelling instruction, we collected answers to a qualitative and open-ended survey at the end of the semester. Answers are reproduced in their integrality in table 51. Positive comments have been identified in green and bold, while negative ones have been identified in red. Each letter represents a particular, anonymous student that attended the novice modelling class in Mechanics.

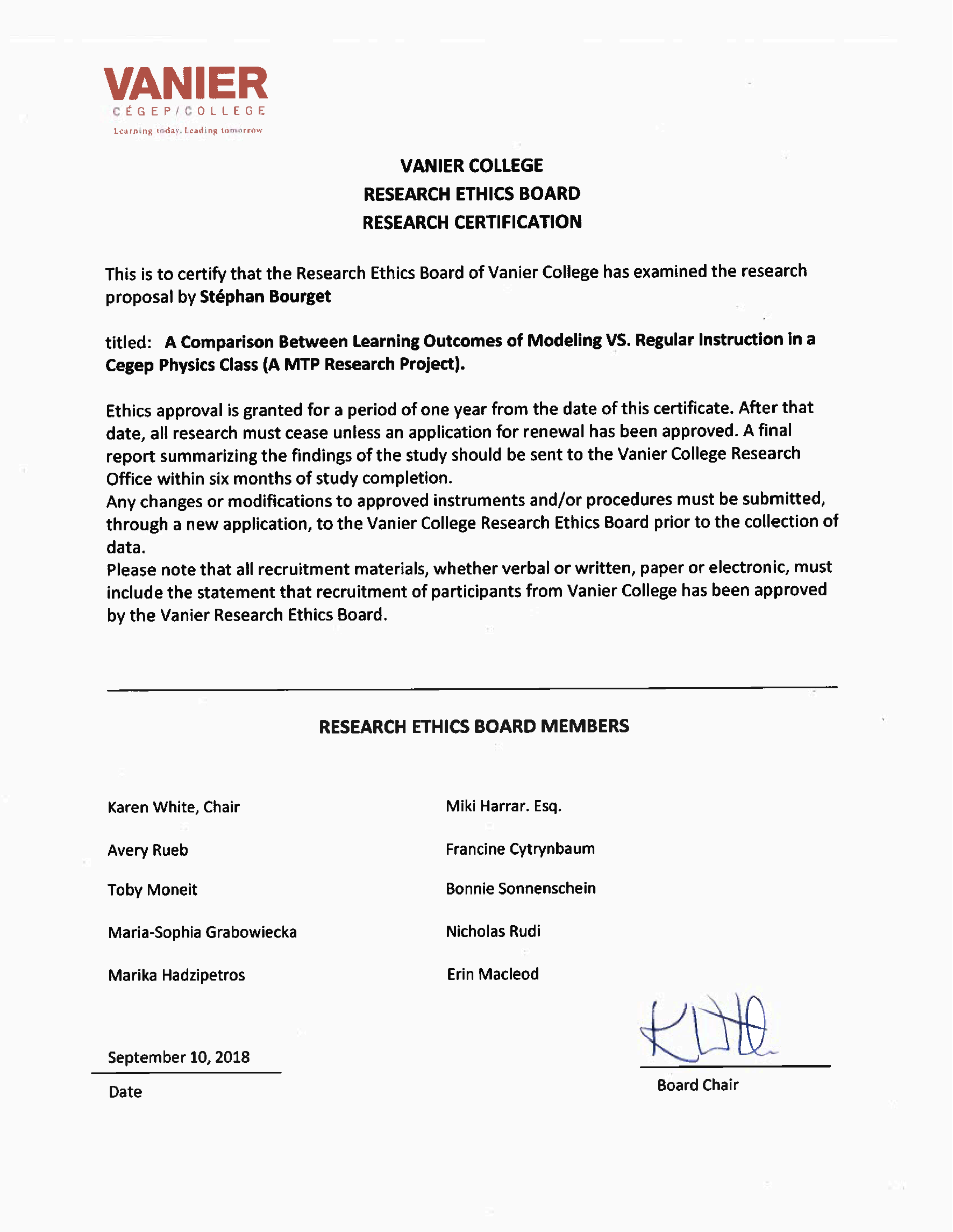
Table 51. Answers to the qualitative survey addressed to Mod-group students.

Question	**Comments** [a]
How did you feel about the modelling approach used in this class?	*(A) Mostly self-taught. Would have preferred some more direct instruction of theory before the modelling (i.e. the labs and whiteboards).* *(B) Terrible.* *(C)* ***I understood the concepts of physics much easily,*** *but had difficulty applying them in problems.* *(D) Bad approach.* ***(E) Interesting approach that I have never seen before. It requires more work than the usual way of teaching a course, but I found I retained the information learned better.*** ***(F) It's fun.*** *(G) Personally, I dislike it.*
What did you like most about the modelling approach used in this course?	***(A) It was interactive and generated discussion*** *(when the students decided to participate).* *(B) Nothing.* ***(C) The group discussions made the course more dynamic and fun.*** ***(D) The discussions in class.*** ***(E) Whiteboarding, discovery through laboratories.*** ***(F) The many labs.*** ***(G) The teamwork involved.***
What did you not like about the modelling approach used in this course?	*(A) If you did not know the material beforehand, there was also little chance of you learning it during the activities where the students had to explain.* *(B) Not my preferred way of being taught Best way: copying notes and having good explanation from teacher.* *(C) I don't like the fact that it relies mostly on experiments rather than theory.* *(D) The white boards because the teacher was not involved enough at the beginning. The material of the course should be explained by the teacher first then put into application by the students and not the opposite.*

Question	**Comments** [a]
	(E) How lab reports were given (I found that while we got a lot freedom, sometimes we were at a loss for where to start for proving certain theories), and how grades are distributed (too little importance given to lab reports in grading scheme). *(F) Less notes.* *(G) The fact that a lot of times I felt as if it were the students teaching themselves, which does not help learning, as we did not have the bases.*
How do you think the modelling approach used in this course impacted your learning?	*(A) I would have learned better in a more traditional course. Perhaps I'm just not used to the modelling approach, but it was a lot more difficult to understand the concepts that were taught.* *(B) I learned nothing; I don't know how to solve challenging mechanics problems.* *(C) I feel like I didn't learn to apply some of the concepts we saw in this course quite well in problems.* *(D) Wrongly. I have the feeling I did not learn that well comparing to my previous years. There was a lot of confusion in the class, a lot of noise during the discussions and the theory was misunderstood and wrongly interpreted.* ***(E) Probably improved my overall learning and retention of information but wasn't as useful for exam preparation and getting a good R-score.*** ***(F) I understand more conceptual physics.*** *(G) Negatively, mechanics was the hardest course for me.*
How does the modelling approach used in this course compare with other science courses (physics or not) you attended in college or high school? Please do not mention any names.	*(A)* ***This course was definitely more interactive than my other science courses, but I*** *performed better in my other classes. My style of learning just doesn't fit the modelling approach too well.* *(B) Other ways of teaching are better than the modelling approach.* *(C)* ***This course allowed discussions in class, which other Science courses rarely allow.*** *Also, this course relies mostly on experiments, while other Science courses relies mostly on theory but with enough experiments. Personally, I prefer the usual way of teaching. I feel like I'm not as prepared as other students who do not have the modelling approach in physics.* ***However, the modelling approach did help me picture and better represent situational problems, which greatly helped me understand and solve problems,*** *though some concepts were not clear to me.* *(D) My learning was highly and badly influenced by this type of approach. Therefore, I did not progress as well comparing to my previous physics classes and I did not understand many concepts viewed this semester. Comparing to my chemistry class where the teacher explained all the material and well leaded the class for the final exam, modelling approach is a bad experience in which I took place this semester.* ***(E) I have only taken 1 physics course prior to this one (Sec V physics), and I like this approach much better than the previous approach. The explanations provided in this course were much more in depth and precise than in my previous course as well.*** ***(F) It's more fun.*** *(G) I highly prefer approaches for the other classes where there were precise class notes to follow along.*

[a] Positive comments are in green and bold; negative ones are in red.

