Les réseaux conceptuels collectifs en enseignement et en apprentissage :

l'usage de TIC pour relier la science scolaire avec la réalité extérieure

Rapport PAREA PA2011-06

Dr. Elizabeth S. Charles, Dr. Nathaniel Lasry, Dr. Kevin Lenton,
Prof. Chris Whittaker, Prof. Michael Dugdale, Prof. Sameer Bhatnagar

Dawson College, John Abbott College and Vanier College

Juin, 2014

Les réseaux conceptuels collectifs en enseignement et en apprentissage : l'usage de TIC pour relier la science scolaire avec la réalité extérieure

La présente recherche a été subventionnée par le ministère de l'Enseignement supérieur, de la Recherche, de la Science et de la Technologie dans le cadre du Programme d'aide à la recherche sur l'enseignement et l'apprentissage (PAREA).

JUIN 2014

Le contenu du présent rapport n'engage que la responsabilité de l'établissement et des auteurs suivants :

Dr. Elizabeth S. Charles (le chercheur principal)
Dr. Nathaniel Lasry (John Abbott College)
Dr. Kevin Lenton (Vanier College)
Prof. Chris Whittaker (Dawson College)
Prof. Michael Dugdale (John Abbott College)
Prof. Sameer Bhatnagar (Dawson College)

DAWSON COLLEGE

Lakeside View Global Vision
CEGEP JOHN ABBOTT COLLEGE

VANIER CÉGEP/COLLEGE

Titre du projet: Les réseaux conceptuels collectifs en enseignement et en apprentissage: l'usage de TIC pour relier la science scolaire avec la réalité extérieure

Subvention PAREA: Aout 2011 – Juin 2014

Rapport soumis le: 30-06-14

Institution Principale: Dawson College
3040 Sherbrooke West
Montréal, QC, H3Z 1A4

Chercheurs:
Dr. Elizabeth S. Charles, Dawson College, echarles@dawsoncollege.qc.ca
Dr. Nathaniel Lasry, John Abbott College, lasry@johnabbott.qc.ca
Professor Chris Whittaker, Dawson College, cwhittaker@place.dawsoncollege.qc.ca
Dr. Kevin Lenton, Vanier College, lentonk@vaniercollege.qc.ca
Professor Michael Dugdale, John Abbott College, michael.dugdale@johnabbott.qc.ca
Professor Sameer Bhatnagar, Dawson College, sbhatnagar@dawsoncollege.qc.ca

Coordonnatrice du projet de recherche:

Dr. Kaila Folinsbee, Coordinator Professional Development & Research, Office of Instructional Development, Dawson College

Dépôt légal — Bibliothèque nationale du Québec, 2014

Dépôt légal — Bibliothèque et Archives Canada, 2014

ISBN 9781-5501674-8-1

RÉSUMÉ

Une préoccupation fondamentale en enseignement des sciences, particulièrement en physique, est de surmonter la difficulté qu'ont les étudiants à acquérir une solide compréhension des principes et concepts de base que l'on nomme généralement le *changement conceptuel*. Il leur est tout aussi difficile de faire ce que l'on appelle un *transfert d'apprentissage*, c'est-à-dire d'appliquer leurs connaissances à d'autres circonstances que celles de l'apprentissage initial. Selon de récentes études, une approche pluraliste serait une solution possible à ce problème. Il s'agirait, par exemple, de mettre les étudiants en situation d'apprentissage collaboratif, de pratiquer une pédagogie fondée sur les faits, et d'utiliser des outils conçus pour soutenir l'apprentissage à l'aide de divers processus sociocognitifs et socioculturels. Le défi est de taille car nous avons beaucoup à apprendre sur la conception de tels outils ainsi qu'à propos de leur intégration à des pédagogies innovatrices. La recherche présentée ici cherche à relever ce défi en examinant l'impact de l'outil d'apprentissage DALITE (*Distributed Active Learning Interactive Technology Environment*).

À la base, DALITE est une mise en application asynchrone de *l'apprentissage par les pairs*, une approche pédagogique qui demande aux étudiants de répondre à des questions conceptuelles, puis de partager leurs réponses avec d'autres étudiants selon une méthode de *réflexion-jumelage-partage*. DALITE est un outil en ligne qui offre un scénario faisant franchir diverses étapes cognitives et sociocognitives aux étudiants : <u>sélectionner</u> la réponse à une question à choix multiples; <u>justifier</u> ce choix par écrit; <u>reconsidérer</u> à la lumière de solutions écrites par des pairs; <u>resélectionner</u> une réponse; <u>voter</u> pour le choix le plus convaincant; et <u>prendre</u> connaissance de la justification d'un expert.

L'étude de DALITE et de son évolution représentent une expérience de *recherche fondée sur la conception* (design-based research, ou DBR), ce qui signifie que la conception de l'outil est basée sur des principes théoriques. Son étude a donc des implications tant pratiques que théoriques. DALITE est conçu de façon à faire appel aux principes d'apprentissages suivants : (1) utiliser le vocabulaire de la discipline en question (son discours); (2) réfléchir aux explications, en comparant la sienne à celles des autres (auto-explications et explications interactives); (3) étudier explicitement la structure conceptuelle (caractéristiques profondes et superficielles); et (4) établir les connaissances préalables requises (échec productif). La présente recherche se divisait en trois études, toutes situées dans le contexte d'un cours collégial d'introduction à la physique, soit Physique NYA.

L'étude 1 employait un modèle de recherche quasi expérimentale et posait deux questions : (1) *L'utilisation de DALITE, intégrée à une méthode pédagogique d'apprentissage actif, favorise-t-elle le changement conceptuel chez les étudiants et améliore-t-elle leurs résultats en apprentissage conceptuel, comparativement à d'autres méthodes pédagogiques qui n'y font pas appel*? Plus précisément, *Les étudiants qui utilisent DALITE acquièrent-ils une compréhension nettement meilleure des concepts que les étudiants qui, avec les mêmes connaissances préalables, n'utilisent pas DALITE? Comment les étudiants qui utilisent DALITE se comparent-ils aux étudiants qui ne l'utilisent pas, mais qui ont recours à l'apprentissage par les pairs en face à face? L'efficacité de DALITE dépend-elle de son utilisation par l'enseignant*? (2) *Quel impact ont les activités effectuées sur DALITE sur les connaissances conceptuelles des étudiants, telles qu'évaluées par les questions conceptuelles généralement posées en examen?* Le

groupe de traitement DALITE (N=168) se divisait en cinq sections, toutes constituées d'étudiants en première année de science, inscrits au cours de Physique NYA. Le sexe et l'âge des participants se situaient dans la norme. Les cinq sections étaient réparties au sein de trois collèges, et quatre enseignants se divisaient les cours, chacun possédant des degrés variables d'expérience des pédagogies d'apprentissage actif. Des devoirs en ligne sur DALITE étaient assignés chaque semaine. Les groupes témoin étaient composés de deux cohortes : (1) un groupe témoin général (N=2000) formé d'étudiants issus d'un large éventail d'expériences pédagogiques; et (2) un groupe d'apprentissage actif (N=190), sélectionnés à l'aide d'une méthode d'*échantillonnage déterminé* afin d'assurer une comparaison authentique entre l'outil et la pédagogie. Nous avons utilisé le *Force Concept Inventory* (ou FCI) pour les prétests et les post-tests. Les résultats démontrent une différence significative entre les groupes DALITE (cinq groupes de traitement) et le groupe témoin général (0.47 ± 0.02 vs 0.35 ± 0.006; $p<0.00001$). La comparaison des groupes DALITE aux groupes d'apprentissage actif n'a révélé aucune différence statistique (0.47 ± 0.02 vs 0.48 ± 0.02; $p=0.84$). La comparaison des cinq groupes de traitement entre eux n'a pas non plus révélé de différence statistique. L'étude 1 utilisait aussi trois évaluations conceptuelles de fin de module, qui comparaient la compréhension conceptuelle entre les cinq groupes de traitement et l'un des groupes témoin en apprentissage actif, soit ce que l'on considère comme des tâches de transfert. Les résultats montrent de légères différences entre les groupes de traitement, ainsi que de légères différences entre le groupe de traitement DALITE et le groupe témoin en apprentissage actif. Ces résultats suggèrent une amélioration modeste des capacités de transfert (c.-à-d. une assimilation plus approfondie du savoir conceptuel) pour la section qui a le plus fréquemment utilisé DALITE.

L'étude 2 effectuait une étude de cas qui incluait des méthodes ethnographiques et posait la question : *Quels sont les impacts de l'activité en ligne avec DALITE et du système entier, incluant le marquage et la production de cartes conceptuelles*? Il s'agissait d'analyser diverses données qualitatives, dont les justifications de DALITE, les données recueillies lors d'exercices de tri, les artéfacts issus de la production de cartes conceptuelles, et les observations en classe. Le rapport documente certaines de ces analyses, qui révèlent des différences entre les cinq sections de traitement DALITE aussi bien qu'au sein de chacune d'elles. Ces différences de capacités et de compréhension croissante de la nature épistémique de la discipline semblent causées par des facteurs tels que : (1) l'intensité du recours à DALITE; (2) l'intensité de la participation de l'étudiant aux activités périphériques, soit la production de cartes conceptuelles; (3) le temps consacré par les groupes d'étudiants, au sein d'une section, à travailler entre eux et avec la classe entière.

L'étude 3 employait une méthode de recherche fondée sur la conception et posait la question : *Comment les étudiants ont-ils perçu le rôle de DALITE, et travaillé avec l'outil, comparativement à ce que visait sa conception*? En plus des données qualitatives décrites ci-dessus, l'étude 3 comprenait aussi des entrevues et des enregistrements vidéo d'étudiants travaillant dans DALITE avec des protocoles de pensée à voix haute. Les résultats démontrent qu'en général, les étudiants travaillaient avec DALITE de la manière prévue, consacrant plus de temps à l'apprentissage : (1) des processus discursifs – utiliser le langage de la physique; (2) des processus métacognitifs – réfléchir aux explications (les siennes et celles des autres); (3) des processus cognitifs de comparaison et de contraste – repérer les similitudes et les différences conceptuellement profondes; et (4) des processus autorégulateurs – repérer les connaissances requises et être mieux préparés pour le cours.

Ce rapport comprend également un survol des découvertes périphériques faites durant ce projet. Certaines d'entres elles sont : l'évaluation de l'efficacité des mesures d'évaluations utilisées pour mesurer l'impact de DALITE et les découvertes faites par les enseignants qui utilisaient le logiciel. Les enseignants s'entendent pour dire que le logiciel fournit une rétroaction immédiate et détaillée et appuie fortement la méthode de classe renversée de même que l'approche d'apprentissage actif en général. De plus, ceux-ci trouvent que les étudiants sont mieux préparés pour le cours et les enseignants peuvent repérer les problèmes conceptuels avant le cours et adapter ce dernier pour mettre l'accent sur ces problèmes précis.

ABSTRACT

A fundamental concern in science education, and physics in particular, is to overcome the difficulty students experience in building robust understandings of core principles and concepts, generally referred to as *conceptual change*. Equally difficult to achieve is the application of knowledge in settings that are different from the original learning, referred to as *transfer of learning*. Recent studies tell us that pluralistic approaches may be an answer. For instance, engaging students in collaborative learning, using evidence-based pedagogy, and using tools designed to support learning through various socio-cognitive and socio-cultural processes. This is no easy challenge and there is much to learn about how to design such tool and how they can work together with innovative pedagogies. This current research is aimed at addressing this challenge by examining the impact of the designed learning tool called the Distributed Active Learning Interactive Technology Environment (DALITE).

To start, DALITE is an asynchronous implementation of *Peer Instruction* (PI) – an instructional approach involving students answering conceptual questions then sharing their answers with others in a *think-pair-share* method. As an online tool, DALITE uses a script that takes students through several cognitive and socio-cognitive steps: select multiple choice answer; write rationale; reconsider based on alternatives; re-select; vote on most convincing; and review an expert's rationale.

The study of DALITE and its evolution is a *design-based research* (DBR) experiment, which means the design of the tool is grounded on theoretical principles therefore researching the tool has both practical and theoretical implications. The specific learning principles used in the design of DALITE are: (1) the use the language of the discipline (discourse); (2) reflections on explanations, one's own and that of others (self-explanation and interactive-explanation); (3) explicit attending to conceptual structure (deep vs. surface features); and (4) eliciting prior knowledge and establishing a need to know (productive failure). This current research was divided into three studies all situated within the context of introductory college physics course - Physics NYA.

Study 1 used a quasi-experimental research design and addressed two questions: (1) *Does the use of DALITE, as part of an active learning pedagogical approach, promote students' conceptual change and increase their conceptual learning outcomes compared to other pedagogical approaches where it is not used*? Specifically, *Do students using DALITE gain statistically significantly more conceptions than students with the same prior knowledge who do not use DALITE? How do students using DALITE compare to students who do not use DALITE but use real-time face-to-face Peer Instruction? And, does the effectiveness of DALITE depend on the instructor using it*? (2) *What is the impact of the DALITE activities on the students' conceptual knowledge, assessed by the common conceptual test questions?* There were five sections of the DALITE treatment (N=168), all first year science majors enrolled in Physics NYA. Gender and ages were standard. These five sections were spread across three colleges and taught by four instructors, each with different levels of experience with active learning pedagogy. DALITE was assigned weekly as homework via the web. The comparison groups were composed of two cohorts: (1) a general comparison group (N=2000) that included students from a wide range of pedagogical experiences; and, (2) an active learning group (N=190), selected using a *purposeful sampling* method to ensure an authentic comparison of the tool

versus the pedagogy. The Force Concept Inventory (FCI) was used as a pre and posttest. Results show a statistically significant difference between DALITE (five treatment groups) and the general comparison group (0.47±0.02 vs 0.35±0.006; $p<0.00001$). Comparing DALITE to the active learning groups showed no statistical difference, (0.47±0.02 vs 0.48±0.02 ; $p=0.84$). Additionally, comparing the five treatment groups to each other showed no statistical difference. Study 1 also used three end-of-unit conceptual assessments, which compared the conceptual understanding between the five treatment groups and one of the active learning comparison classes, what are considered transfer tasks. These results show small differences between the treatment groups, as well as small differences between the DALITE treatment and the active learning comparison group. These results suggest a modest improvement in the transfer capabilities (i.e., deeper processing of conceptual knowledge) in the section that used DALITE the most frequently.

Study 2 used a case study including ethnographic methods and addressed the question: *What is the impact of the DALITE web-based activity and the extended system, including tagging and conceptual mapping activities*? This involved the analysis of a variety of qualitative data including DALITE rationales, data collected from sorting task exercises, concept mapping artifacts and classroom observations. The report documents some of these analyses that show differences both between and within the five DALITE treatment sections. These differences in capabilities and in growing awareness of the epistemic nature of the discipline appear to be based on factors such as: (1) extent to which DALITE was used; (2) levels to which student's engaged with the extended activities – e.g., production of concept maps; (3) degree to which the groups of students, within a section, worked among themselves and with the class as a whole.

Study 3 used a design-based research method and addressed the question: *How did students perceive the role of DALITE, and work with DALITE, compared to what is intended by its design*? In addition to the qualitative data described above, Study 3 also included interviews, and video recordings of student's working with DALITE using think-aloud protocols. Results show that generally speaking students worked with DALITE the way it was intended by providing more practice on: (1) discursive processes - using the language of physics; (2) metacognitive processes - reflection on explanations (self and others); (3) cognitive processes of compare and contrast - identification of conceptually deep similarities and differences; and (4) self-regulatory processes - eliciting a "need to know" and better preparation for class.

This report also includes an overview of what was learned about the effectiveness (strengths and weakness) of the assessment measures designed to capture the effectiveness of DALITE. These results suggest that some of the assessments are very good and could be developed further. Meanwhile, others are less productive. Lastly, we report on the instructor's view of using DALITE. Arguably, it has some important benefits as a practical tool for teaching. It provides immediate and detailed feedback to instructors and greatly supports the flipped classroom method or active learning approach in general. Students are better prepared for class and teachers can identify conceptual issues before class and tailor their class to focus on these specific issues.

ACKNOWLEDGEMENTS

We would like to thank the following individuals for their support and contribution of time and expertise towards the completion of this research project.

PAREA
Caroline Mongrain, Conseillère en innovation sociale
Direction de l'innovation et du transfert
Direction générale de l'innovation et de la recherche
Ministère de l'Économie, de l'Innovation et des Exportations

Office of Professional Development & Research at Dawson College
Kaila Folinsbee
Maeve Muldowney
Suzanne Prevost

Finance Office
Diane Wong
Linda Gregoire

Administrative level support at Dawson College
Robert Kavanagh, Academic Dean
Raymond Bourgeois, Dean of Science, Medical Studies & Engineering
Barbara Freedman, Dean of Instructional Development

Technical support:
Koya Charles, DALITE Programming
Sarah De Guzman, Media Department
Jesse Binstock, Printing

Faculty support:
Jean-François Briere

Research assistants:
Jonathan Guillemette
Chao Zhang
Xihui Wang

Technical assistance:
Hélène Dansereau (translations)

Lastly, we would like to thank the incredible group of students who participated in the research project in the Fall 2012 and 2013, as well as those who participated in the pilot study Fall 2011. These enthusiastic and bright young people are an inspiration to us all and are testament to the depth of learning that is possible at the college level.

Table of Contents

CHAPTER 1
INTRODUCTION

Science educators are increasingly adopting pedagogical approaches that are grounded in empirically based teaching practices (e.g., collaborative inquiry-based, problem-based learning) and current learning theories (e.g., social constructivism, socio-cultural theories). Much of the recent QEP reform, at the elementary and high school levels, is based on such practices and their associated changes to curricula structures. Implementations of these new approaches have been shown to produce improvement in students' learning (Barron, et al., 1998; Charles, Lasry, Whittaker & Trudeau, 2009; Dochy, et al, 2003; Lasry& Aulls, 2007; Schauble, Glaser et al., 1995; Vernon & Blake, 1993). Nonetheless, two aspects of learning remain difficult to improve and require further efforts – (1) transfer of learning; and, (2) willingness and ability to extend context and engage in learning beyond the classroom.

A fundamental concern in science education, and physics in particular, continues to be the difficulty students experience building robust understandings of core principles and concepts along with their ability to use them in different settings – i.e., conceptual change and transfer of learning. Recent studies and practice-based efforts to address these problems tell us that science learning and teaching can benefit from pluralistic approaches (e.g., Treagust & Duit, 2008). These include changes to the ways students engage with the content and each other, the ways teachers orchestrate and use new pedagogical approaches, and the ways we design tools that support these various socio-cognitive and socio-cultural processes.

The Distributed Active Learning Interactive Technology Environment (DALITE) is a web-based tool that aims to promote conceptual learning while working within an asynchronous mode of student engagement. It involves the learners in a variety of tasks including writing explanations for conceptual questions, reflecting on and comparing these explanations to those of peers and experts, and taking part in the social construction of the database repository by voting on the most convincing explanations. It is part of a larger study and system of social constructivist pedagogical practices conceived to promote learning in physics at the postsecondary level. Its design draws on social conceptions of conceptual change, recognition of the role of context (e.g., Engle, 2006) and the success of the practical approach to conceptual learning called Peer Instruction (Mazur, 1997).

This report of the DALITE study will describe the efforts of the research team to both engage in a design-based research (DBR; Hoadley, 2002) as well as test the hypotheses of conceptual change. We start with a full description of the theory and design behind the DALITE intervention. Additionally, we make note that this project is an example of a co-design process (Penuel, Roschelle & Shechtman, 2007). In other words, we recognize the importance of the co-design process that plays a significant role in guiding the work and accomplishments of the researchers/designers and instructor/practitioners team who designed and developed the DALITE learning system over the course of three successive iterations between 2011 and 2013.

1.1 Problem Statement

Transfer of learning, or the ability to use one's prior knowledge across contexts has long been the "Holy Grail" of educators and education but continues to be difficult to achieve (e.g., Bransford & Schwarz, 1999). It is often noted by both researchers and teachers, alike, that students do not make connections between content in one course and content in another. In fact, the literature shows that students experience great difficulty relating the closely tied disciplines of math and physics (e.g., Cui, Rebello & Bennett, 2005). Not to mention, at the specific program and departmental levels we often hear teachers comment on the lack of transfer capability demonstrated by their students both within their discipline course and even within the same course. Take for instance the introductory physics course where students seldom see the relationship between Newton's Third Law and momentum.

Adding to this, there is the problem of students' willingness to extend learning beyond the classroom doors, which has implications on how they view and use their relationship to science (identity and disposition). More specifically, studies have shown that students enter science classrooms with certain beliefs about the relevance of their personal knowledge and experiences as it relates to school science. At the end of a semester of formal science instruction, students almost invariably perceive their personal knowledge and experiences as ***less*** relevant to their science courses (Redish, Saul, & Steinberg, 1998; Perkins, et al, 2004). In other words, classroom science is seen as different and separate from the workings of the world; and these two contexts are often kept disconnected in the student's mind. Additionally, students often do not have the disposition to transfer this school learning to other settings (Bereiter, 1995). It might be further argued that they do not because there is little reason to do so. In fact, Nobel Laureate Carl Wieman contends that the lack of connection is a consequence of how science is taught in our schools (Wieman & Perkins, 2005). He also suggests that this may even explain why many students often rethink their decision to continue in science as a major. Much work is needed to better understand whether and how instructional initiatives can help solve this problem; in the process, help to provide students with the general sense/belief that science is very much part of their everyday lives – what we will describe as developing a "science identity" from hereon.

In addition to the dispositional and volitional perspective, described above, there is evidence suggesting that there is a trajectory of learning – i.e., "preparation for future learning" (Bransford &Schwartz, 1999). These authors also echo the importance of volition or disposition but add the notion of change characterized by the act of "noticing." They claim this is an important developmental step and should be considered as evidence of transfer. They state: "noticing new features is not an act of simply finding common elements between the past and present. Through contrasting cases, one develops the ability to notice finer and finer distinctions. One becomes a connoisseur of the world" (p.92).

Along with noticing, Bransford and Schwartz include the development of a particular type of self-awareness – i.e., the awareness of one's readiness to learn. They characterize a type of epistemic belief change. They state, "Future learning frequently requires "letting go" of previous ideas, beliefs and assumptions. Effective learners resist "easy interpretations" by simply assimilating new information to their existing schemas; they critically evaluate new information and change their views (accommodate) when necessary." (p.93). Such ideas are consistent with the conceptual change literature where ontological and epistemic shifts are described as mechanisms that promote change (e.g., Chi et al., date; Jacobson & Archodidou, 2000). Bendixen and Rule (2004), identify personal epistemology and conditions for change. In short, there may be a sensitization phase during which learners not only begin to take on aspects of the new identity – i.e., "becoming" – but also begin to take up aspects of the new practices or skills. This phase may be characterized by subtle shift in awareness or attention that include changes in belief systems.

Finding ways to close these gaps we turn to twenty-first century interactive information communication technology (ICT). There is growing evidence that such tools can extend context and support learning (Stahl, Koschmann, & Suthers, 2006). But to date, they have not played a central role in bridging between the home and classroom environments. Our interest therefore is to explore how ICT tools can support transfer and development of conceptual change.

In the remainder of this document we report on the investigations we conducted as we explore the role of the ICT tool developed for the purpose of supporting conceptual change and transfer. In addition, we examine how this tool, along with the pedagogical implementations we consider *active learning*, might promote students' awareness of contexts and changes in epistemic beliefs. In doing so, we explore how such educational solutions might prepare students for future learning.

1.2 Research Objectives

The *DALITE System* includes the web-based homework system, supporting curricula materials (concept mapping and tagging activities) and active learning pedagogical approach. In this report we will make reference to both this system of interacting parts, as a whole, as well as to its specific components. The global objectives of this research were to answer questions emerging from the use of the tool and how it promotes a deep understanding of the domain content, and how it influences students' transfer of learning capabilities, as defined earlier. To do this, the research is divided into three studies:

Study 1: investigates whether the treatment condition, DALITE, promotes deeper conceptual understanding and transfer of learning, compared to a control condition (addressed by Research Question 1). In doing so the DALITE treatment groups will be compared to control groups. And, to examine changes in students' attitudes characterized as their epistemic beliefs (Research Question 2).

Study 2: investigates the development of practices emerging from the use of the DALITE treatment (Research Question 3). Specifically, the research focuses on the norms and practices (Cobb, 2002) that enable or prevent students, and the treatment teachers, from fully adopting DALITE.

Study 3: documents the development of DALITE as a *design-based research* (DBR; Cobb et al., 2003) project (Research Question 3).

1.2.1 Research Questions:

1. Does the use of the DALITE system (the experimental treatment condition) promote student's conceptual change and preparation for transfer?
    a. Do students taught with the DALITE condition perform better on a standardized conceptual question tests (e.g., the FCI) compared to students in a control treatment condition?
    b. What was the impact of the DALITE condition on the students' conceptual knowledge, assessed by the common conceptual test questions? How do these results compare to a comparison treatment?
    c. How do students in the DALITE condition perform on immediate transfer tasks (sorting task) using specially designed problem questions?

2. What is the impact of the DALITE web-based activity and the extended system of conceptual mapping?
    a. Is there evidence of conceptual change?
    b. Is there evidence of epistemic belief change?

3. How are DALITE's designed features taken up and how does this match the theoretical underpinning?
    a. What factors facilitate and/or constrain the use of DALITE?
    b. What might learn about conceptual change and preparation for transfer from this design?

1.3 References

Barron, B., Schwartz, D., Vye, N., Moore, A., Petrosino, A., Zech, L., et al. (1998). Doing with understanding: Lessons from research on problem-and project-based learning. *Journal of the Learning Sciences, 7*(3), 271-311.

Bendixen, L. D., & Rule, D. C. (2004). An integrative approach to personal epistemology: A guiding model. *Educational Psychologist*, *39*(1), 69-80.

Bereiter, C. (1995). A Dispositional View of Transfer. In A. McKeough, J. Lupart & A. Marini (Eds.), *Teaching for Transfer: Fostering Generalization in Learning* (pp.21-34). Mahwah, NJ: Lawrence Erlbaum Associates.

Bransford, J. D. & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (Vol. 24, pp. 61–100). Washington, DC: American Educational Research Association.

Charles, E.S., Lasry, N., Whittaker, C. & Trudeau, J. (August 2009). *Technology Supported Collaboration and Learning: How do we build learning environments to build communities & conceptual knowledge?* Technical report for PA2007-014 (ISBN number), for funding agency PAREA, coordinated through Dawson College, QC.

Chi, M. T., Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science*, *18*(3), 439-477.

Cui, L., Rebello, N. S., & Bennett, A. G. (2005). *College Students' Transfer From Calculus to Physics.* Paper presented at the Physics Education Research Conference, Salt Lake City, UT.

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: a meta-analysis. *Learning and instruction, 13*(5), 533-568.

Engle, R.A. (2006). Framing interactions to foster generative learning: A situative explanation of transfer in a community of learners classroom. *Journal of the Learning Sciences*, 15(4), 451-498.

Hoadley, C. (2002). Creating context: Design-based research in creating and understanding CSCL. In G. Stahl (Ed.), *Computer Support for Collaborative Learning,* (pp. 453-462). Mahwah, NJ: Lawrence Erlbaum Associates.

Jacobson, M. J., & Archodidou, A. (2000). The design of hypermedia tools for learning: Fostering conceptual change and transfer of complex scientific knowledge. *The Journal of the Learning Sciences*, *9*(2), 145-199.

Lasry, N., & Aulls, M. (2007). The effect of multiple internal representations on context-rich instruction. *American Journal of Physics, 75*(11), 1030-1037.

Mazur, E. (1997). *Peer Instruction* : A User's Manual. Upper Saddle River, NJ: Prentice Hall

Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process.*Research and Practice in Technology Enhanced Learning*, *2*(1), 51-74.

Perkins, K., Adams, W., Pollock, S., Finkelstein, N., & Wieman, C. (2004). Correlating Student Attitudes With Student Learning Using The Colorado Learning Attitudes about Science Survey. *submitted to PERC Proceedings*.

Redish, E., Saul, J., & Steinberg, R. (1998). Student expectations in introductory physics. *American Journal of Physics, 66*(3), 212-224.

Schauble, L., Glaser, R., Duschl, R., Schulze, S., & John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. *Journal of the Learning Sciences, 4*(2), 131-166.

Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. *Cambridge handbook of the learning sciences*, *2006*.

Treagust, D. F., & Duit, R. (2008). Conceptual change: A discussion of theoretical, methodological and practical challenges for science education. *Cultural Studies of Science Education*. Published online 29 March 2008: doi:10.1007/s11422-008-9090-4.

Vernon, D., & Blake, R. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Academic medicine, 68*(7), 550.

Wieman, C., & Perkins, K. (2005). Transforming physics education. *Physics Today, 58*(11), 36.

CHAPTER 2
THEORETICAL FOUNDATIONS AND RATIONALE

By now, most educators and educational policy makers are familiar with constructivist and social constructivist theories of learning that inform the design of instruction and pedagogical activities. What these theories share is the key assumption that knowledge is constructed from experiences with the world, including use of tools such as language, inscriptions, and disciplinary artifacts and practices that mediate learning (Vygotsky, 1978).

The theory of *situated cognition* adds a further dimension to the thinking about the process of learning and highlights the need for learning activities to be contextualized and embedded in *authentic activity* (i.e., realistic and/or typical to the domain), context and culture of the domain users (Brown, Collins & Duguid, 1989; Lave, 1988; Lave & Wenger, 1999). Accordingly, Brown et al., (1989) suggest that conceptual knowledge is not "neutral" and cannot be abstracted from the situations they are derived from, nor from the domain activities from which they are taken. It is therefore critical to actively think of building everyday contexts into in-class activities.

There are four important aspects to this framework: (1) the purpose of knowledge is understood in its use (this includes knowledge embedded in the tools of a domain): (2) understanding is an active process; (3) learning involves understanding when and how to use knowledge; (4) abstraction of knowledge is induces when learning takes place in multiple contexts -- "[t]his unbinding of knowledge from a specific context fosters its transfer to new problems and new domains" (Collins, Brown & Holum, 1991, p.44).

2.1 Background

Investigating how students learn physics has been a perennial concern not only of physics education research (PER) but also of the learning sciences (e.g., diSessa & Sherin, 1998). The body of research generated by both communities confirms that conceptions of the physical world, such as force, motion, and acceleration, are difficult to change with traditional instruction (e.g., Hestenes, 1992). However, studies of social constructivist instruction, popularly referred to as *Active Learning*, report findings of statistically significant gains in students' conceptual understanding in physics and other science disciplines (Freeman, Eddy, McDonough, Smith, Okoroafor, Jordt, & Wenderoth, 2014; Meltzer & Thornton, 2012). Of particular interest are implementations that focus on promoting conceptual change by placing an emphasis on intentional reflection (Sinatra & Pintrich 2003). Adding to this are questions about the processes involved in self-explanation (Chi, de Leeuw, Chiu & LaVancher, 1994) versus forms of peer explanations such as reciprocal teaching (Palincsar & Brown, 1984), and other collaborative and discursive practices (Stahl, 2006). In fact, it might be argued that there is value in examining the processes of what might be referred to as "interactive explanation" (Ploetzner, Dillenbourg, Preier & Traum, 1999). This interception between explaining to others, as well as reflecting on one's own explanation provides a power nexus for investigation. We propose that such a nexus is found in the variation on Peer Instruction that is at the heart of our designed intervention, DALITE.

2.1.1 Transfer of learning

According to Rebello (2005) traditional views of transfer have described it as a static process where students apply their prior knowledge to new situations, which takes the predefined perspective of the teacher/researcher. From this position transfer is seen as particularly difficult when the "source" features (the material being taught) and the "target" features (the assessment material) are different (e.g., Brown, et al., 1989; Genter & Toupin, 1986). The problem is that learners do not recognize the deep structures in the target materials. Instead, they are distracted by surface similarities or superficial resemblances between source and target materials.

Recent views of transfer describe it as a dynamic process of constructing knowledge within the new context, which brings it into line with current socio-cognitive and socio-cultural theories of learning (see Journal of the Learning Sciences, special issue, 2006). Bransford and Schwartz (1999) suggest that we think of transfer as how prepared learners are to engage in transfer tasks – coined "preparedness for transfer." From a situated theory perspective, scholar such as Greeno and colleagues (Greeno, 1997; Greeno, Smith & Moore, 1993) define transfer "as the extent to which participating in an activity in one situation influences one's ability to participate in another activity in a different situation." They propose that what transfers is the "patterns of participation" across the different contexts and not the knowledge of the task itself. From this perspective the student as actor takes center stage (e.g., Actor-Oriented Transfer, Lobato, 2003). In summary, the critical distinction between the old and this new view of transfer is how the learner is understood as seeing two contexts as similar or different.

Adding to this discussion, Engle (2006) advances the hypothesis that a mechanism for transfer is the relationship (*framing*) between the learning activities, the students' expectations, and the contexts, what she calls *intercontextuality*. Her results show two kinds of framing (i.e., ways of attending and interpreting experiences) that are productive for promoting transfer: (a) when the learning activities are framed as part of a larger ongoing (temporally connected) intellectual conversation actively involving the students; and, (b) when teachers frame students as contributing members of a broader community of learners who are also involved in the same intellectual endeavor. In doing so, students' expectations play a major part in their success in generatively using what they were learning and make better use of the available content-based supports for transfer. This current project defines transfer from this "situated" perspective.

2.2 Role of Identity and Disposition in Learning

The socio-cultural movement in education has brought attention to the role of social, motivational, affective and cultural factors in learning. In doing so, we have become more sensitive to the importance of students' sense of identity, disposition and agency as they enter into the classroom and move beyond. Lave and Wenger (1991) define the process of developing identity, as related to a community, as participation in and alignment with the thinking of the community of practice. Others such as Gee (2000) talk about the importance of the big "D" discourses that individuals engage in as they form a sense of belonging. In the case of academic settings, such Discourse is more than the language and knowledge but includes ways of doing and thinking within a discipline – i.e., the values, practices and epistemic beliefs of the discipline. Accordingly it is these social and cultural factors that mediate learners' participation in the activities of the communities they identify with (Lave, 1996).

In this current research we are interested in disposition and identity as part and parcel of the motivating processes for learning; and, more importantly for being willing to transfer learning beyond the context of the immediate classroom. Looking specifically at the factor of disposition, Gresalfi (2009) describes it as the following:

> [L]earning is a process of developing *dispositions*; that is, ways of being in the world that involve ideas about, perspectives on, and engagement with information that can be seen both in moments of interaction and in more enduring patterns over time (Gresalfi & Cobb, 2006). … In short, dispositions capture not only to what one knows but how he or she knows it; and not only the skills one has acquired, but how those skills are leveraged. (p. 329).

This connection between who we are and what we are willing to do is very strong. In fact, Bereiter (1995) argues that disposition is a critical factor in learners' engaging in transfer of learning. As such, it is incumbent on instructional designers and researchers to consider these social, affective and cultural factors when selecting pedagogy and designing learning, and learning environments to promote transfer. What is becoming clear is that identity (sense of belonging) and disposition are mutually constituted and mediate learning. Thus this current research not only investigate how to promote science identity, but is also guided by these socio-cultural ideals when selecting the pedagogy, the design tools and environment, and assessment of the learning outcomes. Next we will describe the instructional approach of the pedagogy.

2.3 Active Learning as an Instructional Approach

The instructional approach that is consistent with situated learning theory is the *cognitive apprenticeship approach* proposed by Collins et al. (1991) including a "sequence of methods that provide opportunities for students to observe, engage, invent and discover content and learning strategies." More specifically, it involves making thinking visible and providing ways to promote the development of domain competencies and expertise, including transfer learning. It is composed of six component processes: *modeling* of expert performance and practices of the domain; *coaching* consisting of teachers observing students in practice; *scaffolding* of task in action (e.g., just-in-time teaching); *articulation* encouraging time-on-task activities; *reflection* involving metacognitive activities at the individual, group, plenary and between novices and more experienced others – including the *community of practice* (Lave & Wenger, 1991) for the domain in question; and, *exploration,* or time for inquiry.

In designing instruction, modeling, coaching and scaffolding are critical ways to develop the learner's cognitive mechanisms related to attention. That is, they enable the designer/teacher/expert to place filters on experiences, thereby focusing attention, structuring perception, thoughts and emotion (Bransford et al., 2006). Thus these are important factors to consider when assessing the role of the teacher in designing which resources should be made available within the learning environment. Recall that active learning forms the pedagogical approach used as part of the DALITE system. In the next section we will describe how technology can be used to promote transfer learning and connect students to a broader context that provides support for an emerging epistemic belief change.

2.4 Role of Information Technology

Technology-enhanced learning environments are defined as general computer and communication-based systems that encompass a wider class of learning tools or curricula materials that aim to support a variety of individual and collective cognitive and social tasks. This includes authorship and customization of curriculum, registration of teachers and students, storing, retrieving and operating on student-generated learning data, and service of curriculum in classrooms or other educational settings (Zimmerman & Slotta, 2010). In the process, such systems provide new functionality for curriculum design, including expanding the contexts for learning (bringing the world to the classroom and the classroom to the world), providing ways to visualize knowledge (including the modeling of how experts think), providing access to broader social networks, and providing shared workspaces and knowledge resources such as the development of "*third space*" communities (Oldenburg, 1991).

2.4.1 Expanding contexts through the use of ICT tools and networks

Web 2.0 has changed our understanding of how scientific data is collected, how individuals can participate in science, and how data can be mined and organized (e.g., semantic networks of information). Coining the phrase, a *Fourth Paradigm*, Gray (2007) proposes that scientific processes are shifting toward more data-intensive and collaborative methodologies, as well as toward more socially connected processes (e.g., large international collaborations at the Large Hadron Collider in Geneva or the interdisciplinary work of the Intergovernmental Panel on Climate Change).

In the process, this new paradigm is expanding access and creating new pathways to science. Interestingly, an increasing number of mainstream science researchers now use data collected through the cell phones of citizens that have no formal science training. Such citizen science offers new pathways of extending contexts and providing students with new ways to become involved in science. The ubiquity of mobile devices plays a central role in allowing individuals to participate in this scientific research.

Education has taken note of this potential and some initiatives have been explored in the last decade. For instance, in one of the earliest efforts to explore the use of handhelds in learning, Soloway, Norris and Blumenfeld (2001) have applied the principles of Learner Centered Design. They use cell phones as data collection devices that support inquiry learning. Others, such as the WISE research group have used Palm IIIc handhelds interconnected with activities that were part of the WISE curriculum. Still others, such as BioKIDS (Songer, 2006), RAFT (Kravcik et al., 2004) have also designed and investigated ways to use these portable personal devices to promote better learning.

This research adopts and adapts some of these experiences to promote students' understanding of how science extends beyond the classroom walls. At the same time we have designed an ICT enhanced learning environment that allows for the coordination and integration of these personally collected data back into the curriculum. Our research looks at specific questions related to learning, as well as helps inform the design of the learning environment as part of a design-based experiment (Cobb et al., 2003). Next we describe this new ICT enhanced environment, DALITE, designed as part of a co-design collaboration[1].

2.5 Peer Instruction Approach to Learning

Peer Instruction (PI) is an example of an evidence-based pedagogical innovation popularized by Eric Mazur (Mazur, 1997). Its method of engaging students in scientific discourse focuses on acts of explanation, comparison, and reflection that lead to conceptual change. Meltzer (2013) states that, at the postsecondary level, PI is one of the most widely used active learning approaches in North America. No doubt in large part because of the growing body of research supporting claims of its efficacy in producing statistically significant conceptual gains (e.g., Crouch & Mazur, 2001).

In PI implementations, instructors present students with multiple-choice conceptual questions that the students answer using wireless handheld devices, colloquially referred to as *clickers*. These initial polling activities provide instructors with real-time feedback on the status of students' understanding. Answering these questions allows instructors to know whether or not concepts are known, somewhat known or unknown to students. The PI script is described in Figure 2.1. If conceptual understanding falls within the "known" range (correctly answered by more than 70% of students), the teacher can move forward to another concepts and questions. If it is "unknown" (correctly answered by less than 30% of students), the teacher is advised to revisit the ideas. The real peer-to-peer interactions only come into play with the "somewhat known" concepts (30-70% correctly answered). When responses fall within this range, students are asked to turn to their neighbor and discuss their answers and reasoning. It is arguable that these discursive practices allow students to engage in sense making and intentional reflection on these specific concepts.

Some of the most successful implementations of PI have been those found in large lecture halls with hundreds of students. In such settings, rich discussions can arise because of the larger probabilities of having greater diversity among students, which undoubtedly acts to amplify the cognitive dissonance. However, in smaller classrooms there is often less diversity between students' answers and their understandings leading to a paucity of conceptual discussions. In such cases, PI has not always worked well. Adding to this, there is the question of what happens if we were to take PI online. How would it work if peers cannot interact in real-time? With growing interest in active learning pedagogies, which benefit from having students prepared ahead of class work – i.e., the flipped classrooms – there is added pressure on getting design elements for digital and online learning right.

[1] DALITE began as a co-design project between researcher/designers from the Ontario Institute of Studies in

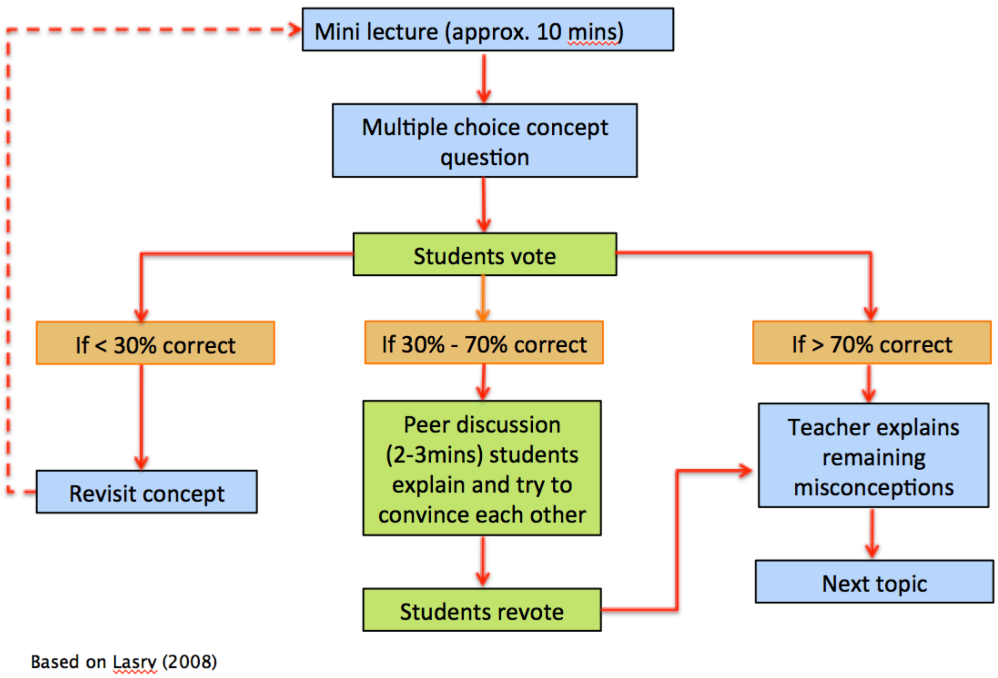


Figure 2.1. Peer Instruction script.

2.6 Digital and Online Instruction of Physics

Computer supported learning environments to promote learning in physics is not new. A major initiative in this area is the Andes project, an intelligent tutoring system for a first year college-level physics course (Gertner & VanLehn 2000). It coaches students through the problem solving process step by step and provides hints should the student get stuck. Andes and similar tutoring systems are very successful and produce significant learning gains compared to traditional instruction. However, some have criticized Andes, and other similar tutoring systems, for failing to get at deep learning. In particular, three weaknesses have been identified as the failure to get students: (1) to use the language of the discipline (i.e., talk science); (2) to reflect more deeply on the learning; and, (3) to work on developing their conceptual knowledge (Graesser, VanLehn, Rosé, Jordan & Harter, 2001). Viewing these as challenges to be overcome when designing an online learning environment, we consider these as the foundation of our design features.

2.7 Summary

Taking the ideas detailed above we extracted four design principles, which form the basis of the DALITE system's environment that the student interacts with; in addition to the PI script described. On the side of the teacher, we were guided by principles of the reflective practitioner and Just-in-Time Teaching (JiTT; Novak, Patterson, Gavrin, & Christian, 1997). The design of the system was intended to both support the ability to review students' work ahead of class as well as customize teaching, making it more student centered. In other words, to determine what concepts were most challenging to the particular cohort of students. The design principles are summarized in the Table 2.1.

We report on what we learned for each of these principles through a series of data collection activities. We discuss these in the methods chapter (chapter 4) but first describe the design of DALITE in the upcoming chapter.

Table 2.1. DALITE's design feature and theoretical relationships

Design feature	Theory
Peer-Instruction	conceptual change – *intentional reflection* (Sinatra & Pintrich, 2003); collaborative learning (Roschelle, 1992; Stahl, 2006)
Writing "Rationales"	self-explanation (Chi, Leeuw, Chiu, & LaVancher, 1994); *interactive* explanation – explaining to self & others (Ploetzner, Dillenbourg, Preier, & Traum, 1999)
Comparing & Contrasting	deep/surface similarity (Gentner, 1989) learning from peers – i.e., modeling
Multi-context use	transfer as *intercontextuality* (Engle, 2006)
Self-directed feedback	creating epistemic agency related to a need to know (Scardamalia & Bereiter, 2003; Damşa, Kirschner, Andriessen, Erkens, & Sins, 2010); changing expectations and "habits of mind"
Contribution to community	sense of community participation and community building (Slotta 2010; Slotta & Najafi, 2010)

2.8 References

Bereiter, C. (1995). A Dispositional View of Transfer. In A. McKeough, J. Lupart & A. Marini (Eds.), *Teaching for Transfer: Fostering Generalization in Learning* (pp.21-34). Mahwah, NJ: Lawrence Erlbaum Associates.

Bransford, J. D., Vye, N. J., Stevens, R., Kuhl, P., Schwartz, D. L., Bell, P., Meltzoff, A., Barron, B. J., Pea, R., Reeves, B., Roschelle, J., & Sabelli, N. (2006). Learning theories and education: Towards a decade of synergy. In P. A. Alexander and P. H. Winne (Eds.), *Handbook of Educational Psychology* (pp. 209-244). Mahwah, NJ: Erlbaum.

Bransford, J. D. & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (Vol. 24, pp. 61–100). Washington, DC: American Educational Research Association.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18, 32-42.

Chi, M. T., Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science*, *18*(3), 439-477.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32, 1: 9 – 13.

Collins, A., Brown, J., & Holum, A. (1991). "Cognitive Apprenticeship: Making Thinking Visible" in *American Educator*, Vol. 15, No. 3, pp. 6-11, 38-46.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, *69*, 970.

Damşa, C. I., Kirschner, P. A., Andriessen, J. E., Erkens, G., & Sins, P. H. (2010). Shared epistemic agency: An empirical study of an emergent construct. *The Journal of the Learning Sciences*, *19*(2), 143-186.

diSessa, A. A. & Sherin, B.L. (1998). What changes in conceptual change? *International Journal of Science Education,* 20 (10), 1155-1191

Engle, R.A. (2006). Framing interactions to foster generative learning: A situative explanation of transfer in a community of learners classroom. *Journal of the Learning Sciences*, 15(4), 451-498.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 201319030.

Gee, J. P. (2000). Identity as an Analytic Lens for Research in Education. *Review of Research in Education, January 2000*(25), 99-125.

Gentner, D. (1989). The mechanisms of analogical learning. *Similarity and analogical reasoning*, *199*, 241.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, 10, 277–300.

Gertner, A. S., & VanLehn, K. (2000, January). Andes: A coached problem solving environment for physics. In *Intelligent Tutoring Systems* (pp. 133-142). Springer Berlin Heidelberg.

Graesser, A. C., VanLehn, K., Rosé, C. P., Jordan, P. W., & Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI magazine*, *22*(4), 39.

Gray, J. (2007). The Fourth Paradigm: Data-Intensive Scientific Discovery. In T. Hey, S. Tansley & K. Tolle (Eds.), *The Fourth Paradigm: Data-Intensive Scientific Discovery* http://research.microsoft.com/en-us/um/people/gray/talks/NRC-CSTB_eScience.ppt.

Greeno, J. G. (1997). Response: On claims that answer the wrong questions. *Educational Researcher*, 26(1), 5–17.

Greeno, J. G., Smith, D. R., &Moore, J. L. (1993). Transfer of situated learning. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on Trial: Intelligence, cognition, and instruction* (pp. 99–167). Norwood, NJ: Ablex.

Gresalfi, M. S. (2009). Taking Up Opportunities to Learn: Constructing Dispositions in Mathematics Classrooms. *Journal of the Learning Sciences*, 18: 3, 327 — 369

Gresalfi, M. S., & Cobb, P. (2006). Cultivating students' discipline-specific dispositions as a critical goal for pedagogy and equity. *Pedagogies*, *1*(1), 49-57.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30, 141-158.

Kravcik M., Kaibel, A., Specht, M., and Terrenghi, L. (2004). Mobile Collector for Field Trips. *Educational Technology & Society*, 7 (2), 25-33.

Lave, J. (1988). *Cognition in Practice: Mind, mathematics, and culture in everyday life*. Cambridge, UK: Cambridge University Press.

Lave, J., & Wenger, E. (1990). *Situated Learning: Legitimate Peripheral Participation*. Cambridge, UK: Cambridge University Press.

Lobato, J. E. (2003). "How Design Experiments Can Inform a Rethinking of Transfer and Vice Versa." *Educational Researcher* 32(1): 17-20.

Mazur, E. (1997). *Peer Instruction* : A User's Manual. Upper Saddle River, NJ: Prentice Hall

Meltzer, D. E., & Thornton, R. K. (2012). Resource letter ALIP–1: active-learning instruction in physics. *American journal of physics*, *80*(6), 478-496.

Meltzer, D. E., & Thornton, R. K. (2012). Resource letter ALIP–1: active-learning instruction in physics. *American journal of physics*, *80*(6), 478-496.

Novak, G. M.. Patterson, E.T., Gavrin, A.D., & Christian, W. *Just-In-Time-Teaching: Blending Active Learning with Web Technology*, Prentice Hall, 1999.

Oldenburg, R. (1991). The Great Good Place. *New York: Marlowe & Company.*

Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and instruction*,*1*(2), 117-175.

Ploetzner, R., Dillenbourg, P., Preier, M., & Traum, D. (1999). Learning by explaining to oneself and to others. *Collaborative learning: Cognitive and computational approaches*, 103-121.

Rebello, N. S., D. A. Zollman, et al. (2005). A Model for Dynamic Transfer of Learning. Annual Meeting of the National Association for Research in Science Teaching, Dallas, TX, NARST Publications.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The journal of the learning sciences*, *2*(3), 235-276.

Scardamalia, M. & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York: Cambridge University Press.

Sinatra, G. M., & Pintrich, P. R. (2003). The role of intentions in conceptual change learning. *Intentional conceptual change*, 1-18.

Slotta, J. D. (2010). Evolving the classrooms of the future: The interplay of pedagogy, technology and community. In Mäkitalo-Siegl, K., Kaplan, F., Zottmann, J. & Fischer, F. (Eds.). *Classroom of the Future: Orchestrating collaborative spaces* (pp. 215-242). Rotterdam: Sense.

Slotta, J. D., & Najafi, H. (2010). Knowledge communities in the classroom. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education, Volume 8* (pp. 189-196). Oxford: Elsevier.

Soloway, E., Norris, C., Blumenfeld, P., Fishman, B., Krajcik, J., & Marx, R. (2001). Log on education: Handheld devices are ready-at-hand. *Communications of the ACM*, *44*(6), 15-20.

Songer, N. B. (2006). *BioKIDS: An animated conversation on the development of curricular activity structures for inquiry science*.

Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Zimmerman, J., & Slotta, J. (2010). *Promoting 21st Century Science: Technology-Enhanced Learning Across Formal and Informal Environments*. NSF funding proposal 2009.

CHAPTER 3
DESIGNING DALITE

DALITE's design is guided by social constructivist principles thus aims to promote students' participation in the technology enhanced learning space by having access to the work of their peers. The focus and content is college-level physics and owes a great debt to the success of several PAREA funded projects undertaken by members of this current research proposal (Charles, 2009; Charles, Whittaker & Lasry, 2010; Lasry, Charles, Whittaker, & Lautman, 2009).

3.1 Design Background

DALITE was conceived as a way to harness the benefits of PI and extend the potential of the approach. In traditional enactments of PI, student conversations disappear into the ether. Though instructors sometimes overhear conversations, no trace remains of students' thinking except for test scores, which are summative in nature. More importantly, instructors seldom know enough about their student's thinking. For instance, what reasoning is convincing to students, in the first place. What discursive elements help students change their answers – whether a change towards or away from the correct answer.

Additionally, there is little know about the context of distributed. For instance, how the wording of questions frames the context. Or, might there be different effects regarding the timing of assigning questions, might there be an issue of the context of delivery or the wording. Can we promote better forms of *intercontextualization* with the sequencing and design of questions? DALITE as a solution provides students with a diversity of explanations for all the possible answer choices. It allows students to interact with rationales from students at different institutions or even "peers" who took the class previously.

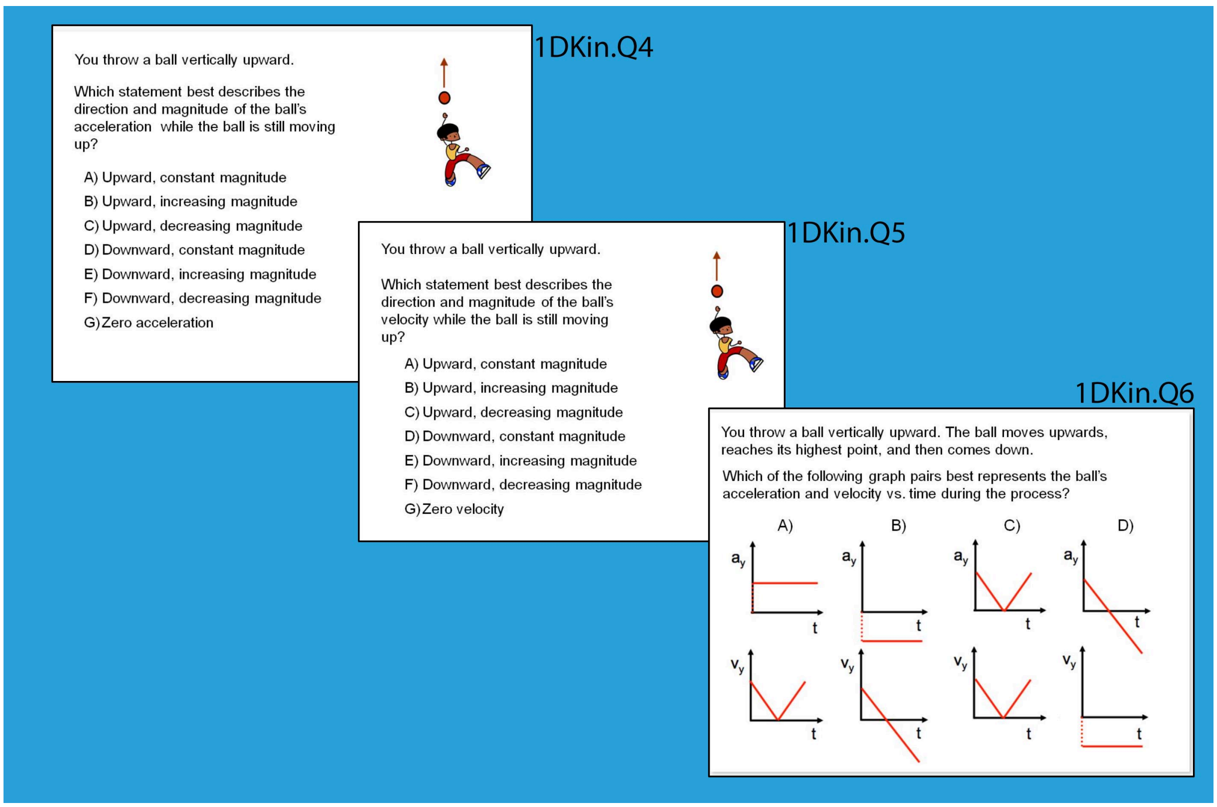
3.2 Components and Script

The DALITE infrastructure is made up of the following components: (1) a student registration and software application management; (2) a framework for data mining and tracking of student interactions in real time including the instructional scripts; (3) a central database or repository; and (4) data displays for instructors. The platform uses "Agile" development practices with the aim of ensuring future availability, scalability, and performance. The database repository is composed of two parts: (1) the curriculum content – conceptual multiple-choice questions (sometimes referred to as *concept test* questions); and, (2) the student-generated answers and rationales for these answers.

To date, the curriculum database contains over 120 questions spread across the three main topics generally covered in an introductory physics course – i.e., kinematics, dynamics, and energy and momentum. These questions are designed to be roughly at the first-year university level. Influenced by the Ohio State concept test questions (Lee, Ding, Reay, & Bao, 2011), many questions are organized into sets of three to four questions on a single concept that progressively increase in difficulty. These sets of increasingly difficult multiple-choice problems are built on similar deep structures with different surface features, or similar surface features with different deep structures. Instructors have control over the selection and assignment of questions using a specially designed teacher portal. Each problem set can be customized to meet the perceived knowledge level of the students.

3.2.1 DALITE's curricula content

The teaching content is influenced by the Ohio State concept test questions (Lee, 2009), each set is sequenced as a series of increasingly difficult multiple-choice problems built on similar deep structures with different surface features, or similar surface features with different deep structures – i.e., addressing issues involved in transfer discussed earlier. Figure 3.1 is an example of such a triplet of questions designed to develop a certain conceptual understanding first with the description in words 1DKin Q4 & Q5, and then in a graphical representation 1DKin Q6. These three questions are *homomorphic*, in the sense that experts do not distinguish them. However, students are known to be affected by the surface-level features of the questions and may not see these questions as applications of the same concept (Chi et al, 1994; Singh, 2008).



<u>Figure 3.1</u>. A triplet from the bank of Ohio State conceptual tests questions.

The second database repository, the student-generated rationales, has been developed through a "seeding" process. That is, the database asks about 20 students to answer the questions and write rationales, without working through the full DALITE script. This process enables the first participants in the system to see other students' rationales. However, it places constraints on the development of new questions entering the system. In addition, because rationales are student-generated we believe it necessary to develop a mechanism of cleaning up and categorizing the database. This has lead to the implementation of a voting system – students have the option of giving a "thumbs up" to the rationale that convinced them. In doing so, these ratings are a design element. In the future a heuristic will be designed to highlight these popular rationales. Lastly, nonsense rationales will be eliminated (e.g., unreadable text, meaningless strings of symbols).

Lastly, the data display for instructors brings DALITE into the classroom. The display provides an interface to allow the instructor to review students' progress in real-time as well as provide a tool for in-class review. It has proven to be more important than we had thought. We discuss this in the upcoming section.

3.2.2 DALITE Scripts – How does it work?

The script for DALITE mirrors much of what we imagine students do when they engage in the discursive practices of PI. It also models and coaches what we hope students' should be doing – i.e., promoting the development of an *agentic* epistemic framework.

3.2.2.1 How is it implemented?

The teacher has control over the selection and upload of the questions using a specially designed teacher portal; thereby he/she can customize the problem set to meet the perceived knowledge level of the students. For instance, teachers can populate DALITE with curricular artifacts: physics problems, conceptual questions, pictures and videos. In this implementation we only used conceptual questions. Next, an assignment is created based on a selection of pre-populated multiple-choice questions. In general, assignments were made up of three to four questions. Next, the assignment is posted to the web for the students to log into.

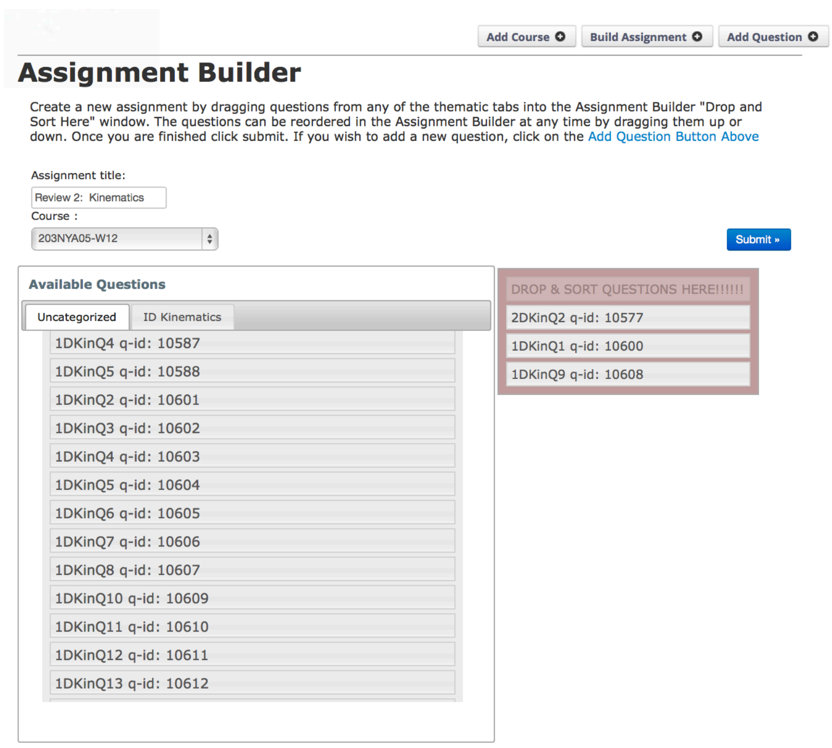


Figure 3.1a. Screen shot of the teacher assignment interface in DALITE.

Once completed, the teacher, both in terms of the first vote and revote, can review the assignment results by class and by individual student (see Figure 3.2). Sometimes revealing particularly difficult concepts and fragile understandings where students can go from the right answer in their first vote to the wrong answer in their second vote (see Figure 3.3)

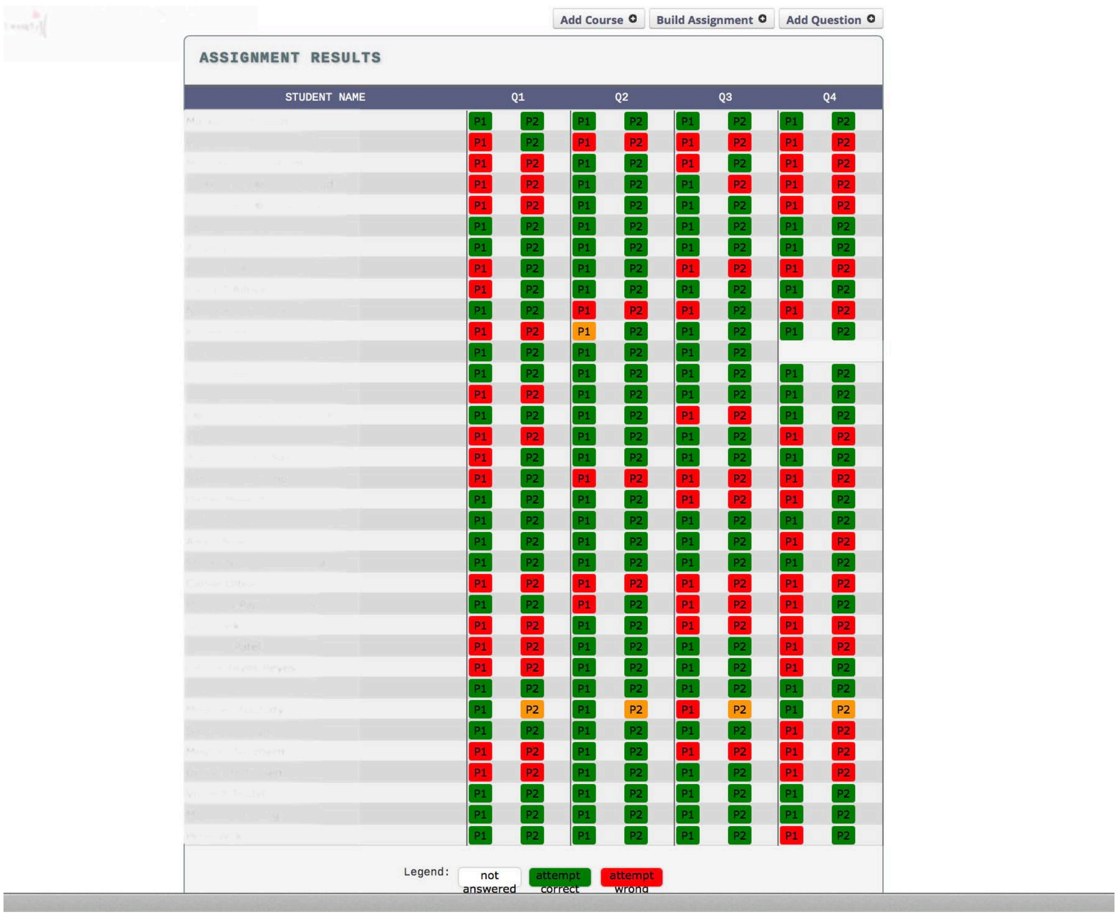


Figure 3.2. Teacher analytics report of vote and revote results for an assignment containing four questions.

Sometimes revealing particularly difficult concepts and fragile understandings where students can go from the right answer in their first vote to the wrong answer in their second vote (see Figure 3.3). Note the third column where the majority of students are persuaded away from the right answer.

Figure 3.3. An example of the right to wrong answers in the teacher analytics screen.

Additionally the actual student's rationale can be reviewed by hovering over the students' vote. A proxy of this is shown in Figure 3.4.

Q4				Q5				Q6			
ANS 1	ANS 2	CORR Ect ANS	RATIONALE	ANS 1	ANS 2	CORR ECT ANS	RATIONALE	ANS 1	ANS2	ECT ANS	RATIONALE
4	4	4	There is always a force (gravity) causing object to accelerate towards the Earth (downwards) at a constant rate. If the acceleration was directed upwards, the ball would fly up faster and faster and go into space. If there was no acceleration at all, the ball would keep moving upwards at a constant rate.	3	3	3	The ball is thrown upwards, so that is the direction of the velocity vector. Because of the gravity pulling down on the ball, the magnitude of the velocity is getting smaller and smaller until it stops and goes back down.	2	2	2	**If we make \down\" negative, then the gravitational acceleration has to be constant and negative**. Since the ball is going upwards at first, the velocity is positive but slows down until it is zero, then changes directions and falls to the ground faster and faster (which means an increasing negative value)."
3	3	4	The ball is slowly increasing and at a certain point decreasing, stops and then falls			3		1	1	2	After the ball leaves the hand, the acceleration remains constant it is 9.8 m/sec^2 directed vertically downwards, and stays at this value until the ball is caught.
4	4	4	When the ball rises, the direction of the ball's acceleration is pointing downward because the object is slowing down, so the acceleration direction is opposite to the velocity direction. Also, the magnitude is constant.	3	3	3	When the ball is going up, the velocity direction is always pointing upward. Then, the magnitude is decreasing because of the gravity.	1	1	2	The acceleration is constant during the process because it's still 9.8m/s/s. Then, the velocity graph must be: when the ball is going up, the velocity is decreasing and when it reaches its highest point, the velocity is 0. Then, when it's going down, the velocity should be increasing.
3	3	4	the ball slows dwn as it rises. The velocity vectors are pointing upward but getting shorter as the ball approaches the top.	3	3	3	the velocity is bigger at the biginning of the upward and shorter at the last velocity upward. the velocity becomes shorter and shorter for the upward	1	2	2	if the ball moves upwards, the velocity becomes shorter and shorter. so the slope is decreasing and when the ball moves downwards, the velocity becomes longer and longer, that means the slope increase. And the acceleration is the same for both, its constant.

Figure 3.4. Teacher analytics report including student's rationales.

3.2.2.2 What are students expected to do?

The DALITE script consist of the following six steps: Select, multiple choice answer; Write, rationale; Reconsider, based on alternatives; Reselect; Vote on most convincing; and Review expert rationale.

In step 1, students are presented with a multiple-choice conceptual question. They are asked to select an answer from the multiple choices. In step 2, they then write an explanation for their choice, what we call rationales. In step 3, they are asked to reconsider their answer based on another possible alternative, the aim is to replicate the experience of the "turn to your neighbor" phase in PI. If their answer is incorrect, they are presented with student rationales for the correct answer as well as rationales from other students on the same incorrect answer they chose. The aim of this comparison is to present the contrast and cognitive dissonance of traditional PI. If their answer is correct, they are presented with rationales from other students on the same correct answer as well as student rationales for the most popular wrong answer; the aim of this comparison is to test for fragile understanding or lucky guessing. In step 4, the student is asked to consider whether one of the rationales was particularly convincing, if yes they are asked to vote it "thumbs up". In step 5, students are asked to re-choose an answer for the original question: either their original answer, or the other answer that was just presented to them, based on the reading of these rationales. Lastly, step 6, they are presented with a normative rationale of an expert, but are not given "the" answer; the aim of this decision being to delay feedback and increase self-regulation of criteria and standards.

What does it look like? Students enter their answers into the DALITE system by first logging in with their unique user IDs and passwords. They are prompted to select an answer to the multiple-choice question presented. Most importantly, asked to provide an explanation (rationale) for this choice (see Figure 3.5). In short, explain their reasoning. They cannot move forward without completing this step. In addition, they must enter a minimum number of characters for this explanation. This feature was intended to prevent students entering brief answers. It cannot prevent them from entering nonsense characters rather than words. However, this has not been a problem in the answers we reviewed.

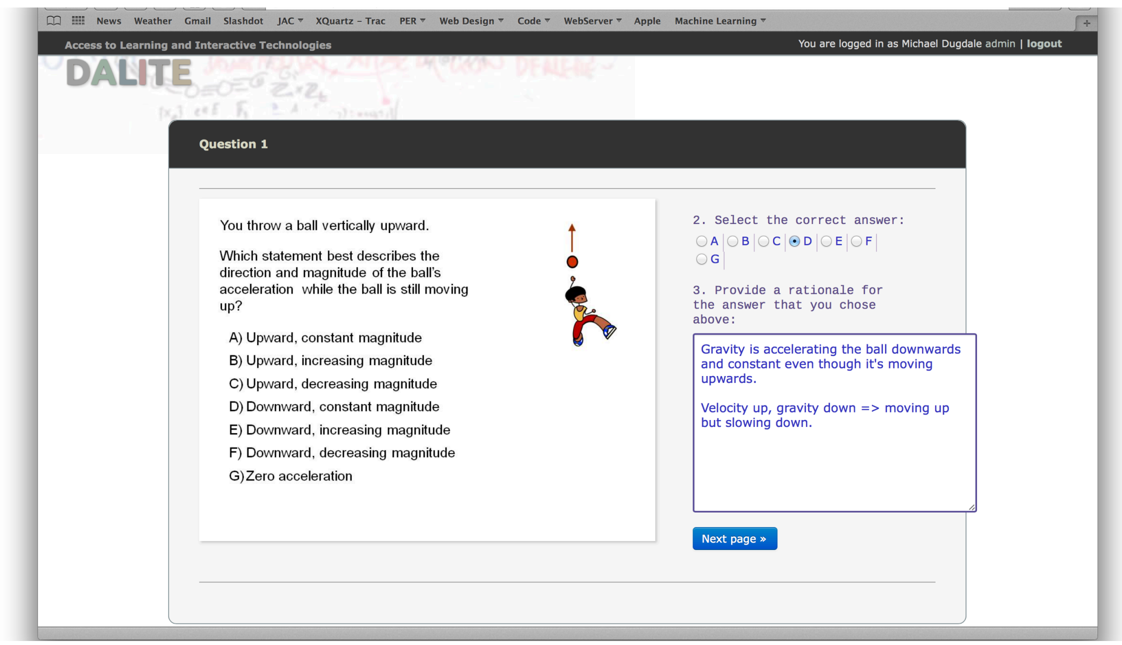


Figure 3.5. DALITE screens for steps 1 and 2: select, vote and write rationale – i.e., *self-explain and explain to others.*

Next, students are shown four rationales that were produced for the same answer. In addition, they are shown a similar set fir an alternative answer to the question. These are drawn from the DALITE database of student-generated answers (Figure 3.6). The alternative answer can be right or wrong. If the student's answer is right, then the alternative is a distractor. In other words, it has been popular with students and earned many votes and produced many explanations that are likely to contain faulty reasoning but very common misconceptions (i.e., a distractor). If the student is wrong, the alternative answer will be the "right" answer and have many explanations that contain good examples of how they should be reasoning. See the close-up for this question where answer D is correct (Figure 3.7).

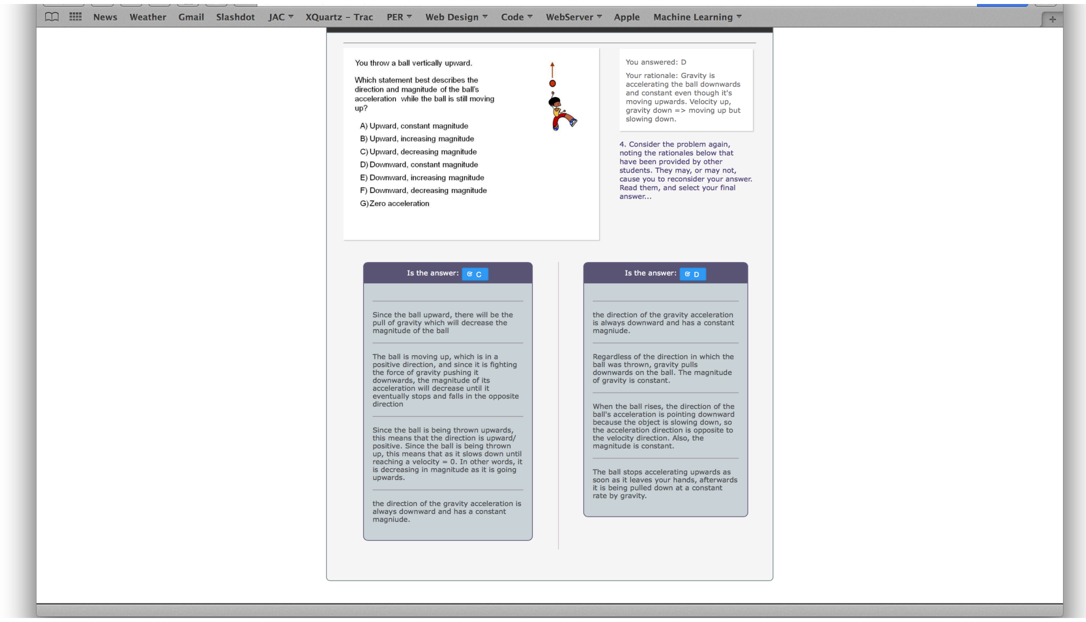


Figure 3.6. DALITE screens for steps 4 – i.e., *compare and contrast*.

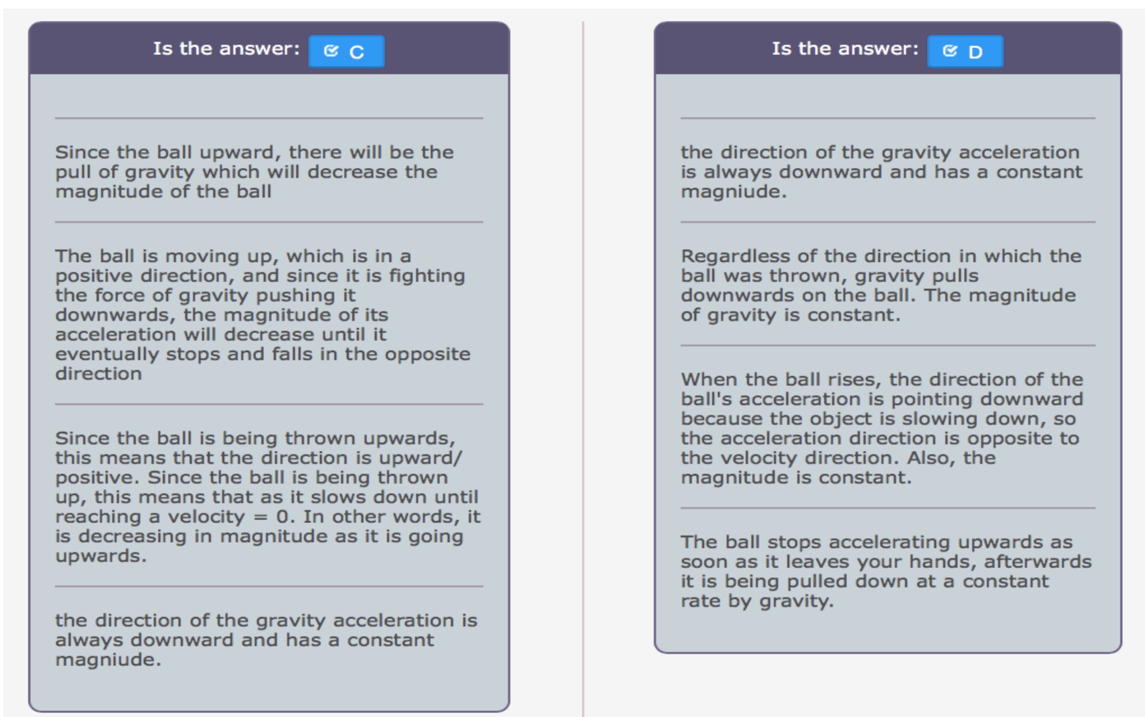


Figure 3.7. Close up of DALITE screens for steps 4, answer D is correct.

Next, the students are asked to "thumbs up" the explanation that was most convincing (Figure 3.8). The reasoning behind this step was to generate a pool of rationales that are students view as persuasive as well as provide them with a sense of contributing to the larger database.

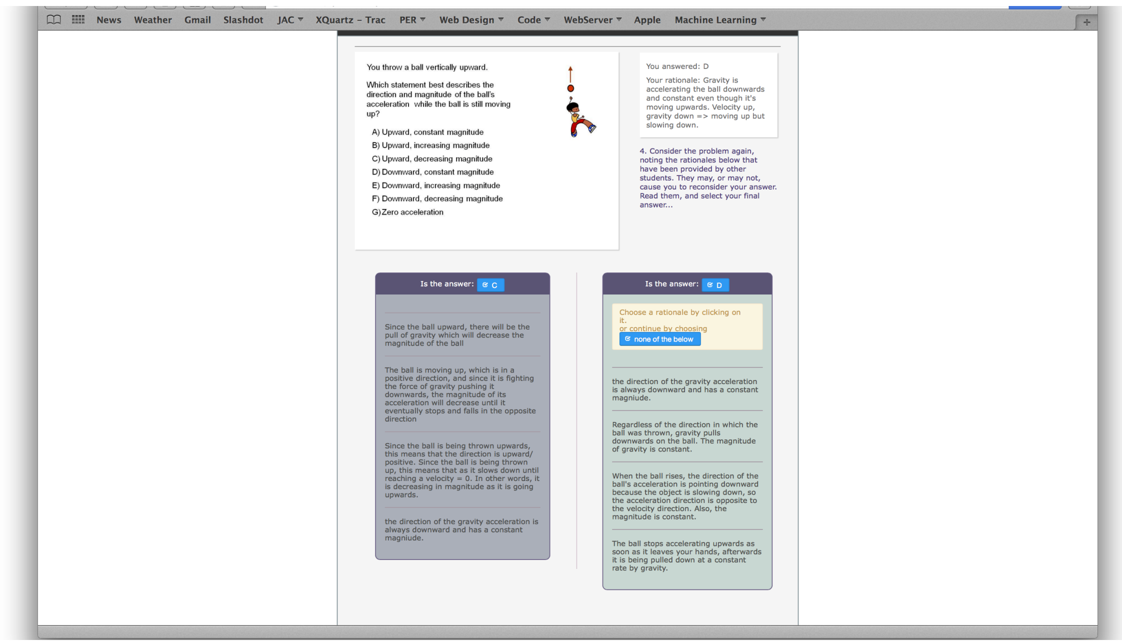


Figure 3.8. DALITE screens for steps 5, thumbs up – *evaluate and contribute*.

The last step in the script involves another comparison task. This time students are shown the explanation of an expert. We are careful to frame it as "an expert" and not "the teacher" in order to promote the development of their epistemic agency. However, the hope is that students will view this rationale as one that should be the model of what it to be expected – i.e., promoting their internal model of the normative ways of writing in physics.

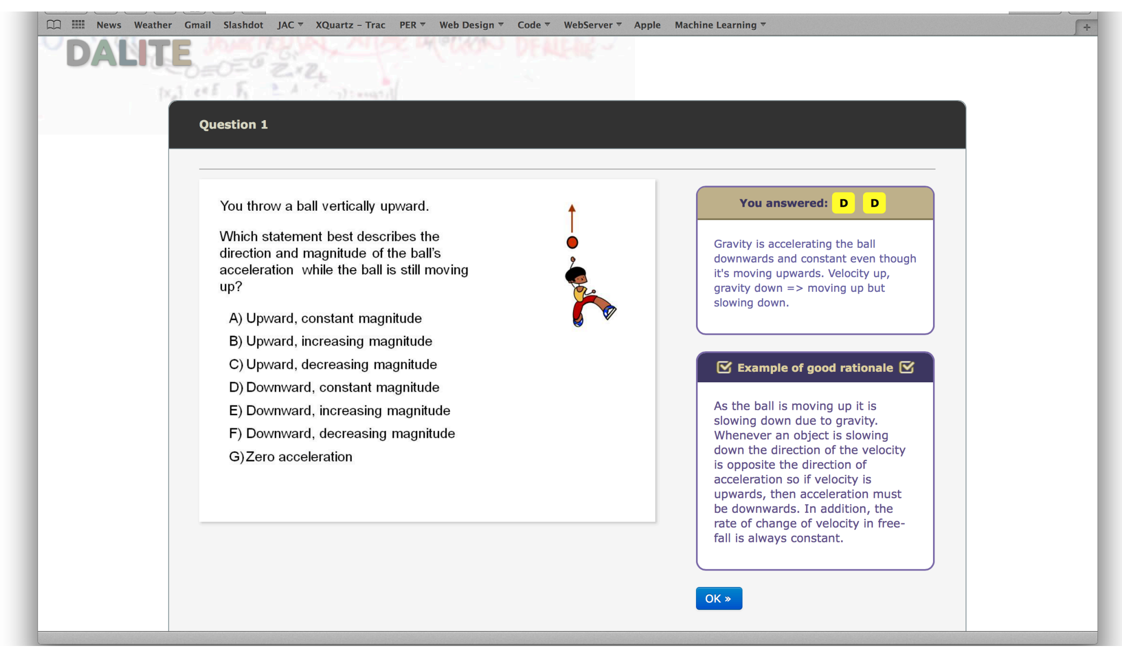


Figure 3.9. DALITE screens for steps 6, expert rationale – *compare and model*.

3.3 Context for using DALITE

3.3.1 Using DALITE at home

While DALITE was designed as an online tool, generally used for homework, it was assigned in one of two modes – (1) pre-instruction, or (2) post-instruction.

The Pre-instruction questions were designed to introduce students to the topic, much like Just-in-Time assignments. Such an assignment was designed to be somewhat easier questions that should have been within the capability of some students but generally still a challenge, leading to a sense of curiosity.

The Post-instruction were designed to help student consolidate their understanding by trying out their explanations on questions that were similar to those covered in the class lessons. Sometimes, these questions were designed to provide a bit more challenge.

3.3.2 Using DALITE in class

The DALITE curriculum was orchestrated (Dillenbourg & Fischer, 2007) to include activities that would require students to revisit or review DALITE questions in class. These activities include a teacher review and opportunities to extend the context of DALITE.

Teacher review: the teacher reviews the questions and provides the students with the answer and a further explanations, if applicable. This also coincided with the use of the teacher display (see back to Figure 3.2 & 3.3) which could show both the summary of the students' selections as well as the rationales, when the arrow hovers over any of the multiple choice responses.

In addition to the online components, we consider DALITE to be embedded into an extended system that includes a Tagging and a Concept Mapping tool. The tagging tool is digital and paperbased. It is designed to prompt students' thinking about the deep structure of the content contained in the DALITE questions. As such, this tagging tool takes students through a series of cascading concepts – from general to specific. It starts with a DALITE question, then asks the students to reflect on and identify/tag key concepts, first individually, then collaboratively in small groups.

Tagging activity: students engage in a designed activity that involves first identifying key concepts in a particular DALITE question. Then discussion and explanation. Two of the sections engaged in this activity (T09 and T10). In all instances students were asked to write rationales for the DALITE questions after the activity.

Concept map activity: students engage in a designed activity that involves using concept being learned in the course unit and placement of the DALITE questions. The concept mapping tool is computer-based, and presently uses the CmapTools, a client-server based software kit developed at the Institute for Human and Machine Cognition (IHMC; Sumi, Etani, Fels, Simonet, Kobayashi, & Mase, 1998). It takes the opposite approach to the tagging tool. It starts by asking students to work collaboratively to identify connections and state relationships between a restricted set of concepts – in the process creating a concept map. It then asks students to add in the DALITE questions to the appropriate area of the map. At the end of this process, students are asked to work on the maps individually, as a reflection exercise. Two of the five sections engaged in this activity (T06 & T08). This is elaborated on in an upcoming chapter.

3.4 Summary

This chapter described how DALITE was designed and how its intended use. Recall that this is a design-based research (DBR) project therefore the match between what is intended and what actually happened is of critical importance and relevance. This reflects the explanation of DBR put forward by Penuel, Roschelle and Shechtman (2007) as they lament on the challenge of matching "what software ought to be able to do" with what it actually does do.

In the remainder of this document we will report on what students learning when using DALITE and how they used it. Before moving to these results, in the following, Chapter 4, we describe the methods and the assessments that were designed for this study.

3.5 References

Charles, E.S. (2009). Learning Through a Community of Practice Approach. *Pédagogie Collégiale*, special issue

Charles, E.S., Lasry, N. (2010). *Who's talking in your classroom? Two sides of the same pedagogical challenge*. Paper presented at 30th annual AQPC symposium: Sherbrooke, QC.

Chi, M. T., Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science*, *18*(3), 439-477.

Dillenbourg, P. & Fischer, F. (2007). Basics of computer-supported collaborative learning. *Zeitschrift für Berufs- und Wirtschaftspädagogik*. 21, pp.111-130.

Lasry, N., Charles, E.S., Whittaker, C. & Lautman, M. (2009). When Talking Is Better Than Staying Quiet. In M. Sabella, C. Henderson & C. Singh (Eds), *American Institute of Physics (AIP) Conference Proceedings, Vol 1179*, pp.181-184. ISBN: 978-0-7354-0720-6

Lee, A. (2009). Development and evaluation of clicker methodology for introductory physics courses. (Doctoral dissertation). Available from comPADRE. PER/document/ ServeFile.cfm?ID=10358&DocID=1838

Lee, A., Ding, L., Reay, N. W., & Bao, L. (2011). Single-concept clicker question sequences. *The Physics Teacher*, *49*(6), 385-389.

Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process.*Research and Practice in Technology Enhanced Learning*, *2*(1), 51-74.

Singh, C. (2008). Assessing student expertise in introductory physics with isomorphic problems. II. Effect of some potential factors on problem solving and transfer. *Physics Education Research*, 4, 010105-5-10.

Sumi, Y., Etani, T., Fels, S., Simonet, N., Kobayashi, K., & Mase, K. (1998). C-map: Building a context-aware mobile assistant for exhibition tours. In *Community computing and support systems* (pp. 137-154). Springer Berlin Heidelberg.

CHAPTER 4
OVERVIEW OF RESEARCH METHODS

In order to answer our research questions we designed three studies, each using a different methodology and focused on slightly different parts of the corpus of data. The majority of the data for all 3 studies was collected as part of normal classroom activity. We elaborate on each of these studies in the individual chapters but briefly describe the design of the studies below.

4.1 Research Design Overview

- **Study 1**: quasi-experimental design – experimental vs. comparison sections. (N.B., we purposefully use the term "section" as opposed to "group" so as to maintain a distinction between the treatments and the groups of students who work within them.
- **Study 2**: case study design using qualitative & ethnographic methods (Garfinkel, 1967).
- **Study 3**: design-based experiment (described later) that uses mixed methods.

4.1.1 Study 1 Research Design

Study 1 was a quasi-experimental design. It collected data from five course section that used the DALITE treatment condition. We refer to these generally as the *Treatment condition.* That is to say, the Treatment sections used the DALITE system for their students' homework as well as introduced the extended activities of explanation, and concept mapping or tagging into their classroom settings. In addition, each teacher used a form of active learning pedagogy. That said, their level of experience with active learning pedagogy varied. These differences are taken into account in the interpretation of the results and discussion.

The *Comparison condition* was made up of two cohorts. Cohort 1 was made up of students who have taken a Mechanics course and completed the Force Concept Inventory (FCI; Hestenes, Wells & Swackhamer, 1992). We will discuss this in greater detail shortly. Cohort 2 was made up of students who have engaged in an active learning pedagogy that included traditional Peer Instruction (PI) as part of their in-class activity. These students were part of two sections, one taught by a teacher in College X and another taught by a teacher at a larger institution. This *purposeful sampling* of the comparison treatment was deemed critical for the authentic comparison of DALITE.

4.1.2 Study 2 Research Design

Study 2 was a case study design (Stake, 1998). It focused on understanding how the treatment is been implemented and how the design affordances are taken up by the teacher and the students. Teachers from the treatment condition were observed, and a sample of their classes documented as part of an ethnographic study of student practices that develop while engaged in using DALITE. Students' group work relating to the extended DALITE activities (concept mapping and tagging) were also recorded.

4.1.3 Study 3 Research Design

Study 3 will use a design-based research approach (Brown, 1992; Cobb et al., 2003). Design-based research (DBR) allows for the design of tools or conceptual models that help us better understand the conditions under which the context and/or the intervention can promote better learning outcomes (see Figure 4.1). In turn, the design can be adapted to support better learning (Anderson & Shattuck, 2012).

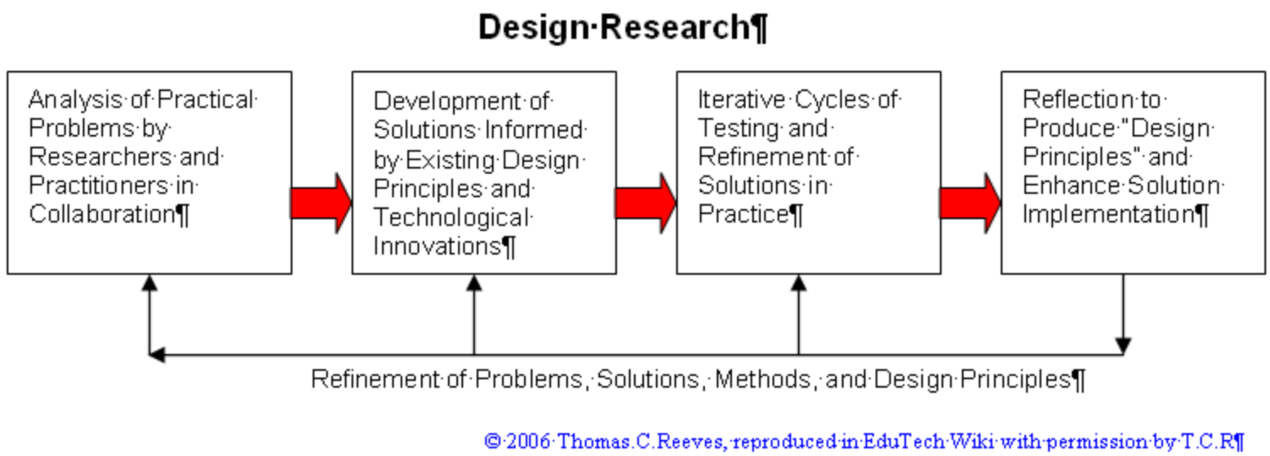


Figure 4.1. Description of processes involved with design-based research.

Data collected were for the purpose of informing us on the design principles related to promotion of conceptual understanding and conceptual change. Mixed-methods were used that include: standardized pre-post questionnaires (i.e., the FCI); and course grades. Qualitative data on student's conceptual understanding is documented in their DALITE rationales. Student interviews were conducted and include video recordings of think-aloud protocols that help to reveal how DALITE was used and how it was perceived as a tool to promote conceptual learning. Classroom observations were also collected to document the ways in which DALITE was used as part of the instructor's system of active learning practices.

4.2 Context and Participants Overview

The three studies were situated within physics classrooms in three English-speaking colleges in Quebec. Four instructors participated, each also a member of the research team (see Table 4.1). Student participants were first year science majors (N=168), ages 17-19, enrolled in one of five sections of a 15-week introductory physics course.

Table 4,1. Description of the five treatment sections – i.e., student numbers, institution, instructor and instructor's experience with active learning, and classroom setting.

Section	T06	T07	T08	T09	T10
Student #s	n=30	n=41	n=36	n=31	n=30
Institution	College1	College 2	College 2	College 3	College 3
Classroom setting	ALC	Hybrid	Hybrid	ALC	ALC
Teacher	T1	T2	T3	T4	T4
Extended DALITE	Concept Map	Concept Map	None	Tagging Activity	Tagging Activity
Teacher's experience with active learning	Moderate	Moderate	Moderate	High	High

The context for all three studies was within the physics NYA course, Mechanics. This course is most often divided into three parts. The first part, Kinematics, builds a model for motion where the acceleration of the object is constant. The second part, Dynamics, examines how forces cause motion. This section goes into the details of Newton's three laws of motion, and their implications. The third part, commonly referred to as Conservation, includes the study of momentum and energy – laws of Conservation of Momentum, and Conservation of Energy are explained. The importance of these three components will become clear as we describe the design and development of the DALITE intervention and the related assessment measures.

Traditionally, the learning activities in a Mechanics course revolves around algorithmic problem solving activities. However, it has been shown that students can solve these conventional problems without having a clear understanding of the underlying concepts (Mazur, 1997, Kim & Pak, 2002). Instructors increasingly complement their traditional problem solving activities with conceptual questions.

4.3 Procedure

DALITE was part of an active learning pedagogical implementation, what we call IT-DALITE. It was assigned weekly as homework via the web with each assignment consisting of between three to five questions. The total number of DALITE questions assigned by teacher range between 48 and 66. Variations between sections were a feature of the teacher's pedagogical style with the most experienced active learning teacher (T09 & T10) assigning the most DALITE questions. Exact usage is described in Table 4.2. Also assigned on a regular basis were readings from the textbook as well as a problem solving online homework system related to the text book or other (e.g., Mastering Physics, LON CAPA).

DALITE was brought into the classroom setting regularly, which consisted of having the instructor follow up with the correct answers and elaboration on questions that were identified as challenging. Additionally, four of the five sections included an extended activity that involved either concept mapping or "tagging" activities as part of their DALITE treatment. One section only use DALITE (recall Table 4.1).

4.4 Assessment Process

4.4.1 Overview of the assessment instruments used in Study 1

4.4.1.2 Conceptual knowledge assessment - FCI

The *Force Concept Inventory* (FCI) was used as pre and posttest to assess student's conceptual understanding. In physics, students may know how to solve problems without having a complete conceptual understanding of the physics involved (Kim & Pak, 2002).