APPENDIX B. RESEARCH CERTIFICATION FROM VANIER COLLEGE’S RESEARCH ETHICS BOARD



VANIER
CÉGEP/COLLEGE
Learning today. Leading tomorrow

VANIER COLLEGE
RESEARCH ETHICS BOARD
RESEARCH CERTIFICATION

This is to certify that the Research Ethics Board of Vanier College has examined the research proposal by **Stéphan Bourget**

titled: **A Comparison Between Learning Outcomes of Modeling VS. Regular Instruction in a Cegep Physics Class (A MTP Research Project).**

Ethics approval is granted for a period of one year from the date of this certificate. After that date, all research must cease unless an application for renewal has been approved. A final report summarizing the findings of the study should be sent to the Vanier College Research Office within six months of study completion.
Any changes or modifications to approved instruments and/or procedures must be submitted, through a new application, to the Vanier College Research Ethics Board prior to the collection of data.
Please note that all recruitment materials, whether verbal or written, paper or electronic, must include the statement that recruitment of participants from Vanier College has been approved by the Vanier Research Ethics Board.

RESEARCH ETHICS BOARD MEMBERS

Karen White, Chair	Miki Harrar. Esq.
Avery Rueb	Francine Cytrynbaum
Toby Moneit	Bonnie Sonnenschein
Maria-Sophia Grabowiecka	Nicholas Rudi
Marika Hadzipetros	Erin Macleod

September 10, 2018
Date

Board Chair

APPENDIX C. RESEARCH RECRUITING SCRIPT

Hello everybody!

My name is ______________ and I'm ______________ here at Vanier.

Mr. Stephan Bourget is currently doing a research project comparing ways of teaching Mechanics, to see if one gives better results than the other. Both are normal, accepted ways to teach Mechanics.

Your participation is important to improve the teaching of physics and Mr. Bourget would like your consent to participate.

Your participation involves allowing your teacher to send the research supervisor Stephen Taylor, who does not work at Vanier, your grades of some regular course activities and assessments, that is, questionnaires done at the start and the end of the semester, plus the results of the common final examination. Mr. Taylor will remove all names and student numbers and record only the data of those students who have agreed to participate. That data will then be given to the Researcher, Mr. Bourget, without any identifying information, after final grades for the course have been submitted at the end of the semester. In this way, Mr. Bourget will never be able to know which students have participated or not, or what the grades of research participants were.

There is also a short, end-of-semester, anonymous survey, only for Mr. Bourget's classes.

This project has been approved by the Vanier Research Ethics Board. Your teacher will send out a summary of the results once the research is completed.

Please read the consent form that I'll distribute, ask any questions you might have about it, and sign it if you agree to participate in this research. Bring back one copy and keep the other one for your records.

APPENDIX C. RESEARCH CONSENT FORM

CONSENT FORM

A COMPARISON BETWEEN LEARNING OUTCOMES OF MODELLING VS. REGULAR INSTRUCTION IN A CEGEP PHYSICS CLASS

Researcher(s)

Stéphan Bourget, M.Ed. candidate (University of Sherbrooke), Physics Department of Vanier College, 418-664-0785, bourgets@vaniercollege.qc.ca

For inquiry, please call Monday to Friday between 10:00 AM and 5:00 PM.

Sponsor(s)

None.

Description of the Research

This research is investigating how students learn physics and how to teach physics better. Since you are in or hope to be in the CEGEP Science Program, we would like you to participate in this research. More precisely, this research seeks to assess the impact of modelling instruction (based on mimicking the process of science through experimental inquiries, whiteboarding, and model development) on learning and understanding, compared to other modes of instruction, when applied to an introductory physics course (Mechanics) in a Quebec CEGEP. This topic is important because students have difficulty mastering and understanding physics concepts. This study proposes to answer three questions. First, how will the introduction of modelling instruction impact students' understanding? Second, how will it impact students' attitudes about science? Third, how will students perceive modelling instruction? No special task other than regular learning activities and the completion of a few questionnaires are asked from participants. Participants' past performance in high school or previous college semesters will not be reviewed. Future use of the research data is not anticipated. Raw data will be destroyed (deleted or shredded) no later than seven (7) years after the end of the research.

Classes receiving the modelling instruction and those who receive the usual ways of teaching this material are being asked to participate in this research so that comparisons can be made.

Your participation would consist of the following;

- For both classes receiving the modelling instruction and those receiving regular instruction, participants will allow the researcher access to the results of several regular course activities and assessments (diagnostic questionnaires at the start and end of the semester, common final exam). Class results for those activities will be provided by your teacher to the Research Supervisor, Stephen Taylor, who will record only the results of those students who have agreed to participate. They will then be given to the Researcher, Stephan Bourget, without any identifying information, after final grades for the course have been submitted.
- For the classes that receive the Modelling Instruction, participants will complete a brief questionnaire at the end of the semester. This questionnaire will be anonymous (no name or identifying information), and the Researcher will only have access to them after final grades have been submitted for the semester.

Potential Harms

Both the classes receiving the modelling instruction and the classes being taught in the 'usual manner' will be receiving recognized effective instruction techniques. Therefore, there will be no major benefit or disadvantage in the learning experience for either group.

Potential Benefits

Participation will help to improve the body of knowledge about physics education and therefore improve the quality of teaching. There will be no direct benefit to individual participants.

Confidentiality

Confidentiality will be respected. No information that discloses your identity will be released or published. Course grades will be given to the Research Supervisor (Stephen Taylor), who will then record only the grades of the students who have signed this consent form. The Researcher (Stephan Bourget) will then receive the grades without ANY identifying information, after the end of the course, when final grades have been submitted. The questionnaires completed at the end of the semester will also only be available to the Researcher after final grades have been submitted. Electronic data will be stored on the researcher's personal computer, which is password-protected and not accessible to unauthorized people. Any hard copies of questionnaires will be stored by the supervisor. Raw data will be destroyed (deleted or shredded) no later than seven (7) years after the end of the research.

Participation

Participation in research must be voluntary. If you choose not to participate, you will continue to have access to quality education. If you choose to participate and later decide to change your mind, you may withdraw your consent and your data up until it has been analyzed. In such a case, you would have to contact *Krista Melanie Riley, pedagogical counsellor – Academic Programs and Innovation, at rileyk@vanier.college or by phone at 514-744-7500 ext. 8241*. If necessary, Krista Riley will contact Stephen G. Taylor to withdraw the consent. Again, you will continue to have access to quality education.

Surveys at the end of the semester are anonymous, identified only by the group the participant belongs to. This data cannot be withdrawn after it is submitted, because of this anonymity.

This project has been approved by the Vanier Research Ethics Board. A summary of the research results will be provided to all teachers whose classes are involved. Your teacher will then send it to you either by posting it on LEA or by MIO.

Statement of Consent

This research aims at comparing results in learning of modelling vs. regular instruction in a CEGEP physics class. I certify that I have read the above information, understand the risks, benefits, responsibilities and conditions of participation as outlined in this document, and freely consent to participate in the project with the assurance that the data will be kept **confidential** and in **no way affect my academic record at CEGEP**. I also consent to my teacher providing the relevant course grades to the Research Supervisor, as outlined above.

DATE: ____________________

PRINT YOUR FIRST NAME: ____
(Given Name)

PRINT YOUR LAST NAME: ____
(Family Name)

STUDENT #: ______________________________

SIGNATURE: ____

TEACHER'S NAME: _____________________________ SECTION #: __________

APPENDIX D. INSTRUCTIONS FOR SCANTRON OPSCAN ANSWER SHEETS

OpScan Answer Sheets – Instructions

Students' participation is voluntary but strongly encouraged to improve the teaching of physics through a representative data analysis.

1) Do **not** write anything on the questionnaire.
2) Mark your answers on the scantron.
3) Make **only one** mark per item.
4) Do **not** skip any questions.
5) Avoid guessing. Your answers should reflect what **you** think.

Students should be informed of the following rules before they enter any marks on the forms:

1) Use a reasonably sharp HB (or harder) lead pencil.
2) **Do not enter any marks except in the designated areas.**
3) Do not write any notes on the form or cross out any area of the form.
4) **Never make a mark that cannot be erased**.
5) Always erase mistakes completely.
6) **Use a medium amount of pressure when entering a mark.**
7) The mark should fill but not exceed the pre-printed boundaries.

Also…

8) Print name but DO NOT blacken corresponding boxes in the student name area.
9) Print student number but DO NOT blacken corresponding boxes in the student number area.
10) Print class section if you have time.
11) For test ID, use 001 for the diagnostic on *attitudes*, 002 for the diagnostic on *translational mechanics*, and 003 for the diagnostic on *rotational mechanics*.

Sheets without a name and/or a student number will be rejected as they can't be tracked to the appropriate consent form.

For answers between 1 and 5, use the corresponding scale A = 1, B = 2, C = 3, D = 4, E = 5.

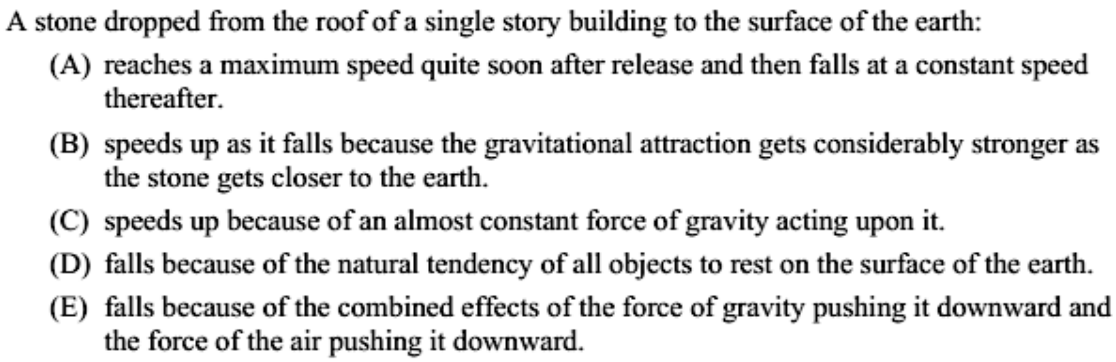
Time allowed

- Diagnostic on attitudes: 8-10 min, up to 15 min if time allows.
- Diagnostic on translational mechanics: 30 min
- Diagnostic on rotational mechanics: 30 min, no explanations/justification needed for answers.

APPENDIX E. STANDARDIZED PRE/POST-TESTS

1. FORCE CONCEPT INVENTORY (FCI)

Sample Question:



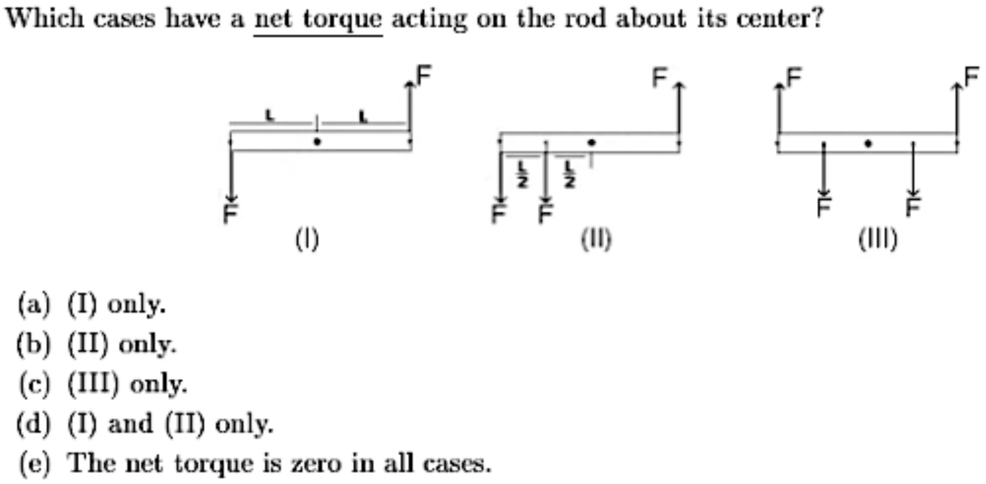
A stone dropped from the roof of a single story building to the surface of the earth:

- (A) reaches a maximum speed quite soon after release and then falls at a constant speed thereafter.
- (B) speeds up as it falls because the gravitational attraction gets considerably stronger as the stone gets closer to the earth.
- (C) speeds up because of an almost constant force of gravity acting upon it.
- (D) falls because of the natural tendency of all objects to rest on the surface of the earth.
- (E) falls because of the combined effects of the force of gravity pushing it downward and the force of the air pushing it downward.

Source: https://www.physport.org/assessments/assessment.cfm?I=5&A=FCI

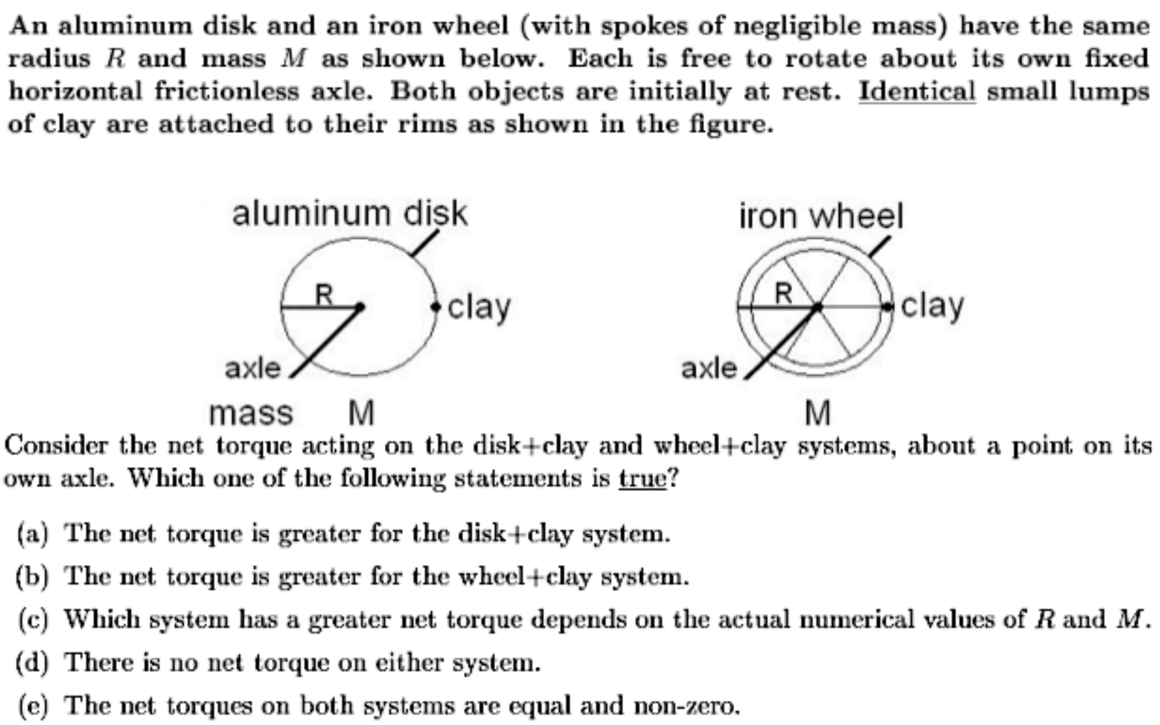
2. ROLLING MOTION CONCEPTUAL SURVEY (RRMCS)

Sample Questions:



Which cases have a <u>net torque</u> acting on the rod about its center?

- (a) (I) only.
- (b) (II) only.
- (c) (III) only.
- (d) (I) and (II) only.
- (e) The net torque is zero in all cases.



An aluminum disk and an iron wheel (with spokes of negligible mass) have the same radius R and mass M as shown below. Each is free to rotate about its own fixed horizontal frictionless axle. Both objects are initially at rest. Identical small lumps of clay are attached to their rims as shown in the figure.

Consider the net torque acting on the disk+clay and wheel+clay systems, about a point on its own axle. Which one of the following statements is true?

(a) The net torque is greater for the disk+clay system.
(b) The net torque is greater for the wheel+clay system.
(c) Which system has a greater net torque depends on the actual numerical values of R and M.
(d) There is no net torque on either system.
(e) The net torques on both systems are equal and non-zero.

Source: https://www.physport.org/assessments/assessment.cfm?I=14&A=RRMCS

3. COLORADO LEARNING ATTITUDES ABOUT SCIENCE SURVEY (CLASS)

Sample Questions:

A significant problem in learning physics is being able to memorize all the information I need to know.

Strongly Disagree 1 2 3 4 5 Strongly Agree

Knowledge in physics consists of many disconnected topics.