The Force Concept Inventory (FCI), a 30-item multiple-choice instrument, is unique in that it asks conceptual physics questions in simple terms and proposes distractors that are compiled from the most prevalent misconceptions given by students in interviews (Halloun & Hestenes, 1985a,b). To answer FCI questions, students do not resort to computations or memorized algorithms but have to identify the accurate concept from a number of "d*istractors*". To expert physicists, the correct answers to FCI questions are straightforward. The gap between what instructors think their students understand and what the FCI shows, has contributed to making the FCI "*the most widely used and thoroughly tested assessment instrument*" in physics (McDermott & Redish, 1999).

4.4.1.3 Conceptual knowledge assessment - transfer and ability to explain

- Common Conceptual Tests (CCTs) : Three common conceptual tests, designed to assess the conceptual understanding of the students. Each test targeted on different content that mapped to that covered in the three units of the curriculum – Kinematics, Dynamics and Conservation principles. Teach was made up of two questions, one requiring a written explanation to a multiple choice question, much like the DALITE questions. The second, requiring students to evaluate the quality of a written explanation and to justify their decisions. In short, they were asked to compare and contrast rationales that were written by others. This activity was an individual assessment and completed in class.

- Concept maps: Three concept mapping activities, designed to capture the students' conceptual understanding. Each activity was designed as an end-of-unit activity and included the content of that unit. This activity was a group assessment and completed in class.

- Sorting task activity: Three sorting task activities, designed to capture students' understanding of the underlying deep structural similarities between questions. This activity was an individual assessment and completed in class.

The upcoming section will elaborate on these assessments. In addition, the implementations will be described in the data analysis chapters, to come.

4.4.2 Overview of the assessment instruments used in Study 2

4.4.2.1 Attitudes and epistemic beliefs

- Student interviews: An open-ended questionnaire was designed – eight questions. Students were interviewed individually and in small groups. Think-aloud protocols were also collected, those data are not presented in this report.

- Classroom observations: An ethnographic approach was used to collect classroom data. This was done on a regular basis for sections T09 and T10. Sections T06 and T08 were also observed on a schedule of four-six times during the semester (every two to three weeks). Section T07 was observed twice. Decisions to observe classes more or less were determined after the beginning of the semester and according to the researcher's availability. It was determined that section T10 made for a good case study after the mid-term and data were collected from this class on a weekly and sometimes twice weekly basis.

4.4.2.2 Teacher pedagogical approach

- Classroom observations: Same as above.

- Teacher written interviews: Teacher interviews were conducted using a written questionnaire. Five questions were designed for this questionnaire.

4.5 Assessment Instrument

4.5.1 Designing the DALITE Activities and Assessments

Building the DALITE activities and assessment measures involved a series of stages. To start, there was the identification and selection of the most general underlying principles, or "big idea" for each of the three course segments – Kinematics, Dynamics and Conservation.

The next stage involved the selection and design of conceptual questions for that content. Later, the other activities and assessments were designed around those segments and "big ideas." The development of three different types of assessments to capture and triangulate the developing student understanding: (1) conceptual test questions, (2) concept mapping/or tagging activity; and, (3) sorting task activity. The final pairing of course content and assessment is represented in Figure 4.2.

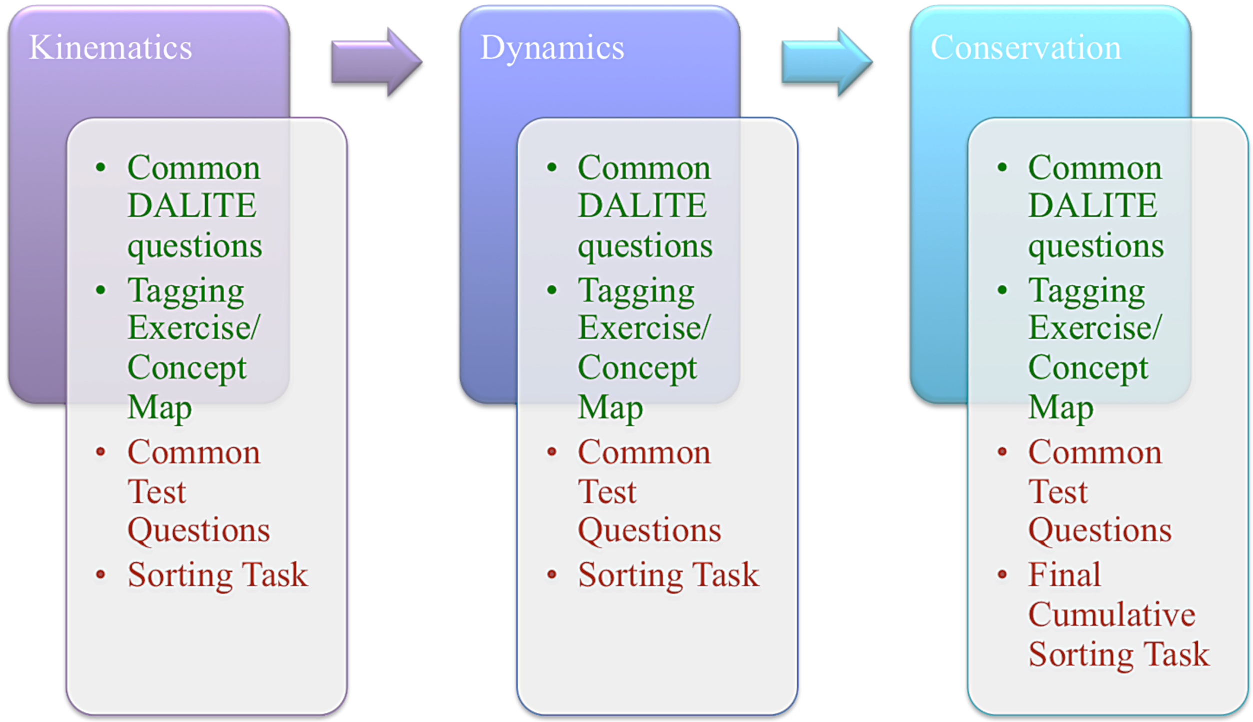


Figure 4.2. The three units of the course Physics NYA, and the data collected for each. Green refers to the content components, red to the assessments.

The design and development of materials for each of these three segments themselves involved a process of identifying specific questions that were in the DALITE database, then using those to design and develop the assessment tools. Thereby, creating a two stage process described below.

Stage 1

Identifying the big ideas was accomplished as a co-design initiative (Penuel et al, 2007), i.e., as a joint enterprise between researcher and practitioner/teacher. Additionally, identification of big ideas was based on our teacher's years of experience as well as the physics education research literature (e.g., Ohio State paper). Dividing up the course content in this manner allowed us to both select appropriate DALITE questions from the database of 120 questions as well as develop the supporting and assessment activities that followed.

Stage 2

Development of the activities and assessments designed to capture and triangulate the developing student understanding was completed in a process similar to that described above. We will elaborate on the assessment and activities in the following order: Common conceptual test questions, concept mapping activity, tagging activity, and sorting activity.

Note that we will focus mainly on the common conceptual test questions and concept mapping activities in this report because they have been analyzed most extensively. The tagging and sorting tasks will be described only briefly. Additionally, our treatment sections had different experiences with these activities. Two of the five were exposed to the concept mapping activity while a different two sections used the tagging activity. One section did neither, and acted as a within-treatment control for these extensions to DALITE.

4.6 Common Conceptual Test (CCT)

As stated, the common conceptual test was developed using the two stage process for each of the three course segments – Kinematics, Dynamics and Conservation. Next we elaborate on the design and development of each of these assessment measures, describing each within its content segment.

4.6.1 Kinematics Assessments

Big ideas targeted by the Kinematics CCT#1 were: (1) that the kinematic formulae only apply when objects are undergoing *constant acceleration*, and (2) the importance of *initial conditions* on determining the trajectory of an object's motion. As described above, the first phase of this process was to identify the DALITE questions that would be assigned (Stage 1), then follow with the development of assessment tool that mapped to these concepts (Stage 2).

Kinematics Stage 1. The DALITE questions assigned for this unit are: 1DKinQ4, Q5, Q6 and Q8 (see Appendix A). These questions address our first "big idea" of kinematics where all equations used apply *only to situations where there is constant acceleration.* Each of these questions is designed to promote students' ability to *de-couple* the notion of velocity from that of acceleration. In other words, the objective is to have students understand the physics involved when an object is in *free-fall*. Students must recognize that the velocity of the object is changing, while its acceleration remains constant. This can be quite challenging because the conflation of velocity and acceleration is among well-documented misconceptions in mechanics (Viennot, 1979; Clement, 1982; Halloun & Hestenes, 1985, Hestenes, Wells & Swackhammer, 1992).

The DALITE questions mapped to this principle were: 2DKinQ1 to Q6 (see Appendix B). These questions were assigned as triplets. Questions 2DKinQ1, 2DKinQ2 & 2DKinQ3 each look at balls, which are falling off a table. Meanwhile, questions 2DKinQ4, 2DKinQ5, & 2DKinQ6 are comparing two canon shells launched from a battleship towards targets at different distances (regardless of the different surface features), all six compare the trajectories of two projectiles, and in each case, the answer is solely dependent on the initial conditions of the launch. Kinematic formulae can be used to show that given a projectile launched from a certain height, the time it takes to hit the ground depends solely on the vertical component of its initial launch velocity. Successfully answering any one of these six problems relies on understanding this principles.

Kinematics Stage 2. The two common test questions were designed to assess whether or not students could apply their developing understanding of the concepts targeted in the DALITE questions to similar questions on a multiple-choice conceptual test. Additionally, assess the state of their ability to explain their rationale for their choice of answer. Both these skills are key design features of DALITE.

4.6.1.1 Kinematics CCT#1 question #1

Kinematic conceptual test question 1 (see Figure 4.3) assess whether students recognize differences between context. In this example, both questions are comparing the time of flight for two projectiles: the only difference is that one is in 1D, while the second is in 2D. Both questions are based on the same "big idea" described earlier.

How does this question differential understanding? The question was designed to differentiate between students with a early understanding of the concept of vertical velocity from those with a high level of understanding of the importance of the initial state. Specifically, the cuing of words such as: "*at some point during its rise.*" Such information is a clear indication that the initial state of the object is moving in a vertical direction.

1. *A hot air balloon is rising vertically upwards. At some point during its rise, a stone is dropped from the basket of the hot air balloon. At the same instant, an airplane flying horizontally also drops a stone, from exactly the same altitude. Which stone will hit the ground first?*

   a. *The stone dropped from the hot air balloon hits the ground first.*

   b. *The stone dropped from the airplane hits the ground first.*

   c. *Both stones hit the ground at the same time.*

   d. *There is not enough information to tell.*

   Please provide a rationale for the answer you selected in the previous question.

Figure 4.3. Question 1, Common Conceptual Test 1 (CCT1.Q1), hot air balloon (1D Kinematics).

4.6.1.2 Kinematics CCT#1 question #2

In the second Kinematics common conceptual test question, students were asked to circle the correct answer, and then evaluate rationales provided (N.B., the rationales provided on common tests were created by a physics teacher who was tasked with creating and reviewing DALITE content). As can be seen in the figure 4.4, students were provided with two rationales for each multiple-choice answer. They are first asked to identify which rationale they agreed with most, then choose a rationale that they felt was incorrect.

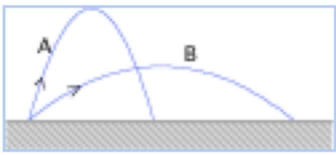
In both cases students were asked to explain why they had selected that particular (correct and incorrect) rationale. The format of this question was meant to assess whether students had developed their ability to read and evaluate good & bad rationales for the correct answer to a question the specific reason why rationales for incorrect answers were flawed.

How does this question differential understanding? The surface features of this question were a lot more like the DALITE questions assigned. Students who understand this question should be able to select the correct rationale, and which rationale is well-explained (answer A2).

At the same time, they should be more likely to also select the incorrect answer as well as identify the explanation that was the unsatisfactory.

2. Two projectiles are fired from a cannon. For projectile A, the cannon is tilted upward at an angle twice that of projectile B. (As usual, neglect air resistance.)

Part I _ Which projectile was in the air longer?



A) Projectile A was in the air longer

B) Projectile B was in the air longer

C) Both projectiles are in the air for the same amount of time

Part II Below is a selection of rationales written by students, justifying each of the answer choices.

- *Which rationale do you agree with the most? (Circle it in the table). Explain why.*
- *Choose a rationale that is incorrect (make an "X" through it in the table). Explain why that rationale is incorrect.*

Answer A	*Answer B*	*Answer C*
As the amount of time a projectile remains in the air is entirely dependent upon the vertical velocity, the one with twice the angle will have a larger vertical component, leading to a longer air-time. A1	*Since B travels further horizontally, it must have been in the air longer.* B1	*The vertical component of acceleration is the same for both projectiles: gravity. Therefore, they will both hit the ground at the same time regardless of launch angle.* C1
Movement in the y-axis determined by acceleration due to gravity. A takes more time to reach a greater height, and will take the same amount of time to come down. A2	*Horizontal velocity is unaffected by the acceleration of gravity. This means that in the horizontal direction, it never slows down. Since B is launched with a lower angle, it has a greater horizontal component to its initial velocity, and will stay in the air longer.* B2	*A goes up higher but lands closer to where it was launched. The angle does not impact time spent in the air* C2

Figure 4.4. Question 2, Common Conceptual Test CCT1.Q1, canon projectile (2D Kinematics)

4.6.2 Dynamics Assessments

The big ideas targeted by the Dynamics common conceptual test (CCT#2) were: (1) the direction of the net force can be determined by the direction of the acceleration; and (2) once the direction of the net force is determined, the relative magnitudes of the individual forces being exerted on that object can be inferred.

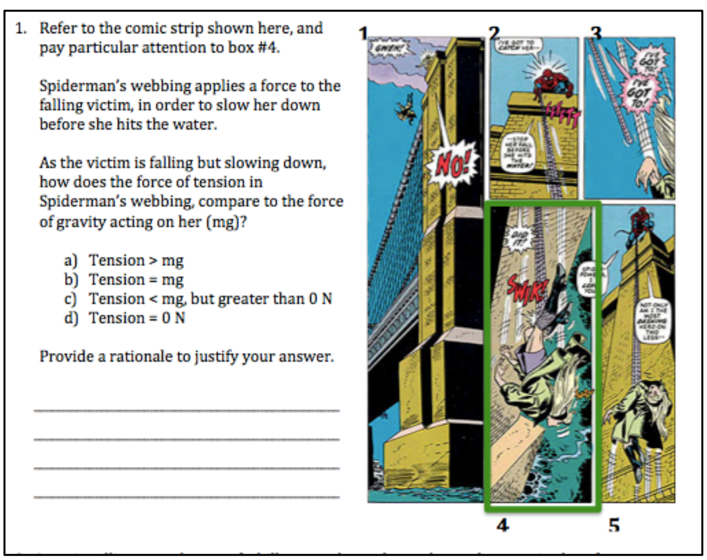
As before, the next steps follow a two stage process. Recall that the first phase was to identify the DALITE questions that would be assigned (Stage 1), then follow with the development of assessment tool that mapped to these concepts (Stage 2).

Dynamics Stage 1. The DALITE questions assigned to all students in the treatment condition, which address these big ideas, were LinDynQ3, Q7, Q16 and Q20 (see Appendix C). Each of these questions, in its own way, is designed to promote students' ability to understand the two big ideas.

Dynamics Stage 2. The two common test questions for Dynamics were designed to assess whether or not students could apply their developing understanding of the concepts targeted in the DALITE questions to similar questions on a multiple-choice conceptual test. Same as described earlier in the Kinematics assessments. Both of these common test questions students address the *big idea* for this section of the course: namely that the direction of the acceleration is needed to determine the direction of the net force on an object.

4.6.2.1 Dynamics CCT#2 question #1

The first question in CCT#2 was that shown in figure 4.5. In this question Spiderman is trying to rescue his falling girlfriend before she reaches the ground. Students are asked to evaluate the tension in the rope he is using, relative to the weight of the mass it is supporting (namely his girlfriend).



1. Refer to the comic strip shown here, and pay particular attention to box #4.

Spiderman's webbing applies a force to the falling victim, in order to slow her down before she hits the water.

As the victim is falling but slowing down, how does the force of tension in Spiderman's webbing, compare to the force of gravity acting on her (mg)?

a) Tension > mg
b) Tension = mg
c) Tension < mg, but greater than 0 N
d) Tension = 0 N

Provide a rationale to justify your answer.

Figure 4.5. Question 1, Common Conceptual Test 2 (CCT2.Q1), Spiderman.

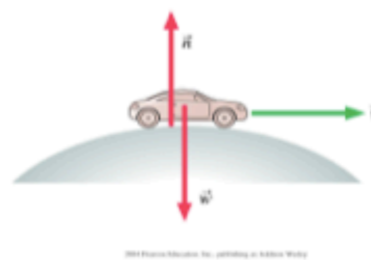
How does this question differential understanding? This question is almost identical to LinDynQ3, from the common DALITE questions, where the student must evaluate the tension in the rope supporting an elevator, relative to the weight of the elevator. A good answer will describe that the woman is moving *downwards* and slowing down. Whereas, the elevator question showed that it was moving upwards and slowing down.

4.6.2.2 Dynamics CCT#2 question #2

The second common question of this second test on Dynamics asks students first to determine the size of the normal force on a car moving over a curve (and therefore accelerating), relative to it's own weight (see Figure 5). This question is similar to LinDynQ20.

How does this question differential understanding? In this question, students must evaluate normal force and weight of a person in an accelerating elevator relative to each other. They must recognize that in this question the car is in a circular motion. Whereas, in the DALITE question, the object is in a linear motion.

2. A car is rolling over the top of a hill at speed *v* as shown (note: the magnitudes of vectors are not necessarily drawn correctly). At this instant,
   a) $n > mg$.
   b) $n = mg$.
   c) $n < mg$.



Below is a selection of rationales written by different students, justifying each of their (different) answer choices.

Make an "X" through all of the rationales that you think are wrong. For each "X" you make, explain, in your own words, what is wrong with that rationale.

Student #1 - Answer	Student #2 - Answer b)	Student #3 - Answer c)
When travelling over a speed bump, the passengers have a tendency to be thrown up off their seats, which is similar to what is happening rolling over this hill. This means that the car is accelerating upwards vertically, and that the net force must be upwards as well (by Newton's Second Law). For an upwards net force, the upwards pointing normal force must be greater than the force of gravity (mg)	The diagram only shows horizontal velocity, so the car is not moving vertically. This means that the acceleration in the y-direction is 0, which means, by Newton's First Law, that the forces in the y-direction must cancel. This means the normal force must equal the force due to gravity.	The hill can be seen as a circular path for the car, where the center is somewhere below the car. For circular motion, the acceleration points towards the center, so the net force here must also point down below the car. For a downwards net force, the force of gravity must be greater than the normal force.

Figure 4.6. Question 2, Common Conceptual Test 2 (CCT2.Q2), car on hill.

It is important to note that in the second part of these common test questions, the format of the rationale evaluation step is different from the previous assessment (CCT#1). In this instance, students choose from three rationales, one rationale for each multiple choice answer.

Then they were asked to mark an "X" through all rationales they thought were *wrong*. They were then asked to explain what was wrong with each rationale they marked with an "X". Our aim was once again to measure if students using DALITE were improving in their ability to detect flaws in erroneous descriptions of physics.

4.6.3 Conservation Assessment

The big idea targeted CCT#3 focused on the symmetries between work and impulse, and between energy and momentum: (1) Work is the force applied multiplied *distance travelled*, and equals the change in *kinetic energy* of a system; and (2) Impulse is the force applied multiplied by the *time interval,* and equals the change in *momentum* of the system. Once again, the next steps follow a two stage process.

Conservation Stage 1. DALITE questions mapped to this unit test are: Momentum Q1, Q7, Q11 and Q15 (see Appendix D). Each of these questions is designed to promote students' ability *to…*. Successfully answering these four questions relies on understanding those principles. Difference between questions from this unit, and those before, is that vectors and diagrams are not as prominent, and so the questions require more reading on the part of the student (and less interpretation of visual representations).

Conservation Stage 2. This first common test question shown figure 4.7 is mapped to DALITE question: MomentumQ1 (shown earlier).

4.6.3.1 Conservations CCT#3 question #1

Contextual differences are that the in DALITE question, the force was applied for *equal amounts of time* to both objects: therefore both end with the same *momentum*. In common test question, the ice boats are subject to the same force for an *equal distance*: therefore they finish with the same *kinetic energy*.

Q1. Two ice boats are in a race, where boat A has twice the mass of boat B. They start at rest, and are both pushed along the frictionless ice surface by the same wind force over the entire distance of the race. At the finish line, which of the following is true:

a) Boat A has a greater momentum
b) Boat B has a greater momentum
c) Both boat A & B have the same momentum
d) Boat A has a greater kinetic energy
e) Boat B has a greater kinetic energy

Provide a rationale to justify your answer.

Figure 4.7. Question 1, Common Conceptual Test 3 (CCT3.Q1), ice boats.

4.6.3.2 Conservation CCT#3 question #2

The second common test question shown below, asks students to evaluate which of the following would have less an effect on you if you were standing on a skateboard: catching a ball thrown at you, or deflecting it back in the direction from where it came. This is mapped to DALITE questions: MomentumQ7 and MomentumQ11 shown earlier. The second part of the question once again asks to evaluate the rationales of hypothetical students, and identify those that are flawed, just like in Test 2. However, in addition to the rationale that they provide, they are asked to *underline the exact words* in the erroneous rationales that they specifically disagree with.

Q2. You are standing on a skateboard, initially at rest. A friend throws a very heavy ball towards you.

You can either catch the heavy ball, or hit it back towards your friend (such that it moves away from you with the same speed as it was originally thrown).

What should you do in order to MINIMIZE your speed on the skateboard?

A) Catch the ball.
B) Hit the ball back.
C) Your final speed on the skateboard will be the same regardless whether you catch the ball or deflect

Part II
Below is a selection of rationales written by students, justifying each of the answer choices.

- Make an "X" through all of the rationales that you think are wrong. For each "X" you make, explain, in your own words, what is wrong with that rationale.
- *In the boxes where you made an "X", underline the parts of the rationale that you specifically disagree with.*

Answer A	Answer B	Answer C
If you catch the ball, it is a perfectly inelastic collision (when masses stick together). This means that the initial positive momentum of the ball being thrown is now shared between the ball and the skateboarder. The total momentum stays the same, but since the mass increased, the speed must be lower than the initial speed of the ball. Had the ball been hit back, the skateboarder's speed would be higher.	The momentum of the system before the "collision" is all in the ball being thrown. Assume that is the positive direction. If you hit the ball back in the opposite direction, the momentum of the ball will be negative after the collision: this will cancel with the initial positive momentum, which will mean that the speed of the person on the skateboard will be minimized. If the ball is caught instead, the skateboarder will absorb that momentum and go faster.	The law of conservation of momentum says that if there are no net external forces on a system, the momentum of that system is conserved. If the skateboarder catches the ball, or hits it back, in either case, these are internal forces, and so momentum is conserved in both cases. And if the momentum is the same in either case before & after the collision, the speed of the skateboarder will not depend on whether the ball is caught or hit back.

Figure 4.8. Question 2, Common Conceptual Test 3 (CCT3. Q2), skateboard.

4.6.3 Procedure

Each CCT was administered at the end of the unit segment as part of the regular mid-term assessment (90min in class section). The mid-term also contained other problem solving questions.

4.7 Sorting Task

The “Sorting Task" activity was designed to promote the development of deep understanding of the conceptually common features among questions regardless of their surface feature differences. This design is based on the expert-novice literature, which suggests that novices have a difficult time recognizing deep structural similarities between questions (phenomena) compared to experts who are not easily distracted by surface similarities.

4.7.1 Design of the sorting task instrument

In each row of a table, there are three problems, two of which are similar to each other, and the third is different in its underlying physical solution. The students must pick the two problems that are most similar to each other for each of the three rows, and write a short answer explaining why each time. We describe Sorting Task #1 as a typical example of the type of questions that make up this assessment (see Figure 4.11).

4.7.1.1 Sorting Task #1

The first row of the Sorting task above, problems A & B may look very similar due to their graphs, however it is problems A & C which are most alike in their physics, as both have *two different phases of constant acceleration* (while problem B only has two different phases of constant *velocity*).

The second row, problems D & F may look most similar because they both include descriptions based on unit vectors. However, F is the problem that is very different from the other two, as the particle experiences a *changing* acceleration. Problems D & E are actually both 2D kinematics problems, each with *constant* acceleration in one of the two directions.

The third row, problems G & H may look the most similar to each other, as they both are about a spinning wheel. However G & I are the ones that have the most in common in their physics, as they both have *constant angular accelerations* (problem H has *zero angular acceleration*, and is more concerned with *radial* acceleration).

The final section of the sorting task asks students to select the four problems, out of the six they already selected, that are most similar to each other. There is no best answer here, and we are looking mostly at how students are able to justify their choice with a rationale.

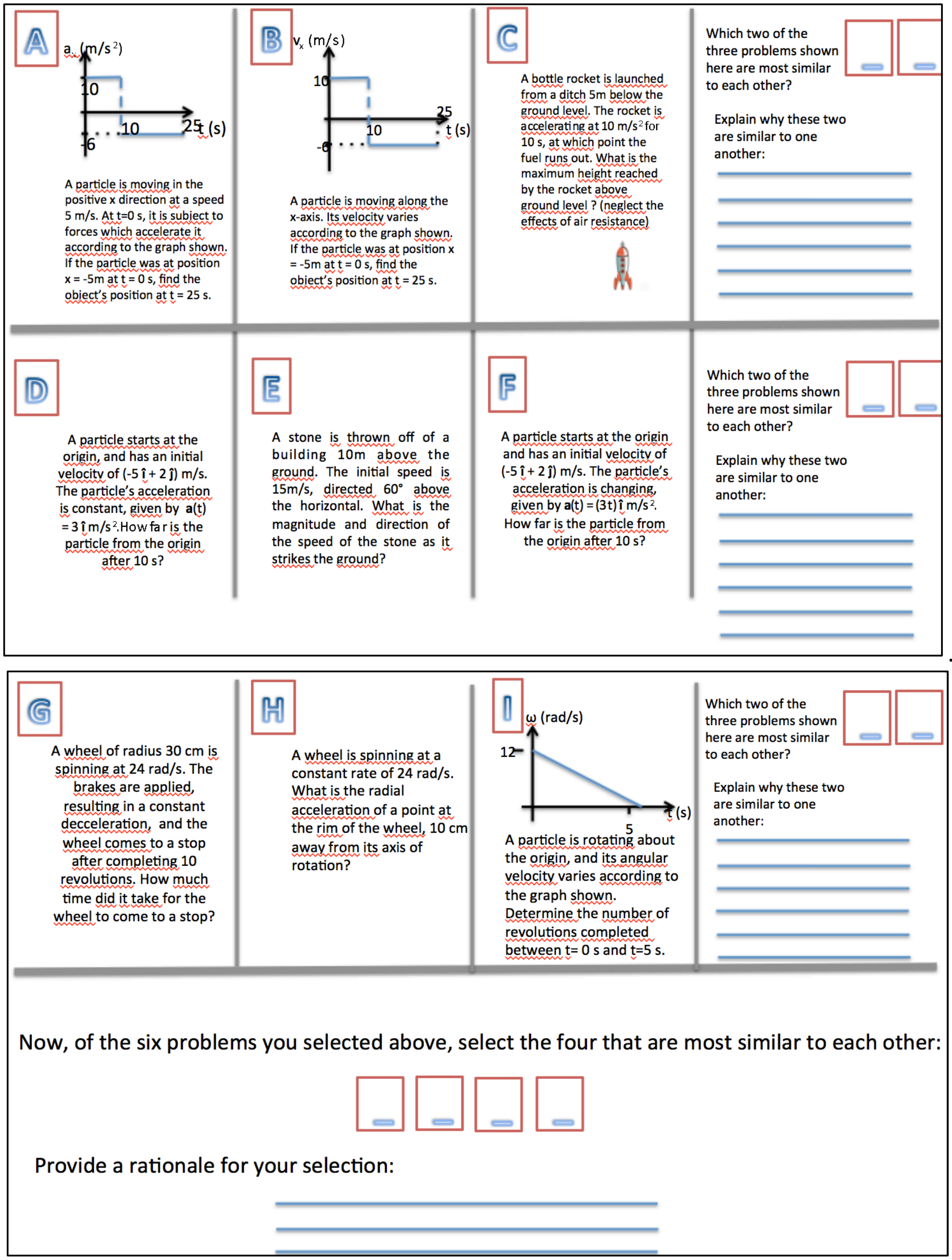


Figure 4.9. Sorting Task #1, Kinematics end-of-unit test.

4.7.1.2 Sorting Task #2

Turning next to Sorting Task #2 we see a similar design of the questions. Note that this time, the content is on Dynamics and once again makes up the end-of-unit assessment.

The first row of this sorting task (Figure 4.12), problems A & B look very similar, based on a quick inspection of their diagrams (horizontal force, diagonal black string and black diagonal ramp), while problem C looks completely different, with its picture of a man mowing his lawn. There are two ways of looking at the features in these problems. While conceptual, there is the more obvious feature of friction (i.e., problems B & C are very similar if we consider their free-body diagrams). Both have a horizontal force, a normal force, and friction along their respective surfaces. The only difference is that in problem B, we are on an inclined plane, while on the lawnmower in problem C is on a flat surface. This is a superficial difference that does not affect the physical solution of the problem significantly. However, A & C both feature the concept of equilibrium, which is the most sophisticated similarity. We will discuss this further in the analysis.

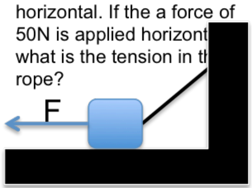
The second row, all three problems involve using Newton's Second Law on a curved surface (where the acceleration is towards the center of the curve). However problems D and F match each other the most when we draw the free body diagrams, as a 2D analysis is required; in problem D, it is a component of the tension force which is responsible for keeping the riders in their circular path, while in problem F it is a component of their normal force which gives the cyclists their centripetal acceleration. Problem E is different, because it is only in 1D, and requires that the normal force actually be set to 0 for a solution (this is often taught in a different discussion in classes, where students are asked to consider the idea of "critical speeds in vertical circular paths").

The third row (directly above), problems G & H both have two interacting objects, as well as a wheel that allows for both to move (the pulley in G, and the roller-skates in H). Also, problem G is typically one that asks students to find the acceleration (which requires only Newton's Second Law), and problem H explicitly asks for the acceleration. These two traits may make the problems look similar. Also, the phrasing of problem I almost guarantees that the student classifies it as a problem based on Newton's Third Law (based on the typically numerous reminders students get about forces exerted by interacting objects being equal). This is all designed to trick students into pairing G & H. However, if the student actually thinks about how they would solve H, they would see that its solution requires Newton's Third Law as well, and that H & I are the best match. The similarities between G & H are actually simply distracting surface features.

Finally, as in sorting task 1 described earlier, there is no best answer for the choice of four problems that go best together. The aim was simply to study the logic the students use in grouping problems together. This exercise, of grouping problems together, was also meant to prepare them for the final Summative Sorting Task, described below.

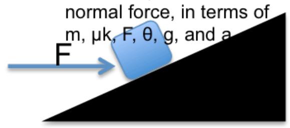
A

A 100kg mass is sitting on a frictionless horizontal surface. It is tied to the wall with a rope, which makes an angle of 40 above the horizontal. If a force of 50N is applied horizont[ally,] what is the tension in th[e] rope?



B

A mass is sitting on an incline, where friction is not negligible. The mass is being pushed up the incline by a horizontal force. If the block is accelerating up the ramp, find an expression for the normal force, in terms of m, μk, F, θ, g, and a



C

A 100 kg lawn mower is being pushed with a force directed at 40° below the horizontal. Treat the lawnmower as if it was sliding, and assume the coefficient of kinetic friction is 0.3 between the lawnmower and the ground, what force must the person apply to keep the lawnmower moving at constant speed?

Which two of the three problems shown here are most similar to each other? [ ] [ ]

Explain why these two are similar to one another:

____

D

A chair and passenger on the « Paratrooper » carnival ride have a combined mass of 85kg. The 2.5 m rope supporting them makes an angle of 30° with the vertical. The ropes are hanging from the supporting platform above which has a radius of 4 m. How many revolutions does the ride complete per minute?

E

Ice skaters going over a circular bump, risk losing contact with the ice iff they go too fast. In terms the mass of the skater and the radius of curvature of the bump, derive an expression for the critical speed at which they lose contact w[ith th]e ice.

F

Olympic indoor cyclists are constantly riding on a banked curve. What forces are causing the centripetal acceleration?

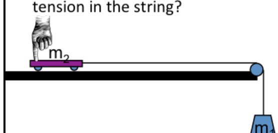
Which two of the three problems shown here are most similar to each other? [ ] [ ]

Explain why these two are similar to one another:

____

G

A cart with mass m_2 is connected to a mass m_1 using a string that passes over a frictionless pulley, as shown below. Initially, the cart is held motionless. What is the tension in the string?



H

Two friends, each wearing frictionless roller-skates, are initially at rest, holding hands. If the 50kg friend pushes off with a force of 100 N, what will be the acceleration of the 40kg friend?

I

A 10kg window pane is struck by a 100g baseball. The window breaks. What is the ratio of the force of the baseball exerted on the window, to the force of the window on the baseball?

a) Greater than 1
b) Less than 1
c) Equal to 1

Which two of the three problems shown here are most similar to each other? [ ] [ ]

Explain why these two are similar to one another:

____

Now, of the six problems you selected above, select the four that are most similar to each other:

[ ] [ ] [ ] [ ]

Provide a rationale for your selection:

____

Figure 4.10. Sorting Task #2, Dynamics end-of-unit test.

4.7.2 Procedure

Each Sorting Task (Kinematics, Dynamics) was administered at the end of its related course unit as part of an in-class activity taking approximately 60 minutes. The script included both individual work followed by a group discussion, followed by an opportunity to edit or modify answers. The latter was a deliberate activity, which asked the students to use a different coloured pen to make the changes clear. Because of the active learning approach used in these classes, the group work was completed by groups of 4-6 students who regularly worked together.

A third sorting task was administered as an online assignment using a special application InterLACE created by researchers at Tufts University, Massachusetts. This activity is not documented in this report because the data are still being analyzed.

4.8 Concept Mapping Activity

The Concept Mapping activity was designed to support student's conceptual development as well as promote their ability to relate the "big ideas" across contexts. This process was intended as a way to promote *intercontextuality* (Engles, 2006) In short, extending the context for understanding and relating the conceptual knowledge of the DALITE questions to the larger ideas of the course content.

Concept mapping activities are recognized as an important knowledge construction tool that teachers can use in the classroom to help students construct the meta-cognition necessary to make sense of complex ideas. It will be described at length in an upcoming chapter.

Because concept maps are graphical tools for organizing and representing knowledge, they represent a snapshot of how a person is thinking and linking concepts and sub-concepts. In the classroom they can be used to help students make the links necessary to deepen understanding, or as an assessment tool to highlight misconceptions and provide feedback to the teacher on students' segmenting of knowledge. It is particularly useful to track the evolution of a student's concept mapping over a semester.

4.8.1 Procedure

The concept mapping activity was administered three time throughout the term. It was always done in groups of three to four students and took place within a class session – between 50-80 mins. Students were provided with a fixed set of terms (nodes) and asked to relate them to each other (links). These links were to be annotated with prepositional statements that describe the relationship between the nodes. These actions thereby creating a concept map. Once completed, students were asked to discuss and position a set of assigned DALITE questions into their maps. The list of terms was associated with the current components of the course content (Kinematics, Dynamics, Conservation) and the "big ideas," both described in the DALITE section seen earlier. Each session added new terms therefore the list grew to include a complete picture of the NYA course. The final activity requires students to produce a concept map with the full course in mind – i.e., a complete conceptual understanding of terms from the three components of the course.

Two of the five treatment sections were involved with the Concept Mapping activities. Both sections use the C-MAP software package. One section also used a hand-drawn concept map for their first activity. An example of a C-MAP concept map is shown in Figure 4.9, and an example of a hand drawn map is shown in Figure 4.10.

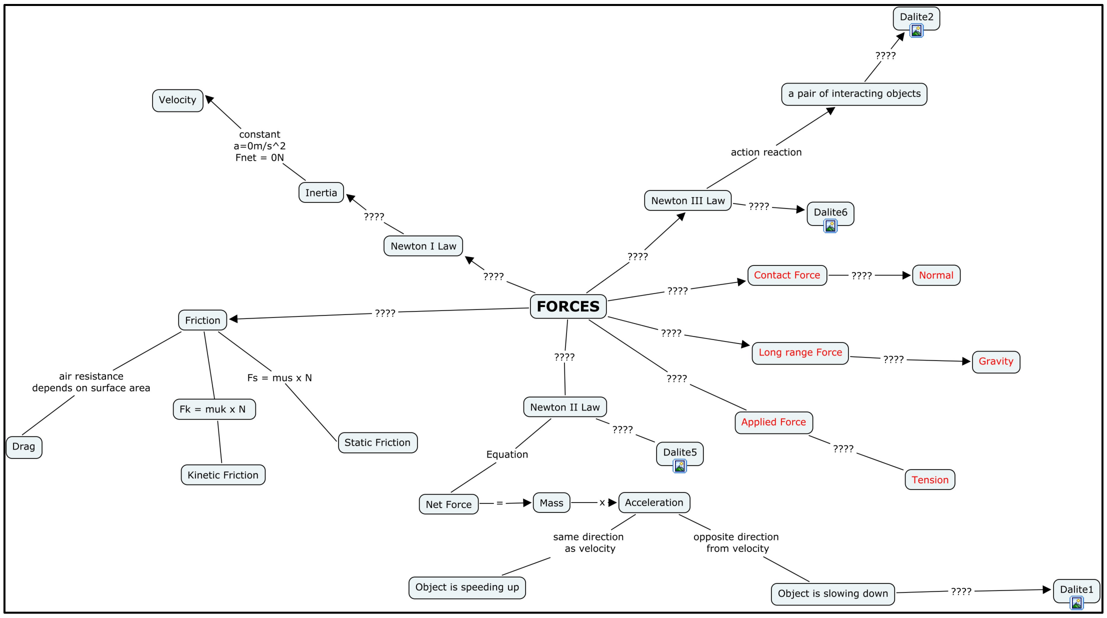


Figure 4.11. Concept map produced in C-MAP software describing the relationship of all major concepts covered in the NYA course – Kinematics, Dynamics and Conservation.

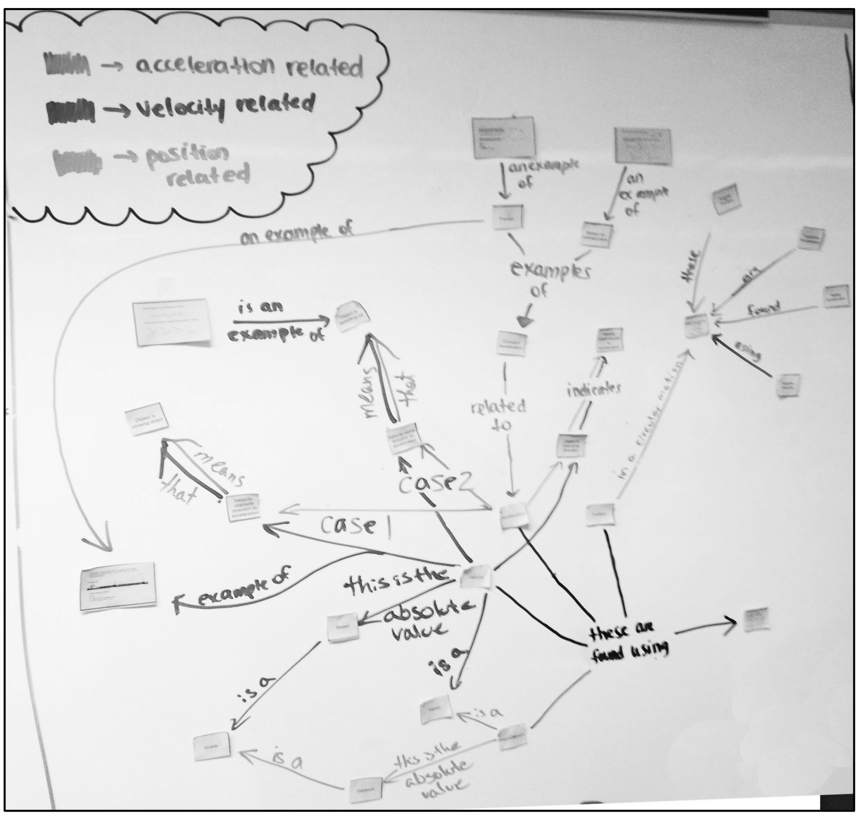


Figure 4.12. Concept map produced as hand drawing describing relationship between concepts in Kinematics.

4.9 Tagging Exercise

The tagging activity was also designed to support the development of conceptual relationships. The purpose of dividing the tags into levels is to scaffold the students into expert-like thinking. Experts are able to make cognitive distinctions between different levels of importance when processing the information surrounding a problem.

The tags are divided into three levels:

- The first asks students to identify what phenomena the question is about.
- The second level asks students to identify which variables they think are important to solving this particular problem.
- The third level asks them to identify what physical laws are the most important to solve this problem.

Questions selected were again based on the "big ideas" set of questions already identified. Tags for those sets of questions from the three course components were designed based on the three levels of tagging described above. See example from 2DKin (Appendix E).

4.9.1 Procedure

Two treatment sections completed this activity. This activity was administered five times during the semester (approximately every 2-3 weeks). Approximately 15-20 minutes of the course was allocated each time. Students worked on the exercise sheet first as individuals for 5mins during which they were asked to identify the concepts that were relevant to answering the question at hand. These were to be identified by using a particular coloured marker – e.g., yellow. They were then asked to talk amongst themselves for another 10 minutes during which they were encouraged to negotiate the best fitting tags, only. If they changed their thinking this should be highlighted in a different colour marker – e.g., blue. They were then asked to write a rationale for the question, individually. Appendix E shows an example of the students' use of the tagging exercise sheet.

The intention of this procedure was to capture the individual understanding as well as any changes that might arise from the collective discourse. Note that the "original" answers to these questions were logged in the DALITE system therefore we could compare the pre-activity and post-activity explanations.

4.10 Summary

The treatments and assessments described above provide a detailed explanation of the methods used in this project. As such the research collected a large and rich data corpus. In the next chapters of the report describe the analyses of several segments of these data (see Table 4.2). However, we note that the data produced are far to numerous to analyze in the course of this project. Therefore we acknowledge that those data will be reserved for future publications. In doing so, we will continue the good work started by this grant project.

Table 4.2. Design of the two investigations that make up Study 1 & 2.

Chapter	Study	Focus	Research Question	Data Collected	Analysis
	Study 1	*Quasi-experimental* research design			
Chapter 5 & Chapter 6		Comparison of DALITE treatment sections and comparison group on their conceptual knowledge gains	*Do students using DALITE gain statistically significantly more conceptions than students with the same prior knowledge who do not use DALITE? How do students using DALITE compare to students who do not use DALITE but use real-time face-to-face Peer Instruction? And, does the effectiveness of DALITE depend on the instructor using it*?	FCI (pretest and posttest)	t-test ANCOVA
Chapter 7		Comparison of DALITE treatment sections to comparison section on three *Common Conceptual Tests*	*What is the impact of the DALITE activities on the students' conceptual knowledge, assessed by the common conceptual test questions?* And, *Whether there is a difference between DALITE students and the comparison group, both on the conceptual tests as well as on other class assessments (i.e., the mid-term tests)?*	1. Kinematics Common Conceptual Test (CCT#1) 2. Dynamics Common Conceptual Test (CCT#2) 3. Conservation principles Common Conceptual Test (CCT#3)	Quantitative and Qualitative analysis
	Study 2	*Case study* research design			
Chapter 8		Change in conceptual understanding	*What is the impact of the DALITE web-based activity and the extended system of conceptual mapping*? *Specifically, Do students who use DALITE and concept mapping demonstrate improvements in their understanding of the course content and epistemic change*?	Concept maps	Qualitative analysis - development of categories and themes

Chapter 9		Epistemic belief related to the use of DALITE	*What can we learn about the DALITE system from how it was used? Specifically, what can the rationales produced tell us about its value as a pedagogical tool? How did rationales differ between sections, between questions, between course units, if any? What might be learned because of this use and these differences?*	1. Classroom observations 2. Interviews	Case study narrative
	Study 3	Design-based research			
Chapter 9 & 10			What factors facilitate and/or constrain the use of IT-DALITE? What types of cognitive (e.g., cognitive load), metacognitive, affective and socio-cultural supports develop or are required in order for students to efficiently use such new IT environments?	DALITE statistics and rationales Student interviews	Qualitative & quantitative case study analysis
		Lessons learned	What are the processes involved in the *design-based experiment* as IT-DALITE is iteratively refined?	DALITE statistics and rationales Student interviews Teacher interviews	Qualitative case study analysis

4.11 References

Anderson, T., & Shattuck, J. (2012). Design-Based Research A Decade of Progress in Education Research?. *Educational Researcher*, *41*(1), 16-25.

Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *Journal of the Learning Sciences*, 2, 141–178.

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*(1), 66-71.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., Schauble, L. (2003). Design experiments in educational research. Educational Researcher, 32, 1: 9 – 13.

Engle, R.A. (2006). Framing interactions to foster generative learning: A situative explanation of transfer in a community of learners classroom. *Journal of the Learning Sciences*, 15(4), 451-498.

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30, 141-158 .

Halloun, I. A., & Hestenes, D. (1985a). Common-Sense Concepts About Motion. *American Journal of Physics, 53*(11), 1056-1065.

Halloun, I. A., & Hestenes, D. (1985b). The Initial Knowledge State of College Physics Students. *American Journal of Physics, 53*(11), 1043-1055.

Kim, E., & Pak, S. (2002). Students do not overcome conceptual difficulties after solving 1000 traditional problems. *American Journal of Physics, 70*, 759.

Lasry, N. (2006). PAREA Report: Implementing Peer Instruction in Cegep

Stake, R.E. (1998). Case Studies. In N.K. Denzin and Y.S. Lincoln (Eds.), *Strategies of qualitative inquiry* (pp. 86-109). Thousand Oaks, CA: Sage Publication.

Mazur, E. (1997). *Peer Instruction* : A User's Manual. Upper Saddle River, NJ: Prentice Hall

McDermott, L., & Redish, E. (1999). Resource letter: PER-1: Physics education research. *American Journal of Physics, 67*, 755.

Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process.*Research and Practice in Technology Enhanced Learning*, *2*(1), 51-74.

Viennot, L. (1979). Spontaneous Reasoning in Elementary Dynamics. *International Journal of Science Education, 1*(2), 205-221.

CHAPTER 5
DALITE AND STUDENTS' CONCEPTUAL LEARNING

This chapter addresses three research questions: *Do students using DALITE gain statistically significantly more conceptions than students with the same prior knowledge who do not use DALITE? How do students using DALITE compare to students who do not use DALITE but use real-time face-to-face Peer Instruction? And, does the effectiveness of DALITE depend on the instructor using it*?

DALITE is a platform that was designed to engage students in a number of interesting ways. Foremost, DALITE can be seen as a means to implement Peer Instruction (Mazur, 1997) asynchronously. DALITE allows the course content to follow students outside of class. Using a computer or mobile device, students can login the system, engage with the concepts and discuss asynchronously with peers. DALITE emphasizes the need for students to clearly express their ideas and master scientific discourse. Indeed, each time students are given a question, they are asked to provide a rationale that will eventually be sent to another student. Students must therefore learn to express themselves clearly in writing and begin to grasp the subtleties within the discourse of the discipline they are learning. DALITE can also ask students to engage in expert like categorization tasks.

While there are many different ways to measure whether DALITE is effective in reaching its explicit goals, in this chapter we focus only on the conceptual learning achieved by students using DALITE. On a very general level, this chapter was designed to ask broad questions. Foremost, active learning is most often portrayed as a process that requires real-time collaborative learning. Hence, the broad question is whether active learning achievable through asynchronous online collaborations or if real-time face-to-face collaboration is a necessary condition.

5.1 Peer Instruction yields greater conceptual learning

Students engaged in real-time face-to-face Peer Instruction have greater conceptual gains than students who are exposed to teacher-centered lecturing. Indeed, Harvard students who learned physics by discussing face-to-face in class had greater conceptual gains than other Harvard students who were given physics lectures (Crouch & Mazur, 2001; Mazur, 1997, 2009). Greater conceptual learning has also been reported in students using face-to-face Peer Instruction in a host of institution types ranging from high schools, US community colleges and universities (Fagen, 2003; Fagen, Crouch, & Mazur, 2002) as well as in our own Quebec Cegep classrooms (N Lasry, 2006, 2008; N Lasry, Mazur, & Watkins, 2008). The question we address in this chapter is whether it is possible to achieve similar conceptual gains without having students interacting face-to-face in real-time. For instance, does an online asynchronous Peer Instruction implementation of DALITE enable students to achieve greater conceptual learning or is the real-time face-to-face interaction a *sine qua non* condition for greater conceptual change?

Answering this question accurately requires two important methodological steps. The first is a reliable way of measuring how students change conceptions. The second is a sizable control section to which we could compare our students.

5.2 Assessing how students change conceptions in physics

A large number of student 'misconceptions' have been documented over the past 30 years (Clement, 1982; Halloun & Hestenes, 1985a, 1985b; Minstrell, 1982; Pfundt & Duit, 1988; Viennot, 1979). Most researchers and instructors use this work on misconceptions to better understand students' thinking and help them move towards more expert-like thinking. To assess students' conceptions in Newtonian physics, the Force Concept Inventory (FCI) stands out as the most widely cited conceptual assessment instrument (Hestenes, Wells, & Swackhamer, 1992).

The FCI is a 30-item multiple-choice instrument that surveys students' conceptions of force and motion. FCI questions are formulated in simple, everyday words and students do not need to make any calculations to answer. Besides the correct answer, the FCI provides incorrect choices that are based on answers that were most frequently given by students during interviews. Therefore, students are likely to find an answer that matches their thinking and not very likely to be guessing randomly at the possible choices offered. Teachers are frequently surprised by the poor results their students get on the FCI, even in the most elite institutions (Mazur, 1997, 2009). Indeed, it is quite surprising that students can solve many complicated numerical problems without fully understanding the basic underlying concepts (Kim & Pak, 2002). This, combined with the simplicity and the reliability (N. Lasry, Rosenfield, Dedic, Dahan, & Reshef, 2011) of the FCI, has made it one of the most widely used and researched instruments in physics education (McDermott & Redish, 1999).

5.3 Measuring conceptual gains: The Hake-gain

To measure conceptual change in students, inventories such as the FCI are usually given to students once at the beginning and then again at the end of a course. This is often referred to as "pre-post testing". In what is probably the most cited paper in physics education research (with over 2500 citations to date), Richard Hake showed that traditional methods, such as lecturing, are associated to sizably and significantly lower gains than student-centered "interactive engagement" pedagogies (Hake, 1998). To compare pedagogies, Hake analyzed data from more than six thousand college and university physics students and defined a measure of conceptual 'normalized' gain as:

$$g = \frac{Post - \mathrm{Pre}}{30 - \mathrm{Pre}} \quad (5.1)$$

In this instance above, Pre corresponds to the FCI score (out of 30) at the beginning of the term and Post, the score at the end of the term. The Post-Pre difference is divided by the maximum increase in score possible (or equivalently, the amount of incorrect answers given on the pre-test). Dividing by the maximum increase possible avoids issues such as ceiling effects that become significant when students have large Pre-test scores. Indeed, if a student scores 25/30 on the pre-test, she cannot gain more than 5 points. However, a student with 10/30 can gain 20 (and gaining more than 5 is quite possible). Hence, the normalized Hake-gain (equation

1) provides a value between 0 and 1 that describes the fraction of the total possible increase in score (how much increase out of a maximum of 5 for the first student and of a maximum of 20 for the second student). *A priori*, the Hake-gain provides a simple quantitative way to measure how students change conceptions.

5.4 Conceptual change can go both ways

In some cases, the post-test score is smaller than the pre-test score. In such cases, the net score decreases and changes are losses rather than gains (Marx & Cummings, 2007). To measure the normalized 'loss' in score, a similar approach to the Hake-gain is taken. The Hake-gain is a value between 0 and 1 that describes the percentage of maximal increase in score. To adjust for when post-test scores are smaller than pre-test scores, Marx and Cummings (2007) proposed a normalized change metric that measures both gains and losses as is defined as follows:

$$g_{change} = \begin{cases} \dfrac{Post - \text{Pre}}{30 - \text{Pre}} & if \quad Post > \text{Pr}\,e \\ \dfrac{\text{Pre} - Post}{\text{Pre}} & if \quad \text{Pr}\,e > Post \end{cases} \qquad (5.2)$$

To understand the effect of the losses, we propose a hypothetical scenario. Suppose a very strong student scores 28/30 on the pre-test. That student only has 2 more questions to get to a maximal Hake-gain of 1 (i.e. 100%). However, it may happen that the student scores a 24/30 at the end of the term. This means that the student has 'lost' 4 point out of 30 from their original score. This loss of 4 should not be compared to the two questions the student had left to gain (which would yield a meaningless Hake-gain of -2). Instead, the loss of 4 questions should be compared to what the student had in the beginning: namely 28 correct questions of which 4 were 'lost' (for a loss of 4/28 or 0.14). Hence, the equation (2) proposed by Marx and Cummings (2007) allows researchers to measure gains, as a fraction of maximum numbers of questions left to gain (a value between 0 and 1), as well as losses as a fraction of maximum possible numbers of questions that can be lost (a value between -1 and 0). We use this metric is a means to measure conceptual change in physics.

5.5 Research Questions

In this chapter, we seek to determine the effectiveness of DALITE. All students using DALITE were given the FCI both at the beginning and at the end of the semester. It has been shown that the greater the prior-knowledge students have, the greater the conceptual gains (Coletta & Phillips, 2005; N Lasry et al., 2008). Hence, we seek to compare DALITE students with in a sample of students that are comparable in prior knowledge to our college student population. We also seek to determine whether the effectiveness of DALITE depends on the instructor implementing it. Our objective for this chapter leads to the following research questions.

1) Do students using DALITE gain statistically significantly more conceptions than students with the same prior knowledge who do not use DALITE?
2) How do students using DALITE compare to students who do not use DALITE but use real-time face-to-face Peer Instruction?
3) Does the effectiveness of DALITE depend on the instructor using it?

5.6 Methods

To determine the expected Hake-gain (Hake-g), of non-DALITE students as a function of their pre-test score, we appealed to physics education researchers from our network of contacts and obtained anonymized data for 13,422 students from various institutions across the world. All students were assessed on the FCI both at the beginning and at the end of an introductory physics course. Most students took the FCI in high school courses (N_{HS} = 10 007). We also obtained data from one US public university (N_{PU} = 1560), three elite universities (N_{RIU} = 884) and three Cégeps (N_{TYC} = 971).

The resulting database does not include data identifying students, their gender or indications of the various pedagogies used. Of these, 2912 students had incoming pre-test scores that were of 9 or 10 on 30, just like our students. This subset of 2912 students was used as the non-DALITE control section. Second, we determine whether face-to-face interactions are necessary of if asynchronous Peer instruction can work as well. We identified two instructors who used Peer Instruction but not DALITE and were willing to share their pre-post FCI data. This "active learning" control consisted of 188 students who were exposed to real-time Peer Instruction in their classes. Finally, to determine the variation of effectiveness between instructors, we compare the conceptual gains across the five DALITE sections studied.

5.7 Results

5.7.1 Conceptual gains of DALITE students vs matched controls

We compare the 137 students having used DALITE with the 2912 students who have not and find that the DALITE students learn statistically significantly more than the controls (0.47±0.02 vs 0.35±0.006; $p<0.00001$).

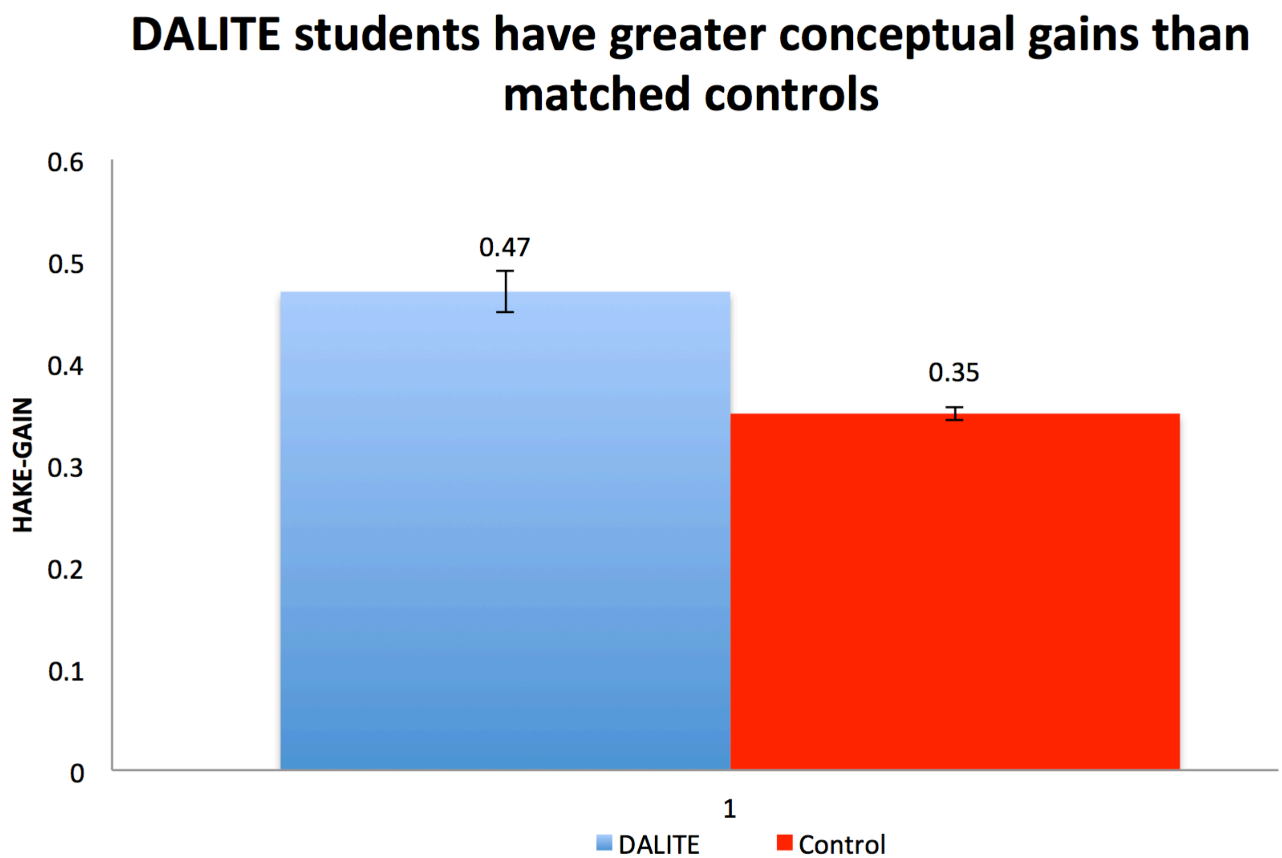


Figure 5.1. Shows that students using DALITE (n=137) in their college courses had significantly higher conceptual gains at the end of the semester ($p< 0.00001$) when compared to controls (n=2912)

5.7.2 Conceptual change in DALITE vs real-time Peer Instruction

We compare the 137 students having used DALITE with the 188 students having used real-time Peer Instruction in their classrooms. Figure 3.2 shows that, with respect to conceptual gain, the use of real-time Peer Instruction in class is not statistically significantly different to using the asynchronous Peer Instruction platform DALITE (0.47±0.02 vs 0.48±0.02 ; p=0. 84).

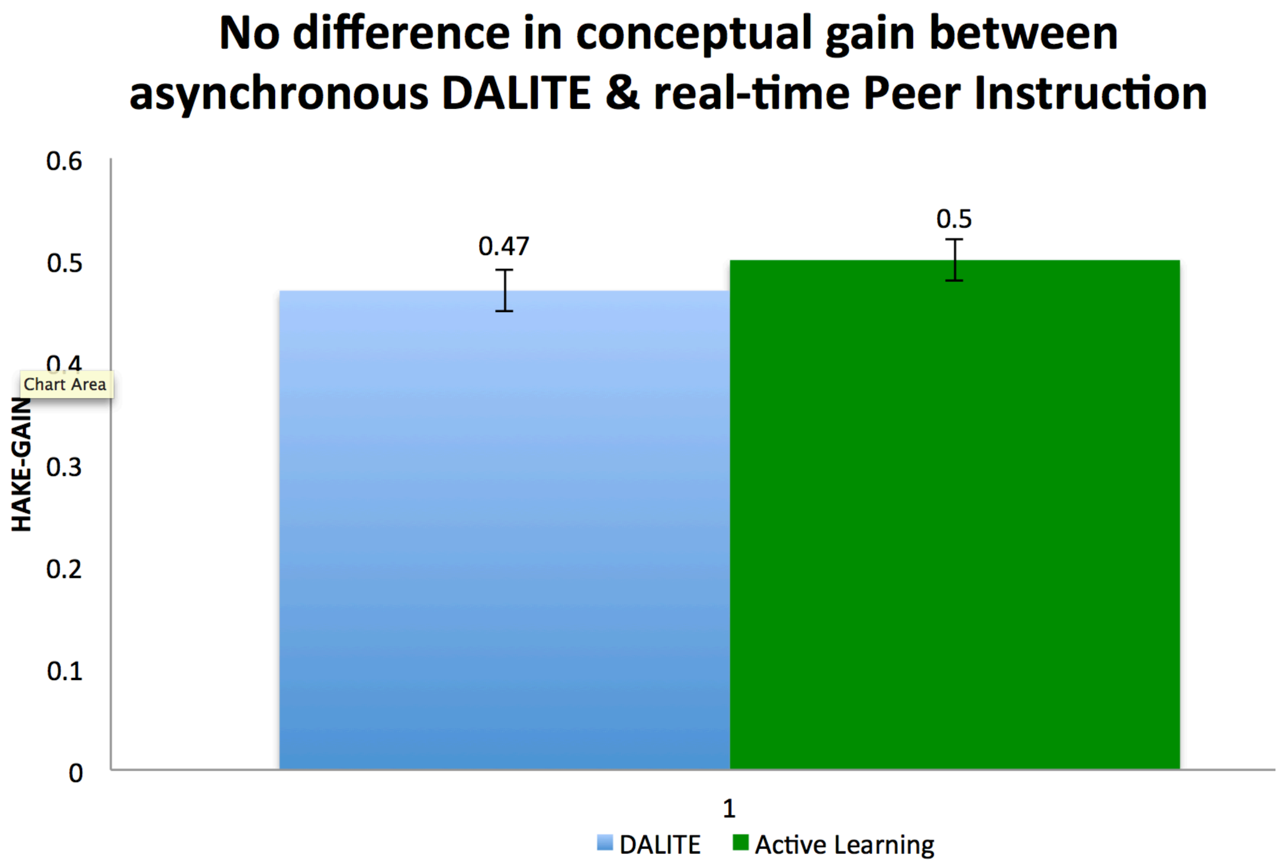


Figure 5.2. Shows that students using DALITE (n=137) in their college courses do not differ significantly in conceptual gains (p= 0.38) than students who used real-time Peer Instruction (n=188)

5.7.3 Comparing conceptual change between DALITE sections

We compare the conceptual gains achieved in the 5 different sections having used DALITE. We find a very surprising similarity between four of the five groups and a small difference with a fifth. The overall differences are between groups are not statistically significant ($g_1 = 0.50$; $g_2 = 0.50$; $g_3 = 0.47$; $g_4 = 0.48$; $g_5 = 0.38$; $p = 0.06$) with four of the five groups being extremely similar and close to all the variation residing in the fifth group. We will later return to this group for deeper analysis.

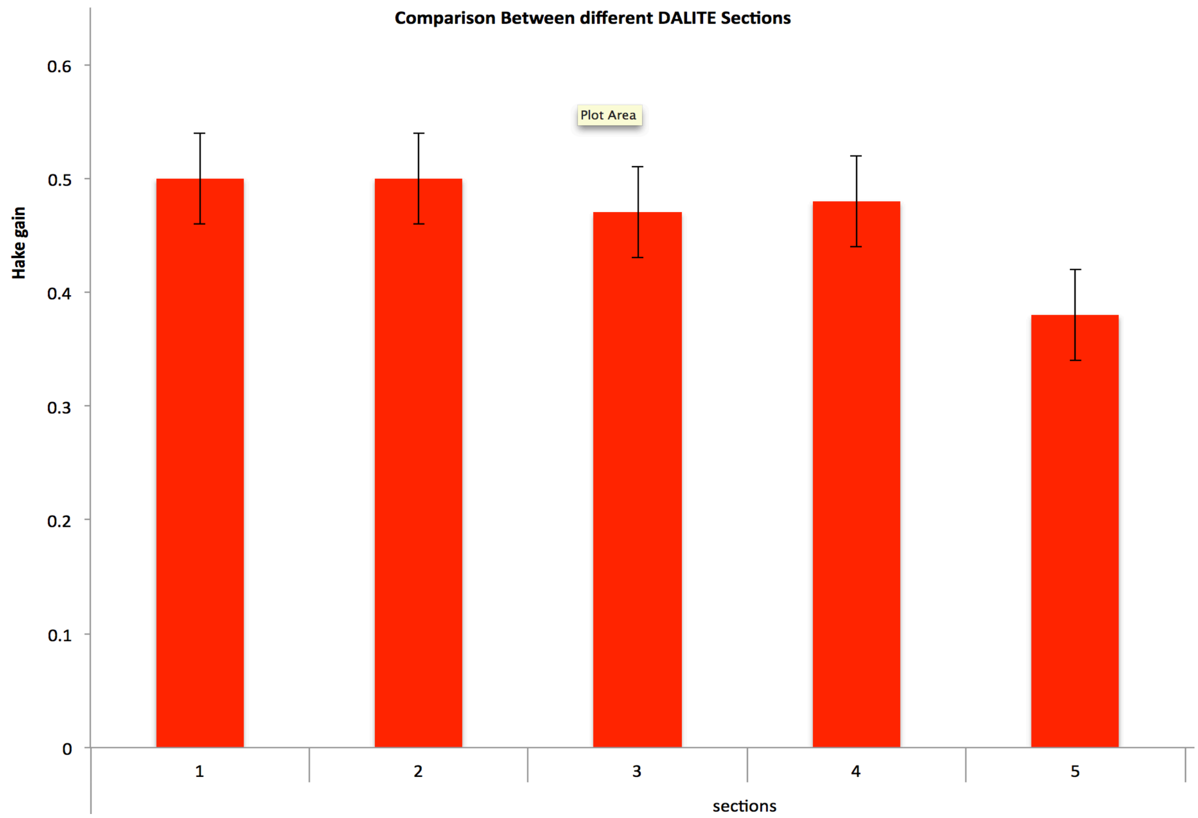


Figure 5.3. Shows that different sections using DALITE (n=137) do not differ significantly from each other in conceptual gains (g_1 = 0.50; g_2 = 0.50; g_3 = 0.47; g_4 = 0.48; g_5 = 0.38 ; p= 0.06)

5.8 Discussion

We set out to measure whether DALITE is a platform that enables teachers to bring student-centered active learning outside of the classroom. DALITE provides students with a means to discuss concepts with each other asynchronously. In this chapter we assess whether DALITE works better than lecture-based instruction, if it differs from real-time face-to-face Peer Instruction and whether its effectiveness depends on the instructor using it.

5.8.1 DALITE students have greater conceptual gains than controls

We find that DALITE works well. It enables students to achieve significantly greater conceptual gains than students who were matched for background knowledge and did not use active learning. This finding can be viewed in two ways. First, it can be viewed as being consistent with meta-analytic findings that show how active learning enables students to achieve greater conceptual change (Freeman et al., 2014; Hake, 1998). However, viewing this finding as a replication assumes that DALITE successfully implements active learning asynchronously. The second way to view these findings is to be agnostic on whether active learning is possible in asynchronous collaborative modalities and test whether DALITE achieves the conceptual change benchmarks expected from of active learning approaches. From this perspective, our findings suggest that DALITE achieves the active learning benchmarks in conceptual learning because DALITE students achieve significantly greater conceptual gains than students with similar prior knowledge.

5.8.2 DALITE is as good as face-to-face Peer Instruction

We next address whether DALITE students achieve as much conceptual gains that the original Peer Instruction students. Figure 2 shows that there is no significant difference between DALITE students and students using real-time Peer Instruction in their courses. This finding suggests that DALITE successfully enables active learning through asynchronous Peer Instruction. More generally, these data show that active learning does not require face-to-face real-time collaboration and that asynchronous collaborations can be equally effective.

5.8.3 Effectiveness of DALITE is similar between instructors

Our data show that the use of DALITE is almost independent of the instructor using it. In four of the five sections that implemented DALITE, the conceptual gains found were with 3 percent of each other (from 0.47 to 0.50). One of the five sections however had a sizably (though not significantly) different conceptual gain fro the four others (0.38). We identified may qualitative differences in this section and will return to document the particularities of this section in the upcoming chapters.

5.9 Conclusion

In this chapter we look at the effectiveness of DALITE by comparing groups of students that have used it or not. We find that DALITE students have greater conceptual gains than controls, that DALITE is as good as face-to-face Peer Instruction and that the effectiveness of DALITE is similar between instructors. However, the focus of this chapter is entirely on comparing groups of students. Hence, this chapter is silent with respect to the individual students and their individual conceptual change. In the following chapter, we change the unit of analysis and document how individual students change conceptions on a student-by student basis and on a concept-by-concept basis.

5.10 References

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*(1), 66-71.

Coletta, V., & Phillips, J. (2005). Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability. *American Journal of Physics, 73*, 1172.

Crouch, C., & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics, 69*(9), 970-977.

Fagen, A. (2003). *Assessing and Enhancing the Introductory Science Course in Physics and Biology: Peer Instruction, Classroom Demonstrations, and Genetics Vocabulary.* (PhD), Harvard University, Cambridge, MA.

Fagen, A., Crouch, C., & Mazur, E. (2002). Peer Instruction: Results from a Range of Classrooms. *PHYSICS TEACHER, 40*(4), 206-209.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 201319030.

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.

Halloun, I. A., & Hestenes, D. (1985a). Common-Sense Concepts About Motion. *American Journal of Physics, 53*(11), 1056-1065.

Halloun, I. A., & Hestenes, D. (1985b). The Initial Knowledge State of College Physics Students. *American Journal of Physics, 53*(11), 1043-1055.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher, 30*(3), 141-158.

Kim, E., & Pak, S. (2002). Students do not overcome conceptual difficulties after solving 1000 traditional problems. *American Journal of Physics, 70*, 759.

Lasry, N. (2006). PAREA Report: Implementing Peer Instruction in Cegep

(L'enseignement par les pairs au cégep) *PAREA Report* (pp. 69). Montreal, QC: John Abbott College.

Lasry, N. (2008). Une mise en œuvre au cégep de la méthode d’apprentissage par les pairs de Harvard. *Pédagogie Collégiale, 21*(4), 21-28.

Lasry, N., Mazur, E., & Watkins, J. (2008). Peer instruction: From Harvard to the two-year college. *American Journal of Physics, 76*(11), 1066-1069. doi: Doi 10.1119/1.2978182

Lasry, N., Rosenfield, S., Dedic, H., Dahan, A., & Reshef, O. (2011). The puzzling reliability of the Force Concept Inventory. *American Journal of Physics, 79*, 909.

Marx, J. D., & Cummings, K. (2007). Normalized change. *American Journal of Physics, 75*, 87.

Mazur, E. (1997). *Peer instruction : a user's manual*. Upper Saddle River, N.J.: Prentice Hall.

Mazur, E. (2009). Education Farewell, Lecture? *Science, 323*(5910), 50-51. doi: Doi 10.1126/Science.1168927

McDermott, L., & Redish, E. (1999). Resource letter: PER-1: Physics education research. *American Journal of Physics, 67*, 755.

Minstrell, J. (1982). Explaining the’’at rest’’condition of an object. *The Physics Teacher, 20*, 10.

Pfundt, H., & Duit, R. (1988). Bibliography. Students' alternative frameworks and science education.

Viennot, L. (1979). Spontaneous Reasoning in Elementary Dynamics. *International Journal of Science Education, 1*(2), 205-221.

CHAPTER 6
EVOLUTION OF STUDENTS' CONCEPTIONS

Analyzing whether a pedagogical approach is effective inherently depend on the measures used to establish effectiveness. Each metric has its level of granularity. In the previous chapter, we document the effectiveness of DALITE by comparing the average conceptual gains achieved by different groups of students. To document the effectiveness of DALITE we used groups of students as our unit of analysis and compared averages of aggregate scores between groups of students. This enabled us to show that DALITE helps groups in the way that meta-analyses show that active learning helps groups of students (Freeman et al., 2014; Hake, 1998). In the present chapter we change the unit of analysis and document the effectiveness of DALITE by looking at how individual students change their conceptions. Thus, instead of analyzing aggregate scores for groups of students we now analyze the evolution of single concepts in individual students.

When measuring conceptual change, the Hake-gain (as expressed in equation 5.1) can be found for an entire class by subtracting the average pre-test score of the group from their average post-test score. Alternately, the Hake-gain can also be found for each student by subtracting the pre-test from the post-test scores. Interestingly, the average of all the individual gains in a class can be surprisingly different from the average gain for the class (Bao, 2006). Although this is a clearly documented and well-known difference in computing conceptual change, what we propose is to radically change the scale at which the conceptual change is being measured. In both instances of Hake-gain, the resulting gain is a measure of the change in total FCI score, that is, an aggregate of the answers on *all* the questions. We seek to determine how students evolve with respect to each one of the conceptions surveyed at the beginning of the semester.

6.1 Gains and Losses Student-by-Student & Question-by-Question

A finer grained way to look at how students move towards or away from Newtonian conceptions is to analyze students' answers on each question both at the beginning and at the end of the course (Miller et al., 2010). We treat each question as a dichotomous variable (i.e. can be a right Newtonian answer or a wrong non-Newtonian answer). Comparing the answers given by students on the pre-test and the post-test results in four possible transitions: right-to-right (RR), right-to-wrong (RW), wrong-to-wrong (WW) and wrong-to-right (WR). Only two transitions affect the score: WR increases score and RW decreases it. Looking at responses question-by-question, the score change (Post-Pre) of the Hake-gain in equation 1 can be rewritten as the difference WR-RW and the Hake-gain as:

$$g = \frac{WR - RW}{W_{PRE}} = \frac{WR}{W_{PRE}} - \frac{RW}{W_{PRE}} \qquad (6.1)$$

The numerator in equation 6.1 expresses the change in score and the denominator W_{PRE} describes the total number of incorrect answers given on the pre-test, or equivalently the maximum increase in score possible. Written this way, the Hake-gain appears to be the difference of a normalized gain and loss. However, while the first term of equation 6.1 can be viewed as a gain, the second term is *not* a properly normalized loss. The Marx & Cummings normalized change metric given in equation 5.2 shows how losses must be compared to questions that are initially correct. Hence, normalized losses should be expressed as the fraction of initially correct answers (R_{PRE}) that are lost, such that gains (G) and losses (L) can be expressed as follows:

$$G = \frac{WR}{W_{PRE}}; L = \frac{RW}{R_{PRE}} \qquad (6.2)$$

Note that the last term in equation 6.1 is a ratio of RW to questions that are initially incorrect (W_{PRE}). This is at best difficult to interpret and at worse meaningless. Therefore, using this reformulation of the Hake-gain on a question-by-question basis (equation 3) is ***only*** valid if the number of RW is approximately zero. If the fraction of RW is large, the Hake-gain does not accurately represent the changes in individual student conceptions. A fine-grained conceptual change metric should take into account appropriately normalized Gains (G) and Losses (L) such as defined in equation 4.

Conceptual losses are particularly interesting in instruments like the FCI. These tests are designed to avoid guessing because all the incorrect answers are answers that are most frequently given by students in interviews (Hestenes, Wells, & Swackhamer, 1992). However, the metric given by equation 5.2 does not account for losses on individual questions (Marx & Cummings, 2007). Indeed, losses can occur even when the overall score increases, provided some questions answered correctly in the beginning of the course are answered incorrectly at the end. Looking at these kinds of losses on a question-by-question basis can provide new finer grained insight into how students are changing conceptions.

6.2 Research Questions

In this chapter, we seek to determine whether students using DALITE as a means to participate in asynchronous Peer Instruction learn more concepts than comparable peers who most often do not use active learning pedagogies. It has been shown that the greater the prior-knowledge students have, the greater the conceptual gains (Coletta & Phillips, 2005; N Lasry, Mazur, & Watkins, 2008). Hence, we are interested in a sample of students that are comparable in prior knowledge to our college student population. Our objective for this chapter is to dig deeper into students' conceptions. We focus on how individual students' change in each of the conceptions surveyed and essentially the questions posed in the previous chapter. This leads to a similar analysis carried out of a different level of granularity. The questions we pose in this chapter are:

1) What are the expected individual student's Hake-gains (Hake-g) as well as the concept-by-concept Gains (G), Losses (L) of DALITE and non-DALITE students as a function of their prior knowledge (measured using pre-test FCI scores)?
2) Do students using DALITE differ statistically significantly in Hake-g, G and L from students with similar prior knowledge?
3) How do students using DALITE compare to students using real-time face-to-face Peer Instruction?

6.3 Methods

To determine the expected Gains (G), Losses (L) and Hake-gains (Hake-g), of non-DALITE students as a function of their pre-test score, we obtained annoynimized data for 13,422 students from various institutions. All students were assessed on the FCI both at the beginning and at the end of an introductory physics course. Most students took the FCI in high school courses (N_{HS} = 10 007). We also obtained data from one US public university (N_{PU} = 1560), three elite universities (N_{RIU} = 884) and three Cégeps (N_{TYC} = 971). The resulting database does not include data identifying students, their gender or indications of the various pedagogies used.

6.4 Results

6.4.1 What are the expected Gains, Losses & Hake-gains?

We use the large data set (n= 13 422) collected from our network of non-DALITE users and use the Pre-FCI score to account for prior knowledge. We find on average that 46% of questions (G=0.46) students answered incorrectly in the beginning of the term were answered correctly at the end. Surprisingly, we also find that 30% of questions answered correctly in the beginning of the term were answered incorrectly at the end (L=0.30).

To illustrate the dependence of our measures on prior knowledge, Figure 6.1 shows the dependence on Pre-FCI scores of G, L, the G – L difference and Hake-g.

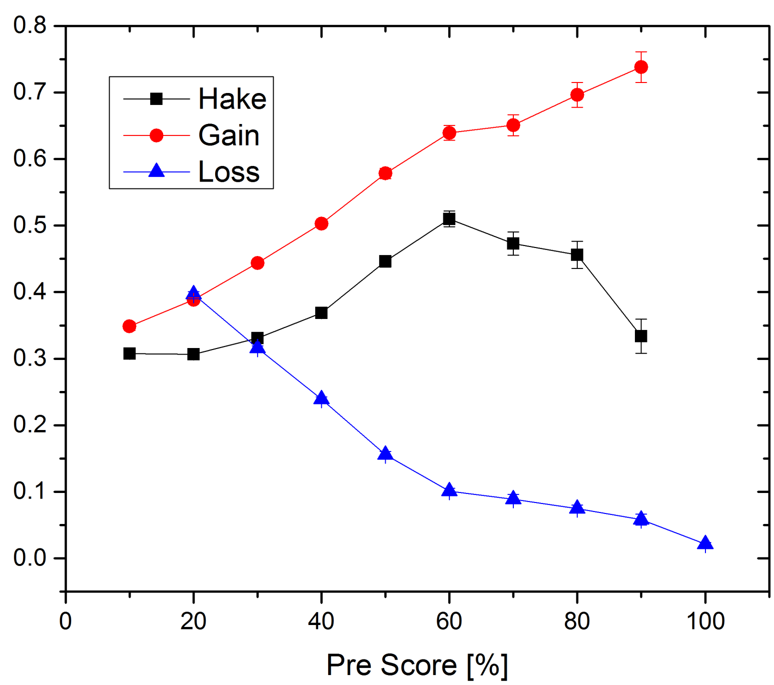


Figure 6.1. As background knowledge increases (as measured by pre-FCI score), appropriately normalized gains (G) increase and losses (L) decrease. The Hake-gain has a non-obvious relationship to background knowledge.

Gains (G) increase while losses (L) decrease with growing background knowledge. The lines of best fit for gains are described by G = 0.550x + 0.282 while losses are described by L= -0.438x + 0.435. In contrast, Hake-g is fairly independent of background knowledge at lower scores (10-30%) and has an inverted U-shape relationship with Pre scores between 40-90%. The Hake-g appears to be independent from G, or L.

6.4.2 Comparison of DALITE vs non-DALITE students

The lines of best fit obtained from figure 1 allow us to compare how well our DALITE students did when compared to students with comparable Pre-test FCI scores in the large data set. In our college population, the average pre-FCI score for DALITE students was 33.1%. Of the 13 422 students for which we collected pre and post FCI data, a total of 2912 students had pre-FCI scores ranging between 25%-35%. Using the best fitting line equation, Table 1 compares the expected gains and losses for students coming in with a 33% pre-FCI score.

Table 6.1. DALITE students have larger gains and smaller losses than matched controls.

	Expected (n=2912)	DALITE (n=137)	Effect-size
G	0.464 ±0.003	0.59 ±0.02	0.6
L	0.291 ±0.003	0.19 ±0.02	0.5
Hake-g	0.347 ±0.003	0.47 ±0.02	0.4

These data shows that DALITE students have larger gains and smaller losses than what would be expected from students with similar prior knowledge. The size of this difference is moderate (effect sizes of 0.5-0.6) and statistically significant ($p<0.05$).

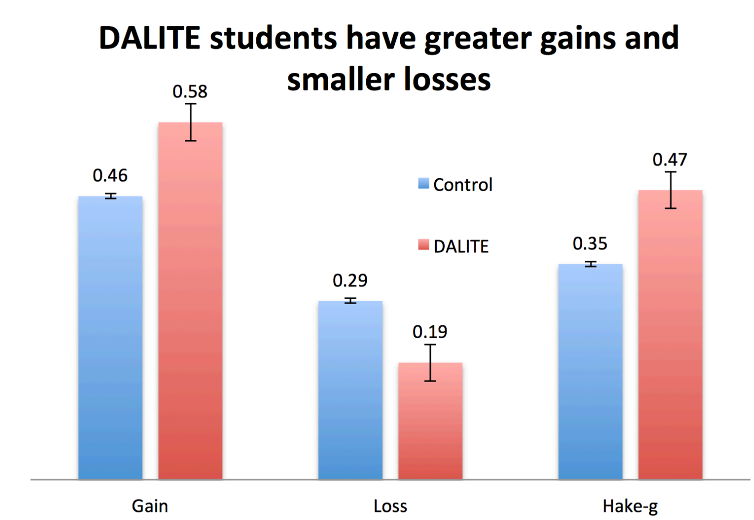


Figure 6.2. Gains and losses for DALITE and Control sections.

Although DALITE students do perform better than expected, we are still left with the question of whether the asynchronous Peer Instruction is as effective as face-to-face implementations.

6.4.3 Comparison of asynchronous DALITE vs real-time Peer Instruction

We measure the effect of active learning in the asynchronous modality on DALITE and compare it to the effectiveness of real-time Peer Instruction as it is used in classes.

Table 6.2. Comparison of asynchronous DALITE vs. real-time Peer Instruction.

	Real-time Peer Instruction (n=188)	DALITE (n=137)	p-value
G	0.57 ±0.01	0.59 ±0.02	< 0.0001
L	0.21 ±0.01	0.19 ±0.02	0.001
Hake-g	0.48 ±0.02	0.47 ±0.02	0.4

We find no difference if Hake-gain (as reported in the previous chapter) and find virtually no difference in the concept-by-concept gains or losses either.

6.5 Discussion

In this chapter we propose new metrics for measuring conceptual change in students by changing scales to a concept-by-concept unit of analysis. We present a set of finer-grained metrics that enable teachers and researchers to assess how individual students change their answers on each conception surveyed by the FCI. Our results are both interesting and intriguing.

6.5.1 Conceptual learning does not progress linearly

We find that students change close of half of the incorrect answers to Newtonian answers (G= 0.46) at the end of their introductory physics courses. However, of all correct answers provided by students in the beginning of the term, roughly 30% are answered incorrectly at the end of the term. This large change *away* from Newtonian conceptions after a full semester of instruction is intriguing and has never been documented in this way. We recently published this unexpected result the high-impact Nature Physics journal (Lasry, Guillemette, & Mazur, 2014). We used these findings to suggest that learning progresses in a non-linear way. Students gain some concepts and lose others. This proposition that has provided the paper with its title: “Two steps forward, on step back”.

6.5.2 DALITE students have greater gains and smaller losses

We used the concept-by-concept analysis of change to determine the difference between DALITE students and students who did not use any active learning. Our findings suggest that learning progresses in a non-linear way for both populations. Although DALITE students do exhibit both gains and losses, students in the control section exhibit statistically less gains and more losses.

6.5.3 Asynchronous DALITE as good as real-time Peer Instruction

One of the main questions in this chapter and the previous is whether active learning requires real-time collaboration or if asynchronous collaboration can work equally well. We find that the Hake-gain, as well as the normalized gain and loss found in the DALITE group are not significantly different from the ones found in the real-time Peer Instruction group. Hence, active learning may be as effective when carried out in an asynchronous modality.

6.6 References

Bao, L. (2006). Theoretical comparisons of average normalized gain calculations. *American Journal of Physics, 74*, 917.

Coletta, V., & Phillips, J. (2005). Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability. *American Journal of Physics, 73*, 1172.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 201319030.

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.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher, 30*(3), 141-158.

Lasry, N., Guillemette, J., & Mazur, E. (2014). Two steps forward, one step back. *Nature Physics, 10*(6), 402-403. doi: 10.1038/nphys2988

Lasry, N., Mazur, E., & Watkins, J. (2008). Peer instruction: From Harvard to the two-year college. *American Journal of Physics, 76*(11), 1066-1069. doi: Doi 10.1119/1.2978182

Marx, J. D., & Cummings, K. (2007). Normalized change. *American Journal of Physics, 75*, 87.

Miller, K., Lasry, N., Reshef, O., Dowd, J., Araujo, I., & Mazur, E. (2010). *Losing it: The Influence of Losses on Individualsí Normalized Gains*.

CHAPTER 7
COMMON CONCEPTUAL TEST

This chapter addresses the research question: *What is the impact of the DALITE activities on the students' conceptual knowledge, assessed by the common conceptual test questions?* And, *Whether there is a difference between DALITE students and the comparison group, both on the conceptual tests as well as on other class assessments (i.e., the mid-term tests)?*

Recall that there were three Common Conceptual tests, described earlier, each composed of two conceptual questions. Each test was intended to assess the key concepts covered in that particular part of the course, Kinematics (CCT1), Dynamics (CCT2) and Conservation (CCT3) principles. Both questions required students to write explanations for their choice of answer. These tests were administered to all five DALITE treatment sections as well as the AL comparison group as part of the regular mid-term assessments. In addition, all six groups wrote similar mid-term tests covering the same three course units. We will report on both these assessments. Note, however, we will not report on CCT3 in this document.

7.1 Common Conceptual Tests results

The three Common Conceptual Tests (CCT) provided a mix of results and showed that while students gained some conceptual knowledge, they did not always know how to transfer this knowledge to a new context, even with the extra support of the DALITE system (5 treatment groups) or active learning activities (the comparison group). This lead us to look more closely at rationales from the individual questions and at individual groups thereby making this analysis a qualitative investigation of these data. Under this close observation we are able to begin to identify important differences that may be a key to unlocking some of the misery around the processes of conceptual change, and more importantly, about transfer of learning. We start this reporting with the results of the first conceptual test, CCT1 then move on to CCT2 where we go even deeper into these subtleties.

7.1.1 Common Conceptual Test 1 (CCT1)

7.1.1.1. CCT1 question 1 (CCT1.Q1)

The results of CCT1.Q1 as presented in Figure 7.1 show that T10 outperformed the other five sections (T06, 07, 08, 09 and T-comp) who were given this end-of-unit test. However, the results also show that students in all six groups had difficulty correctly answering this question, the average of the groups was below 20% correct. Examining the students' rationales we see that many students overlooked the detail that the hot air balloon was "rising" upwards and therefore the stone continued to have an upward vertical velocity after it was released.

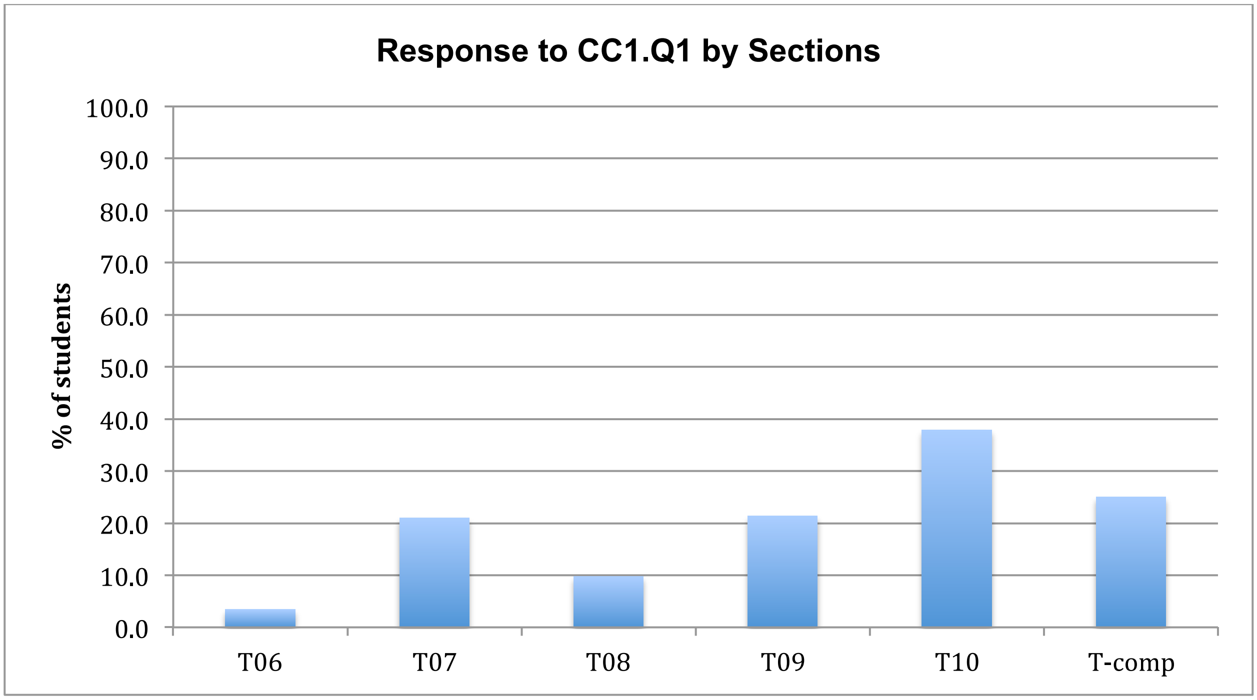


Figure 7.1. Results of CCT1.Q1 for each of the five treatment and one comparison group.

Examples of good explanations for the correct answer "B" are:

> *CCT1.Q1_B-1: since they are both dropped from the same height, the stone with less initial vertical velocity (assuming the positive direction is upwards) hits the ground first. Since the airplane is flying horizontal its initial vertical velocity is zero. Since the balloon is rising its initial vertical velocity is greater than zero.*
>
> *CCT1.Q1_B-2: The stone from the airplane hits the ground first because the stone has a velocity vertical going up directly so it takes more time to go down compared to the one from the airplane and has no initial vertical velocity.*

Examples of good explanations but for the wrong answer "C" are:

> *CCT1.Q1_C-1: They both hit the ground at the same time because they are both in free falls and during a freefall, horizontal velocity (plane) does not change the time it takes for the stone to hit the ground. Both stones are affected by gravity and are accelerating at the same constant vertical rate to the ground so they will to ground at the same time (see diagram).*
>
> *CCT1.Q1_C-2: I'm assuming we're ignoring the air resistance, both the hot air balloon and the airplane have the same vertical component acting on it (gravity) and they are at the same height, so they will hit the ground at the same time.*

> *CCT1.Q1_C-3: If we neglect the air resistance, both stones will hit the floor at the same time because its only acceleration is gravity and they are both dropped from the same height.*

Examples of explanations for the incorrect answer “A” are:

> *CCT1.Q1_A-1: the hot air balloon is moving vertically, so it's component is zero and the ball being dropped will also have an x-component of zero, whereas the plane is moving horizontally, so dropping the ball means taking both the x and y component into account, which makes the resultant velocity smaller than the one from the hot air balloon. It will go slower because the initial x-component.*

A likely explanation for the incorrect choice of answer “C” is that it represents the student’s misreading of the question. Note the quality of the examples of explanations for answer “C” presented above. However, these results also show how the subtlety of wording, which is clear to physics experts, is not the same for students. In fact, the results confirm that only the most capable of students in each of the sections were able to correctly answer this question. Based on such results it could be argued that such category of questions give us a clue as to what students do not focus on in their reading of physics questions. As well, such wording suggests that novices may not differentiate between small differences in the initial state of motion, which is a hallmark expert-like thinking in physic.

7.1.1.2. CCT1 question 2 (CCT1.Q2)

Recall that question CCT1.Q2 had two parts, the first was to have students select the correct answer, then asked students to select “*the rationale they agreed with most*” among an assortment of two for each of the three answers. The results of this question presented in Figure 7.2 shows that all groups performed equally well in selecting the correct answer, with the exception of group T07 (see Figure7.2). Additionally, the figure shows that the most groups, with the exception of T07 and T-comp were equally capable of selecting what the experts identified as the most well-written explanation, which is also for the correct answer “A1”.