Strongly Disagree 1 2 3 4 5 Strongly Agree

Source: https://www.physport.org/assessments/assessment.cfm?A=CLASS&S=1

APPENDIX F. ANONYMOUS QUALITATIVE SURVEY FOR TREATMENT (MODELLING) GROUP

Anonymous Qualitative Survey

Description of the Research

This research seeks to assess the impact of modelling instruction (based on mimicking the process of science through experimental inquiries, whiteboarding, and model development) on learning and understanding, compared to other modes of instruction, when applied to an introductory physics course (Mechanics) in a Quebec CEGEP. This topic is important because students have difficulty mastering and understanding physics concepts. This survey should take 10 to 15 minutes to fill out.

Potential Harms

Both the classes receiving the modelling instruction and the classes being taught in the 'usual manner' will be receiving recognized effective instruction techniques. Therefore, there will be no major benefit or disadvantage in the learning experience for either group.

Potential Benefits

Participation will help to improve the body of knowledge about physics education and therefore improve the quality of teaching. There will be no direct benefit to individual participants.

Confidentiality

All survey data will be anonymous. Anonymous data will be accessible by the researcher Stephan Bourget after final grades have been submitted for the semester. Electronic data will be stored on the researcher's personal computer, which is password-protected and not accessible to unauthorized people. Any hard copies of questionnaires will be stored by the supervisor Stephen G. Taylor. Raw data will be destroyed (deleted or shredded) no later than seven (7) years after the end of the research.

Participation

Participation in research must be voluntary. If you choose not to participate, you will continue to have access to quality education. Even if you consented to participate in this project, you are not obliged to fill out this survey. You can hand it in blank, or not answer specific questions. You may also fill out the survey if you did not sign a consent to participation earlier in the semester. A summary of the research results will be provided to all teachers whose classes are involved. Your teacher will then send it to you either by posting it on LEA or by MIO.

By completing this survey, you acknowledge that you consent to participate in the study as described above.

Survey

In this class, I used an approach called 'modelling,' where I invited you to explore phenomena in teams, to communicate your thoughts on whiteboards and to construct new models though interactions with the whole class. You then applied those constructed models in various situations to ultimately challenge them to build better ones capable of explaining a larger spectrum of observations. The following survey asks you to reflect on how you felt about that approach.

1. How did you feel about the modelling approach used in this class?

2. What did you like most about the modelling approach used in this course?

3. What did you not like about the modelling approach used in this course?

4. How do you think the modelling approach used in this course impacted your learning?

5. How does the modelling approach used in this course compare with other science courses (physics or not) you attended in college or high school? Please do not mention any names.

APPENDIX G. QUESTIONNAIRE FOR TEACHERS (ASSESSMENT OF TEACHING METHOD)

Assessment of Teaching Method

(Mechanics NYA, A2018)

Because the teaching method is going to affect the interpretation of results collected from diagnostic tests, it is necessary to better distinguish between the ways participating teachers have taught. Therefore, I kindly ask you to fill out the questionnaire below and to send the document back to me, Stephan Bourget, by email at bourgets@vaniercollege.qc.ca

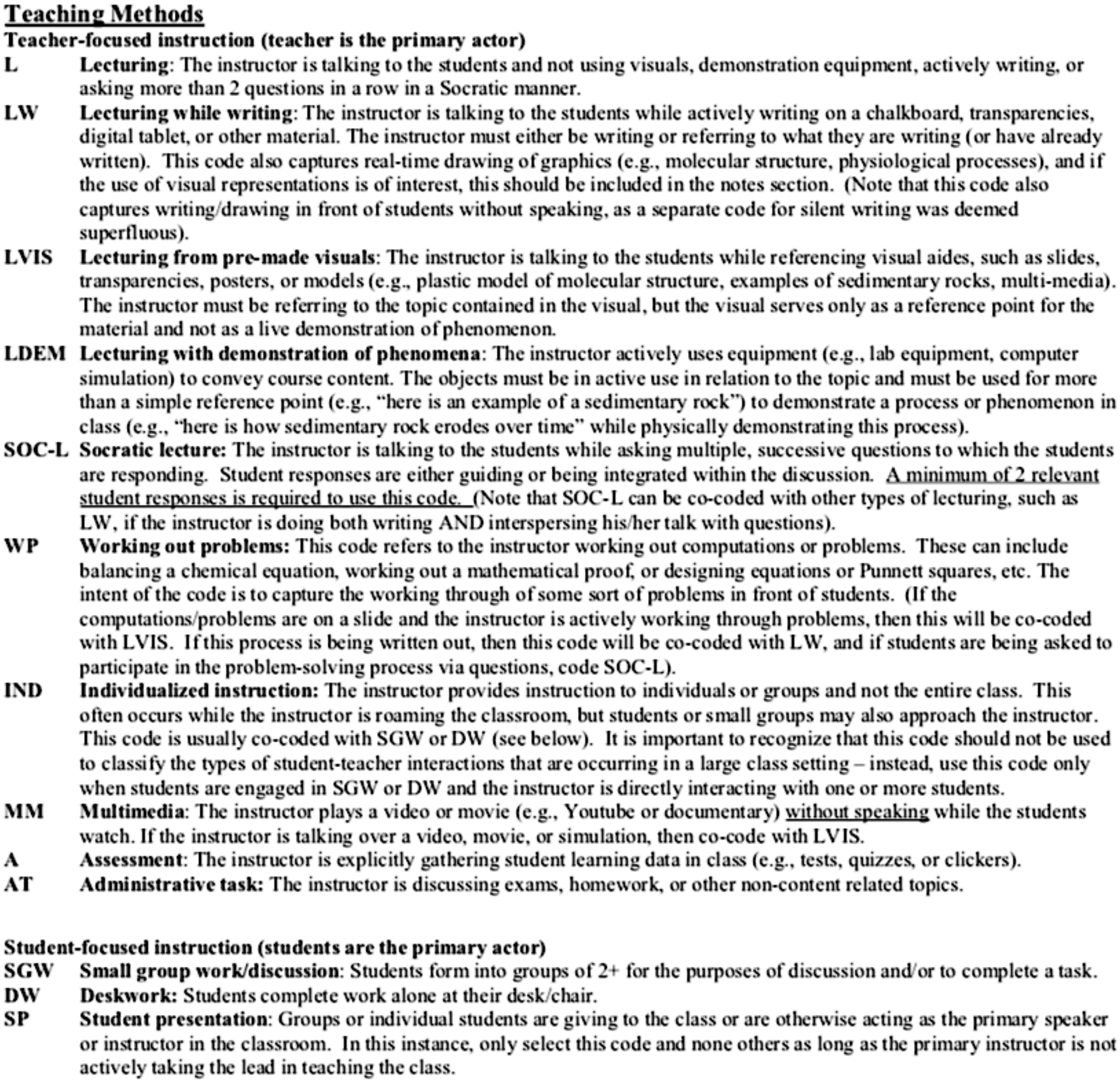
1. Teacher's name:

2. Years teaching this course:

3. Technologies used in class:

4. Describe your typical methodology to teach Physics NYA this semester:

5. Over the following pages, circle (in Word, using the Draw tab) any code that corresponds to a typical characteristic of your classes. You can also just check out the codes that apply below.



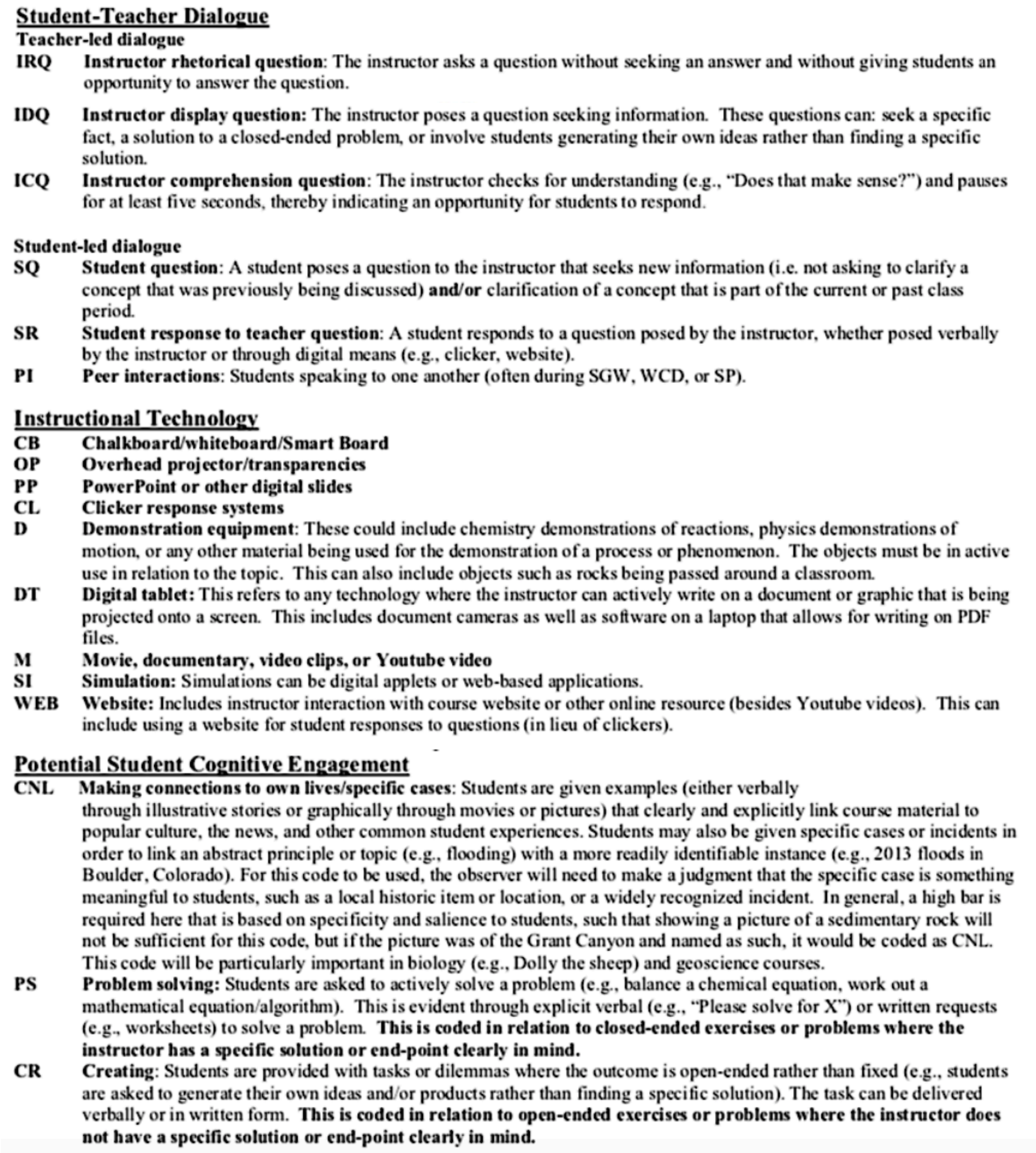
<u>Teaching Methods</u>

Teacher-focused instruction (teacher is the primary actor)

L	**Lecturing**: The instructor is talking to the students and not using visuals, demonstration equipment, actively writing, or asking more than 2 questions in a row in a Socratic manner.
LW	**Lecturing while writing**: The instructor is talking to the students while actively writing on a chalkboard, transparencies, digital tablet, or other material. The instructor must either be writing or referring to what they are writing (or have already written). This code also captures real-time drawing of graphics (e.g., molecular structure, physiological processes), and if the use of visual representations is of interest, this should be included in the notes section. (Note that this code also captures writing/drawing in front of students without speaking, as a separate code for silent writing was deemed superfluous).
LVIS	**Lecturing from pre-made visuals**: The instructor is talking to the students while referencing visual aides, such as slides, transparencies, posters, or models (e.g., plastic model of molecular structure, examples of sedimentary rocks, multi-media). The instructor must be referring to the topic contained in the visual, but the visual serves only as a reference point for the material and not as a live demonstration of phenomenon.
LDEM	**Lecturing with demonstration of phenomena**: The instructor actively uses equipment (e.g., lab equipment, computer simulation) to convey course content. The objects must be in active use in relation to the topic and must be used for more than a simple reference point (e.g., "here is an example of a sedimentary rock") to demonstrate a process or phenomenon in class (e.g., "here is how sedimentary rock erodes over time" while physically demonstrating this process).
SOC-L	**Socratic lecture:** The instructor is talking to the students while asking multiple, successive questions to which the students are responding. Student responses are either guiding or being integrated within the discussion. <u>A minimum of 2 relevant student responses is required to use this code.</u> (Note that SOC-L can be co-coded with other types of lecturing, such as LW, if the instructor is doing both writing AND interspersing his/her talk with questions).
WP	**Working out problems:** This code refers to the instructor working out computations or problems. These can include balancing a chemical equation, working out a mathematical proof, or designing equations or Punnett squares, etc. The intent of the code is to capture the working through of some sort of problems in front of students. (If the computations/problems are on a slide and the instructor is actively working through problems, then this will be co-coded with LVIS. If this process is being written out, then this code will be co-coded with LW, and if students are being asked to participate in the problem-solving process via questions, code SOC-L).
IND	**Individualized instruction:** The instructor provides instruction to individuals or groups and not the entire class. This often occurs while the instructor is roaming the classroom, but students or small groups may also approach the instructor. This code is usually co-coded with SGW or DW (see below). It is important to recognize that this code should not be used to classify the types of student-teacher interactions that are occurring in a large class setting – instead, use this code only when students are engaged in SGW or DW and the instructor is directly interacting with one or more students.
MM	**Multimedia**: The instructor plays a video or movie (e.g., Youtube or documentary) <u>without speaking</u> while the students watch. If the instructor is talking over a video, movie, or simulation, then co-code with LVIS.
A	**Assessment**: The instructor is explicitly gathering student learning data in class (e.g., tests, quizzes, or clickers).
AT	**Administrative task:** The instructor is discussing exams, homework, or other non-content related topics.

Student-focused instruction (students are the primary actor)

SGW	**Small group work/discussion**: Students form into groups of 2+ for the purposes of discussion and/or to complete a task.
DW	**Deskwork:** Students complete work alone at their desk/chair.
SP	**Student presentation**: Groups or individual students are giving to the class or are otherwise acting as the primary speaker or instructor in the classroom. In this instance, only select this code and none others as long as the primary instructor is not actively taking the lead in teaching the class.



Student-Teacher Dialogue

Teacher-led dialogue

IRQ	**Instructor rhetorical question**: The instructor asks a question without seeking an answer and without giving students an opportunity to answer the question.
IDQ	**Instructor display question:** The instructor poses a question seeking information. These questions can: seek a specific fact, a solution to a closed-ended problem, or involve students generating their own ideas rather than finding a specific solution.
ICQ	**Instructor comprehension question**: The instructor checks for understanding (e.g., "Does that make sense?") and pauses for at least five seconds, thereby indicating an opportunity for students to respond.