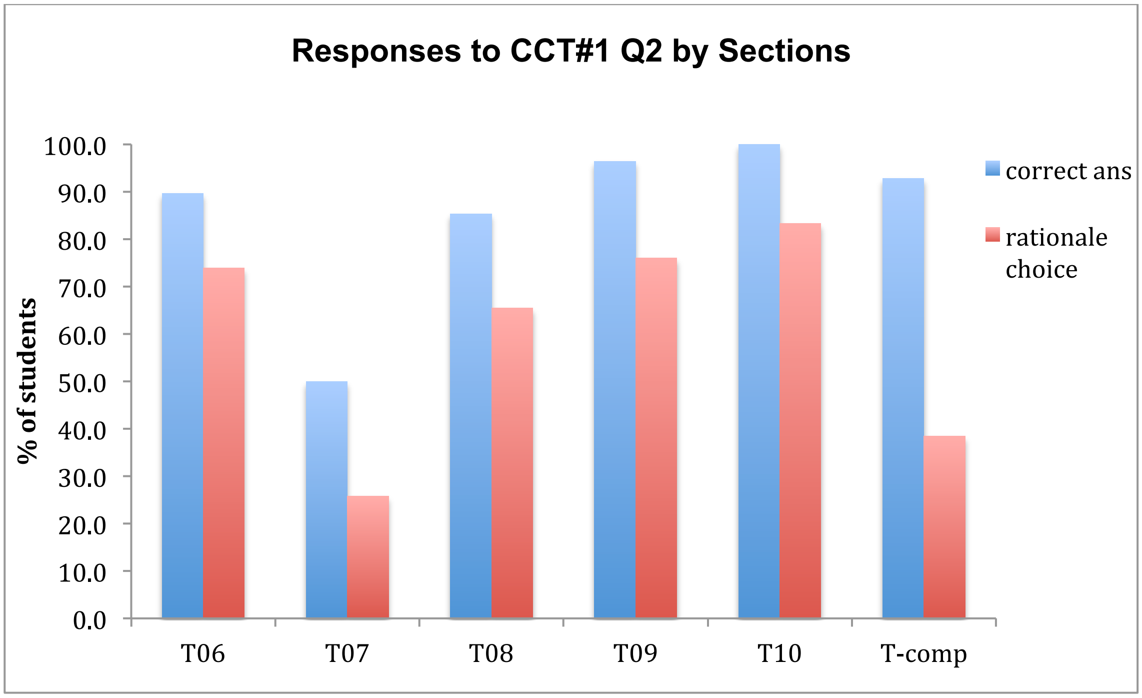


Figure 7.2. Results of CCT1.Q2 for each of the five treatment and one comparison group.

Recall that this assessment was based on material covered by the DALITE assignments therefore this is a transfer task. As such, these results suggest that all the DALITE groups, except for T07 have been able to see the similarity in the context between the assignments and the test. We can explain T07 only if we look at their use of DALITE, which shows that fewer students in that group completed the DALITE assignments. In other words, their completion rates were lower than other groups. The results also shows, however, that the comparison group T-comp who were not assigned DALITE did equally well. A possible explanation for these results are that this group also covered these concepts in their peer-instruction. Therefore, taken as a whole, these results suggest that students could recognize the deep structural similarities across the different contexts of their online assignments (and in-class peer instruction), and this end-of-unit test. The limitation to this claim comes when the students do not complete the assigned homework, which sometimes happens.

Looking at the results for the second part of the CCT1.Q2, the selection of the rationale, this time we see that both T07 and T-comp score lower than the other groups. Given that both T07 and T-comp would have had less or no exposure to DALITE, these results support the claim that the activity script that includes reading rationales may help students to evaluate better between the quality of explanations.

7.1.2 Common Conceptual Test 2 (CCT2)

7.1.2.1. CCT2 question 1 (CCT2.Q1)

The results of CCT2.Q1 as presented in Figure 7.3 show that this time around T08, T10 and T-comp outperformed the other two sections (T06, 09, note T07 data is missing). Recall that this question series was designed based on a similar set of questions in DALITE (the elevator and Martian lander questions) therefore it was expected that DALITE students should be able to perform well on this test. Given the results, it is clear that some groups were better prepared to complete this transfer task. And, that the comparison group T-comp is equally prepared by their active learning instruction that involves in-class Peer Instruction activities that likely used similar questions to that used in the DALITE system.

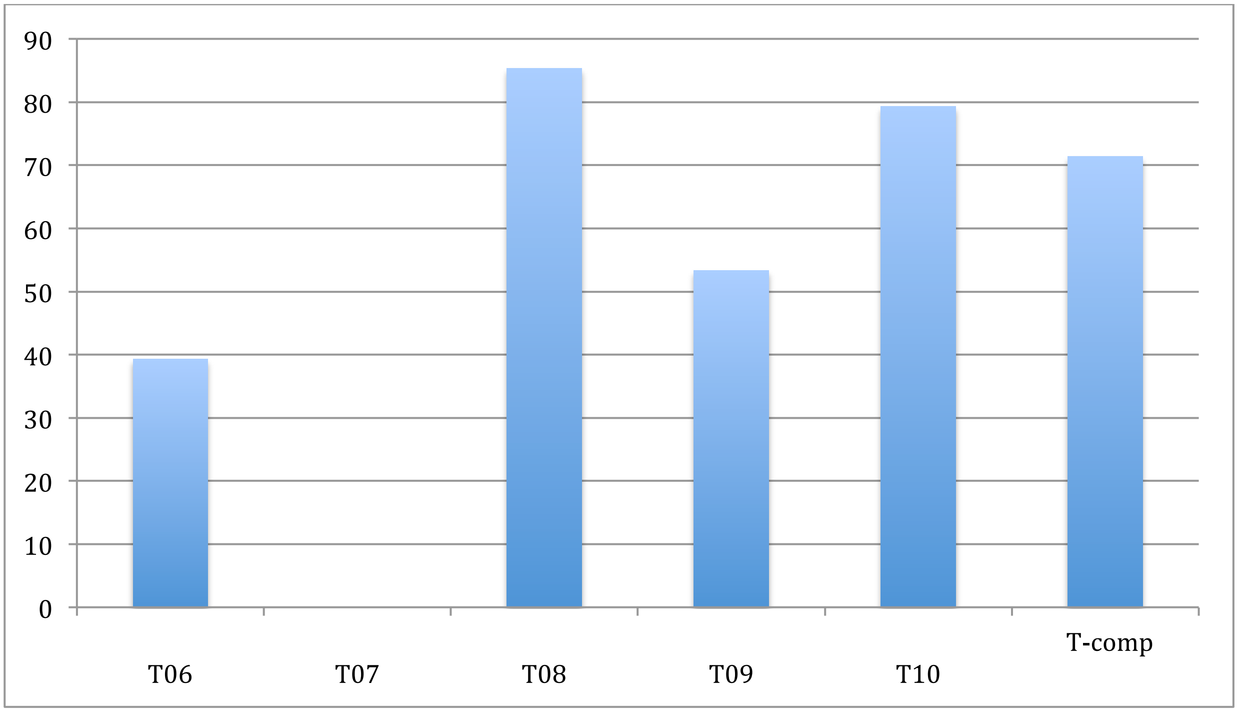


Figure 7.3. Results of CCT2.Q1 for each of the five treatment and one comparison group.

Examining the students' rationales we see that students in T08 and T10 who answered correctly, generally, were able to identify the deep structure concept that is common between this and the DALITE assignment. Additionally, we see no difference between explanations provided by the T-comp students who select the correct answer and those in the DALITE groups who select the right answer.

Examples of good explanations for the correct answer "A" are:

> *CCT2.Q1_A1: the tension has to be bigger than mg because the falling woman was falling down and the webbing slows her. This means that the acceleration has to be in the opposite direction of the woman which is up. If the acceleration is up, the Fnet has to be up. So the tension which points up has to be bigger than the mg.*

CCT2.Q1_A2: tension must be greater than the force of gravity because her velocity is downwards, and since she is slowing down the acceleration must be upwards using Newtons 2nd law, Fnet=ma the Net force must be in the same direction as the acceleration. Therefore tension must be greater than (mg)

What is different, however, is when we turn our attention toward the students who answer incorrectly. And, when we look at their explanations. What we see there allows us to create another research question, *What might be happening differently for the lower end students in T10 compared to T.comp*? We address this emergent question using a cross-case study comparison. And, conducted a fined-grained case study analysis of the rationales written by these sets of students and compared them for common test questions #2. A description of this study follows.

7.2 Cross-Case Comparison of Common Conceptual Test #2 (CCT2)

Using the data from the Common test #2, question 1, we designed two case studies which would allow us to compare the answers of the students in sections T10 and T.comp (the comparison section). Recall that the question from this second unit test of the semester involved an assessment of the tension in the rope required for Spiderman to save his falling girlfriend, Gwen (see Figure 7.3).

The question revolves around a comic-book scenario in which Spiderman catches a falling Gwen Stacy with a web. The student is asked, "As the webbing slows the falling woman, how does the force of tension in Spiderman's webbing compare with the force of gravity acting on her (mg)?" and given the following three choices:

(a) Tension > mg (correct choice)

(b) Tension = mg

(c) Tension < mg

1. Refer to the comic strip shown here, and pay particular attention to box #4.

   Spiderman's webbing applies a force to the falling victim, in order to slow her down before she hits the water.

   As the victim is falling but slowing down, how does the force of tension in Spiderman's webbing, compare to the force of gravity acting on her (mg)?

   a) Tension > mg
   b) Tension = mg
   c) Tension < mg, but greater than 0 N
   d) Tension = 0 N

   Provide a rationale to justify your answer.

   ____

   ____

   ____

   ____

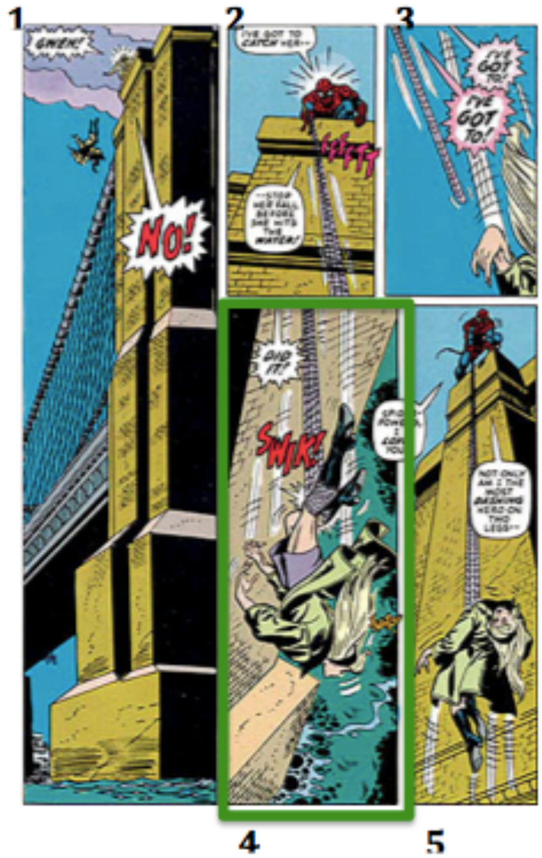


Figure 7.4. Spiderman question from Common test #2.

The students are then asked to provide a rationale to justify their answer. A complete and correct answer involves the following:

- Woman is falling (her velocity is downwards) but slowing down (acceleration is opposite the velocity, implying acceleration is upwards).
- Acceleration is proportional to the net force, so F_{net} must point upwards
- F_{net} (upwards) is the vector sum of the force of gravity mg (downwards) and the tension in the webbing (upwards). Thus, the tension must be larger than mg.

Table 7.1. Distribution of responses to multiple-choice component of Spider-Man problem for the T10 and T.COMP groups.

	T10	T.COMP
T > mg (correct)	23 (79.3%)	20 (71.4%)
T = mg	5 (17.2%)	3 (10.7%)
T < mg	1 (3.4 %)	5 (17.9%)

The responses from groups T10 and T.comp are summarized in Table 7.1. While small differences between the groups are evident (the T10 group does slightly better), these quantitative differences are not statistically significant due to the small number of respondents and the relatively high rate of correct answers.

The difference in mean length of student rationale (40±2 for T10, 36±3 for T.comp) is not statistically significant. Moreover, there were no significant differences observed in the *quality* of rationales submitted by students who correctly answered the multiple-choice.

Qualitative differences between the two groups are, however, evident when examining the rationales of students who chose the incorrect multiple-choice answer.

7.2.1 CCT2 case study by section

7.2.1.1 Case Study section T10

Responses of students answering "T=mg".

In the group T10, 5 of the 6 incorrect responses (and one of the students with a correct multiple-choice response!) involved mis-identification of the problem as one involving constant acceleration.

> *"at the moment she is caught, she changes direction so the acceleration is zero, meaning F_{net}=0, this means that the forces are balanced, therefore, tension is equal to mg"*

> *"the tension and the F_g are the same because the woman will stop moving (equilibrium) so that means that the two opposite forces have to be equal"*

> *"at this instant Spiderman catches the woman, stopping her from falling, if she is not moving the forces acting on her are in equilibrium, therefore F_T=mg"*

"because there's no acceleration in the tension side, so the net force must be 0"

"the tension is equal to mg b/c the woman is no longer falling which means she has an acceleration of 0. this therefore means that is a=0, F_{net} will also equal zero. Because both forces in the y direction have to be equal"

In each of these cases, the student explicitly mentions either that the system is in equilibrium or, equivalently, that the acceleration is zero. As such, the only error in reasoning evident in these responses is that the student has not recognized that a downwards-falling object that is slowing down is accelerating upwards. This could be due to a misconception.

Responses of students answering "T<mg".

Only one student chose multiple-choice option (c), T < mg. Compared with the well-formed (albeit incorrect), rationales above, this student's writing suggests an incoherent and fragmented understanding of Newton's Laws of motion.

"since because of the web, the falling slows down and does not stay at rest, the net force is still downward, these force of gravity is bigger, but tension is definitely more than 0, or else she would be falling down and die."

7.2.1.2 Case Study section T10

Responses of students answering "T=mg"

The rationales provided for the incorrect responses in the T.COMP class, tell a different story.

"since he is stopping her, that means that she doesn't move, then she is at rest and the T is equal to F_g (weight)."

"the tension is going to be equal to the weight because she will not more down, she will remain in the same place."

These two rationales might be considered similar to those provided by students in T10's class for this option choice. In both classes, they are using the idea of equilibrium to motivate the supposed balance of tension and gravitational forces. However, unlike the rationales from the other section, neither student makes *explicit* the idea that acceleration is zero of dynamical equilibrium. It's not clear whether the students are attributing the balance of forces to a zero acceleration (consistent with a Newtonian framework) or to zero velocity (which would be consistent with undifferentiated concepts of velocity and acceleration).

A clear example of this conflation of velocity and acceleration is found in the rationale of one student who (incorrectly) answered T = mg.

> *"it's a web and in order to catch her at rest, then the tension must be equal to the weight. If she was dragged as … tension> weight. If she fall in, weight > tension."*

Responses of students answering "T=mg".
Students answering T < mg were more frequent in this class than in T10. Of the 5 students drawing this conclusion, 4 of them have rationales that suggest that the students aren't differentiating between velocity and acceleration, at least when applied to Newton's second law:

> *"the victim is still falling, but . If tension was bigger than the weight, it would pull the victim up, and if they were equal the victim would remain in equilibrium. The tension is more than 0N because it slows the victim."*

> *"the tension must be smaller than her weight so that F_{net} is still pointing down (still falling) but must be greater than 0N so that she is not in free fall (she is)"*

> *"C because as she is , her weight is larger causing her to still fall, however, T≠0 because she is slowing so T is gradually increasing but will be smaller than weight until she completely stops at which point $T=F_g$."*

> *"the victim still has a velocity in the negative direction, however, her fall is , so an acceleration is the positive direction thanks to the tension. The tension is smaller than her weight because if it was greater, she would have both a positive velocity and acceleration. It is not zero either as in this case. She would continue falling and not slow down."*

The remaining incorrect rationale suggests clear thinking about the relationship between velocity and acceleration. The student has, however, elected to use inconsistent mathematical formulas instead of clear conceptual reasoning to arrive at a conclusion.

> *"if she is still falling but she has an acceleration with is opposite to her original acceleration. $T-F_g=-ma$ and $T-ma=F_g$."*

7.3 Summary of CCT results

The small number of students and high rate of correct responses to this question make statistical comparisons between groups meaningless. There are, however, clear qualitative differences between these groups when looking at the rationales provided for the incorrect responses.

In the T10 class, all but one of the incorrect rationales can be traced to the misidentification of zero acceleration, but with an otherwise consistent, Newtonian, explanation for their thinking.

In contrast, many of the incorrect rationales from the T.COMP group, indicate that these students had more serious misconceptions, in particular conflation of velocity and acceleration as applied to Newton's second law of motion.

7.4 Sorting Tasks Assessment

There were two sorting task administered to the five groups of DALITE treatment students. For the most part, these assessments accomplished their goals (1) to determine whether students could identify fundamental similarities between problems; and (2) to promote the message that conceptual understanding means reading past surface similarities to get at the nature of the physics within the problem. Generally, results showed that students were relatively successful at these tasks. However, the collection of rationales and observational data from the classroom discussions allowed us to see beyond the mere answers. Therefore, from this assessment we learned some valuable information that we elaborate on next.

7.4.1 Analysis of Sorting Task #1 (ST.1) Kinematics

Recall that the task involves three series of three problems each. The students were asked to identify which of the three problems were most alike, and to indicate their reasoning. The students were then instructed to discuss their responses in groups, and allowed to make any changes that they wanted, but in a different colour of ink so that changes pre- to post-discussion could be tracked.

Sorting Task 1 (ST.1), the first row is shown in Figure 7.5. A & B may look very similar due to their graphs, however it is problems A & C which are most alike in their physics, as both have *two different phases of constant acceleration* (while problem B only has two different phases of constant *velocity*).

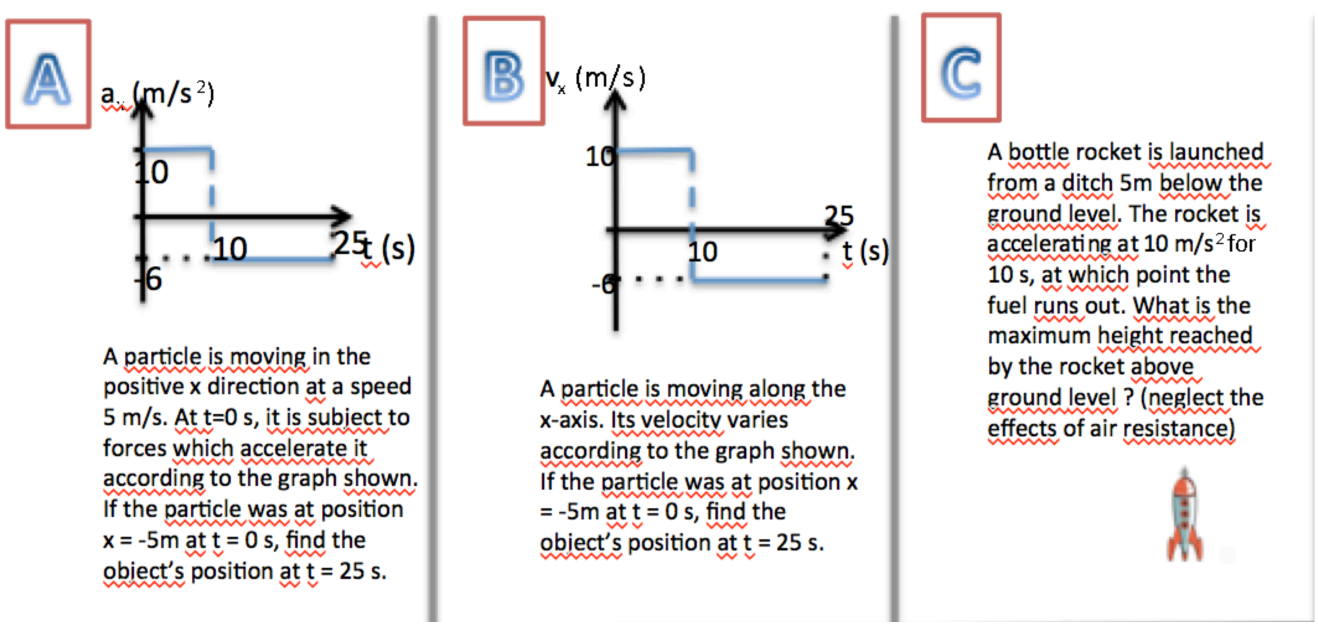


Figure 7.5. Sorting Task 1 – Kinematics, row 1.

The second row problems D & F may look most similar because they both include descriptions based on unit vectors (see Figure 7.6). However, F is the problem that is very different from the other two, as the particle experiences a *changing* acceleration. Problems D & E are actually both 2D kinematics problems, each with *constant* acceleration in one of the two directions.

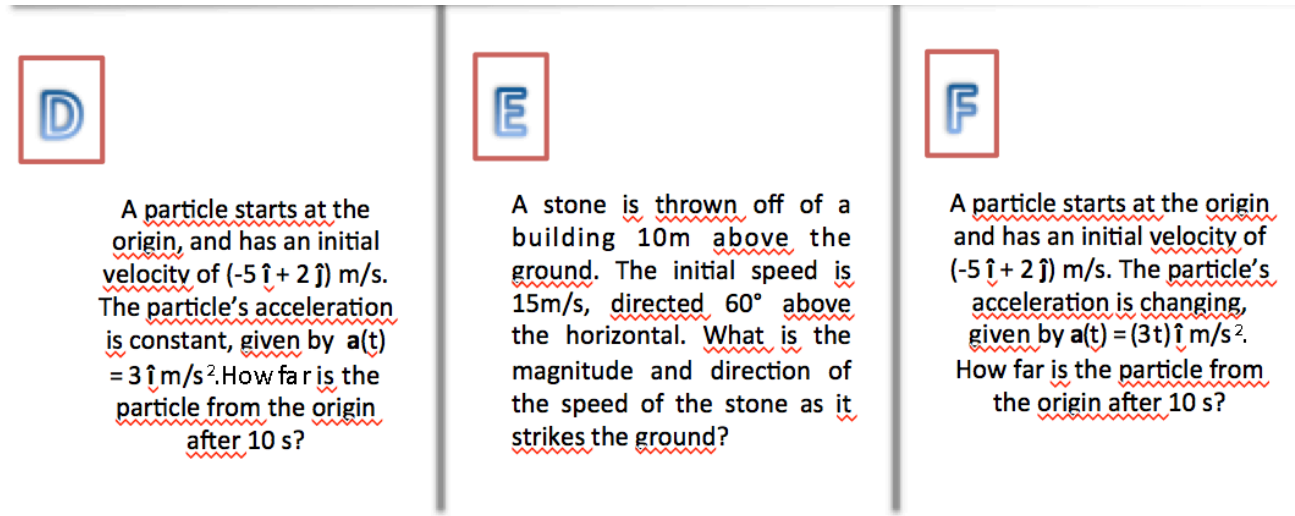


Figure 7.6. Sorting Task 1 – Kinematics, row 2.

The third row problems G & H may look the most similar to each other, as they both are about a spinning wheel (see Figure 7.7). However G & I are the ones that have the most in common in their physics, as they both have *constant angular accelerations* (problem H has *zero angular acceleration*, and is more concerned with *radial* acceleration).

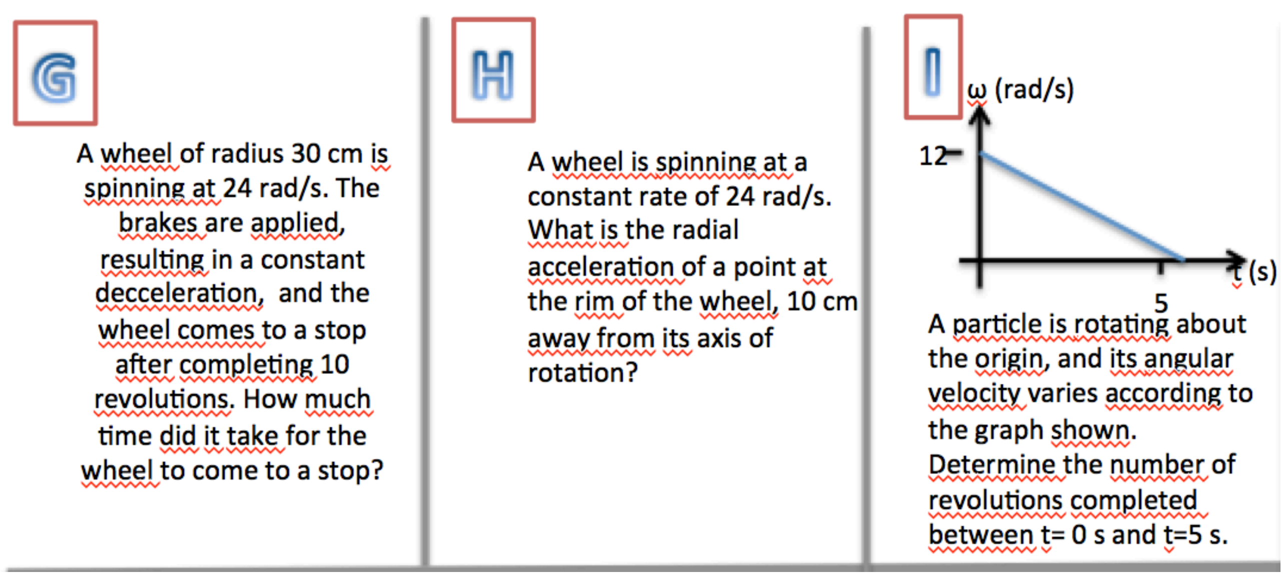


Figure 7.7. Sorting Task 1 – Kinematics, row 3.

Results of the sorting task, ST1-R1 show that sections T07, T08 and T09 easily identified the structural similarities between problems A and C (Figure 7.2). That said, sections T07 and T09 also showed a considerable amount of openness to change with 32% and 26% of students changing their answers after discussion. Surprisingly, sections T06 and T10 show very similar results in this instance, given that in other assessments they look like opposites. When we look at the rationales from the students, however, we see there is a difference between these two.

Table 7.2. Results of Sorting Task 1 assessment row 1 (ST1-R1), for five treatment sections.

Row 1	T06	T07	T08	T09	T10
A/C first	48.1	67.6	77.1	70.4	55.2
A/C review	14.8	32.4	14.3	25.9	17.2
total right	63.0	100.0	91.4	96.3	72.4

Results of ST1-R2, however, show a very different story (Figure 7.3). This time around, section T08 and T10 are the ones where students start out with a better identification of the questions that share a common physics principle. Once again, T06 is the weaker section and has a similar number of students who are willing to change their answers. Most noteworthy, is the change we see in sections T07 and T09 where nearly 40% of the students are convinced to change their answers after the group discussions.

Table 7.3. Results of Sorting Task 1 assessment row 2 (ST1-R2), for five treatment sections.

Row 2	T06	T07	T08	T09	T10
D/E first	37.0	35.1	68.6	44.4	69.0
D/E review	18.5	37.8	11.4	40.7	20.7
total right	55.6	73.0	80.0	85.2	89.7

Generally, ST1-R3 was easy for all sections (Figure 7.4), sectionsT07, T08 and T10, in particular. Surprisingly, section T06 also does well this time around.

Table 7.4. Results of Sorting Task 1 assessment row 3 (ST1-R3), for five treatment sections.

Row 3	T06	T07	T08	T09	T10
G/I first	81.5	70.3	91.4	77.8	100.0
G/I review	14.8	29.7	5.7.	18.5	0.0
total right	96.3	100.0	97.1	96.3	100.0

While it is clear that some of the distractors in the sets of problems were known to be the obvious surface features (e.g., similar appearance of the question), what we learned was that other students who get the wrong answers are focusing on another type of similarity. In these cases the students are seeing similarities in what we might call the "grammar" of the question, i.e., the "givens" and "unknowns."

For instance in ST1-R1 such students stated the following:

> *Both of these problems are similar because they both start at the same position at t=0s and both want to know the position at t=25s*
>
> *both problems deal with parameters varying along the x-axis. Therefore, one must use Δx to find the position of each problem (A & B)*
>
> *These two problems are similar because both involve acceleration and velocity during certain intervals of time*

And, in ST1-R2 they stated the following:

> *Both have constant acceleration. In D the y acceleration is 3 m/s2 and in E since it is a projectile only the y-value accelerates*
>
> *D and F asks for distance, it has i and j velocity. E is about projectile motion and an angle is known, he question is also different*
>
> *Both problems start with the same velocity and want to know the position after 10s. Both can be solved usng a velocity graph and finding the area under the curve.*

7.4.1.1 Findings for Sorting Task #1 (ST.1) Kinematics

These findings are important because it tells us something about why students might not be able to transfer their learning. Unfortunately, they may be taking away the wrong message from instruction. They become used to applying a set of rules and procedures in a static fashion and without understanding that these are simplified "rules of thumb" that are not always appropriate. We might consider such action as a "didactic epistemic" approach. Instead, the sorting task requires students to break from that view and adopt a more sophisticated epistemic stance, what we might consider as an "adaptive epistemic" approach. The results of CCT1 and CCT2 shows that students in this study can begin to use an adaptive epistemic approach to their problem solving. And, that the work they do in with the DALITE system, which includes the active learning environment,

7.4.2 Analysis of Sorting Task 2 (ST.2) Dynamics

The ST2-Dynamics assessment provided some interesting insight into how students identify important ideas within physics problems. Recall that Sorting Task 2 (ST.2) also had three rows of problems. In this instance, problems A & B look very similar, based on a quick inspection of their diagrams (horizontal force, diagonal black string and black diagonal ramp), while problem C looks completely different, with its picture of a man mowing his lawn. In this instance, problems A and C are the correct choices.

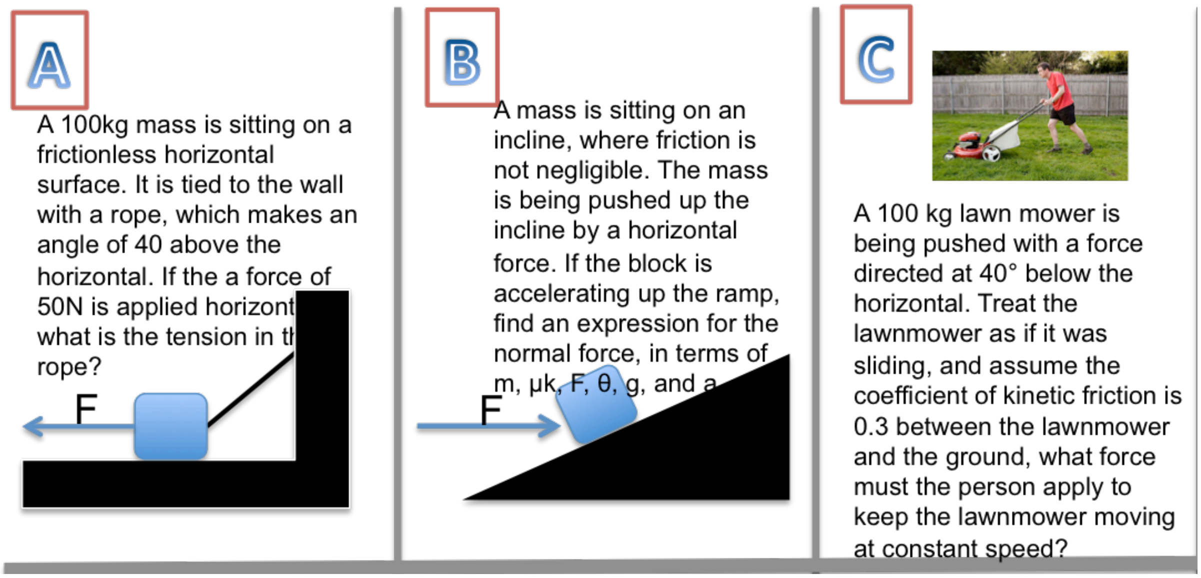


Figure 7.8. Sorting Task 2 – Dynamics, row 1.

The second row ST2-R2, all three problems involve using Newton's Second Law on a curved surface (where the acceleration is towards the center of the curve). However problems E and F are the best matches because both problems involve finding the centripetal force and radial acceleration.

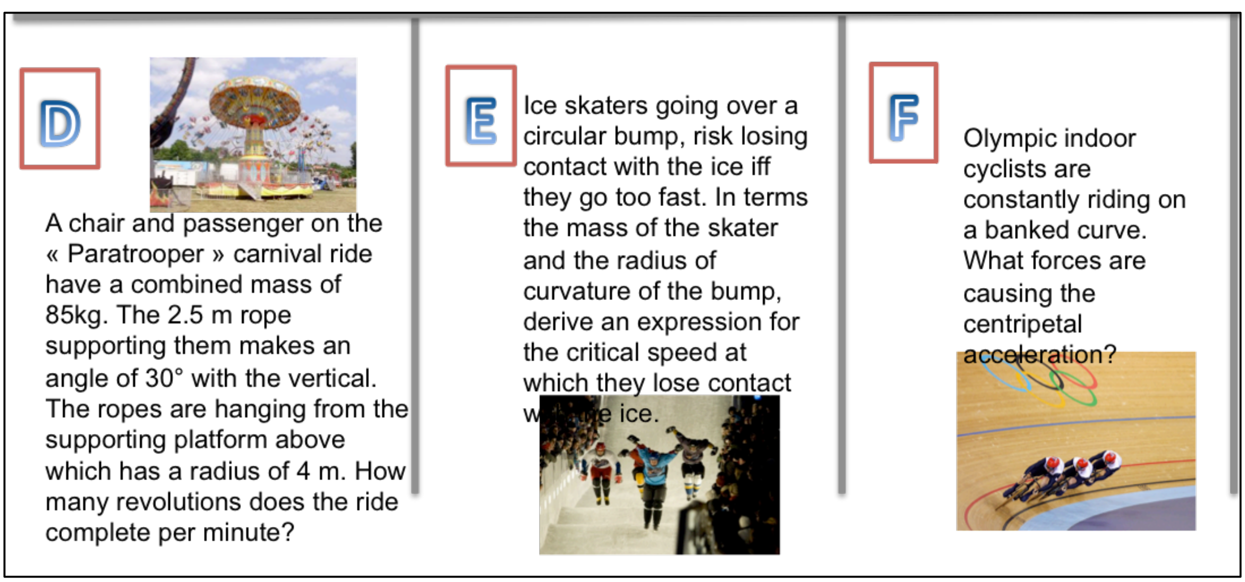


Figure 7.9. Sorting Task 2 – Dynamics, row 2.

The third row ST2-R3, H and I are the best matched when it comes to the physics because they both are based on Newton's Third Law.

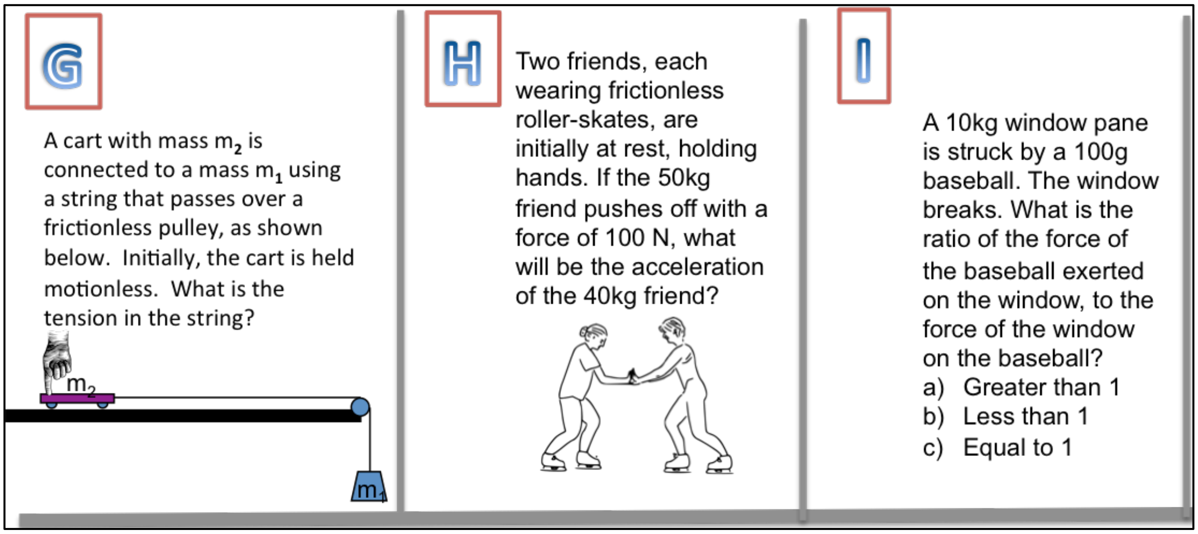


Figure 7.10. Sorting Task 2 – Dynamics, row 3.

7.4.2 Results of Sorting Task 2 (ST.2) Dynamics

Results of the sorting task, ST2-R1 is presented below (Tables 7.5 – 7.7). We note from these results that section T10 did particularly well on all three sets of problems. In fact, section T10 is the only one that managed to get ST2-R1 right on the first try with over 60% of students (Table 7.5). On the other end, sections T06 and T09 did particularly poorly. This trend is consistent for T06 but not for T09. And, once again we are seeing that T07 is more flexible in changing their answers after discussion.

Table 7.5. Results of Sorting Task 2 assessment row 1 (ST2-R1), for five treatment sections.

Row 1	T06	T07	T08	T09	T10
A/C first	8.0	35.1	26.8	17.2	64.3
A/C review	4.0	32.4	19.5	3.4	0.0
total right	12.0	67.6	46.3	20.7	64.3

The trend of flexibility and willingness to change holds up with ST2-R2 (Table 7.6) where 27% are convinced by 19%. This time T10 students also demonstrate a high degree of change, relative to the numbers who got the answer right at the first.

Table 7.6. Results of Sorting Task 2 assessment row 1 (ST2-R2), for five treatment sections.

Row 2	T06	T07	T08	T09	T10
E/F first	48.0	18.9	31.7	20.7	39.3
E/F review	8.0	27.0	17.1	13.8	21.4
total right	56.0	45.9	48.8	34.5	60.7

Once again, the third set of problems is relatively easy for all sections (Table 7.7). And, with the exception of T06 all sections are above 75% in the first answer phase.

Table 7.7. Results of Sorting Task 2 assessment row 1 (ST2-R3), for five treatment sections.

Row 3	T06	T07	T08	T09	T10
H/I first	64.0	97.3	82.9	93.1	75.0
H/I review	8.0	2.7	12.2	0.0	14.3
total right	72.0	100.0	95.1	93.1	89.3

These results we perplexing and required that we take a closer look at the detail of the results and what else might be happening. Below we elaborate on not only the correct choice of answers but also the incorrect ones as well and present a mini-case study featuring T10.

7.4.3 Deeper analysis of ST2-Dynamics taken as a whole

Recall that in ST2-R1 the desired answer, that problems A and C are most similar, was designed around the idea that both cases exhibit *equilibrium* of forces (static equilibrium in case A, dynamic equilibrium in case C). In contrast, the object in case B is said to be accelerating and cannot be in equilibrium. We note, however, that very few (17 of the 157 responses = 10.8%) correctly reasoned that the idea equilibrium A and C as most similar because of equilibrium.

The primary "surface" distractor seems to have been noting that situations B and C both have friction acting, whereas friction is absent in case A. Considering the data in tables (7.8 and 7.9), we note that about half (79/157 = 50.3%) of the responding students identified "friction" as a primary common feature. Some interesting patterns emerge from the data.

Table 7.8: Response patterns of students not identifying friction as a primary differentiator.

	Response: (pre-discussion) -> (post-discussion)								
	AB->AB	AB->AC	AB->BC	AC->AB	AC->AC	AC->BC	BC->AB	BC->AC	BC->BC
CW09	1	0	0	0	5	0	0	0	5
CW10	0	0	0	0	15	0	0	0	4
KL	0	0	0	0	2	0	0	0	6
MD	1	1	0	0	11	3	0	0	5
NL	0	2	0	0	12	0	0	4	1
Total	2	3	0	0	45	3	0	4	21

Table 7.9: Response patterns of students identifying friction as a primary differentiator

	Response: (pre-discussion) -> (post-discussion)								
	AB->AB	AB->AC	AB->BC	AC->AB	AC->AC	AC->BC	BC->AB	BC->AC	BC->BC
CW09	0	0	0	0	0	0	0	1	17
CW10	0	0	0	0	3	0	0	0	6
KL	0	0	0	0	0	1	0	0	16
MD	0	0	1	0	0	1	0	2	16
NL	0	0	0	0	1	0	0	4	10
Total	0	0	1	0	4	2	0	7	65

Students identifying A and C as similar

The results show that 54 of the 157 students chose situations A and C as being most similar to each other. Of these, 17 (10.8% of total, 31.5% of students choosing A & C) made explicit reference to equilibrium ideas. The rest focused on “surface features” such as:

- the objects in both situations are in contact with a horizontal surface
- in both cases, there are forces directed at an angle of 40° to the horizontal
- the free-body diagrams for these situations are similar

Interestingly, however, of these 17 cases, 16 (94.1%, 10.2% of total) chose A and C as most similar to each other. Peer discussion changed these numbers somewhat, with 63 choosing A and C as most similar, of whom 25 used the concept of equilibrium as a primary differentiator. Of the 79 students identifying friction as a primary differentiating factor, only 6 initially chose cases A and C as similar.

Students identifying B and C as similar

In this instance, the results show that 97 of the 157 students (61.8%) initially chose situations B and C as most similar. Of these, 72 (74.2% of the 97, 45.9% of total) made explicit mention of friction as a primary differentiator in their rationales.

Equilibrium (deep structure) vs. friction (surface feature).

Of the 79 students identifying friction as a primary feature, 72 (91.1%) chose case B and C as being most alike (reasonable, as the situation represented by A was declared to be frictionless) in their first responses (before discussion). After discussion with their peers, 68 of the 79 (86.1%) students retained the choice B & C.

Of the 78 students *not* identifying friction as primary feature, 48 (61.5%) initially chose situations A and C as being most similar. That this response agrees with the intended answer should be treated carefully: few (17) of these rationales explicitly mentioned equilibrium (zero acceleration) as a common feature.

Mini-case study of section T10

As already noted in different contexts throughout this report, the CW10 group stood out from the others in its character. Of the 28 students, 9 (32.1%) identified friction as the important differentiator. Compared with the other groups (79/157 = 50.3%), this represents a significantly (p=0.033) lower frequency. Moreover, the fraction of students choosing A and C as most similar (18/29 = 62.1% for CW10 vs 52/157 = 34.4% overall) is significantly (p=0.00024) higher.

Of the 19 students not identifying friction in their rationales, 15 of chose A and C as similar, and 4 chose B and C. This rate was not significantly different than the rest of the groups (p=0.073). Interestingly, unlike the other groups, no students changed their answers after the discussion with peers.

7.4.4 Conclusions

The deep feature we hoped would manifest in the students writing was lower than expected. The primary distractor was that friction is acting in cases B and C, but not in A. Students who focused attention on this (79/157 = 50.3%) were strongly led to choosing B and C as the most similar problems. The majority (61.5%, 30.6% of total) of students who didn't focus on friction chose A and C as the most similar.

7.5 Mid-Term tests across all six sections

Averaging the three tests, the results across sections shows that students in T06, T07, T08 and T09 were similar. However, there is small difference between these and sections T10 and the comparison section (T.comp).

Table 7.10. Results of three mid-term tests for five DALITE sections and Comparison section.

Section		CCT1	CCT2	CCT3	Average
T6	mean	64.65	63.39	66.22	64.75
	SD	15.62	12.80	19.61	
T7	mean	65.45	64.52	68.43	66.13
	SD	16.61	15.48	18.09	
T8	mean	64.38	69.39	64.68	66.15
	SD	17.46	23.18	19.87	
T9	mean	71.43	66.24	71.09	69.59
	SD	12.64	15.01	11.84	
T10	mean	75.90	74.54	72.45	74.30
	SD	12.19	14.05	12.12	
T.comp	mean	81.8	70.4	65.8	72.67
	SD	14.25	19.20	21.84	

Interestingly, looking at the Table 7.10, the within section trends across the three tests show minor fluctuations between CCT1, CCT2 and CCT3 for the DALITE sections. However, the Comparison section shows a stead decline in the students' averages. Additionally, the standard deviations for the Comparison section also increase across time. By contrast, results for section T10 show little change with time and a consistent standard distribution suggesting that students were maintaining their level of knowledge across the three sections. And, the distribution between high and low students remained consistent. In other words, generally speaking, there were fewer differences between students.

Figures 7.11 and 7.12 graphically illustrate these differences between sections T10 and T.comp on both Common Test #2 and #3, respectively. They show this increasing difference between students at the lower end of the spectrum. And Figure 7.6 is even more striking with the difference between sections beginning to show up in the middle ranged students.

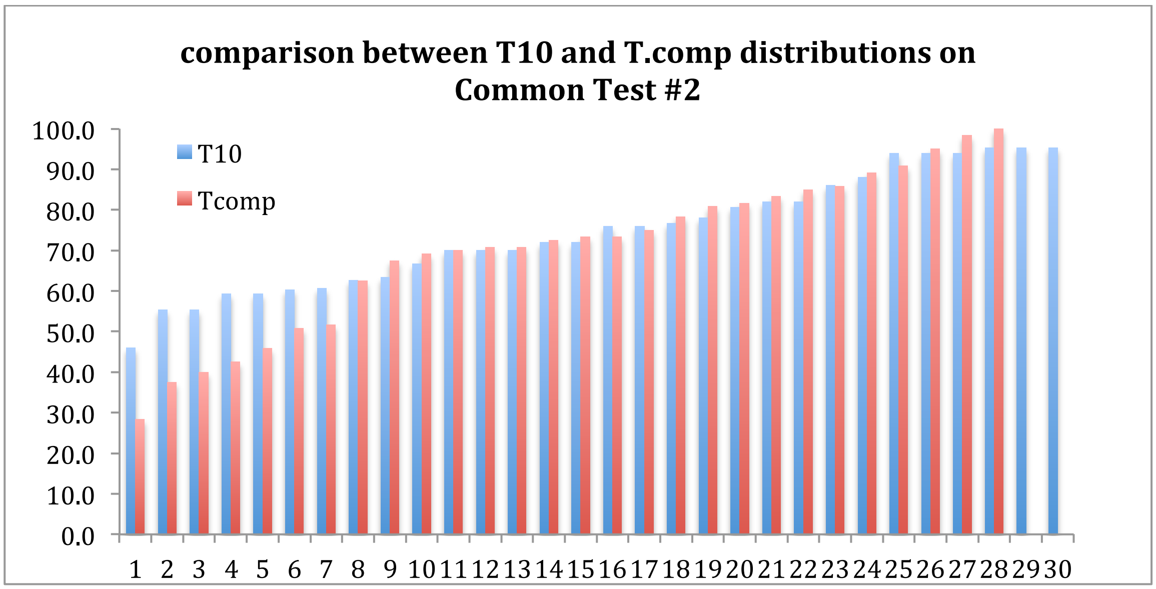


Figure 7.11. Comparison of results on Common Test#2 for sections T10 and T.comp by student.

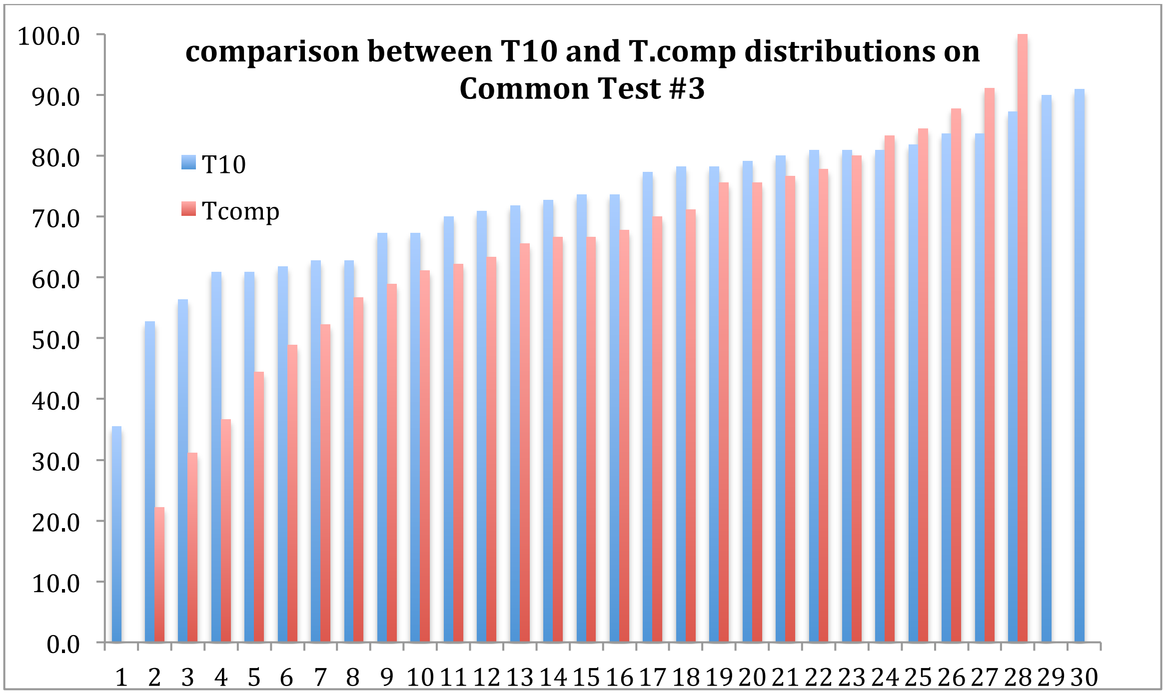


Figure 7.12. Comparison of results on Common Test#3 for sections T10 and T.comp by student.