Student-led dialogue

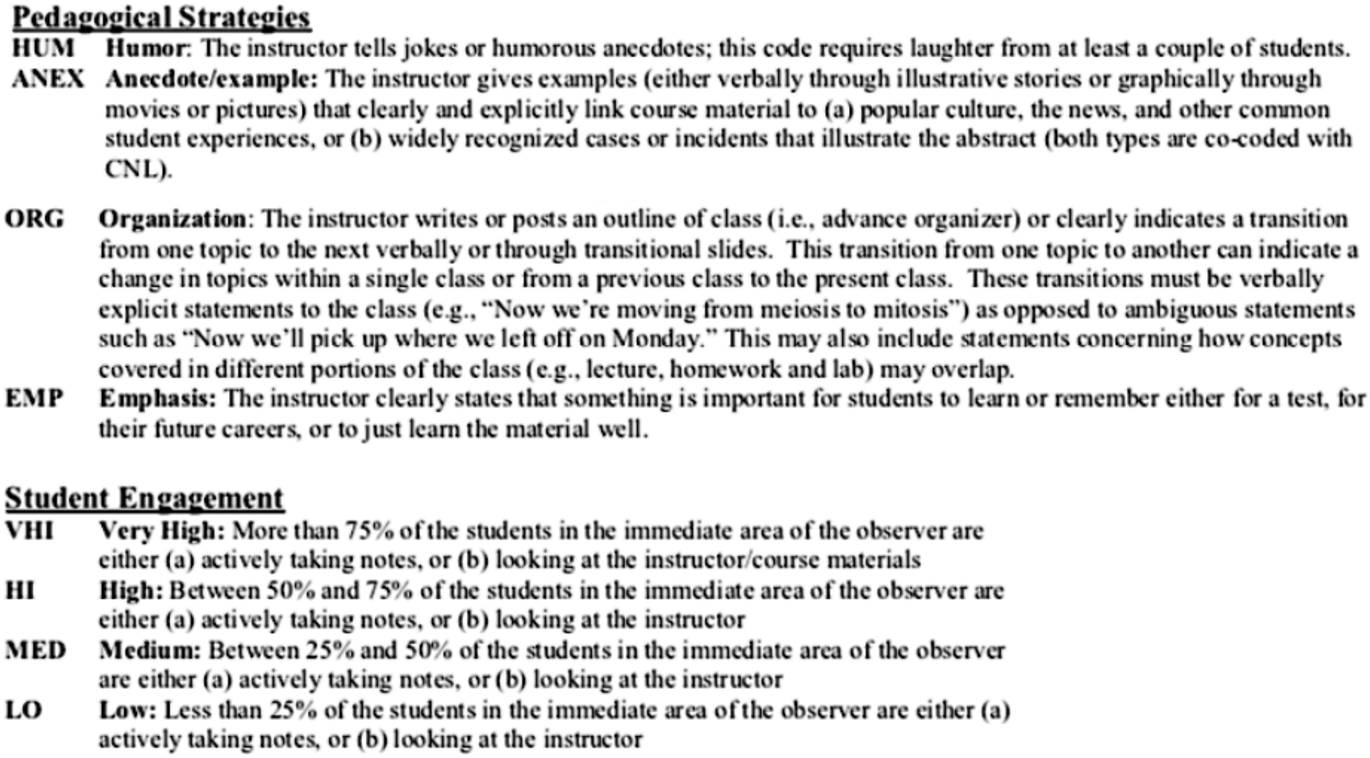
SQ	**Student question**: A student poses a question to the instructor that seeks new information (i.e. not asking to clarify a concept that was previously being discussed) **and/or** clarification of a concept that is part of the current or past class period.
SR	**Student response to teacher question**: A student responds to a question posed by the instructor, whether posed verbally by the instructor or through digital means (e.g., clicker, website).
PI	**Peer interactions**: Students speaking to one another (often during SGW, WCD, or SP).

Instructional Technology

CB	**Chalkboard/whiteboard/Smart Board**
OP	**Overhead projector/transparencies**
PP	**PowerPoint or other digital slides**
CL	**Clicker response systems**
D	**Demonstration equipment**: These could include chemistry demonstrations of reactions, physics demonstrations of motion, or any other material being used for the demonstration of a process or phenomenon. The objects must be in active use in relation to the topic. This can also include objects such as rocks being passed around a classroom.
DT	**Digital tablet:** This refers to any technology where the instructor can actively write on a document or graphic that is being projected onto a screen. This includes document cameras as well as software on a laptop that allows for writing on PDF files.
M	**Movie, documentary, video clips, or Youtube video**
SI	**Simulation:** Simulations can be digital applets or web-based applications.
WEB	**Website:** Includes instructor interaction with course website or other online resource (besides Youtube videos). This can include using a website for student responses to questions (in lieu of clickers).

Potential Student Cognitive Engagement

CNL	**Making connections to own lives/specific cases**: Students are given examples (either verbally through illustrative stories or graphically through movies or pictures) that clearly and explicitly link course material to popular culture, the news, and other common student experiences. Students may also be given specific cases or incidents in order to link an abstract principle or topic (e.g., flooding) with a more readily identifiable instance (e.g., 2013 floods in Boulder, Colorado). For this code to be used, the observer will need to make a judgment that the specific case is something meaningful to students, such as a local historic item or location, or a widely recognized incident. In general, a high bar is required here that is based on specificity and salience to students, such that showing a picture of a sedimentary rock will not be sufficient for this code, but if the picture was of the Grant Canyon and named as such, it would be coded as CNL. This code will be particularly important in biology (e.g., Dolly the sheep) and geoscience courses.
PS	**Problem solving:** Students are asked to actively solve a problem (e.g., balance a chemical equation, work out a mathematical equation/algorithm). This is evident through explicit verbal (e.g., "Please solve for X") or written requests (e.g., worksheets) to solve a problem. **This is coded in relation to closed-ended exercises or problems where the instructor has a specific solution or end-point clearly in mind.**
CR	**Creating**: Students are provided with tasks or dilemmas where the outcome is open-ended rather than fixed (e.g., students are asked to generate their own ideas and/or products rather than finding a specific solution). The task can be delivered verbally or in written form. **This is coded in relation to open-ended exercises or problems where the instructor does not have a specific solution or end-point clearly in mind.**



Pedagogical Strategies

HUM	**Humor**: The instructor tells jokes or humorous anecdotes; this code requires laughter from at least a couple of students.
ANEX	**Anecdote/example:** The instructor gives examples (either verbally through illustrative stories or graphically through movies or pictures) that clearly and explicitly link course material to (a) popular culture, the news, and other common student experiences, or (b) widely recognized cases or incidents that illustrate the abstract (both types are co-coded with CNL).
ORG	**Organization**: The instructor writes or posts an outline of class (i.e., advance organizer) or clearly indicates a transition from one topic to the next verbally or through transitional slides. This transition from one topic to another can indicate a change in topics within a single class or from a previous class to the present class. These transitions must be verbally explicit statements to the class (e.g., "Now we're moving from meiosis to mitosis") as opposed to ambiguous statements such as "Now we'll pick up where we left off on Monday." This may also include statements concerning how concepts covered in different portions of the class (e.g., lecture, homework and lab) may overlap.
EMP	**Emphasis:** The instructor clearly states that something is important for students to learn or remember either for a test, for their future careers, or to just learn the material well.

Student Engagement

VHI	**Very High:** More than 75% of the students in the immediate area of the observer are either (a) actively taking notes, or (b) looking at the instructor/course materials
HI	**High:** Between 50% and 75% of the students in the immediate area of the observer are either (a) actively taking notes, or (b) looking at the instructor
MED	**Medium:** Between 25% and 50% of the students in the immediate area of the observer are either (a) actively taking notes, or (b) looking at the instructor
LO	**Low:** Less than 25% of the students in the immediate area of the observer are either (a) actively taking notes, or (b) looking at the instructor

6. Other comments:

Thanks a lot for your participation and collaboration. Preliminary results will be shared with you as soon as they become available, and then you can share with your students through MIO or email as you see fit.

APPENDIX H. PHYSICS NYA COURSE OUTLINE (MODELLING GROUP)

PHYSICS NYA COURSE OUTLINE

Course Title:	Mechanics
Course Number:	203-NYA-05
Section number:	00025/00026/02001
Semester:	Fall 2018
Ponderation:	3-2-3
Pre-requisites:	Sec 5 Physics PH 504 or 203-001 & Sec 5 Math TS/SN 506 or 201-015-50
Co-requisites:	201-NYA
Teacher:	Stephan Bourget
Office:	B406
Telephone:	N/A
Office Hours:	Send a MIO for an appointment, or check on my office's door

COURSE DESCRIPTION

This first semester physics course introduces students to the basic concepts and principles of Newtonian Mechanics. Topics include vectors, one-dimensional and two-dimensional motion, Newton's laws, work, energy, momentum, laws of conservation, rotation and gravitation. See Appendix A for the Objective 00UR and Appendix B for a list of the textbook sections that are relevant to the course. Motion, force, work, energy, momentum and laws of conservation are the most fundamental concepts in science. They are used not only in later physics courses, NYB and NYC, but also in other science disciplines, such as Chemistry and Biology.

TEXTBOOK AND OTHER REQUIRED MATERIAL

- **Required Textbook**
  - *University Physics*, by OpenStax. Free. Will be posted onto LEA.
  - *Mechanics*, author: Luc Tremblay. Free at http://physique.merici.ca/mechanics.html.
- **Bibliography and additional references**
  - *Fundamentals of Physics,* 10th Ed., authors: Halliday & Resnick, Walker, J., J. John Wiley & Sons Publishers, 2010. It costs about $100. The same book is used in NYA (*Mechanics*), NYC (*Waves and Modern Physics*), NYB (*Electricity and Magnetism*), and HTK (*Engineering Physics*) at Vanier College, in the regular fall and winter semesters.
  - *Physics for Scientists and Engineers*, authors: Serway & Jewett.
  - *Physics for Scientists and Engineers*, author: Knight.
  - *University Physics*, authors: Young & Freedman.
- **Other material:** Ruler, protractor, scientific calculator. They cost about $20.

TEACHING METHODS

It is strongly recommended that students come to every class prepared and participate actively during each class. Discussion of the material is an important aspect of the course. It is essential that students keep up to date with the course material, try to understand each topic as it is collaboratively discussed in class, and plan to regularly work on this material outside of class time.

Physics is an experimental science. Lab work is an important part of this course. The weekly lab periods are usually used for lab experiments, problem solving, and/or group work. Students will be expected to prepare detailed reports of selected labs. See **Appendix C** for the lab report style guide.

GRADING SCHEME

The final mark for the course is based upon the following evaluation scheme:

Course Work	Laboratory Activities[*] (lowest result discarded)	15 %	Ongoing
	WeBWorK assignments[*] (lowest result discarded)	5 %	Ongoing
Class Testing	Class Tests[†]: 1: Kinematics (?) 2: Forces (?) 3: Energy & Work, Momentum (?)	30 %	TBA TBA TBA
Final Summative Assessments[††]	Laboratory Exam	10 %	TBA
	Final Exam	40 %	Exam Period

[*] *Refer to Omnivox for the dates and weighting of Lab Activities and online assignments.*
[†] *If the final exam grade is higher than the grade of one or more tests, the lowest test grade will be replaced by the final exam grade. One of the 3 class tests will be a group exam; your best preparation is to participate actively in the learning activities and the modelling cycles.*
[††] *If the average obtained in the category Final examination is not at least 60, the final grade will not exceed 60.*

Check LEA for due dates. Contents of the class tests will be confirmed in class. The final exam is comprehensive and compulsory. The final exam grade can replace the lowest test grade if improvement is shown.

FORMULA SHEET USAGE

A formula sheet is allowed in the final exam. The teacher will either distribute printed formula sheets and/or allow students to prepare their own memory aid. In the latter case (unless otherwise stated in class), the memory aid must fit on a 3"x5" index card. Both sides of the card may be used. Students can write <u>only equations</u>, no words, derivation, graphs or diagrams, etc.

At teacher's discretion, a formula sheet may be allowed in class tests. In that case, the teacher will either distribute printed formula sheets and/or allow students to prepare their own memory aid. In the latter case (unless otherwise stated in class), the memory aid must fit on a 3"x5" index card. Only one side of the card may be used; for final exams, both sides are allowed. Students can write <u>only equations</u>, no words, derivation, graphs or diagrams, etc. The teacher will announce the decision whether to provide printed formula sheets or allow students to make their own, at least two weeks prior to the first class test.

ATTENDANCE

Students are strongly recommended to attend all lectures. Attendance in labs is compulsory and a mark of zero will be given for missed lab work. Students are responsible for obtaining all material covered, including handouts, and announcements made during their absence. Except for absences for religious holy days (Vanier policy 7210-20), there will be no make-up tests, exams, or labs without a well-documented medical note that complies with college policy.

LATE ASSIGNMENTS/MISSED TESTS

The deadlines for assignments and laboratory activities will be clearly communicated in class. Any work that is submitted on the due day but after the specified time will receive a penalty of 5% and later submissions will be subject to a penalty of 10% of the total available marks per day (up to 5 days). Work that submitted more than 5 days late, or not at all, will receive a grade of zero.

All students are expected to complete all assessment activities as scheduled. If a student must miss a test, for a valid reason, the weight of the missed test will be added to the weight of the final exam.

RESOURCE CENTRES

- Physics Study Area: B400, physics teachers may be available.
- Math & Science Centre: F540, student coaches may be available.
- Learning Centre: B205.

COLLEGE POLICIES

The following college policies will be strictly followed. Consult *The Vanier Student Guide*, the *Vanier College Catalogue*, the *Student Handbook* and your teacher for more information.

General Academic Policies:
It is the student's responsibility to be familiar with and adhere to the Vanier College Academic Policies. These policies can be found online on the Vanier College website. Your attention is drawn in particular to the following policies. A brief summary of each is included.

Student Academic Complaints (Policy number 7210-8): The Vanier College Student Academic Complaints Policy and procedures puts an emphasis on mediation as the primary means to resolve complaints in the academic area. If you have a problem with a teacher and have been unable to resolve it by talking with him or her, you may wish to enlist the help of the Faculty Mediation Committee. The committee member names and contact information are available in Student Services or through the office of the Faculty Dean.

Cheating and Plagiarism (Policy number 7210-31): Any form of cheating or plagiarism will result in a grade of zero on the test or assignment and a letter from the teacher will be placed in your file. A repeated offence may lead to even more serious consequences. Please consult the Vanier Student Writing Guide, the Vanier College Catalogue, the Student Handbook, and your teacher for more information.

Student Misconduct in the Classroom (Policy number 7210-19): This policy provides guidelines for handling cases of student misbehaviour in the classroom and other instructional settings. Such cases may include conduct that is abusive to the teacher and/or other students, or disruptive to the teaching/learning process. This policy does not limit the teacher’s or the College’s right to take immediate action in cases of imminent danger to persons or property.

Zero Tolerance (Policy number 7110-2): The following disruptive behaviours will not be tolerated in any degree on campus: Unauthorized Use of Alcohol/Illegal Drugs; Violence against Persons or Property; Possession of Weapons; Verbal or Written Abuse or Intimidation; Theft or Gambling.

Student Absences for Religious Holy Days (Policy number 7210-20): Students whose religious obligations require them to be absent from the College on a holy day not formally recognized in the College calendar must inform their teachers, in writing, during the first week of classes, of the particular date(s) and times of the religious holy days on which they must be absent. Absences approved in this manner are considered to be excused absences. Students are responsible for material covered in the classes and labs they miss.

Appendix A: Objective 00UR

Field of Studies:	Science
Discipline:	Physics
Course Code:	203-NYA-05
Weighting:	3-2-3
Number of Credits:	2⅔

OBJECTIVE 00UR

MECHANICS

GENERAL FRAMEWORK

This course in Mechanics is designed as a first college-level physics course for students with adequate secondary school backgrounds. Students are expected to be concurrently following the first course in Calculus (201-NYA-05).

The attainment of the objective requires an understanding of the laws and major principles of conservation regulating the motion of bodies.

The learning situations should facilitate the development of the scientific spirit as well as the ability to solve problems and conduct laboratory experiments. The students will be presented with technical applications drawn from daily life and technology. They are also encouraged to use differential calculus in accordance with their previous academic profile. It is suggested that the major stages in the development of mechanics be situated within their historical context.

Students will be required to use a calculator, a computer and other appropriate instruments. A procedure will be outlined for the experimental component.

SPECIFIC INDICATIONS (CONTENT)

Scalar and vector quantities: units and dimensions.
Kinematics of the various aspects of rotation and translation: position, displacement, linear and angular velocity and acceleration.
Force, dynamics of translation and rotation.
Energy and mechanical work.
Principle of conservation of energy and momentum.

PERFORMANCE CRITERIA

Appropriate use of concepts, laws and principles.
Adequate representation of situations in physics.
Use of appropriate terminology.
Graphic component and mathematical expressions adapted to the nature of the problem.
Justification of steps in the analysis of situations.
Rigorous application of Newton’s laws and the principles of conservation.
Critical analysis of results.
Interpretation of the limits of the models.
Meticulous experimentation.
Laboratory report in line with established standards.