These differences lead us to ask the question, what might be happening differently for the lower end students in T10 compared to T.comp? With that as our goal we set about to conduct a fined-grained case study analysis of the rationales written by these sets of students and compared them for common test questions #2. A description of this study follows.

7.5 Summary

It is clear that uncovering conceptual understanding from written explanations on a test is difficult. However, we have shown that there are some ways of analyzing these data that will begin to tell us about the process of conceptual understanding and possible trajectories for learning in this complex topic. What we are beginning to see from these analyses is that different ways of supporting conceptual learning may benefit students, as a whole class, differently. Some ways may be benefiting the larger group while others benefit those who are at the top end more so than those at the bottom. We will return to this discussion about possible impact of the class and the way groups function to support the whole in a other chapters.

CHAPTER 8
CONCEPT MAPPING ACTIVITY

This chapter addresses the research questions: *What is the impact of the DALITE web-based activity and the extended system of conceptual mapping*? *Specifically, Do students who use DALITE and concept mapping demonstrate improvements in their understanding of the course content and epistemic change*?

The Concept mapping activities were designed to mirror the Tagging activity, which took place in sections T09 and T10. The concept mapping activity is a deductive process (big ideas to smaller phenomena), whereas the tagging activity was designed as an inductive process (small to big). The experiment focused on the trajectory of student epistemology during the semester. Can Concept Mapping be used to document changes in student thinking while using DALITE?

8.1 Introduction and Theory

Practitioners use concept maps in the classroom because they can be used to measure prior knowledge, and as a tool to reflection on how concepts are linked metacognitively. In addition they can promote self-regulation by the student and provide feedback to the teacher about student learning in a meaningful way.A concept map shows the relationships among concepts, with linking phrases or rationales. A concept map is a physical representation of the mental maps we carry with us and use to problem-solve (Wheeldon, 2009).

Concept maps have been used in pedagogy for almost 40 years since Joseph Novak(Novak, 1990) used concept maps as a means of representing the emerging science knowledge of students. They are used as a dual-purpose tool to increase meaningful learning as well as to represent the expert knowledge of individuals and teams.

Because concept maps are visual tools for organizing and expressing knowledge, they embody a portrait of how a person, or group, is thinking by linking concepts (Dalgarno, 2001). In the classroom they can be used to help students make the links necessary to deepen understanding, or as an assessment tool to highlight misconceptions and provide feedback to the teacher on the segmenting of knowledge.

One of the drawbacks of using concept maps in the classroom is that they are very difficult to analyze (Ruiz-Primo, 1996) in a meaningful way especially for physics. Every concept map is unique to the individual and often appears as a work of art, almost impossible for observers to integrate and assess without some reflection. This is because of varying content and concepts which are often multiply cross-linked giving rise to intricate structures.

Qualitative assessment of concept maps starts by looking at their structure and hierarchy. High scoring concept maps are rich, hierarchical, correct, and exhibit wide thinking outside the course. Low scoring concept maps are not hierarchical, disorganized, and shallow, with errors and misconceptions. Other, more quantitative assessment techniques include counting the number of links and cross-links(Ruiz-Primo, 1996). Yet other techniques evaluate the structure of the map as the key measure, without looking at the content (Huo, 2010). Although many of these techniques do give valuable information about student misconceptions they are difficult to interpret.

We have been using a particular approach for measuring differences in concept maps: the Pathfinder algorithm. Three different concept maps were completed by groups of students in Section 6. What is described here is the use of the Pathfinder network analysis as a tool for analyzing concept maps. This is done by have a restricted concept map, meaning the participants were provided with a standard set of nodes and linking phrases to organize into a map.

8.1.1 The Pathfinder Algorithm

Pathfinder networks are used in the study of expertise, knowledge acquisition, knowledge engineering, citation patterns, information retrieval, and data visualization (Schaneveldt 1989). The networks are potentially applicable to any problem addressed by network theory.

Several methods start by looking at the relative proximity of data points, or concepts in a concept map and can determine the structures of the underlying data organization. For instance, data clustering, multidimensional scaling and other methods based on graph theory. However, Pathfinder networks are calculated from data describing how close together two things are i.e. the proximity of one to the other: proximity data. Proximities can be obtained from similarities, correlations, distances, conditional probabilities, or any other measure of the relationships among entities. In the Pathfinder network, the entities correspond to the nodes of the generated network, and the links in the network are determined by the patterns of proximities.

In the work described here, the entities are concepts, and the proximities describe how concepts are linked in the map, but they can be anything with a pattern of relationships. The links in the network will be undirected if the proximities are symmetrical for every pair of entities. Symmetrical proximities mean that the order of the entities is not important, so the proximity of i and j is the same as the proximity of j and i for all pairs i,j. If the proximities are not symmetrical for every pair, the links are directed. Normally links in concept maps are directional, but for this work, the links were treated as being symmetric, because students did not consistently use directional linking.

The Pathfinder algorithm uses two parameters.

(1) The q parameter constrains the number of indirect proximities examined in generating the network, that is to say the different paths around a map to get from one particular node to the other. In a small concept map, there may only be one path through the map. For a large map with many cross-links (common in Physics concept maps), the number of possible paths may be large. The q parameter is an integer value between 2 and $n - 1$, inclusive where n is the number of nodes. Basically q is the number of degrees of freedom.

(2) The r parameter defines a value used for computing the distance of paths. The r parameter is a real number between 1 and *infinity*, inclusive.
In the literature, a network generated with particular values of q and r is called a PFnet(q, r). Both of the parameters have the effect of decreasing the number of links in the network as their values are increased. The network with the minimum number of links is obtained when $q = n - 1$ and $r = \infty$, i.e., PFnet($n - 1$, ∞).

Essentially, Pathfinder networks identify the shortest possible paths between any two nodes, for all possible nodes, that is to say it goes through a process to minimize the network. The PFnet($n-1, \infty$) will give the minimally connected network for the links defined by the proximity data if such a network exists.

The Pathfinder algorithm is conveniently available as a Java software available from Interlink Inc. To use this program, data has to be formatted in a particular way as described in the documentation files. These files are termed "proximity files" and are given the .prx suffix. The software can generate a "similarity index", which is a score of how similar two particular maps are in terms of the minimized networks. In this study, Pathfinder was used to calculate similarities between student maps and the teacher-expert concept map.

8.2 Methods

8.2.1 Participants and Classroom Implementation

The participants were students in Section 6 as described previously. The students worked collaboratively in groups of 2-4 to produce three sets of concept maps during the semester. The three concept maps were on 1D motion, Forces/Newton's Laws, and a course summary concept map at the end of the semester. The concept maps were given close to the end of instruction on each topic, in preparation for a class test or final exam. The concept maps are therefore negotiated representations of the group cognition. Each concept map activity took the groups about 45 minutes to complete on average.

Although students learn how to do concept maps very quickly, they are often confused the very first time. Therefore there was a pre-activity to introduce concept maps by having the class work together on a neutral concept map- 'What makes for great ice cream?'
For the concept map activities themselves, students used Cmap, a concept map creation software, which was loaded onto the computers in the classroom. Students started the activity by downloading an initial file which contained all the concept nodes that the students would use. Using the computers at their stations, and in some cases Smart Boards, students completed the maps and then submitted their electronic files through the class management platform.

The teacher of the class also did concept maps using the same list of nodes as the students, which was deemed the "expert map". In addition, for the 1D motion concept maps, six other guest experts (either physics teachers, or graduate students) created additional expert concept maps individually, either using Cmap or on paper. Any paper-based concept maps were transcribed into Cmap digital format.

Generation and Treatment of the Concept Maps

The participants generated the concept maps using Cmap (Cañas, 2004, see Figure 1 for an example), a concept mapping software. The nodes were preloaded into Cmap, and then the Cmap file was distributed to the participants. Participants were also given a list of recommended linking phrases, which they were encouraged to use in their maps (see Table 8.1).

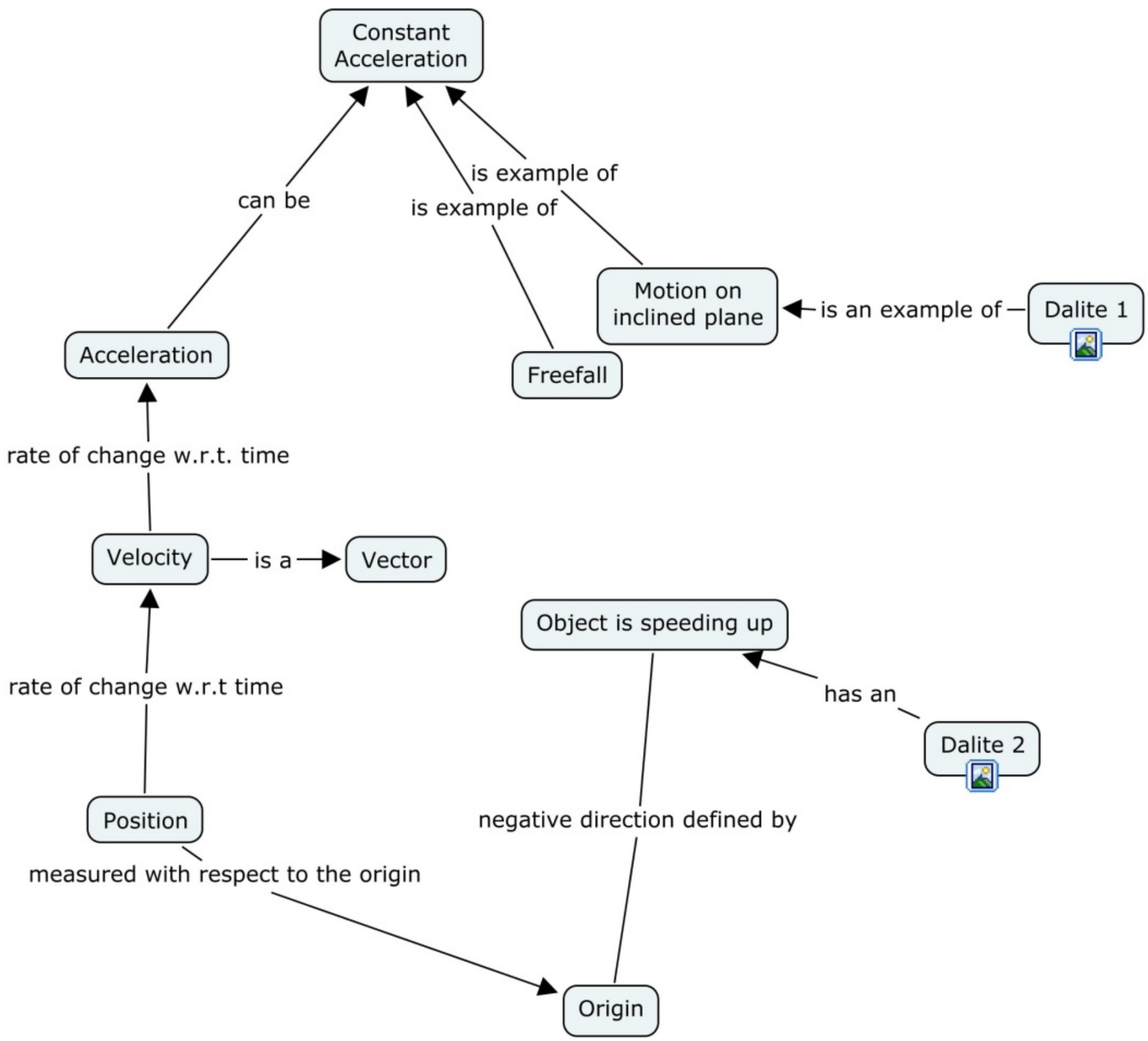


Figure 8.1. A typical 1D Motion concept map showing the clear link between velocity and acceleration.

The node names were chosen to try and match, as much as possible, the tags used in the tagging activity in section 10. Although participants were given a list of nodes, and this restriction is necessary to compare concept maps, they were allowed to either add nodes they felt were relevant, or exclude nodes they did not feel were relevant.

Table 8.1. Recommended nodes and linking phrases for the 1D Motion Concept Maps.

Recommended Nodes	Recommended Linking Phrases
Constant Acceleration Position Object is slowing Average Velocity Displacement Motion on inclined plane Freefall Average Acceleration Scalar DALITE 1 Velocity Distance Object is speeding up Origin Speed Instantaneous Acceleration Instantaneous Velocity DALITE 2 Acceleration Vector	difference of values gives sum of values gives the area under the time graph slope of time graph Rate of change w.r.t. time derivative of gravity is only force vel.&acc. are in opposite directions vel.&acc. are in same directions magnitude only magnitude and direction is a is an example of

The resulting Cmap files were collected and exported into a text file as a list of the propositions. In this file, each line reproduces one proposition, with a node name, the linking phrase and the target node. A Python script was written, which parsed the generated text files to createproximity lists, and a list of node names, as required by the Pathfinder software. The weights between nodes were given equal weighting. For this analysis, only the node names were analyzed because there was two much variation in the linking phrases the students used. Also, in two cases, students changed the node names, for instance from 'Kinematics' to 'KINEMATICS', or added central nodes with non-standard node names, which confused the analysis. These nodes were changed or removed without changing the structure of the rest of the maps themselves.

Table 8.1, showing the recommended nodes and linking phrases for the 1D Motion Concept Maps. Note that these were recommended and participants could add nodes of their choosing or decide to not use particular nodes. DALITE 1 and 2 refer to the DALITE conceptual questions the participants were asked to include in their concept maps.

data
similarity
52 nodes
0 decimal places
1 minimum value
10 maximum value
list
23 pairs

symmetric
12 1 10
18 14 10
22 3 10
22 21 10
24 4 10
27 13 10
31 29 10
33 25 10
40 25 10
44 9 10
44 17 10
44 19 10
47 1 10
47 20 10
48 3 10
48 18 10
48 24 10
48 27 10
48 31 10
48 40 10
48 44 10
50 1 10
50 29 10

Figure 8.2. An example of the contents of a Forces Concept Map proximity file ready for analysis by Pathfinder. The headers describe the type and structure of the data. The nodes are contained in a separate file as an ordered list, and simply referenced by their order number. For instance, the first link described here is a link between node #12 and node #1. For this study all links were given equal weight, 10, but this could be varied if required.

8.3 Results and Discussion

Although the student-group membership was supposed to remain constant throughout the semester, in practice there was some movement between groups and some students did not participate in all of the concept mapping activities. In fact, only six of eight groups had both consistent membership and complete concept maps, and the results focus on these data.

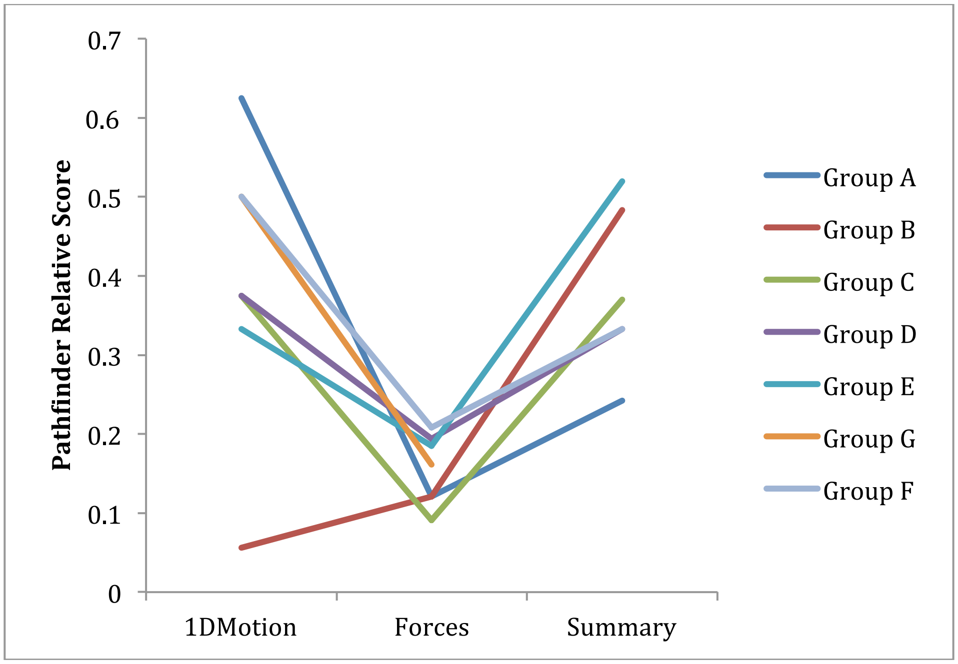


Figure 8.3. The similarity scores relative to the expert concept map, for each group, for the three concept maps produced during the semester.

Figure 8.3 shows the similarity data produced by Pathfinder for the concept maps produced by the class, relative to the expert concept map. The similarity data is fractional, with a score of 1 being equivalent to the expert concept map. The first observation is that the scores are all quite low, nothing above 0.625 compared with the expert. The averages for the class are: 0.38±0.17, 0.14±0.06 and 0.37±0.10 for the 1DMotion, Forces and Course Summary Concept Maps respectively. The guest expert maps (maps done by other experts) had an average of 0.35±0.18, compared with the teacher of this section, indicating that the absolute values, even by other experts are highly varied. Even experts rarely agree on what constitutes an appropriate concept map for a given subject, and the guest experts were more likely to exclude nodes they did not think were particularly relevant. This reduces their similarity pathfinder score, and highlights a weakness of the Pathfinder method: it does not tell *how* the maps are different, just that they are.

Rather, it is the *relative* ranking of the student maps that is the focus here. That is to say, how close are the student maps to their main source of physics epistemology: their own teacher?

Assessment data was available for these students including performing on the pre- and post-FCI, test scores, DALITE participation rates, DALITE Scores. Correlations between the Pathfinder similarity scores and these data were calculated. Because of the very small number of data, six points, only correlation coefficients above 0.85 (R^2=0.72) were considered significant.

8.3.1 Effect of the group strength

The assessment data is of individuals, but the concept maps are a group effort. Although a first approach to reconcile this is to calculate assessment averages for all the individuals in the group for all the available data, arguments can be made that the best elements of each student work to pull the group up, or that the group falls to the lowest common denominator. Analysis of the data reveals that by far the best correlations come from taking the *best* scores of each group for each assessment. For instance, the correlation coefficient between the *change in normalized concept map* score (see below) and performance in the final exam was 0.89 (considered significant) for the best score in the group, and only 0.64 for the group average (considered non-significant). In general, using average and lowest in the group scores increased scatter and decreased correlation coefficients.

For this reason only the best scores in the group for each assessment category were used. It should be noted that it was rare that the same person scored the highest in each category. Rather, different group members had different strengths that they could bring to the table, for instance, different conceptual or problem solving skills.

This is an important result: the maps are not equal to a map constructed by the "best" person in the group; but, negotiated maps are constructed by the sum of the best elements in the group. That is to say the *best contributions* from each individual pull the group up, instead of tending to the average or the lowest common denominator pulling the group down.

Each of the concept map activities resulted in an important or surprising result and they are described below, followed by an analysis of the concept map trajectories. Correlations between concept maps and the assessments not mentioned below were considered non-significant.

8.3.1.1 Course Summary Concept Maps and Success on the Final Exam

The Course Summary Concept Map activity was done in the final week of the semester, shortly before the final exam. The goal of the course summary concept map was to help students understand how all the components of the course fit together. Perhaps not surprisingly there is a correlation between this final concept map score and success in the final exam. In particular the map scores correlated best with the normalized (z) scores on the problem-solving component of the final exam. These z normalized scores are calculated using the entire class, including the students who were not in the six included groups, so the correlations for the raw grades are not the same (they are lower). In the final exam there was a correlation between success on the multiple choice portion of the exam and the problem solving portion of the exam, so that is also true for the concept map data, but the correlation is lower.

Correlation is not causation. Although it probably did help students' cognition to organize the course concepts into a map, it is unlikely that this activity itself led to success on the final exam. The concept map was acting as a diagnostic, giving a snapshot of the students' mental model. A better mental model is an indicator of overall understanding.

What is surprising is that the correlation is quite strong. The deductive reasoning that concept maps bring to the table is a useful activity for students to engage, and provides real information to the teacher about student cognition.

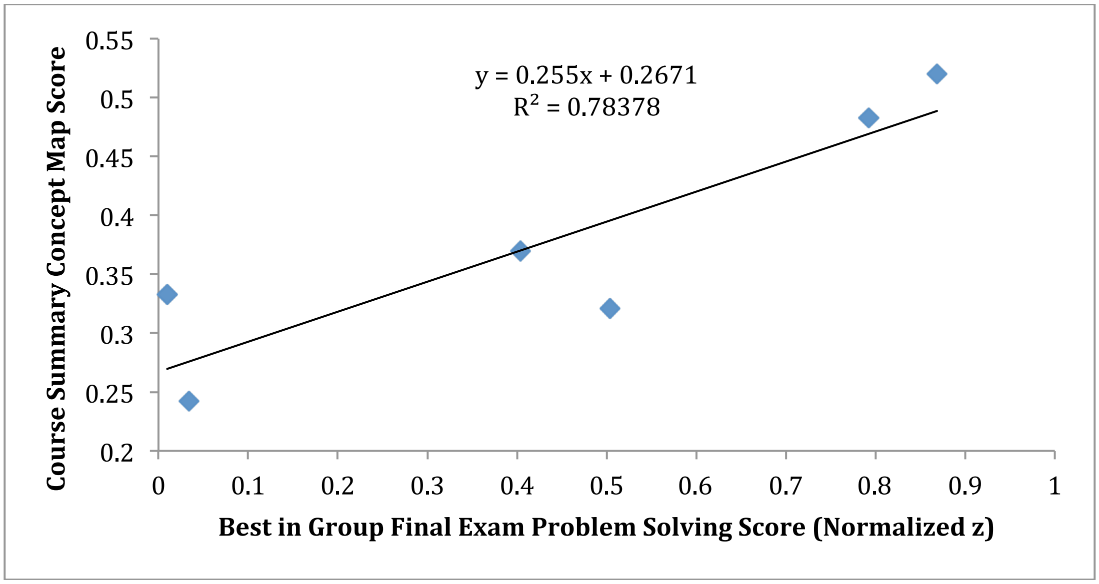


Figure 8.4. Showing the correlation between similarity scores on the final course summary concept map and normalized final exam problem solving scores.

8.3.1.2 Forces Concept Map and Wrong-to-Right Answers in DALITE

The Forces and Newton's Laws Concept Map had sharply decreased scores relative to the other mapping activities. Students clearly found the subject difficult and disjointed. In particular students had difficulty correctly connecting Newton's Laws together, and largely treated then as independent concepts. Because of this, the Forces Concept Map may be the most discriminating activity that the students participated in.

A most surprising result is the correlation between scores on the forces concept map and the rate at which students changed their DALITE answers from wrong to right. This is shown in Figure 8.5. The DALITE wrong-to-right is described as the fraction of attempted questions that the students answered (i.e. normalized to the number attempted).

This parameter, wrong-to-right, is perhaps the most important of the DALITE data. Students start by not knowing the answer, and by engaging with the process of comparing and reflecting on other student rationales, they *change* to the correct answer. This is perhaps the golden key of active learning education: conceptual change through engagement.

Although students may be able to cheat DALITE by always choosing the incorrect answer, no student consistently did this; therefore it was not consciously done.

Nor is this an example of the best prepared and performing students doing well with whatever method is thrown at them. Those students would presumably start correct and stay correct, and there are examples of this in the class. However, there were no other significant correlations with the DALITE data, for instance with initial- or final-Right scores.

Furthermore, the concept map activity was performed after the DALITE assignments. This could be interpreted that what this DALITE parameter could be measuring, *engagement for change*, is an indicator of real conceptual understanding, which is demonstrated by the relative scores in the subsequent concept maps.

It should be noted that students did more dynamics DALITE questions than any other topic, so any effect of DALITE should be strongest for this subject, and explains why correlations with the other concept maps were not significant.

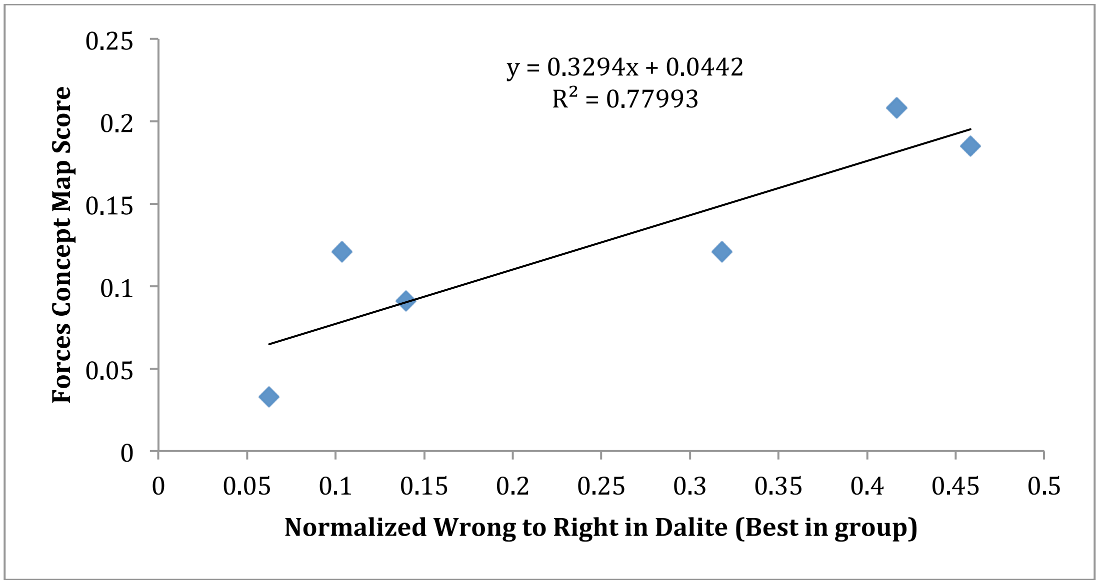


Figure 8.5. Showing the correlation between the Forces Concept Map Scores and fraction of questions that went from wrong to right in DALITE.

1D Motion Concept Maps and Performance on the Final Exam

The 1D Motion Concept Map was done very early in the course (week 3) before students had had much time to really assimilate the concepts. They were largely relying on their previous high school learning, which proved successful for some students, because some of the scores were the highest for any of the concept map activities. Even so, common high school misconceptions showed up on all the 1D motion concept maps. For instance: a tendency to classify parameters as scalars or vectors, something that experts see largely as a surface feature; or an inability to recognize the link between velocity and acceleration. This last observation is something that experts find critical, and all of the expert concept maps stressed this link.

Figure 8.6 shows the correlation between scores on the 1D Motion Concept Map and the best in group final exam performance on the multiple choice questions on the final exam (normalized z score for group). There is a strong *negative* correlation. That is to say that students who scored highest in the group on the final multiple-choice section, in fact belonged to a group who performed poorly on this concept map activity. As before, students who did well on the multiple choice section of the exam, also did well on the problem solving component, but the correlation between 1D Motion Score and problem solving was less significant.

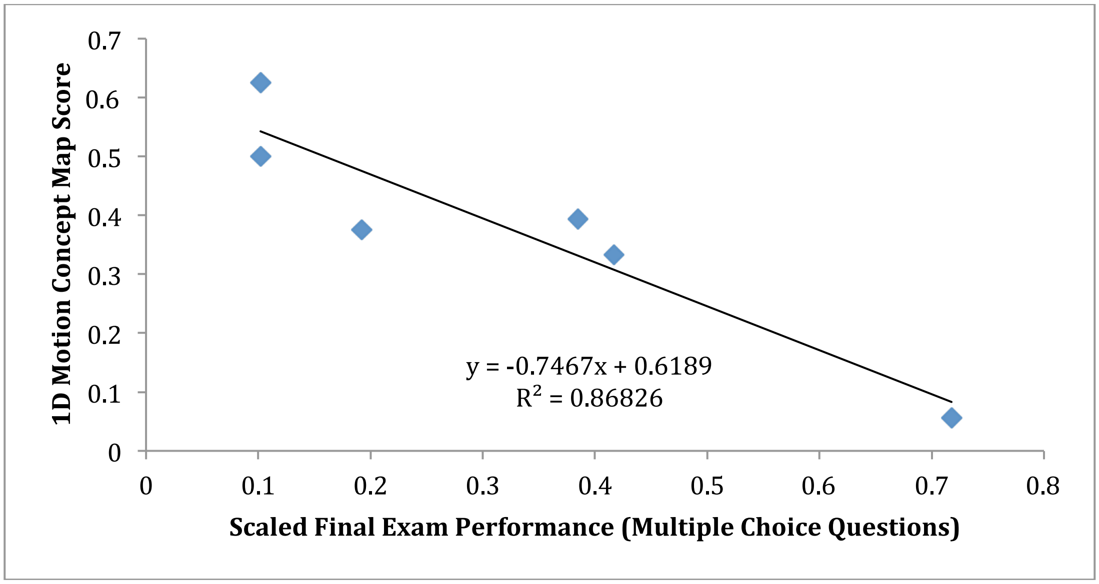


Figure 8.6. Showing the correlation between scores on the 1D Concept Map and normalized final exam performance on the multiple-choice questions (best in group). Here again, the best score in the group gave the best correlation.

This data is quite puzzling to say the least. Initially apparently-strong and well-prepared students do not follow-through to the end of the course. But conversely, initially weak maps correlated with success in the course.

One might expect the first concept map activity, early in the semester, to give less quality data because the students have to learn how to do concept maps themselves, an activity that is complex and has high cognitive load, and this may have indeed influenced this particular result. However, this should be taken together with the comparison of the Course Summary Map with the final exam scores, where there was a positive significant correlation (see Figure 4). Assuming the negative correlation at the beginning of the semester is not an artefact, this is a signal of significant change in conceptual thinking for some students during the semester.

This was a relatively weak group, with only a couple of students who consistently performed well throughout the semester, which makes this concept of change particularly attractive.

An hypothesis is that, only students who were open to *conceptual change* had greater cognitive growth, and were able to overtake students who were not able to change, who were not able to break from their previous misconceptions and training.

8.3.2 Conceptual Change during the Semester: Trajectories of Concept Maps

To investigate this notion of conceptual change during the semester, an analysis was done of how the concept maps changed during the semester. Clearly each activity very different concept maps and the average scores for each activity are inconsistent, for example the Forces concept map scores are much lower than the others.

So, to be able to compare data, normalized z scores ($z = \frac{(x-\bar{x})}{\sigma_x}$) were calculated for each data point i.e. scores were converted into scores relative to the average and distribution of that particular data set, the number of standard deviations from the mean.

Figure 8.7 shows how the groups changed relative to their peers over the three concept mapping activities. Basically, two groups improved over the semester, two groups stayed approximately the same, and two groups decreased.

What is particularly interesting are two groups that were followed more closely with observational and interview data. One of the groups, Group A, started strong, but decreased relative to the class. The other group, Group B, started weak, but came through in the finish.

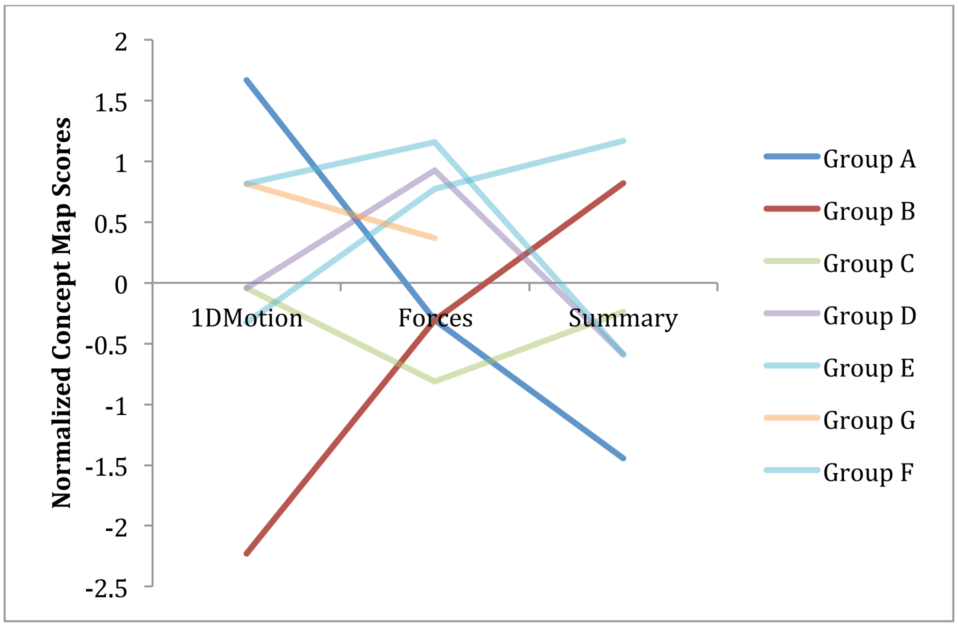


Figure 8.7. Showing how the concept maps scores changed over the semester. The values are normalized z scores. The data in bold, corresponding to Group A and B, are the groups for which there is interview and observational data.

To try and systemize this trajectory of the groups over the semester, the "slopes" of the trajectories were calculated. The slopes are basically the averages of the changes, one concept map to the next, for each group. The mean deviations of these values are relatively big, as shown in Table 8.2, so this exercise is more to give an idea of the changes rather than rigorously significant data.

Figure 8.8 shows this data in graph format, as a function of best in group performance on the final exam. This graph looks quite similar in form to Figure 8.4, which shows the correlation between the Course summary concept map and the final exam data.

Table 8.2. How the groups changed relative to the class.

	Average Change	Mean Deviation
Group A	-1.56	0.42
Group B	1.53	0.40
Group C	0.75	0.35
Group D	-0.10	0.67
Group E	-0.70	1.05
Group F	-0.39	1.49

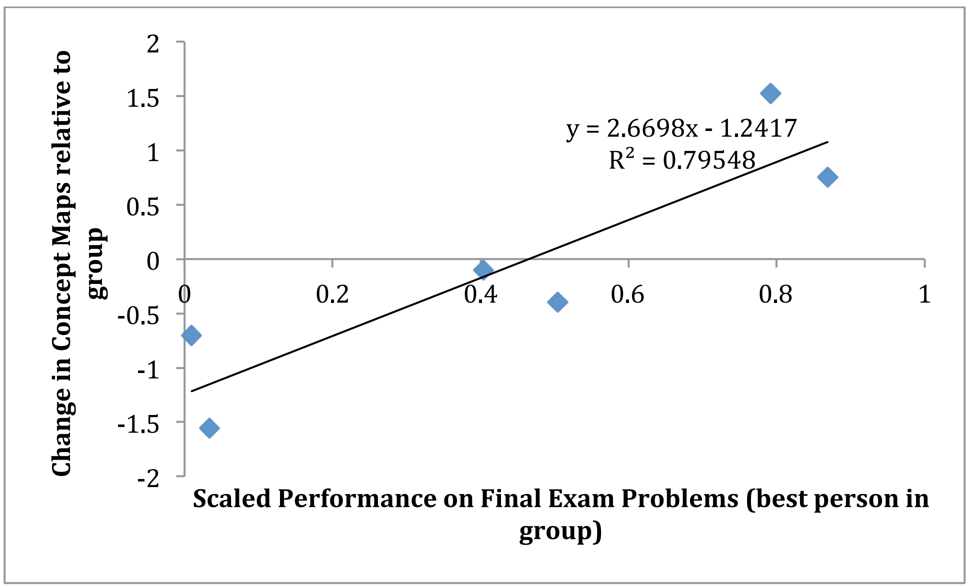


Figure 8.8. Showing the average Change in Concept Map Score, for each group correlated with the normalized performance on the final exam problems.

8.4 Conclusion

Pathfinder can be used as an effective tool for analyzing concept maps in terms of how similar they are to a particular standard, the teacher concept map. This is particularly important given the often nebulous and unique character of concept maps.

Using this similarity data and comparing with other assessment data from the class, important associations could be made. The first is that the concept map closest to the final exam, the course summary concept map, was an indicator for success on the final exam. This is further evidence that successful concept maps are indicators of wider knowledge assimilation.

The second, and perhaps most important conclusion is that conceptual change is an important indicator for success.

There are two points where change is recognized as a parameter: Students who were more likely to change from wrong to right answers in DALITE, an indicator of engaged change, had better Forces concept maps; and groups who had improved concept maps over the semester had better success on the final exam.

For the practitioner, concept mapping, which is already a very valuable tool in the classroom, can now be leveraged for even better and more efficient learning.

8.5 References

Cañas, A. J., Hill, G., Carff, R., Suri, N., Lott, J., Eskridge, T.,& Carvajal, R. (2004). CmapTools: A knowledge modeling and sharing environment. In Concept maps: Theory, methodology, technology. *Proceedings of the first international conference on concept mapping* (Vol. 1, pp. 125-133).

Dalgarno, B. (2001), Interpretations of constructivism and consequences for Computer Assisted Learning. *British Journal of Educational Technology*, 32: pp.183–194.

Hao, J-X., Kwok R., Lau R., Yu, Y.A. (2010) Predicting problem-solving performance with concept maps: An information-theoretic approach, *Decision Support Systems*, Volume 48(4), pp. 613-621,

Ruiz-Primo, M., & Shavelson, R. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33 (6) pp. 569-600.

Shaffer, D., Hatfield D., Svarovsky G., Nash P., Nulty A., Bagley E., Frank K., Rupp A., Mislevy R. (2009). Epistemic Network Analysis: A Prototype for 21st-Century Assessment of Learning. *International Journal of Learning and Media* 1(2) pp33-53.

Novak, J.D., & Wandersee, J.D., (Eds.), (1990). Perspectives on concept mapping. *Journal of Research in Science Teaching*, 20 (10); Special Issue.

Wheeldon, J. P., &Faubert, J. (2009). Framing experience: concept maps, mind maps, and data collection in qualitative research. *International Journal of Qualitative Methods*, 8(3), pp52-67.

Schvaneveldt R.W, Durso F.T. & Dearholt (1989) Network Structures in Proximity Data *The Psychology of Learning and Motivation* Vol 24 pp249-284.

CHAPTER 9
DALITE USE AND RATIONALES

This chapter addresses the research question: *What can we learn about the DALITE system from how it was used? Specifically, what can the rationales produced tell us about its value as a pedagogical tool? How did rationales differ between sections, between questions, between course units, if any? What might be learned because of this use and these differences?*

9.1 Results of the DALITE implementation

DALITE was implemented in five sections (T06, 07, 08, 09 and 10) of the first year college course in Mechanics – Physics NYA. Sections sizes ranged between minimums of 30 students to a maximum of 41 making for a total of 168 first year science students. Seventy-seven questions, from the DALITE database of 120 questions, were selected for this experiment. Over 80% of these overlapped between our five treatment sections. Topics covered in these questions reflect the three units of the Mechanics course – Kinematics (1DKin and 2DKin), Dynamics (LinDyn) and Conservation principles (Momentum and Energy). DALITE questions were arranged in groupings of three to five questions per assignment. These assignments were distributed as either pre-class or post-class homework. In each case students were required to answer the question and produce rationales according to the DALITE script described in the earlier chapter. A total of 7182 student-generated rationales were produced. The distributions are described in Table 9.1.

Table 9.1: Descriptive statistics on the rationales generated by students in the five sections, across the 3 colleges.

Sections	T06	T07	T08	T09	T10
# of students	n=30	n=41	n=36	n=31	n=30
# of questions assigned	48	50	48	66	66
completion rates	0.75	0.80	0.70	0.78	0.88
average # quest. answered	36	40	34	51	58
SD	11.62	11.65	14.77	10.96	12.24
Mode	48	50	48	58	66
Median	39	45	36	56	63
Total questions answered	1081	1637	1206	1235	1678

Looking first at the total number of questions answered between sections we see that section T06 had the lowest number with 1081 questions. Meanwhile, sections T07 and T10 were high with 1637 and 1678 questions answered, respectively. However, it is more important to note the average number of questions answered per student. Looking at these numbers we see that sections T09 and T10 are at the high end with 51 and 58 questions, respectively. Whereas, sections T06 and T08 had 36 and 34 questions answered per student, respectively.

There were also variations in the completion rates range with a low of 70% for section T08 and a high of 88% for section T10. Suffice to say that the main difference is 13-18 point spread between the low (T06 & T08) and the high (T10). Investigating these differences more closely we see that in T10 70% of students completed 90% or higher of DALITE questions. Meanwhile in T08 only 42% of students completed 90% or higher of DALITE questions. And, one third (31%) completing less than 50% of the 48 questions.

Looking at the mode, we see that of these five sections, T10 answered the most questions per student with a high of 66, which was also the maximum number assigned; with the next closest being T09 with a mode of 58 questions per student. Note that the same teacher taught these two sections. This teacher is the most advanced at active learning pedagogy – we will refer to him as DALITE active learning high expertise (D_{High}). D_{High} not only was able to assign more questions but also had a higher DALITE participation rate with his students. On the other end, we see that two of the other sections had lower modes, T06 and T08, each with 48 questions per student. These teachers both had moderate experiences with DALITE active learning but taught in different settings – we consider them $D_{Low\ Active}$ and $D_{low\ Non\text{-}Active}$, respectively.

Later we will discuss these differences more closely. In particular, we will look at how the make up of the class and the context might influence how DALITE is used and the benefits gained. Next we continue with the general results of the DALITE implementation.

9.2 Student performance on DALITE questions

How did students perform on the DALITE questions? In general, sections T09 and T10 answered more questions correct because they completed more questions compared to sections T06, 07 and 08. Their rate of success on these questions, however, is average. The only section that is clearly different is T06, the same section that also had lower FCI results (see chapter 6).

Table 9.2: Descriptive statistics on the DALITE use and student performance.

		# q's correct	success rate	W-->R	R-->W	Wd-->Wd
T06	Mean	20.27	0.55	5.13	3.63	2.40
	SD	7.61	0.15	5.48	2.47	1.85
T07	Mean	26.80	0.67	3.54	2.27	1.71
	SD	9.21	0.11	2.35	2.19	1.40
T08	Mean	24.03	0.70	2.31	1.25	0.47
	SD	11.85	0.14	2.08	2.02	1.18
T09	Mean	31.77	0.62	6.81	3.35	1.39
	SD	8.37	0.10	4.13	2.68	1.76
T10	Mean	37.33	0.65	7.97	4.13	1.87
	SD	14.90	0.15	5.67	3.57	2.29

9.2.1 Success rates for the DALITE unit segments

The results of the first analysis lead to a sub-question: *Are there differences between the content units*? To answer this question we look at the results for each of the three unit segments – Kinematics, Dynamics and Conservation.

9.2.1.1 Kinematics principles

Number of correctly answered Kinematic DALITE questions suggests three difficulty levels. Of the 27 Kinematics questions, 10 produced under 50% correct answers, 9 were between 51-79%, and 8 were 80% and above, suggesting three levels of difficult ranging from tough, moderate and easy questions, respectively.

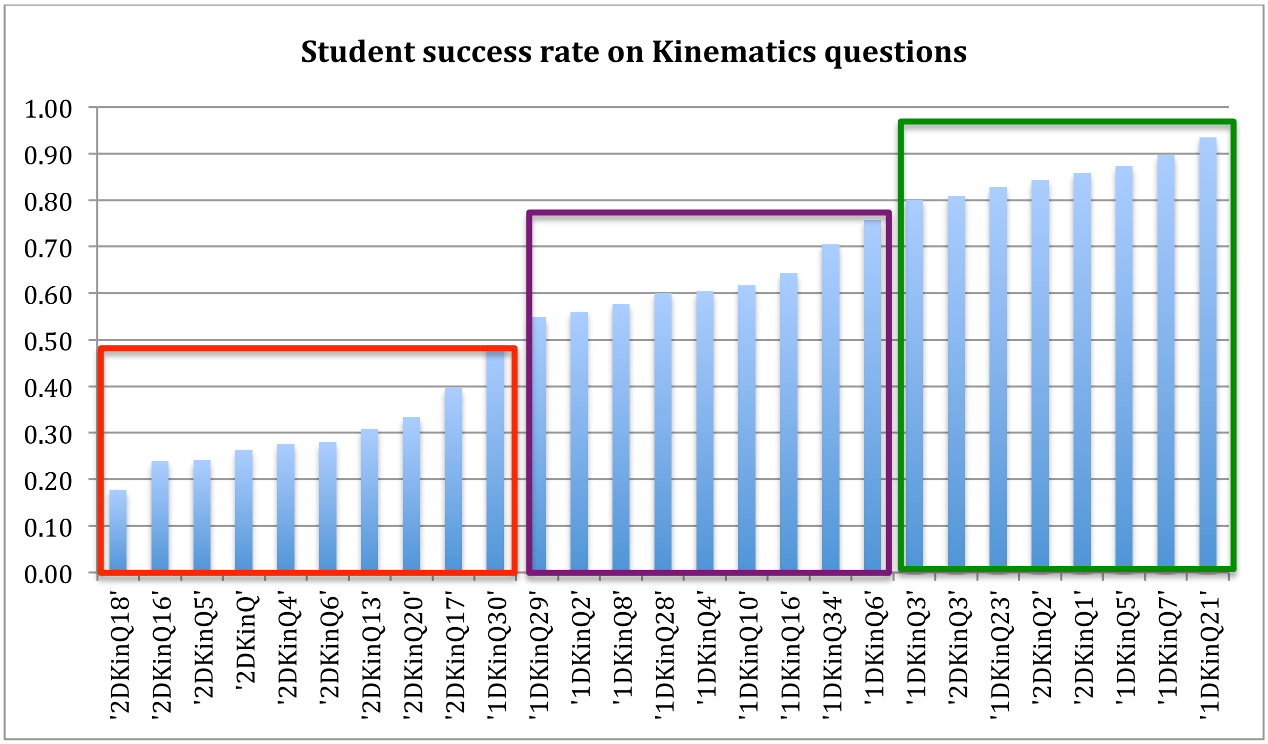


Figure 9.1. Student success rate on Kinematics questions.

What is most important about these results is that with DALITE rationales, we finally have a way to examine whether it is the wording of the questions or the actual concepts, both, or perhaps even something else that causes students to find these questions difficult or easy. We will continue this investigation later on.

9.2.1.2 Dynamics principles

There were also 27 DALITE questions on the topic of Dynamics. This time there are only 5 questions with a success rate of under 50%, 15 between 51-79%, and 7 questions with

80% and above. These results raise the questions: (1) do students find the Dynamics (LinDyn) concepts easier? (2) are the LinDyn questions themselves easier? (3) was it the timing of the questions? Did the instructors give more of these questions as post-instruction compared to the Kinematics assignments? Or, (4) are students becoming familiar with the format of the activity and doing more preparation?

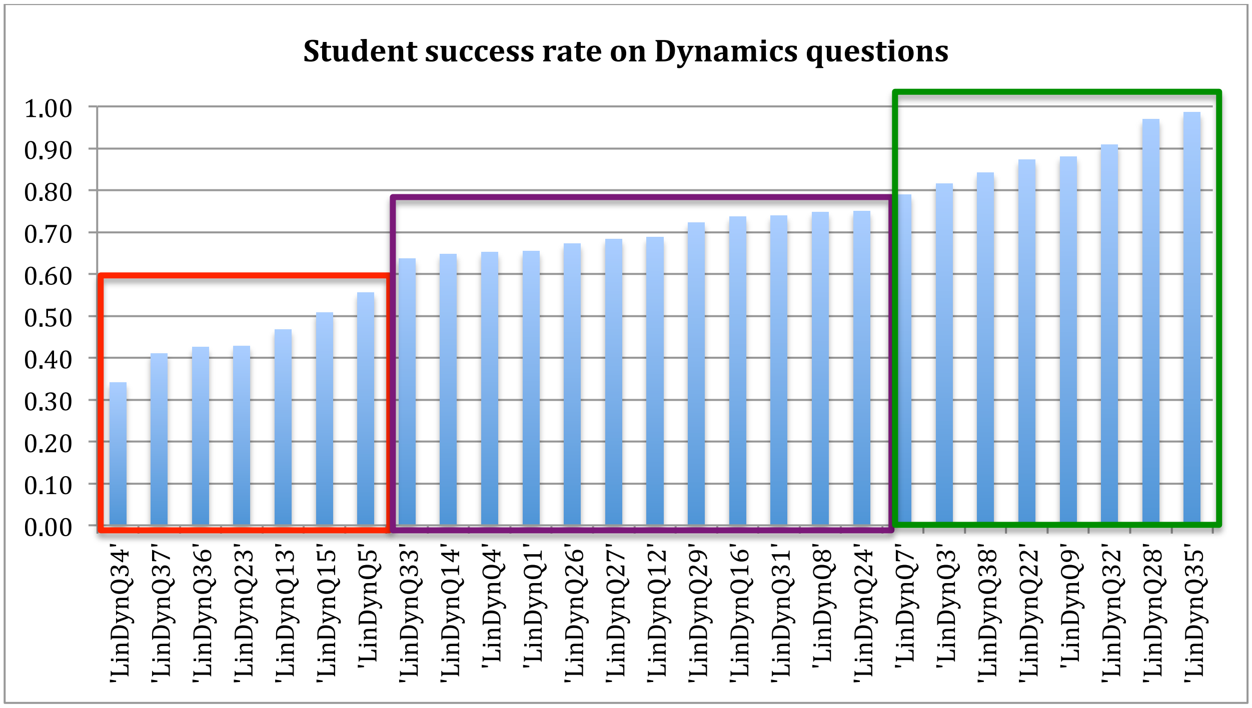


Figure 9.2. Student success rate on Dynamics questions.

9.2.1.3 Conservation principles

There were 15 DALITE questions on the topic of Momentum/Energy. Only 2 questions provided a high level of challenge and produced success rates of below 50%, 6 questions were between 51-79%, and 7 were at 80% and above. These results, once again, cause us to ask the question as to whether this topic is well understood by the students or whether there is something about the wording of the questions, or when they were assigned – pre or post-instruction.

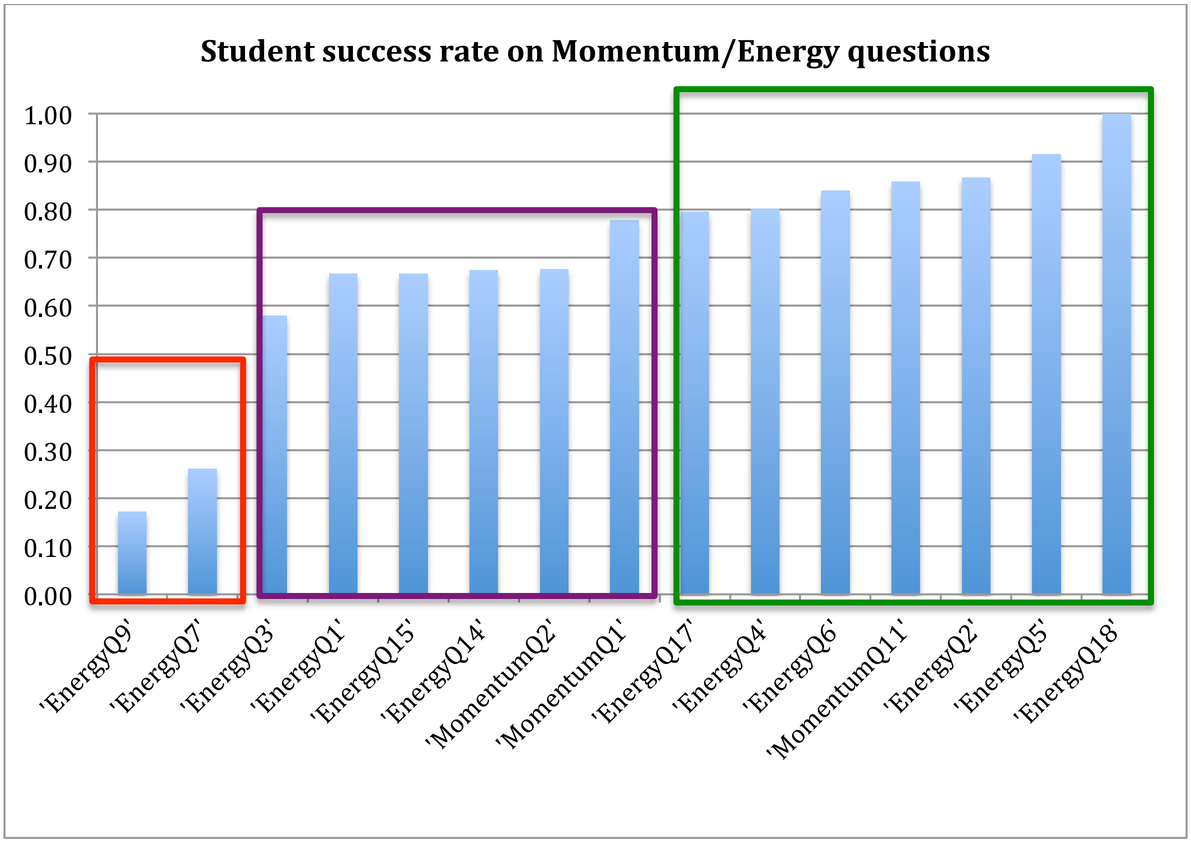


Figure 9.3. Student success rate on Conservation Principles questions.

9.2.2 Meaning of these results

Exploring the meaning of these results we take each of our sub-questions in turn. (1) do students find these concepts easier? (2) are the questions themselves easier? (3) is it the timing of the questions – i.e., pre versus post instruction? (4) are students becoming familiar with the format of the activity and doing more preparation? Or, is there something else to be considered?

First, do students find the Dynamics (LinDyn) and Conservation (Momentum/ Energy) concepts easier? Looking back at the results from the earlier chapters, it seems unlikely that students were finding these concepts easier. In fact, it could be argued that Conservation principles are the most challenging because these are completely new ideas to most students.

Second, are theses questions themselves easier? In reviewing the questions they are very much like the same types of questions that have been used for in-class peer instruction. The questions are therefore no easier or more difficult than those in the Kinematics segment of the course.

Third, is it the timing of the questions – i.e., pre versus post instruction? We have investigated this question and determined timing does have a small impact on the success rates of questions. However, it alone does not explain why students had more success with the Dynamics and Conservation principles. Therefore we turn to the next question.

Fourth, are students becoming familiar with the format of the activity and doing more preparation? To answer this question we rely on the student interviews. From these we learn that indeed students were becoming more familiar with the format. Not only were they learning to use the system itself, but also they were learning how to read the “grammar” of the physics. For instance, the following quote from one of the girls we interviewed:

> ***G3**: ... whenever I have to read an English websites with all the terms, I would not understand it at all. And also just the wording, the way the concepts [are] presented was totally different… with the rationales we had to write, I kind of see the structure that's behind them, so it really helps me to understand better the overall concepts... So I read [physics text] better now. I find I can really now see the information better than just many scientific terms everywhere.*

In addition, some students referring to how they were “learning how to learn” when doing their DALITE assignments. Here are some examples of their comments that illustrate these points.

> ***G2**: Usually I look through my book to look at the theory to see this is right ... You have to look to your notes to try to get more understand… at the same time, it forces me to read, not just look at and then think myself yeah I know why, and makes me like, it forces me to read the rationale and try to understand why it is that answer.*

> ***G11**: I used to write short rationales just thinking why I thought this was the answer, but now I explain the concept behind it and everything, so I give more detailed rationales… at first I found it like all over the place and choppy, but then I got used to it like being somebody's thinking, so it's easier to read now…. Since you have to present it, you have to say “ok this is what we think and why.” It organizes your thoughts.*

Indeed, these quotes suggest that there may have been some learning to learn that helped to explain why questions in Dynamics and Conservation were answered more successfully. And, of course there may be other reasons. Some of those reasons might be the epistemic changes we will discuss in the next chapter. As well, some of these reasons may be explained by the impact of active learning where there is a greater likelihood that students learn from their conversation in their groups. This idea too will be discussed in the upcoming section.

9.3 The Nature of Students' DALITE Responses

Providing evidence of student effort and engagement is difficult at the best of times. With large data copra, such as those produced by DALITE and other CSCL environments, it is necessary to develop other means to approximate effort and engagement. In this case we developed a word count procedure to act as a proxy for student effort and engagement. While we do not suggest that this tells us about the quality of the rationales we explore the possibilities that how much a student writes might be associated with effort and reflect the learning that is taking place. In other words, we attempt to answer the question: what is the relationship between the number of words generated for an explanation and learning outcomes? In addition, we use this method to tell us something about the nature of the questions themselves and the context of the assignment – pre-instruction or post-instruction. Lastly, we use it to identify specific students and specific questions that produced significantly different results – i.e., behaviors that vary from the norm (z-scores). We report on these results below.

Figure 9.4 displays the average number of words produced for each DALITE question in the order in which the questions were assigned. What we learn from examining this figure is that while there is a small trend towards greater number of words being produced for first unit of questions (1DKin & 2DKin), generally speaking the timing of the question was not a major factor in the number of words produced for the explanations. In fact, certain questions in LinDyn and Momentum produce above average word counts as well. Those questions are listed in Table 2 in the order in which they were assigned. We will investigate these questions more fully.

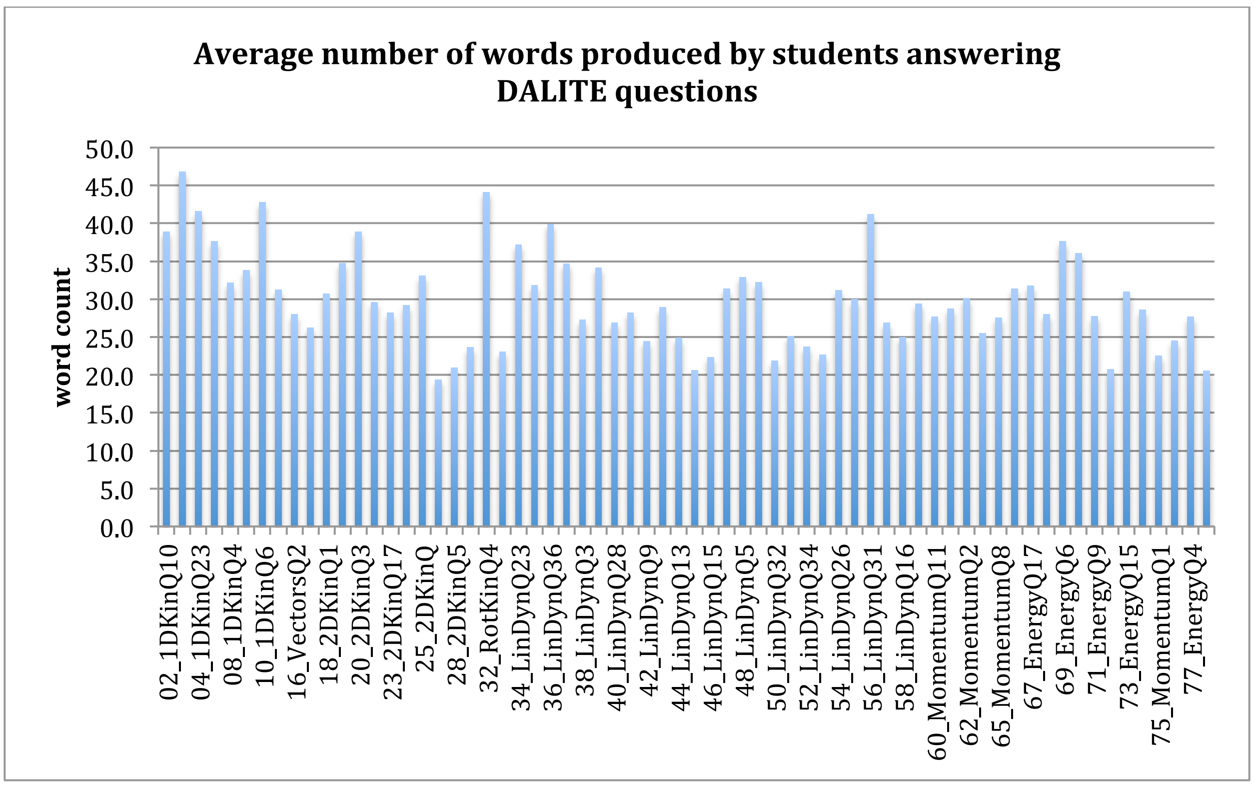


Figure 9.4. Average of words produced by all sections of students answering DALITE questions.

Organizing these data by their z-scores allows us to see the questions that were particularly good at eliciting explanations as well as those that were not. In fact, we can see that 40 of the 66 questions produced above average word counts. Twelve, in particular, elicited greater than average responses that is over one standard deviation (above 1 which represents between 37-46 words).

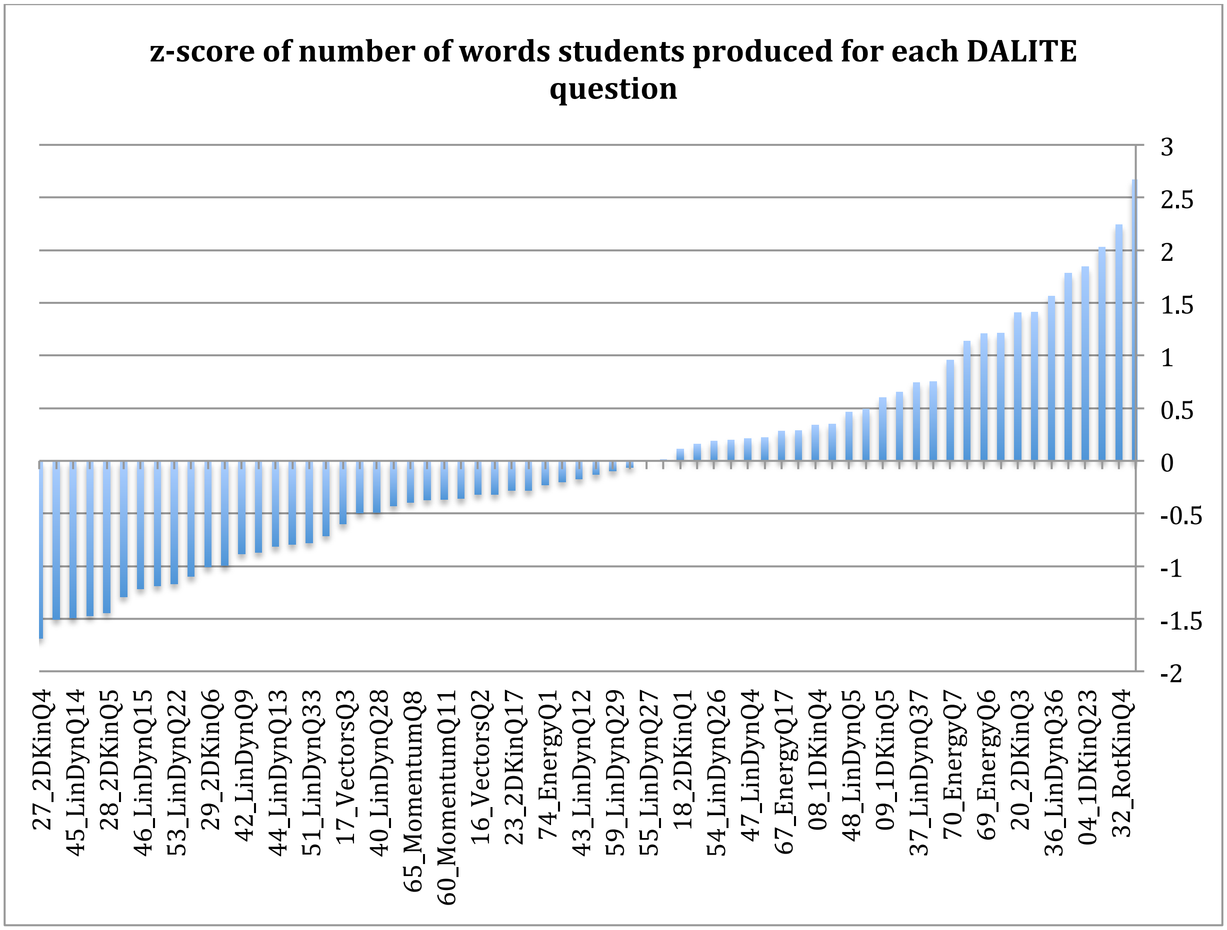


Figure 9.5. Z-score of number of words students produced for each DALITE question.

9.4 Timing of DALITE Questions

Does Timing of the DALITE questions impact the nature of the answers? Looking at these results from another perspective – i.e., for the three course units – we see some small variations (Table 9.3). In these we note that there is a small reduction in the number of words produced as the course proceeds. In essence, Kinematics principles produces more words (27.4) that the Conservation principles (24.6).

Table 9.3. Word count by course unit.

Word count	Kinematics	Dynamics	Momentum/Energy	All questions
Mean	27.4	25.1	24.6	25.7
SD	5.2	4.8	3.4	4.8
Median	26.2	24.2	24.8	25.1

However, when we compare these to the questions that produced the highest word counts the difference pointed to above is mediated by the actual question itself (see Table 9.4). Therefore, we suggest that timing within the course is not the only factor determining whether or not students will engage deeply with the question. Rather, it is the question in and of itself that promotes engagement.

Table 9.4. DALITE questions that elicited a higher than average word count.

DALITE question	z-score of word count
02_1DKinQ10	1.41
03_1DKinQ21	2.67
04_1DKinQ23	1.84
05_1DKinQ34	1.21
10_1DKinQ6	2.03
20_2DKinQ3	1.41
32_RotKinQ4	2.24
34_LinDynQ23	1.14
36_LinDynQ36	1.57
56_LinDynQ31	1.78
69_EnergyQ6	1.21
70_EnergyQ7	0.96

What is noteworthy about these results is that for the 1DKin questions those with graphical representations generate more words than those that are pictorial in nature. See for instance 1DKinQ10 (figure 9.6) and 1DKinQ6 (figure 9.7). It might lead us to believe that the graphical representations require more cognitive effort, which in turn means that more interpretation, leading to more explaining. We will look for closely at the actual words to see whether this conjecture holds up.

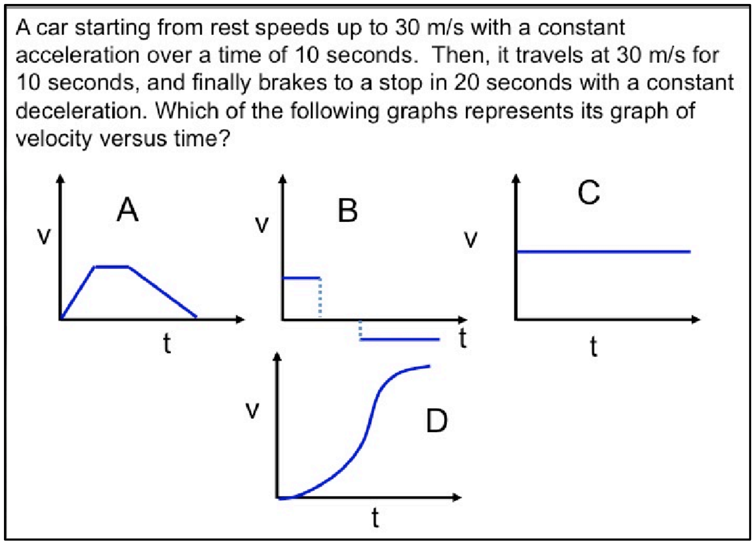


Figure 9.6. Questions 1DKinQ21, which produced a high number of words – 2.67 z-score.

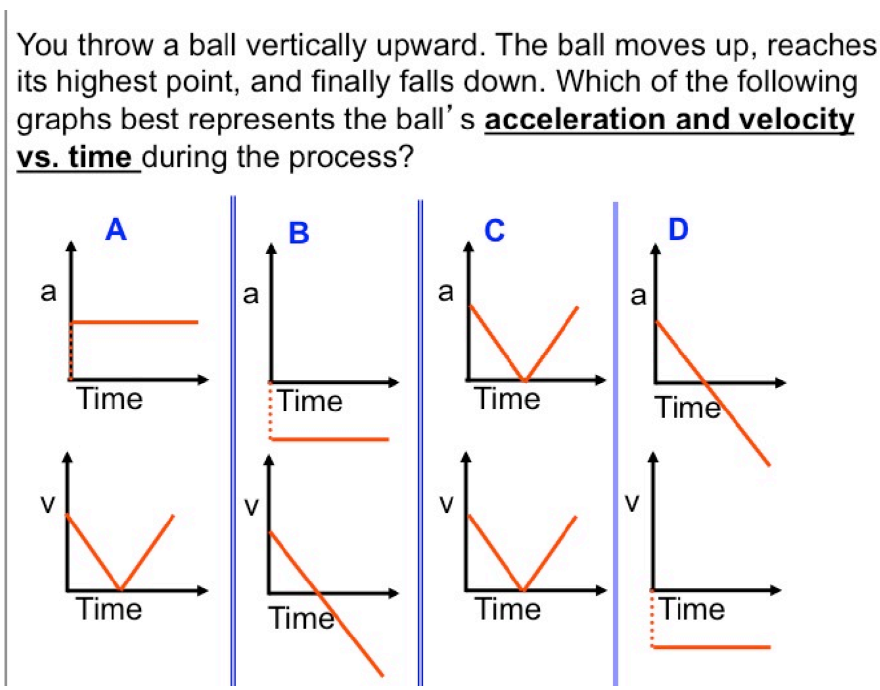
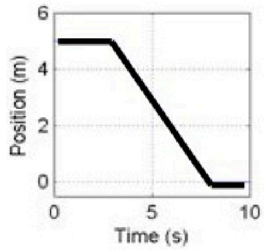


Figure 9.7. Question 1DKinQ6, which produced a high number of words – 2.06 z-score

Below is a graph of a balls motion. Which of the following gives the best interpretation of the ball's motion?



a. The ball moves along a flat surface. Then it moves forward down a hill, and then finally stops.

b. The ball doesn't move at first. Then it moves forward down a hill and finally stops.

c. The ball is moving at constant velocity. Then it slows down and stops.

d. The ball doesn't move at first. Then it moves backwards and then finally stops.

e. The ball moves along a flat area, moves backwards down a hill and then it keeps moving.

Figure 9.8. Questions 1DKinQ23, which produced the second high number of words – z-score 1.84.

9.5 Case Study Question 1DKin Q23

If we were to look more closely at the student's rationales, what might we learn? Taking account of the results that show question 1DKinQ23 (figure 9.8) produced a high number of words, we look more closely at the two cohorts of students, those producing the greatest number of words and those producing the smallest number of words. The objective is to document what these differences reveal. What is worth noting is that both these questions are graphical representations compared to the pictorial or word problems.

9.5.1 Word production for two questions

9.5.1.1 Cohort 1 – the greatest number of words

Below is an example of a student who regularly writes a large number of words, and also regularly gets the question correct.

> ***St#1***: *The correct answer must be one of the options that says that the ball doesn't move at first, because for the first 3 or so seconds, the ball's position doesn't change. That leaves B and D. Had the ball been moving forward down a hill, it would have experienced an acceleration due to gravity, and the curve between 3 and 7 seconds wouldn't be linear. That leaves only option D, which makes sense because the ball does in fact move backwards as its position is decreasing.*

Another example of a student who regularly writes a large number of words, and also regularly gets the question correct. However, in this instance, the student got the answer wrong on their first try but changed to the right answer.

> ***St#2****: It is the most valid and most logical answer. The fact that the position slope is 0 in the beginning and at the end declares that the velocity at the beginning and at the end of the motion is 0, therefore describing an absence in movement. Because of this, we can quickly eliminate answers a, c, and e. After that, we can simply realize that d is also invalid because the ball is not moving backwards since the slope is towards the right. If it was going backwards, it would be going against the x axis. Therefore, the answer is b.*

Examples of fewer numbers of words but also correct rationales, produced by students who regularly averaged above the norm for their word production. These are reported below:

> ***St#3:*** *The flat area is the balls position over the time interval, it is flat because the ball did not move during that time. When the ball does move, its moves closer to the origin which means that the ball is moving in a negative direction.*

> ***St#4:*** *The ball reamains at a position for a few seconds because it isn't moving (it has no velocity). Then the ball must be moving down a hill forward because it's position begins to decrease but it still remains to be positive. Lastly, the ball has the same position for the reamining seconds,so again it reaches a stop.*

> ***St#5:*** *The ball is at rest in the first interval, eliminating answers A, C & E. Then in the second interval its moving backwards because its approaching the x axis, so the answer would be D.*

By contrast, students who regularly produced a small number of words, but also got the answer correct, wrote the following:

> ***St#6:*** *a horizontal line mans its not moving and the diagonal line going downwards means its backing up*

> ***St#7:*** *the position decreases so it's moving backwards also the slope would be negative meaning it'll head to the left.*

Students who regularly produced a small number of words, but got the answer wrong, wrote the following:

> ***St#8:*** *at the fisrt 3 seconds the ball is in the same place, it did not move and then it slow down and stop at the end*

> ***St#9:*** *The line is flat which means the object is not moving then it moves backwards on the downward slope and then the line is flat again as the object stops*

9.6 Other factors involved with using DALITE

Our efforts to identify contributing and co-occurring factors between the ways DALITE was used and other outcomes produced very noisy results. By this we mean there is some evidence of trends for some of the sections but these did not hold up when tested against other. The two sections that appear to have the most robust correlations were T10 and T06, that were somewhat mirror opposites. These correlations will be explored further because these two sections in many ways illustrate how the impact of tools such as DALITE are dependent on students taking up learning opportunities in meaningful ways. And, it also points to one of the interesting outcomes of the active learning instruction, the impact of the group. We discuss this next.

9.6.1 The Effect of the Group

One of the clearest trends in the DALITE rationale within sections is the “peer-effect” within the working groups. Interestingly, the ways that the group members complete DALITE assignments becomes normative for the group. In other words, student working with DALITE were more like their group members than they were like others in the class. Thereby, groups where two or more students began with a high participation rate in DALITE lead to the group having a similar higher than average rate. And, vice versa. There was even some evidence to suggest that individual student’s valuing of the writing demonstrated by their consistent longer text entries, seemed to “rub off” on their group mates.

9.7 Summary

The DALITE rationales produced a great deal of data. For this report we could only analyze a small portion of it to provide some of the broader findings that can be gleaned from early analysis, which are somewhat crude. Clearly, as we begin to dig deeper into the analysis of the rationales themselves we will be able to uncover answers to some of our critical questions about conceptual change and possibly even about the transfer of learning.

In the meantime, we have learned that students can take the exercise of answering questions and producing explanations seriously. On average, over 75% of the students completed the assignments. And, this completion is often tied to certain teachers (teacher values and use of the tool) or instructional styles (more or less active learning). We also learned that with time the students’ success rates increased and this may be explained by a variety of factors including a change in their conceptual understanding and a shift in their learning how to learn. And, that students can write a considerable amount to explain themselves, even when it is initially just for their own benefit. However, there is still much to learn about how to craft good questions but we are beginning to understand that more abstract and graphical representations made have advantages over more wordy ones.

CHAPTER 10
INTERVIEWS

In an effort to understand more about how students responded to the DALITE activity we interviewed them using the open-ended questionnaire (see Appendix F). What is apparent is that students' epistemic beliefs have an impact on how they perceived DALITE activities and what they got out of it. In particular, what they got out of writing and reading the rationales as well as the other components of the DALITE system. We map these on to the design principles we set out earlier in the document. We did this by coding the interviews into the four categories of the design principles. It is evident that the epistemic beliefs play a role here. We present them below but we start with a reminder of what is meant by epistemic beliefs.

10.1 Background on Epistemic Beliefs

Epistemic beliefs refer to individuals' beliefs about the nature of knowledge and knowing (Hofer & Pintrich, 1997). A better understanding of this construct has been gained through decades of research in the field of personal epistemology—an umbrella term used to capture the various labels that researchers have used together with epistemic beliefs (Bendixen, Schraw, & Dunkle, 1998; Schraw, Bendixen, & Dunkle, 2002), such as epistemological beliefs (Jehng, Johnson, & Anderson, 1993; Qian & Alvermann, 1995; Schommer, 1990), reflective judgment (King & Kitchener, 1994, 2004), ways of knowing (Belenky, Clinchy, Goldberger, & Tarule, 1986), epistemological reflection (Baxter Magolda, 1992, 2004), epistemological theories (Hofer & Pintrich, 1997), and epistemological resources (Hammer & Elby, 2002).

Evidence has been accumulating that epistemic beliefs have direct and indirect effects on learning, including the types of learning strategies that university students use (Muis & Franco, 2009; Schommer, Crouse, & Rhodes, 1992), their self-regulated learning (Muis, 2008), their use of metacognitive strategies (Muis & Franco, 2010), their readiness for conceptual change (Franco et al., 2012; Murphy, Alexander, Greene, & Hennessey, 2012), and more generally their cognitive processing (Kardash & Howell, 2000).

Literature supports that epistemological development is a crucial educational outcome (Perry, 1970). Normally, with education and experience, an individual develops from an absolutist to an *evaluatist*. Specifically, the former sees knowledge as black and white, and as static fragmented facts that are handed down by experts or authorities, whereas the latter sees knowledge as interconnected, constructed and should be weighed or evaluated with evidence (Hofer & Pintrich, 1997).

However, epistemic change does not come easily, and teachers need to create a more educationally productive environment to enhance epistemological development in novice learners. For example, certain epistemological tools (Collins, in press; Tsai, 2004) or epistemic artifacts (Hansen, 2012) may assist learners in applying effective strategies to initiate and sustain epistemic change.

10.2 Mapping student's perceptions to the design

What were the perceptions of the design? We provide an overview of the coding in table 9.1. Afterward, we provide samples of the student's quotes that were placed in those categories.

Table 10.1. Coding of interviews into the four design categories.

Total Interviewees	1 promoting the use of the language of the discipline	2 promoting reflections on explanations, one's own and that of others	3 promoting deeper understandings of conceptual structure; and awareness of different contexts with similar		4 promoting students' agency and responsibility for examining their peers' arguments and assessing their correctness and quality; training of students' expectations of the classroom script				
			1	2	TOTAL	Take actions	Students' agency	Examining ideas	Students' expectations
23	15	16	19	16	16	2	7	8	3

10.2.1 Promoting the use of the language of physics

> ***B5***: *I find that it helps you out to write it down, because it's much easier to say "oh yeah I understand that" but then to try and explain it in words, to be concise, it really shows you understand the matter. It helps out a lot.*
>
> ***G2***: *Writing a rationale is becoming easier for sure. You try to put in your own words. To like write your own definition, and so that it stays in your mind.*
>
> ***G9***: *I tend to go step by step now, and say well "this is this and this is this, so this", whereas before it was a jumble of my thoughts.*
>
> ***G11***: *I used to write short rationales just thinking why I thought this was the answer, but now I explain the concept behind it and everything, so I give more detailed rationales… at first I found it like all over the place and choppy, but then I got used to it like being somebody's thinking, so it's easier to read now…. Since you have to present it, you have to say "ok this is what we think and why." It organizes your thoughts.*

These quotes support the claim that students can begin to value the writing process and the use of language of physics. For instance, B5 specifically acknowledges that writing your explanation allows you to confirm that you really understand the concept. G2 recognizes that by writing you are in fact writing your own definitions and by doing so it makes it your own. Meanwhile G9 and G11 state that they have learned how to express their thoughts so that they are no longer jumbled and better organized.

The comments of G11 also take us to the next observations from these interviews, the idea that you learn to "read" how other people write. Additionally, how other people's writing can help you to learn how to write yourself. In particular, the comment "*first I found it like all over the place and choppy, but then I got used to it like being somebody's thinking"*. This is a very profound observation that the rationales could be a form of reading someone else's thinking. Other students also referenced that it became easier to read other people's rationales. Some examples are below.

> ***B8***: *It's easier to understand what people are writing now. I don't know if that just because we're farther into the year of physics, that everybody knows how to explain themselves a little bit more, and we're all using the same kinds of words, in the same kind of way as they're talking in class, whereas in the beginning of the year everyone was just running off of their pre-11 physics education, so they had different words and different ways of explaining stuff, so as it got through it did get easier. I think that may have been due to how we'd re-learnt everything, you know. But yeah, people explain themselves better.*

> ***G5***: *it becomes easier now [that] I understand the concept more, because I struggled at the beginning of this semester with this class… as I understood the concepts more, definitely became a lot easier to explain my thoughts processes and why I got the answer…*

B8 makes an insightful comment with the reference to the convergence that is happening with the explanations. As the students' language take on more of the normative uses of physics their writing becomes easier to understand. At the same time, their understanding of concepts allows them to understand more, as G5 points out. Together, these factors makes the tasks set out by DALITE value in the students' learning.

10.2.2 Comparing and explaining to yourself

A major design feature of DALITE was the idea of comparing and contrasting. In the response to the interviews there was ample evidence that students recognized this feature and appeared to value it as something that helped them to learn. Below are some excerpts that illustrate this point.

> ***G3:*** *You can compare your answer with like somebody else answer, and you can see why you are wrong, where did you make mistake*

> ***G8***: *I read it and I compare it to mine, cause usually if the answers at the end is the same, and I have the same thought process, and I know that I did the question right and I understand the concept. But if it is wrong, by going over the rationale I can see what [my teacher's] thought process was, and I can understand his point of view, how he interpret the problem and helps me to understand the situation, and like, how to interpret it myself if I were to have a test on it or something.*

> ***B9:*** *So it requires thinking, analyzing other people's answers and thinking for yourself to makes more sense.*

There are also good examples of

> ***B2***: *Yeah... when you try to explain [to] yourself and you're still not sure, and then you give your answer and you can read through everybody's explanations, you're able to make sense of what you're saying, and see where your thought process might have been wrong or what the other people's thought process is. And, you can look at what answers actually make more sense to you. So I guess it helps because you're seeing other people's point of view and sometimes you like theirs better.*

> ***G2***: *Reading other people's rationales sometimes feels like you are missing things, and I am never sure if it is right. Sometimes, it kind of persuades me to go to the other answer, but … there is always doubt. But sometimes, there are some rationales [that] are really good, and you are like, yes, I remember this. This is correct. This is what I was missing. I still read every rationale, and try to understand what people are trying to explain.*

> ***B1:*** *You can read through everybody's explanations, you're able to kind of make sense of what you're saying a little bit more and see where your thought process might have been wrong or what the other people's thought process is, and you can look at what answers actually make more sense to you.*

10.2.3 Promoting conceptual conflict and context awareness

> ***G1_T09***: *I want to know which one makes more sense… there is one side [that] convinces you so much, and you like, ok, it must be that [answer]. But then in the back of your head, you know these other people make a good point. So then you get conflicted.*

> ***G2***: *If I'm pretty sure for my answer, I'll read the rationales for my answer, a few for the opposite answer. But if I'm not sure, I'll read for both [answers] 'cause like the rationales help me decide… I find that it helps because then we discuss it in our groups then as a class, so again it's just more like learning everyone else's perspectives on things, and you're doubting what you think or you're getting confidence in what you think.*

> ***G3***: *when you're going over your classmate's rationale, it's sort of similar to what we do in class, which is discuss. And, one person can mention one small detail and it will change what you think.*

The above quote shows two ideas. First, the students' demonstrate a growing awareness of the conceptual conflict they are experiencing. In other words, they are expressing a conscious reflection on their understanding. In addition, Both G2 and G3 make an explicit statement about their awareness of context. For G2, she bring the DALITE assignment back into her classroom discussions "*I find that it helps because then we discuss it in our groups then as a class, so again* *it's just more like learning everyone else's perspectives on things".* Meanwhile, for G3, she recognizes the contextual similarities between the in-class activities and what she is doing at home while doing the DALITE assignments.

Then there were students who talked about conceptual understanding and context together. The example of G13 who recognized that the concepts were similar across several questions, were the context was different. Surprisingly, she recognized this not during one of the planned extended DALITE activities but during the Sorting Task activity, which was intended to be an assessment. This is a great example of how a good assessment tool can also be a teaching/learning opportunity.

> ***G13****: I feel like you do all these, individual problems for each chapter, but then it's like when we first got that first one, and like it was like completely I didn't even realize how similar the problems actually were, I think it was after the kinematics, and like the energy, and like the circular motion, I did not realize that it is almost the same problem, it's just like I find it's really helpful, and like put it all together like that.*

10.2.4 Promoting students' agency and responsibility

One of the final designed features of DALITE was to promote student's sense of agency and wish to know more. In other words, we provided some modeling and coaching in features such as the compare and contrast as well as voting for most convincing explanation. But we purposefully left it up to students to figure out if their answer was correct. The results of this design are captured in the statements below.

> ***G2****: Usually I look through my book to look at the theory to see this is right … You have to look to your notes to try to get more understanding… at the same time, it forces me to read, not just look at the rationales and then think to myself, yeah I know why [that's the answer]. And makes me like, it forces me to read the rationale and try to understand why it is that answer.*

> ***G5****: they force you to think, and then it kind of helps you sort out your thoughts, like I think that is, in my opinion, that was the goal of all of these, to help you take out, like useless information, like to help you get down to what's actually important. It is a long process, but I think it is like slowly gain in there*

> ***B4****: I also think that just knowing what we're expected to write, just kinda getting used to the whole system made them better overtime.*

> ***G10****: I mean there was a problem when I first started doing DALITE. I was wondering where the answer was. I was all over frustrating because I didn't realize that there was an explanation and I should read it. I couldn't find it and I was like, well how am I suppose to know if I got it right or wrong? But then I clicked. It was a while into the semester. It clicked that I should probably read the reasoning [in the expert rationale]. And if I had the same thing then I know I was right. So actually I find it was better than just having the answer… if I can read in words [the explanation], I can understand the [thinking] process and it makes me understand the problem better.*

G2: it's kind of like, you are on your own, you have friends, but you are on your own in your learning.

10.4 References

Baxter Magolda, M. B. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development*. San Francisco, CA: Jossey Bass.

Baxter Magolda, M. B. (2004). Evolution of a Constructivist Conceptualization of Epistemological Reflection. *Educational Psychologist, 39*(1), 31-42. doi: 10.1207/s15326985ep3901_4

Belenky, M., Clinchy, B., Goldberger, N., & Tarule, J. (1986). *Women's ways of knowing: The development of self, voice, and mind*. New York: Basic Books.

Bendixen, L. D., Schraw, G., & Dunkle, M. E. (1998). Epistemic Beliefs and Moral Reasoning. *Journal of Psychology: Interdisciplinary and Applied, 132*(2), 187-200.

Collins, A. (In press). In M. S. Khine & I. Saleh (Eds.), *Models and modeling: Cognitive tools for scientific enquiry*. Dordrecht, The Netherlands: Springer.

Franco, G. M., Muis, K. R., Kendeou, P., Ranellucci, J., Sampasivam, L., & Wang, X. (2012). Examining the influences of epistemic beliefs and knowledge representations on cognitive processing and conceptual change when learning physics. *Learning and Instruction, 22*(1), 62-77. doi: 10.1016/j.learninstruc.2011.06.003

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 169-190). Mahwah, NJ: Erlbaum.

Hansen, F. T. (2012). Epistemic artifacts: The potential of artifacts in design research. http://www.flemmingtvede.dk

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140.

Jehng, J. J., Johnson, S. D., & Anderson, R. C. (1993). Schooling and students beliefs about learning. *Contemporary Educational Psychology, 18*, 23-35.

Kardash, C. M., & Howell, K. L. (2000). Effects of epistemological beliefs and topic-specific beliefs on undergraduates cognitive and strategic processing of dual-positional text. *Journal of Educational Psychology, 92*, 524-535. doi: 10.1037/0022-0663.92.3.524

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco, CA: Jossey-Bass.

King, P. M., & Kitchener, K. S. (2004). Reflective Judgment: Theory and Research on the Development of Epistemic Assumptions Through Adulthood. *Educational Psychologist, 39*(1), 5-18. doi: 10.1207/s15326985ep3901_2

Muis, K. R. (2008). Epistemic profiles and self-regulated learning: Examining relations in the context of mathematics problem solving. *Contemporary Educational Psychology, 33*(2), 177-208.

Muis, K. R., & Franco, G. M. (2009). Epistemic beliefs: Setting the standards for self-regulated learning. *Contemporary Educational Psychology, 34*(4), 306-318.

Muis, K. R., & Franco, G. M. (2010). Epistemic profiles and metacognition: support for the consistency hypothesis. *Metacognition and Learning*, 1-19.

Murphy, P. K., Alexander, P. A., Greene, J. A., & Hennessey, M. N. (2012). Examining epistemic frames in conceptual change research: implications for learning and instruction. *Asia Pacific Educational Review, 13*, 475-486. doi: 10.1007/s12564-011-9199-0

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme.* New York: Holt, Rinehart, and Winston.

Schommer, M. (1990). Effects of Beliefs About the Nature of Knowledge on Comprehension. *Journal of Educational Psychology, 82*(3), 498-504.

Schommer, M., Crouse, A., & Rhodes, N. (1992). Epistemological Beliefs and Mathematical Text Comprehension: Believing It Is Simple Does Not Make It So. *Journal of Educational Psychology, 84*(4), 435-443.

Tsai, C.-C. (2004). Beyond cognitive and metacognitive tools: The use of the Internet as an “epistemological” tool for instruction. *British Journal of Educational Technology, 35*(5), 525-536. doi: 10.1111/j.0007-1013.2004.00411.x

CHAPTER 11
INSTRUCTOR PERCEPTIONS OF DALITE

Our final data source and analysis focuses on the instructor's perceptions of DALITE. The instructor display of student results has proven to be a very important feature of DALITE. Arguably, it the most practical tool for instructors. It provides immediate and detailed feedback to instructors and greatly supports the flipped classroom method or active learning approach in general. Students are better prepared for class and teachers can identify conceptual issues before class and tailor their class to focus on these specific issues. Also it allows students to focus on both the collective and the individual – what does the class know, what do specific individuals need to know (where are they falling short), a feature we only identified in student interviews.

When we interviewed the instructors, we learned more about their perceptions about DALITE and how they thought it was useful. Additionally, how they see it being improved. Five of the six interview questions and their responses are reported on.

1. **How did you use DALITE?**

All four teachers used DALITE as a homework component of their course. It was assigned either pre-instruction or post-instruction and incorporated in the following fashions:

1) DALITE [was used] outside the classroom and on a limited basis inside the classroom. Outside the classroom, it was used as a pre-instruction activity as well a post-instruction activity. When used as a pre-instruction activity it was meant to complement the reading assignments, assess some prior knowledge and push the level of understanding a bit deeper. As a post-instruction activity it was used to complement the work done in class. Inside the classroom we used DALITE as a tagging activity.

2) DALITE was incorporated into the student's weekly assignment. In this class the students had a weekly online assignment using Webwork, usually 10-12 problems per week. The first question of the assignment was set up with instructions, about which assignment to do, deadlines etc. and a link to the DALITE site. Students simply clicked on the link, which took them to the DALITE login.

3 &4) DALITE [was used] both in a JiTT/Flipped context (pre-instruction homework) as well as part of review (post-instruction) assignments. Typically, the students would have 2–3 DALITE questions in an assignment.

2. **How did it fit into your other classroom activities, if it did?**

As a summary, in all five sections, DALITE was a replacement for Peer Instruction (clicker questions). That said, all four instructors mentioned that there was a problem with a segment of the DALITE questions when the system was undergoing a minor upgrade. This resulted in a greater need for these DALITE questions to be brought into the classroom and for the instructor to ensure students' understanding of these concepts. For one instructor it also resulted in his students becoming less motivated and trusting of the system. On the other hand, another instructor emphasized the need for the teacher to "value" the use of the system so as to show the students that their efforts were warranted. These specific responses are below:

1) I brought the pre-instruction DALITE activities into the classroom by reviewing the results and spending time on questions that were problematic (questions that had been answered poorly). Often this was done just after reviewing the reflective writing assignments. As for the post-instruction DALITE assignments I would review them in class only when there were problems.

In order for students to use a system like DALITE, it's important that the work they do be valued beyond just assigning grades. I would have liked to do this a little bit more but the system is still not really fluid for this. I would like to have a system where it was easier to review questions, votes and rationales.

2) DALITE was linked to three other classroom activities: in-class clickers, concept maps, and test questions. DALITE largely replaced the clicker questions that I normally used, particularly as the DALITE questions came from the same databases. But I did use clickers occasionally. If I saw that a DALITE intervention was appropriate, a large number of continuing red answers, or students going to wrong from right (I would display the green/red answer page, with the names off the screen, to show to students that there was a problem), I would go through the DALITE assignment as a clicker activity. That is to say, I used the DALITE questions in an in-class peer instruction activity. This gave students another opportunity to correct any misconceptions they had, and gave me an opportunity to talk about the questions. On other occasions, where students were largely getting it right, I talked about how to write rationales. Or more specifically what I look for when I write rationales, and what to look for when reading other peoples' rationales.

There was an issue of the incorrect answers being programmed into DALITE, so it showed that students had mostly got it wrong, when in fact they had got it right. I became aware of this through the in-class activity as described, just as it was reported by the other teachers. This was unfortunate because students started to mistrust the system, and run under the conclusion that the system was faulty.

Another activity that involved DALITE, was concept mapping. Over the semester, students worked on several different concept maps. On a couple of activities, students would take their concept maps and insert DALITE questions into their maps.

Finally, I made it clear to students that there would be multiple choice questions on the tests and final exam, and the best way for them to prepare for this, was to do and review the DALITE questions. I prepared a file with the questions used for distribution to the students (there is no review mode in DALITE) for study purposes. Students requested that I go over some of them in class, as part of the test and exam review.

3 &4) In class, we used Peer Instruction with clickers, which is very similar to DALITE. Often we'd choose a PI slide as a follow-up to the previous night's DALITE homework.

Also, in cases where there were a lot of incorrect DALITE responses, we would take time in class to explore those questions more thoroughly.

In addition, as part of the PAREA project, I worked to incorporate previously-seen DALITE questions into other contexts. In particular, my students were asked to link several DALITE questions into concept maps that they produced for the course.

3. How did it change or facilitate your ability to teach, if any?

This question generated the most important results. It showed that the instructors saw several real benefits to using DALITE. The three that are most significant are: (1) creating more time in class so that other active engagement activities could replace Peer Instruction (clicker questions); (2) providing a better gauge of students' conceptual understanding and misconceptions; and (3) promoting a change in students' motivational and belief systems – e.g., self-regulation, epistemic beliefs, values.

1) DALITE didn't change how I teach but it did let me take some of the things I would normally do inside the classroom – like clicker questions - and transfer them outside of the classroom. This was important because it let me engage students in classroom activities that were a little more involved or elaborate.