COMPETENCY: TO ANALYZE VARIOUS SITUATIONS AND PHENOMENA IN PHYSICS USING THE BASIC PRINCIPLES OF CLASSICAL MECHANICS.	
Specific Elements of the Competency	Standard of Performance: The student must be able to:
1. To describe the translation of bodies in one dimension.	1.1 define and use the concepts of position, displacement, and average velocity over a time interval 1.2 understand and use the concept of instantaneous velocity as the instantaneous rate of change of position 1.3 understand and use the concept of average acceleration over a time interval 1.4 understand and use the concept of instantaneous acceleration as the instantaneous rate of change of velocity 1.5 recognize the limitations of the equations for constant acceleration 1.6 know and use the equations for constant acceleration, in simple applications including freely falling bodies 1.7 use and interpret graphs to understand relationships among variables, including graphs of position and velocity versus time.
2. To describe the translation of bodies in two dimensions.	2.1 distinguish between scalars and vectors. 2.2 sketch the addition of vectors. 2.3 find the components of a vector. 2.4 add vectors by components. 2.5 find the magnitude and direction of a vector from its components. 2.6 express a vector in unit vector notation. 2.7 understand and use the concepts of position, displacement, average velocity, instantaneous velocity, average acceleration and instantaneous acceleration, in 2 dimensions 2.8 recognize and use the vector nature of acceleration, as being involved in any change of speed or direction or both 2.9 understand the concept of relative velocity (Galilean relativity) 2.10 understand projectile motion in terms of separate horizontal and vertical components. 2.11 calculate the position and velocity of a projectile at any time given its initial velocity and position. 2.12 understand how a body in uniform circular motion accelerates because the change in direction of its velocity. 2.13 know and use the relation between acceleration, speed and radius of motion.
3. To describe the rotation of bodies.	3.1 define angular position, displacement, velocity and acceleration by analogy with the corresponding linear quantities. 3.2 know the relations between linear and angular quantities. 3.3 know and use the equations for rotational kinematics with constant angular acceleration.
4. To apply the concepts and laws of dynamics to analysis of the translation of bodies.	4.1 define inertia and distinguish between mass and weight. 4.2 know the common contact forces (normal, tension, friction) and recognize situations where they appear. 4.3 draw labelled isolation diagrams for typical situations. 4.4 calculate the unknown forces in the transitional equilibrium of a point object. 4.5 state Newton's Laws and explain qualitatively how they apply in common situations. 4.6 apply Newton's Laws quantitatively to situations including friction, apparent weight and uniform circular motion. 4.7 state Newton's Law of Universal Gravitation and apply it to point and spherical masses interacting with each other and in circular orbit around a fixed mass.

COMPETENCY: TO ANALYZE VARIOUS SITUATIONS AND PHENOMENA IN PHYSICS USING THE BASIC PRINCIPLES OF CLASSICAL MECHANICS.	
Specific Elements of the Competency	Standard of Performance: The student must be able to:
5. To measure the amount of work and energy involved in simple situations.	5.1 define and calculate the scalar product of two vectors. 5.2 define work as a scalar product and calculate the work done by a constant force and by a force proportional to the displacement (elastic force). 5.3 define and calculate the kinetic energy of an object. 5.4 state the work-energy theorem. 5.5 distinguish between conservative and non-conservative forces. 5.6 define and calculate the potential energy of simple systems (gravitational and elastic). 5.7 define and calculate the mechanical energy of simple systems. 5.8 know the definition of power.
6. To apply the principles of conservation in mechanics.	6.1 understand the principle of conservation of mechanical energy in friction-free systems. 6.2 recognize forms of non-mechanical energy (heat, electrical, chemical, etc.). 6.3 understand and apply general energy conservation to simple systems. 6.4 define and calculate linear momentum as a vector quantity. 6.5 understand the connection between conservation of linear momentum and Newton's Third Law. 6.6 apply momentum conservation to collisions and explosions in one and two dimensions. 6.7 distinguish between elastic collisions, in which both momentum and kinetic energy are conserved, and inelastic collisions.
7. To apply the concepts and laws of dynamics and angular momentum conservation to the analysis of the rotation of bodies.	7.1 define torque as $r\,F\,sin\theta$ and calculate it for simple planar geometries. 7.2 use the concept of symmetry to find the centre of gravity of simple objects. 7.3 calculate the unknown forces or distances for systems in rotational equilibrium. 7.4 define the moment of inertia for a rigid body and calculate it for point masses and for a ring. 7.5 know and use the definition of rotational kinetic energy. 7.6 understand Newton's Second Law for rotation by analogy with the corresponding translational case. 7.7 apply Nll to the rotation of a rigid body about a fixed axis. 7.8 define angular momentum of a rigid body about a fixed axis for a given moment of inertia and angular velocity. 7.9 understand that the conservation of angular momentum applies to a torque-free system. 7.10 apply angular momentum conservation to rotation about a fixed axis.
8. To verify, experimentally, a number of laws and principles in mechanics.	8.1 perform and submit reports on a number of lab experiments, for example, in kinematics, vectors, dynamics, energy conservation, and momentum conservation. 8.2 know how to analyze and present experimental data and draw appropriate conclusions.

Appendix B: Relevant sections in the textbook, *University Physics* by OpenStax

The actual order of topics covered may depart from the possible sequence shown for pedagogical reasons related to the mode of instruction adopted.

Specific elements of competency	**Topics**	**Sections in textbook**
Vectors (Estimated time: 1 week)	Vectors	2.1 – 2.4
Motions (Estimated time: 3 weeks)	1-D kinematics 2-D kinematics Rotational kinematics	3.1 – 3.6 4.1 – 4.5 10.1 – 10.3
Newton's Laws and Applications Including Gravitation (Estimated time: 3.5 weeks)	Newton's laws of motion Applications of Newton's laws Gravitation	5.1 – 5.7 6.1 – 6.3 13.1, 13.2
Work, Energy, Principles of Conservation (Estimated time: 3.5 weeks)	Work and kinetic energy Potential energy Energy conservation Momentum conservation, impulse	7.1 – 7.4 8.1 8.2, 8.3 9.1 – 9.5
Rotational Dynamics and Torque (Estimated time: 2 weeks)	Moment of inertia, rotational kinetic energy, torque Rotational Equilibrium	10.4 – 10.7 12.1 – 12.2
Angular momentum (Estimated time 1.5 week)	Angular momentum	11.2, 11.3

Relevant sections in the textbook, *Fundamentals of Physics*, 10th ed., by Halliday, Resnick, and Walker

The following is a list of topics that will be studied in order to meet the terminal objectives of the course competencies. For each of the topic headings in the syllabus, references are given for the appropriate sections from the course textbook. The content for this course is divided into four Units that build upon each other sequentially. The Final Exam will cover material from all four Units and will be given during the Exam Period. The actual order of topics covered may depart from the possible sequence shown for pedagogical reasons related to the mode of instruction adopted.

Approximate Timelines	**Unit Topics**	**Chapter References (WHR 10th ed.) (Tremblay 2018)**
UNIT I Weeks (1-5)	Introduction to Physics, Mechanics and Measurement	[WHR 1 all sections]
	1-Dimensional Kinematics	[WHR 2 all sections] [T 1 all sections]
	Introduction to Vector Quantities	[WHR 3 all sections] [T 2.1]
	2-Dimensional (vector) kinematics	[WHR 4 all sections] [T 2 all sections]
UNIT II Weeks (5-8)	Concept and Definition of Forces & Newton's laws applied to linear translation	[WHR 5 all sections] [T 3.1, 3.4, 3.5, 3.6; 4.2, 4.3, 4.4, 4.5; 5.3; 7.1, 7.2, 7.4, 7.5]
	Friction	[WHR 6.1] [T 5.1]
	Universal Gravitation	[WHR 13.1 – 13.3] [T 4.1; 6.5]
	Newton's laws Applied to Circular Motion	[WHR 6.3] [T 2.3, 2.4; 6.1, 6.2, 6.3, 6.4; 7.3]

UNIT III Weeks (8-10)	Work and Energy	[WHR 7.1 – 7.4, 3.3] [T 8.1 – 8.2]
	Conservation of Energy	[WHR 8 all sections] [T 9.1 – 9.6]
	Power / Efficiency	[WHR 7.6] [T 8.3]
	Collisions and Linear Momentum	[WHR 9.2 – 9.8] [T 10.1 – 10.8]
UNIT IV Weeks (11-15)	Centre of Mass	[WHR 9.1] [T 11]
	Rotational Kinematics (fixed axis)	[WHR 10.1 – 10.3, 11.1] [T 12.1]
	Rotational Kinetic Energy & Moment of Inertia	[WHR 10.4 – 10.5, 10.8] [T 12.2 – 12.4, 12.7]
	Torque	[WHR 10.6, 3.3, 11.2 - 11.4] [T 12.5]
	Rotational Dynamics & Static Equilibrium	[WHR 10.7, 12.1, 12.2] [T 12.6; 13 all sections]
	Angular Momentum	[WHR 11.5 – 11.8] [T 12.8]

Appendix C: Lab report style guide

Refer yourself to documents posted on LEA.