Using DALITE as a pre-instruction tool also let me motivate students to do the reading in preparation for class and it let me get a better sense of their prior knowledge. As a post-instructional tool it let me push the classroom material to try and get students to understand at a deeper level.

2) The main way it changed how I taught is I did not do nearly as many clicker activities as I did previously. I appreciated this as the clicker set up at [my College] is clunky, often crashes, and thus interrupts the classroom flow. DALITE allowed students to get a peer instruction experience, without is taking up as much class time as clickers.

I particularly appreciated DALITE when students initiated questions based on the DALITE questions, which happened on both the individual and class level. In other words, DALITE allowed access or entry points into the subject and opportunities for students to self-regulate.

3) DALITE had several impacts on my teaching. For one, it gave me insight into the flexibility/rigidity of student's conceptions and misconceptions. It's quite striking to see the occasions where students who initially answered correctly were easily pushed off of the correct answer. This is a dimension that isn't captured by traditional homework.

Another aspect is, that by putting such a focus on conceptual understanding, communicating and reflecting, it provided an opportunity to communicate more clearly to the students that I, as a teacher, value conceptual understanding more than the ability to perform certain calculations by rote.

4. How did it change what your students were able to do, if any?

Generally, the instructors saw DALITE as a tool that supported their students' conceptual knowledge construction, which in turn supported their efforts to implement an active learning pedagogy in their classroom. Two point of growth were mentioned: (1) beginning to think more deeply about the material; and (2) ability to read and write better, in English.

1) I think we were able to do more and it provided a better way for the students to engage with the material – getting them to do a reading assignment only goes so far in really getting a student to begin thinking about the material. I don't think it let me cover more material but it did let us get a bit deeper in terms of the depth of understanding. Being able to extend the kinds of thinking, questioning and evaluating that goes on in a classroom peer instruction activity to an out-of-class activity is great.

2) I cannot for sure say it changed what students were able to do, mostly because they did DALITE outside of the classroom. However, I believe it helped students read and write in English. More importantly, I believe it helped in the progression of some students from a high school mentality to a more self-aware learning style.

3) The F2013 class had very good gains on the FCI, and performed well on the final examination. It's not crystal clear what, apart from how it affected my teaching (see answer to question #3), is a direct result of DALITE and how much of it has to do with other active pedagogies. I would say DALITE contributed as an important part of the whole pedagogy.

I'll take this moment to re-emphasize a point made earlier: it gave me insights into student conceptions and misconceptions (particularly how in a few cases, students were easily shifted away from a correct answer). Since this is a new insight for me, it's hard to compare this group's performance with other groups.

6. What other observations about DALITE or about the uptake of DALITE do you wish to share?

Both instructors who answered this question suggest improvements for DALITE. One of the two also makes suggestions about how it should be assigned.

1) I like DALITE a lot but there's still a lot of work to do to make it easier to use. On the teacher side, there are still issues with creating assignments and using/viewing the results – especially in class. On the student side, I'd like to have the voting (thumbs-up & thumbs-down) feature working and I still think we need to work on the "tagging" to see if it can be useful and effective. Lastly, we have to work on the issue of how to deal with new questions (what to do when there are no rationales) and how to select the rationales that are shown to students.

2) I think less is more, for DALITE, and some students reported spending much more time on it than I would have felt warranted. This is something we could perhaps pull out of the metadata. I think it may have more impact if students had done fewer questions, but had really taken time to work out their rationales and read other students' rationales.

A few quirks of the implementation worked as impediments to students: no review mode meant that students could not effectively review their answers for study, nor could they interrupt a session to complete it later. Some students give up at the slightest hint of an issue with the platform, and it was difficult to get some students back on track when the system mucked up.

CHAPTER 12
CONCLUSIONS

This research tells us several things about how to promote conceptual change and transfer of learning. It also tells us about the design of tools to support and assess such learning. Lastly, and most importantly, it tells us about the impact the design features and how these could further the understanding of the theory they are built on. We elaborate on these findings next.

12.1 Meaning of the FCI results

First and foremost, DALITE, as part of varied levels of active learning pedagogy, was shown to produce a statistically significant increase in students' conceptual knowledge when compared to average instruction. When compared to an active learning pedagogy alone, however, the results show that DALITE and active pedagogy are equivalent. When the DALITE treatments were compared to each other, there was also no statistical difference between groups.

Taken together, these results suggest that DALITE is a good addition to an active learning pedagogy. It can be an effective tool for supporting the development of students' conceptual knowledge. Additionally, its effectiveness is not dependent on the quality of the implementation of the active learning. In short, its use with various implementations of active learning shows no difference between groups. However, while not statistically significant there is a trend that suggest something was different for treatment group T06. This difference was explored more closely in the other assessments and analyses.

12.2 Meaning of other assessments measures

It is clear that identifying evidence of students' learning is a difficult task. Sometimes it requires looking closely between the way students explain their answers, sometimes it means looking at the differences between those who are still incorrect. In doing so, we posit that it reveals the state of misconceptions, thereby a glimpse of a possible trajectory for this learning.

The common conceptual tests showed no difference between DALITE treatment groups and the active learning comparison group. These results suggest that the DALITE treatments we as good as an active learning implementation. However, when we looked more closely at these tests as individual assessments, we see differences between the treatment and the comparison, as well as between treatment groups. To start, some of the DALITE groups appear to have a better understanding of Dynamics concepts compared to the comparison group. We established this through a fine grained analysis whereby we determined that students in the comparison group not only misinterpreted how they were suppose to answer but held deeper misconceptions. In particular, the conflation of velocity and acceleration as applied to Newton's second law of motion.

When we looked at the results of the mid-term tests, they showed that the DALITE groups were more likely to maintain their level of growth even when the questions and concepts get more difficult. However, the comparison group show that as the concepts get more difficult, there is a growing difference between the high and low students. In other words, the data suggest that the DALITE group are moving forward

Lastly, the Sorting Tasks (ST#1 & #2), completed by the DALITE treatment groups only, shows that students were better able to see common themes at the deep similarities, more often than not. However, these analyses helped to identify specific differences in capabilities between groups.

12.3 Meaning of extended DALITE tasks

These many differences begged the question, what was the cause of these differences, which in turn lead us to closely examine specific treatment groups and how students used DALITE. Therefore, it lead us to look closely at how DALITE was used by the different groups. And, the types of rationales produced, which allowed us to identify even more specific capabilities that might be attributed to working with DALITE.

12.3.1 Written rationales

We learned that students generally did engage in the DALITE activity and with time their success rates increased and this may be explained by a variety of factors including a change in their conceptual understanding and a shift in their learning how to learn.

There is still much to learn about how to craft good questions but we are beginning to understand that more abstract and graphical representations made have advantages over more wordy ones.

12.3.2 Concept maps

We learn how to analyze the concept mapping activity so that we could identify students who understand the conceptual knowledge, from those who do not. To that end, the data showed a high correlation between students who produce good maps and those who also did well in the course. Lastly, students report that making the concept maps helped them consolidate their understanding of the course content, and how the DALITE questions fit in.

12.4 Meaning of the interviews

Question 3 in the instructor interviews generated the most important results. It showed that the instructors saw several real benefits to using DALITE, namely: (1) creating more time in class so that other active engagement activities could replace Peer Instruction (clicker questions); (2) providing a better gauge of students' conceptual understanding and misconceptions; and (3) promoting a change in students' motivational and belief systems – e.g., self-regulation, epistemic beliefs, values.

12.5 Importance of the research to the College Network

This research addressed the issue of transfer of learning, which is a major concern raised by educational policy makers, educational administrators (school boards), researches and teachers alike, (e.g., Québec Education Program (QEP) Secondary School, Cycle One, 2004; Frenay & Bédard, 2004; Tardif, 1998).

The research team has already begun to disseminate our findings by presenting at several local and international conferences. Additionally, we have been busy publishing the results of our earlier studies, which also relate to the challenge of conceptual change. See list below for publication written since the beginning of PA2011-006 grant. We believe this is a significant contribution to the College network and way to pay back our debt to the generous support of the Quebec government.

Publications by the PAREA team:

Charles, E.S, Whittaker, C., Lasry, N., Dugdale, M., Lenton, K., Bhatnagar, S., Guillemette, J. (2014). *Taking DALITE To the Next Level: What have we learned from a web-based peer instruction application*. Polman,P.L., Kyza, E.A., O'Neill, D.K., Tabak, I., Penuel, W.R., Jurow, A.S., O'Connor, K., Lee, T., & D'Amico, L. (Eds.). *Learning and Becocming in Practice: ICLS2014 Conference Proceedings, Volume II*, pp.982-986. International Society of the Learning Sciences.

Lasry, N., Charles, E., & Whittaker, C. (2014). When teacher-centered instructors are assigned to student-centered classrooms. Physical Review Special Topics-Physics Education Research, 10(1), 010116.

Lasry, N., Guillemette, J., & Mazur, E. (2014). Two steps forward, one step back. Nature Physics, 10(6).

Lasry, N., Dugdale, M. & Charles, E. (2014). Just in Time to Flip Your Classroom. *The Physics Teacher,* 52(1): 34-37.

Lasry, N., Dugdale, M. et Charles, E. (2014). Zut! J'ai renversé ma pédagogie. *Pédagogie Collégiale*, 27, hiver.

Charles, E., Lasry, N., et Whittaker, C. (2013). L'adoption d'environnements sociotechnologiques comme moteur de changement pédagogique. *Pédagogie Collégiale,* 26 (3).

Lasry, N., E. Charles, C Whittaker, H Dedic, S Rosenfield (2013). Changing classroom designs: Easy; Changing instructors' pedagogies: Not so easy. *AIP Conf. Proc.* (1513), AIP Press, Melville NY.

Charles, E., Lasry, N., & Whittaker, C. (2012). *Redesigning Classroom Learning Spaces: When technology meets pedagogy and when they clash*. van Aalst, J., Thompson, K., Jacobson, M. J., & Reimann, P. (Eds.) (2012). *The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012) – Volume 2. pp. 207-211.* International Society of the Learning Sciences: Sydney, NSW, Australia.

Lasry, N., Charles, E., Whittaker, C., Dedic, H. & Rosenfield. S. (2012). *Changing Classroom Designs: Easy; Changing Instructors' Pedagogies: Not So Easy.* Paper presented at the annual Physics Education Research Conference (PERC), August, Philadelphia, PA.

Charles, E.S., Tissenbaum, M., Whittaker, C., Lui, M., Dugdale, M., & Slotta, J.D. (2011). *Co-design of Collaborative Collective Knowledge Environment.* In Spada, H., Stahl, G., Miyake, N., Law, N. (Eds.) *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings. Volume II.* pp. 641-645. International Society of the Learning Sciences.

Charles, E.S., Park, J., Whittaker, C., & Lasry, N. (2011). Exploring the Role of Technology-Supported Peer Instruction in Student Understanding and Interaction in College Physics Classrooms. In Spada, H., Stahl, G., Miyake, N., Law, N. (Eds.) *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings. Volume III.* pp. 866-867. International Society of the Learning Sciences.

Conferences by the PAREA team:

Charles, E.S., Whittaker, C., & Lasry, N. (2014). *Taking DALITE To the Next Level: What have we learned from a web-based Peer Instruction application.* Paper presented at 33rd annual AQPC symposium: Laval, QC.

Charles, E., Whittaker, C., & Lasry, N. (2014). *Active Learning Classroom Use in College Level Physics: The impact of technology design and adaptive orchestration.* Symposium entitled: Space and Technologies for Learning in Schools, Museums and Workplaces: Recent Approaches in Design-Based Research. Paper presented at the American Educational Research Association (AERA): Philadelphia, PA.

Charles, E., d'Apollonia, S., Orjuela-Laverde, M., Whittaker, C. (2013). *If we build them... : Can active learning classrooms promote changes to teaching practice?* Paper presented at National Association for Research in Science Teaching (NARST), April: Puerto Rico, U.S.A.

Whittaker, C. & Charles, E.S. (2013). *The Perfect Marriage - Technology and Pedagogy in a Next-Generation Active Learning Classroom*. Paper presented at 32nd annual AQPC symposium: Montreal, QC.

Charles, E., & Whittaker, C. (2012). *Change Pedagogy or Change Learning Spaces, What's Most Important?* Paper presented at Canadian Association of Physics, June: Calgary, AB.

Seiler, G., Charles, E., Whittaker, C., & Kunicki, S., (2012). : *Using Technology and Innovation to Improve Science Education. Paper presented at the* Scholarship in Teaching & Learning in Higher Education (STLHE) annual conference, June: Montreal, QC.

Charles, E.S., Tissenbaum, M., Whittaker, C., Lui, M., Dugdale, M., & Slotta, J.D. (2011). *Co-design of Collaborative Collective Knowledge Environment.* Paper presented at the 9th International Conference on Computer Supported Collaborative Learning, Hong Kong, China, July 6th 2011.

Charles, E.S., Whittaker, C. & Lasry, N. (2011). *Smart Classroom Smart Students: Leveraging new learning environments to improve learning*. Paper presented at 30th annual AQPC symposium: Quebec, QC.

REFERENCES

Adams, W., Perkins, K., Podolefsky, N., Dubson, M., Finkelstein, N., & Wieman, C. (2006). A new instrument for measuring student beliefs about physics and learning physics: the Colorado Learning Attitudes about Science Survey. *Physics Review ST-PER, 2*(1), 010101.

Anderson, T., & Shattuck, J. (2012). Design-Based Research A Decade of Progress in Education Research?. *Educational Researcher*, *41*(1), 16-25.

Bao, L. (2006). Theoretical comparisons of average normalized gain calculations. *American Journal of Physics, 74*, 917.

Barbeau, D. (2006). *Meta-analysis of the effectiveness of interventions for academic success in college*. *Rapport de recherche, Association pour la recherche au collégial, 121 p.* Research report, Association for Research in college, 121 p.

Barrette, C. (2007). Successful integration of ICT: An *action guide for more precise. Click,* No. 63, January.

Barrette, C. (2005). Towards a meta-synthesis of the impact of ICT on learning and teaching in institutions of the college system in Quebec. *Mise en perspective, "Click,* No. 57, March 2005, p. *18-24.*

Barron, B., Schwartz, D., Vye, N., Moore, A., Petrosino, A., Zech, L., et al. (1998). Doing with understanding: Lessons from research on problem-and project-based learning. *Journal of the Learning Sciences, 7*(3), 271-311.

Baxter Magolda, M. B. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development*. San Francisco, CA: Jossey Bass.

Baxter Magolda, M. B. (2004). Evolution of a Constructivist Conceptualization of Epistemological Reflection. *Educational Psychologist, 39*(1), 31-42. doi: 10.1207/s15326985ep3901_4

Belenky, M., Clinchy, B., Goldberger, N., & Tarule, J. (1986). *Women's ways of knowing: The development of self, voice, and mind*. New York: Basic Books.

Bendixen, L. D., & Rule, D. C. (2004). An integrative approach to personal epistemology: A guiding model. *Educational Psychologist*, *39*(1), 69-80.

Bendixen, L. D., Schraw, G., & Dunkle, M. E. (1998). Epistemic Beliefs and Moral Reasoning. *Journal of Psychology: Interdisciplinary and Applied, 132*(2), 187-200.

Bereiter, C. (1995). A Dispositional View of Transfer. In A. McKeough, J. Lupart & A. Marini (Eds.), *Teaching for Transfer: Fostering Generalization in Learning* (pp.21-34). Mahwah, NJ: Lawrence Erlbaum Associates.

Bernard, R. M., Abrami, P. C. & Wade, C. A. (2007). A Summary of "Review of E-learning in Canada: A Rough Sketch of the Evidence, Gaps and Promising Directions". *Horizons,* 9(2), 32-36.

Bransford, J. D. & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (Vol. 24, pp. 61–100). Washington, DC: American Educational Research Association.

Bransford, J. D., Vye, N. J., Stevens, R., Kuhl, P., Schwartz, D. L., Bell, P., Meltzoff, A., Barron, B. J., Pea, R., Reeves, B., Roschelle, J., & Sabelli, N. (2006). Learning theories and education: Towards a decade of synergy. In P. A. Alexander and P. H. Winne (Eds.), *Handbook of Educational Psychology* (pp. 209-244). Mahwah, NJ: Erlbaum.

Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in

Creating Complex Interventions in Classroom Settings. *Journal of the Learning Sciences*, 2, 141–178.

Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. In K.McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229-270). Cambridge, MA: MIT Press/Bradford Books.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 18, 32-42.

Cañas, A. J., Hill, G., Carff, R., Suri, N., Lott, J., Eskridge, T.,&Carvajal, R. (2004).CmapTools: A knowledge modeling and sharing environment. In Concept maps: Theory, methodology, technology. *Proceedings of the first international conference on concept mapping* (Vol. 1, pp. 125-133).

Charles, E.S. (2009). Learning Through a Community of Practice Approach. *Pédagogie Collégiale*, special issue

Charles, E.S., Lasry, N. (2010). *Who's talking in your classroom? Two sides of the same pedagogical challenge*. Paper presented at 30th annual AQPC symposium: Sherbrooke, QC.

Charles, E.S., Lasry, N., Whittaker, C. & Trudeau, J. (August 2009). *Technology Supported Collaboration and Learning: How do we build learning environments to build communities & conceptual knowledge?* Technical report for PA2007-014 (ISBN number), for funding agency PAREA, coordinated through Dawson College, QC.

Choi, J., Rosen, J., Maini, S., Pierce M., & Fox, G., (2008). Collective Collaborative Tagging System. *IEEE Grid Computing Environments Workshop, GCE '08*, November 12-16, Austin.

Chi, M. T., Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science*, *18*(3), 439-477.

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*(1), 66-71.

Cobb, P. (2002). Reasoning with tools and inscriptions. *The Journal of the Learning Sciences*, 11(2&3), 187-215.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., Schauble, L. (2003). Design experiments in educational research. Educational Researcher, 32, 1: 9 – 13.

Coletta, V., & Phillips, J. (2005). Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability. *American Journal of Physics, 73*, 1172.

Collins, A. (In press). In M. S. Khine & I. Saleh (Eds.), *Models and modeling: Cognitive tools for scientific enquiry*. Dordrecht, The Netherlands: Springer.

Collins, A., Brown, J., & Duguid, P. (1989). "Situated Cognition and the Culture of Learning" in *Educational Researcher*, Vol. 18, No. 1, pp. 32-42.

Collins, A., Brown, J., & Holum, A. (1991). "Cognitive Apprenticeship: Making Thinking Visible" in *American Educator*, Vol. 15, No. 3, pp. 6-11, 38-46.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, *69*, 970.

Cui, L., Rebello, N. S., & Bennett, A. G. (2005). *College Students' Transfer From Calculus to Physics.* Paper presented at the Physics Education Research Conference, Salt Lake City, UT.

Dalgarno, B. (2001), Interpretations of constructivism and consequences for Computer Assisted Learning. *British Journal of Educational Technology*, 32: pp.183–194.

Dedic, H., Rosenfield, S., Alalouf, E. & Klasa, J. (2004). *Calculus and Computer-supported Collaborative Learning*. Final report submitted to PAREA. ISBN 2-9211024-64-0

Dillenbourg, P. & Fischer, F. (2007). Basics of computer-supported collaborative learning. *Zeitschrift für Berufs- und Wirtschaftspädagogik*. 21, pp.111-130.

diSessa, A. A. & Sherin, B.L. (1998). What changes in conceptual change? *International Journal of Science Education,* 20 (10), 1155-1191

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: a meta-analysis. *Learning and instruction, 13*(5), 533-568.

Engle, R.A. (2006). Framing interactions to foster generative learning: A situative explanation of transfer in a community of learners classroom. *Journal of the Learning Sciences*, 15(4), 451-498.

Fagen, A. (2003). *Assessing and Enhancing the Introductory Science Course in Physics and Biology: Peer Instruction, Classroom Demonstrations, and Genetics Vocabulary.* (PhD), Harvard University, Cambridge, MA.

Fagen, A., Crouch, C., & Mazur, E. (2002). Peer Instruction: Results from a Range of Classrooms. *PHYSICS TEACHER, 40*(4), 206-209.

Franco, G. M., Muis, K. R., Kendeou, P., Ranellucci, J., Sampasivam, L., & Wang, X. (2012). Examining the influences of epistemic beliefs and knowledge representations on cognitive processing and conceptual change when learning physics. *Learning and Instruction, 22*(1), 62-77. doi: 10.1016/j.learninstruc.2011.06.003

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 201319030.

Frenay, M. & Bédard, D. (2004). Des dispositifs de formation universitaire s'inscrivant dans la perspective d'un apprentissage et d'un enseignement contextualisés pour favoriser la construction de connaissances et leur transfert. In A. Presseau & M. Frenay (Eds.), *Le transfert des apprentissages,* (p. 239-267). Québec, Presses de l'Université Laval.

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Gee, J. P. (2000). Identity as an Analytic Lens for Research in Education. *Review of Research in Education, January 2000*(25), 99-125.

Gentner, D. (1989). The mechanisms of analogical learning. *Similarity and analogical reasoning*, *199*, 241.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, **10**, 277–300.

Graesser, A. C., VanLehn, K., Rosé, C. P., Jordan, P. W., & Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI magazine*, *22*(4), 39.

Gray, J. (2007). The Fourth Paradigm: Data-Intensive Scientific Discovery. *In* T. Hey, S. Tansley & K. Tolle (Eds.), The Fourth Paradigm: Data-Intensive Scientific Discovery http://research.microsoft.com/en-us/um/people/gray/talks/NRC-CSTB_eScience.ppt.

Greeno, J.G. (2006). Authority, accountable positioning and connected, general knowing: Progressive themes in understanding transfer. *Journal of the Learning Sciences*, 15(4), 539-547.

Greeno, J. G. (1997). Response: On claims that answer the wrong questions. *Educational Researcher*, 26(1), 5–17.

Greeno, J. G., Smith, D. R., &Moore, J. L. (1993). Transfer of situated learning. In D. K.

Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 99–167). Norwood, NJ: Ablex.

Gertner, A. S., & VanLehn, K. (2000, January). Andes: A coached problem solving environment for physics. In *Intelligent Tutoring Systems* (pp. 133-142). Springer Berlin Heidelberg.

Gresalfi, M. S. (2009). Taking Up Opportunities to Learn: Constructing Dispositions in Mathematics Classrooms. *Journal of the Learning Sciences*, 18: 3, 327 — 369

Gresalfi, M. S., & Cobb, P. (2006). Cultivating students' discipline-specific dispositions as a critical goal for pedagogy and equity. *Pedagogies*, *1*(1), 49-57.

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.

Halloun, I. A., & Hestenes, D. (1985a). Common-Sense Concepts About Motion. *American Journal of Physics, 53*(11), 1056-1065.

Halloun, I. A., & Hestenes, D. (1985b). The Initial Knowledge State of College Physics Students. *American Journal of Physics, 53*(11), 1043-1055.

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 169-190). Mahwah, NJ: Erlbaum.

Hansen, F. T. (2012). Epistemic artifacts: The potential of artifacts in design research. http://www.flemmingtvede.dk

Hao, J. X., Kwok, R. C. W., Lau, R. Y. K., & Yu, A. Y. (2010). Predicting problem-solving performance with concept maps: An information-theoretic approach. *Decision Support Systems*, *48*(4), 613-621.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30, 141-158 .

Hoadley, C. (2002). Creating context: Design-based research in creating and understanding CSCL. In G. Stahl (Ed.), *Computer Support for Collaborative Learning,* (pp. 453-462). Mahwah, NJ: Lawrence Erlbaum Associates.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140.

Jacobson, M. J., & Archodidou, A. (2000). The design of hypermedia tools for learning: Fostering conceptual change and transfer of complex scientific knowledge. *The Journal of the Learning Sciences*, *9*(2), 145-199.

Jehng, J. J., Johnson, S. D., & Anderson, R. C. (1993). Schooling and students beliefs about learning. *Contemporary Educational Psychology, 18*, 23-35.

Kardash, C. M., & Howell, K. L. (2000). Effects of epistemological beliefs and topic-specific beliefs on undergraduates cognitive and strategic processing of dual-positional text. *Journal of Educational Psychology, 92*, 524-535. doi: 10.1037/0022-0663.92.3.524

Kim, E., & Pak, S. (2002). Students do not overcome conceptual difficulties after solving 1000 traditional problems. *American Journal of Physics, 70*, 759.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco, CA: Jossey-Bass.

King, P. M., & Kitchener, K. S. (2004). Reflective Judgment: Theory and Research on the Development of Epistemic Assumptions Through Adulthood. *Educational Psychologist, 39*(1), 5-18. doi: 10.1207/s15326985ep3901_2

Kravcik M., Kaibel, A., Specht, M., and Terrenghi, L. (2004). Mobile Collector for Field Trips. Educational Technology & Society, 7 (2), 25-33.

Lasry, N. (2006). PAREA Report: Implementing Peer Instruction in Cegep (L'enseignement par les pairs au cégep) *PAREA Report* (pp. 69). Montreal, QC: John Abbott College.

Lasry, N. (2008). Une mise en œuvre au cégep de la méthode d'apprentissage par les pairs de Harvard. *Pédagogie Collégiale, 21*(4), 21-28.

Lasry, N., Guillemette, J., & Mazur, E. (2014). Two steps forward, one step back. *Nature Physics, 10*(6), 402-403. doi: 10.1038/nphys2988

Lasry, N., Mazur, E., & Watkins, J. (2008). Peer instruction: From Harvard to the two-year college. *American Journal of Physics, 76*(11), 1066-1069. doi: Doi 10.1119/1.2978182

Lasry, N., Rosenfield, S., Dedic, H., Dahan, A., & Reshef, O. (2011). The puzzling reliability of the Force Concept Inventory. *American Journal of Physics, 79*, 909.

Lasry, N., & Aulls, M. (2007). The effect of multiple internal representations on context-rich instruction. *American Journal of Physics, 75*(11), 1030-1037.

Lasry, N., Charles, E.S., Whittaker, C., & Lautman, M. (2009). When Talking Is Better Than Staying Quiet. In M. Sabella, C. Henderson & C. Singh (Eds.), *American Institute of Physics (AIP) Conference Proceedings, Vol 1179*, pp. 181-184. ISBN: 978-0-7354-0720-6.

Lave, J. (1988). *Cognition in Practice: Mind, mathematics, and culture in everyday life*. Cambridge, UK: Cambridge University Press.

Lave, J., & Wenger, E. (1990). *Situated Learning: Legitimate Peripheral Participation*. Cambridge, UK: Cambridge University Press.

Lee, A. (2009). Development and evaluation of clicker methodology for introductory physics courses. (Doctoral dissertation). Available from comPADRE. PER/document/ServeFile.cfm?ID=10358&DocID=1838

Lee, A., Ding, L., Reay, N. W., & Bao, L. (2011). Single-concept clicker question sequences. *The Physics Teacher*, *49*(6), 385-389.

Lobato, J. E. (2003). "How Design Experiments Can Inform a Rethinking of Transfer and Vice Versa." Educational Researcher **32**(1): 17-20.

Lobato, J. (2006). Alternative Perspectives on the transfer of learning: history, issues, and challenges for future research. *The Journal of the Learning Sciences, 15*(4), 431–449.

Lui, M., Tissenbaum, M., & Slotta, J.D. (2011 to be presented) *Helping Students Solve Physics Problems Conceptually: The Impact of Collaborative Tagging in a Smart Classroom Environment*. Paper submitted to the annual meeting of the American Educational Research Association, New Orleans, LA.

Marx, J. D., & Cummings, K. (2007). Normalized change. *American Journal of Physics, 75*, 87.

Mazur, E. (1997). *Peer instruction : a user's manual*. Upper Saddle River, N.J.: Prentice Hall.

Mazur, E. (2009). Education Farewell, Lecture? *Science, 323*(5910), 50-51. doi: Doi 10.1126/Science.1168927

McDermott, L., & Redish, E. (1999). Resource letter: PER-1: Physics education research. *American Journal of Physics, 67*, 755.

Meltzer, D. E., & Thornton, R. K. (2012). Resource letter ALIP–1: active-learning instruction in physics. *American journal of physics*, *80*(6), 478-496.

Miller, K., Lasry, N., Reshef, O., Dowd, J., Araujo, I., & Mazur, E. (2010). *Losing it: The Influence of Losses on Individualsí Normalized Gains*.

Minstrell, J. (1982). Explaining the''at rest''condition of an object. *The Physics Teacher, 20*, 10.

Muis, K. R. (2008). Epistemic profiles and self-regulated learning: Examining relations in the context of mathematics problem solving. *Contemporary Educational Psychology, 33*(2), 177-208.

Muis, K. R., & Franco, G. M. (2009). Epistemic beliefs: Setting the standards for self-regulated learning. *Contemporary Educational Psychology, 34*(4), 306-318.

Muis, K. R., & Franco, G. M. (2010). Epistemic profiles and metacognition: support for the consistency hypothesis. *Metacognition and Learning*, 1-19.

Murphy, P. K., Alexander, P. A., Greene, J. A., & Hennessey, M. N. (2012). Examining epistemic frames in conceptual change research: implications for learning and instruction. *Asia Pacific Educational Review, 13*, 475-486. doi: 10.1007/s12564-011-9199-0

Novak, G. M.. Patterson, E.T., Gavrin, A.D., & Christian, W. *Just-In-Time-Teaching: Blending Active Learning with Web Technology*, Prentice Hall, 1999.

Novak, J.D. & Wandersee, J.D., (Eds.), (1990).Perspectives on concept mapping. *Journal of Research in Science Teaching*, 20 (10); Special Issue.

Oldenburg, Ray (1991). *The Great Good Place.* New York: Marlowe & Company.

Palincsar, A.S., & Brown, A.L., (1984). Reciprocal teaching of comprehension monitoring activities. *Cognition and Instruction* 1:117-175.

Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process.*Research and Practice in Technology Enhanced Learning*, *2*(1), 51-74.

Perkins, K., Adams, W., Pollock, S., Finkelstein, N., & Wieman, C. (2004). Correlating Student Attitudes With Student Learning Using The Colorado Learning Attitudes about Science Survey. *submitted to PERC Proceedings*.

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme.* New York: Holt, Rinehart, and Winston.

Peters, V. L., & Slotta, J. D. (2010). Scaffolding knowledge communities in the classroom: New opportunities in the Web 2.0 era. In M. J. Jacobson & P. Reimann (Eds.), *Designs for learning environments of the future: International perspectives from the learning sciences* (pp. 205-232). Secaucus, NJ: Springer.

Pfundt, H., & Duit, R. (1988). Bibliography. Students' alternative frameworks and science education.

Ploetzner, R., Dillenbourg, P., Preier, M., & Traum, D. (1999). Learning by explaining to oneself and to others. *Collaborative learning: Cognitive and computational approaches*, 103-121.

Rebello, N. S., D. A. Zollman, et al. (2005). A Model for Dynamic Transfer of Learning. Annual Meeting of the National Association for Research in Science Teaching, Dallas, TX, NARST Publications.

Redish, E., Saul, J., & Steinberg, R. (1998). Student expectations in introductory physics. *American Journal of Physics, 66*(3), 212-224.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change.*The journal of the learning sciences*, *2*(3), 235-276.

Ruiz-Primo, M. and Shavelson, R. (1996).Problems and issues in the use of concept maps in science assessment.*Journal of Research in Science Teaching*, 33 (6)pp. 569-600.

Trigwell, K., Prosser, M. & Ginns, P. (2005) Phenomenographic pedagogy and a revised Approaches to Teaching Inventory, *Higher Education Research and Development*, 24 (4) 349-360.

Scardamalia, M. & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York: Cambridge University Press.

Schauble, L., Glaser, R., Duschl, R., Schulze, S., & John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. *Journal of the Learning Sciences, 4*(2), 131-166.

Schommer, M. (1990). Effects of Beliefs About the Nature of Knowledge on Comprehension. *Journal of Educational Psychology, 82*(3), 498-504.

Schommer, M., Crouse, A., & Rhodes, N. (1992). Epistemological Beliefs and Mathematical Text Comprehension: Believing It Is Simple Does Not Make It So. *Journal of Educational Psychology, 84*(4), 435-443.

Schvaneveldt R.W, Durso F.T. & Dearholt (1989) Network Structures in Proximity Data *The Psychology of Learning and Motivation* Vol 24 pp249-284

Shaffer, D., Hatfield D., Svarovsky G., Nash P., Nulty A., Bagley E., Frank K., Rupp A., Mislevy R. (2009). Epistemic Network Analysis: A Prototype for 21st-Century Assessment of Learning. *International Journal of Learning and Media* 1(2)pp33-53

Singh, C. (2008). Assessing student expertise in introductory physics with isomorphic problems. II. Effect of some potential factors on problem solving and transfer. Physics Education Research, 4, 010105-5-10.

Sinatra, G. M., & Pintrich, P. R. (2003). The role of intentions in conceptual change learning. *Intentional conceptual change*, 1-18.

Soloway, E., Norris, C., Blumenfeld, P., Fishman, B., Krajcik, J., & Marx, R. (2001). *Devices* are Ready-at-Hand. Communications of the ACM (June).

Songer, N.B. (2006) BioKIDS: An Animated Conversation on the Development of Complex Reasoning in Science. In R. Keith Sawyer, (Ed.) *Cambridge Handbook of the Learning Sciences.* New York: Cambridge University Press. P. 355-369

Slotta, J. D. (2010). Evolving the classrooms of the future: The interplay of pedagogy, technology and community. In Mäkitalo-Siegl, K., Kaplan, F., Zottmann, J. & Fischer, F. (Eds.). *Classroom of the Future: Orchestrating collaborative spaces* (pp. 215-242). Rotterdam: Sense.

Slotta, J. D., & Najafi, H. (2010). Knowledge communities in the classroom. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education, Volume 8* (pp. 189-196). Oxford: Elsevier.

Slotta, J. D., & Linn, M. C. (2009). *WISE science: Web-based inquiry in the classroom*. New York: Teachers College Press.

Slotta, J. D., & Aleahmad, T. (2009). WISE technology lessons: Moving from a local proprietary system to a global open source framework. *Research and Practice in Technology Enhanced Learning,4*(2),169-189.

Songer, N. B. (2006). *BioKIDS: An animated conversation on the development of curricular activity structures for inquiry science*.

Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.

Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. *Cambridge handbook of the learning sciences*, *2006*.

Stake, R.E. (1998). Case Studies. In N.K. Denzin and Y.S. Lincoln (Eds.), *Strategies of qualitative inquiry* (pp. 86-109). Thousand Oaks, CA: Sage Publication.

Sumi, Y., Etani, T., Fels, S., Simonet, N., Kobayashi, K., & Mase, K. (1998). C-map: Building a context-aware mobile assistant for exhibition tours. In *Community computing and support systems* (pp. 137-154). Springer Berlin Heidelberg.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285.

Tissenbaum, M. & Slotta, J. D. (2009, June). *A new framework for smart classroom research: Co-designing curriculum, research and technology*. Poster presented at the annual Computer Supported Collaborative Learning (CSCL) conference CSCL 2009, Rhodes, Greece.

Treagust, D. F., & Duit, R. (2008). Conceptual change: A discussion of theoretical, methodological and practical challenges for science education. *Cultural Studies of Science Education*. Published online 29 March 2008: doi:10.1007/s11422-008-9090-4.

Tsai, C.-C. (2004). Beyond cognitive and metacognitive tools: The use of the Internet as an "epistemological" tool for instruction. *British Journal of Educational Technology, 35*(5), 525-536. doi: 10.1111/j.0007-1013.2004.00411.x

Vernon, D., & Blake, R. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Academic medicine, 68*(7), 550.

Viennot, L. (1979). Spontaneous Reasoning in Elementary Dynamics. *International Journal of Science Education, 1*(2), 205-221.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*, Cambridge: Cambridge University Press.

Wheeldon, J. P., &Faubert, J. (2009). Framing experience: concept maps, mind maps, and data collection in qualitative research.*International Journal of Qualitative Methods*, 8(3), pp52-67.

Wieman, C. & Perkins, K. (2005). Transforming physics education *Physics Today,* 58, 26–41. Downloaded from *menem.com/~ilya/wiki/images/f/fc/Wieman-perkins-05.pdf (January 15, 2011).*

Wieman, C., & Perkins, K. (2005). Transforming physics education. *Physics Today, 58*(11), 36.

Zimmerman, J., & Slotta, J. (2010). *Promoting 21st Century Science: Technology-Enhanced Learning Across Formal and Informal Environments*. NSF funding proposal 2009.

APPENDICES FOR REPORT PA2011-006

Appendix A - DALITE questions from 1D Kinematics (1DKin)

1DKinQ4:

You throw a ball vertically upward. Which statement best describes the ***direction and magnitude*** of the ball's ***acceleration*** while the ball is still moving up?

1. Upward, constant magnitude
2. Upward, increasing magnitude
3. Upward, decreasing magnitude
4. Downward, constant magnitude
5. Downward, increasing magnitude
6. Downward, decreasing magnitude
7. Zero acceleration

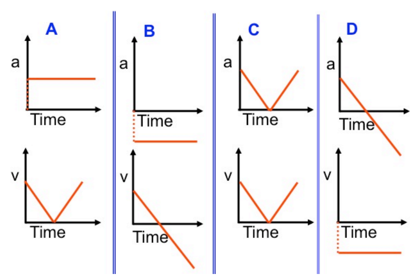
1DKinQ5

You throw a ball vertically upward. Which statement best describes the ***direction and magnitude*** of the ball's ***velocity*** while the ball is still moving up?

1. Upward, constant magnitude
2. Upward, increasing magnitude
3. Upward, decreasing magnitude
4. Downward, constant magnitude
5. Downward, increasing magnitude
6. Downward, decreasing magnitude
7. Zero velocity

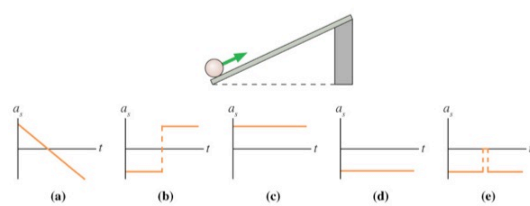
1DKinQ6:

You throw a ball vertically upward. The ball moves up, reaches its highest point, and finally falls down. Which of the following graphs best represents the ball's **acceleration and velocity vs. time** during the process?



1DKinQ8:

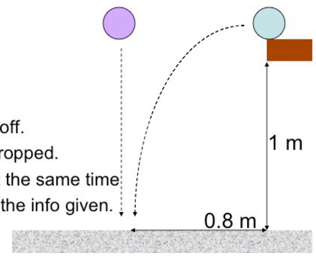
The ball rolls up the ramp, then back down. Which is the correct acceleration graph?



Appendix B - DALITE questions from 2D Kinematics (2DKin)

2DKinQ1

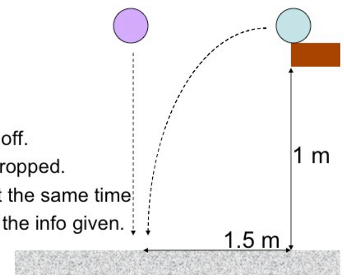
A ball is rolled off the edge of a table that is 1 meter high. It lands on the floor 0.8 meters away from the edge of the table. At the exact moment another ball rolls off the edge of the table, ball X is **released from rest** at a height of 1.0 meters from the floor. Which ball will hit the floor first?



1. The ball that rolled off.
2. The ball that was dropped.
3. They will both hit at the same time
4. You can't tell from the info given.

2DKinQ2

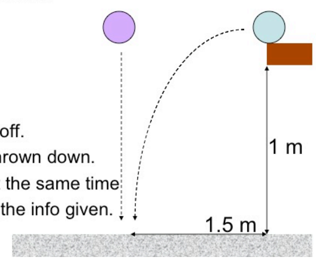
The same ball is rolled off the edge of the 1 meter high table again. This time it is rolling faster so it lands on the floor farther away from the table at 1.5 meters away from the edge of the table. If another ball is again **released from rest** at a height of 1.0 meters the instant the first ball rolls off the table, which ball will hit the floor first?



1. The ball that rolled off.
2. The ball that was dropped.
3. They will both hit at the same time
4. You can't tell from the info given.

2DKinQ3

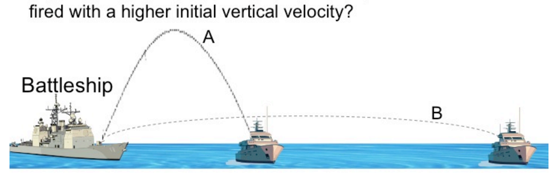
The same ball is rolled off the edge of the 1 meter high table again. This time it is rolling faster so it lands on the floor farther away from the table at 1.5 meters away from the edge of the table. If another ball is again **thrown** straight down from a height of 1.0 meters the instant the first ball rolls off the table, which ball will hit the floor first?



1. The ball that rolled off.
2. The ball that was thrown down.
3. They will both hit at the same time
4. You can't tell from the info given.

2DKinQ4

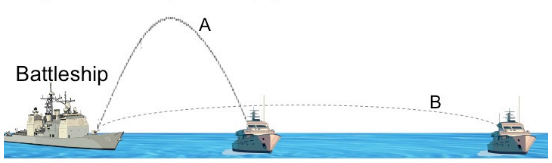
A battleship simultaneously fires two shells with the **same** initial speed at enemy ships. If the shells follow the parabolic trajectories shown, which trajectory corresponds to the shell fired with a higher initial vertical velocity?



1. A
2. B
3. Both have the same initial vertical velocity.
4. Not enough info is given.

2DKinQ5

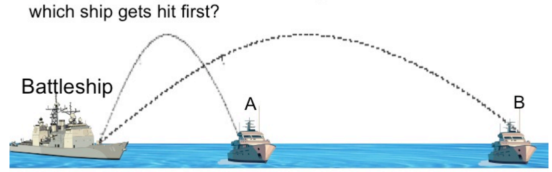
A battleship simultaneously fires two shells with the **same** initial speed at enemy ships. If the shells follow the parabolic trajectories shown, which ship gets hit first?



1. A
2. B
3. Both are hit simultaneously.
4. Not enough info is given.

2DKinQ6

A battleship simultaneously fires two shells with **different** initial speeds at enemy ships. If the shells follow the parabolic trajectories with **same** maximum height as shown below, which ship gets hit first?

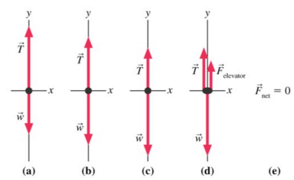


1. A
2. B
3. Both are hit simultaneously.
4. Not enough info is given.

Appendix C - DALITE questions from Dynamics (LinDyn)

LinDynQ3

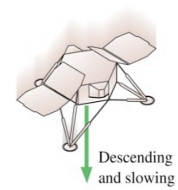
Q: An elevator suspended by a cable is moving upwards and slowing to a stop. Which FBD is correct? Friction is negligible.

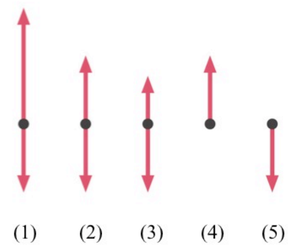


A. A
B. B
C. C
D. D
E. E

LinDynQ7

A Martian lander is approaching the surface. It is slowing its descent by firing its rocket motor. Which is the correct free-body diagram for the lander?





LinDynQ16

A cart with mass m_2 is connected to a mass m_1 using a string that passes over a frictionless pulley, as shown below. Initially, the cart is held motionless. After the cart is released, the tension in the string

1. Increases.
2. Decreases.
3. Remains the same.
4. Cannot tell from the information given.

LinDynQ20

Consider a person standing in an elevator that is accelerating upward. The upward normal force *N* exerted by the elevator floor on the person is …

a) larger than
b) identical to
c) smaller than

the downward weight *W* of the person.

Appendix D – DALITE questions from Momentum & Energy

MomentumQ1

A tow truck is pulling a stalled, massive SUV. A second identical tow truck is pulling a stalled compact light-weight car. Both tow trucks pull with the same amount of horizontal force. Both the SUV and the compact car start at rest. After the tugboats have been pulling for the same amount of time, which of the following is true about the SUV & the compact car?

a. The SUV will have a greater magnitude of momentum
b. The compact car will have a greater magnitude of momentum
c. Both ships will have the same magnitude of momentum
d. Both ships will have the same kinetic energy
e. The SUV will have a greater kinetic energy

MomentumQ7

A cart collides and sticks to a stationary cart of equal mass and they both travel at speed v. If the cart were to instead bounce off the stationary cart, the stationary cart would move at a speed:

A) greater than v
B) less than v
C) equal to v
D) equal to zero
E) don't know

MomentumQ11

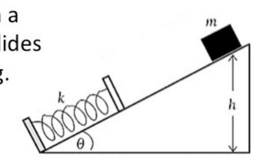
At an amusement park, you decide to play a game where you throw balls at a block in order to knock it over. You are given the choice to throw one of two balls. Each ball has the same mass, however one will stick to the block and the other will bounce.

Which should you choose to throw in order to give yourself the best chance of knocking the block over? Assume that you will throw either ball at the same speed.

a) The ball that will stick
b) The ball that will bounce
c) Either ball will give you equal chances of knocking the block over.
d) Depends on if block is lighter or heavier that balls.

EnergyQ15

A block is released from rest on a frictionless incline, as shown. It slides down and compresses a spring. Which statement is true?



A. None of the below
B. The kinetic energy of the block just before it collides with the spring will be equal to *mgh.*
C. The kinetic energy of the block when it has fully compressed the spring will be equal to *mgh.*
D. The kinetic energy of the block just before it collides with the spring will be $\frac{1}{2}kx^2$, where x is the maximum compression of the spring.

Appendix E – Tagging Exercise worksheet

Kinematics Name:____________________

A battleship simultaneously fires two shells with the **same** initial speed at enemy ships. If the shells follow the parabolic trajectories shown, which ship gets hit first?

1. A
2. B
3. Both are hit simultaneously.
4. Not enough info is given.

What phenomena are we seeing?

- Free-fall
  - Air drag negligible
  - Air drag not negligible
  - Near Earth's surface
  - Not near Earth's surface

- Motion on an inclined surface
  - Friction negligible
  - Friction not negligible
  - Gravity is only force pulling
  - Some force (other than gravity) pushing/pulling?

- Motion on a horizontal surface
  - Friction negligible
  - Friction is not negligible

- Motion described as a function of time
  - One dimensional
  - Multidimensional

- Projectile motion (in two dimensions)
  - Air friction negligible
  - Air friction not negligible
  - Something (other than gravity) pushing/pulling

- Circular motion
  - Horizontal circle
  - Vertical circle
    - Gravity is only force pulling
    - Something (other than gravity) pushing/pulling
      - Parallel to path
      - Perpendicular to path
      - Has components both parallel and perpendicular

Important quantities, definitions and special cases

- Position
  - ✓ *Equal to location of dot/point on graph or motion diagram*
  - Horizontal
    - Initial horizontal position is zero
    - Initial horizontal position of both objects are the same
  - Vertical
    - Initial vertical position is zero
    - Initial vertical position of both objects are the same
  - Angular
    - Initial angular position is zero
    - Initial angular position of both objects are the same

- Distance
  - ✓ *Equal to sum of the displacement values*

- Displacement
  - ✓ *Equal to the difference in position*
  - ✓ *Equal to the area under a velocity graph*

- Speed
  - ✓ *Related to distance between dots on a motion diagram*
  - ✓ *Equal to steepness of position function*

Kinematics   Name:_______________________

- ✓ *Speed decreases when velocity and acceleration are in opposite directions*
- ✓ *Speed increases when velocity and acceleration are in the same direction*

- Velocity
  - Horizontal
    - Initial horizontal velocity is zero
    - Initial horizontal velocity of both objects are the same
    - Initial horizontal velocity of both objects are different
  - Vertical
    - Initial vertical velocity is zero
    - Initial vertical velocity of both objects are the same
    - Initial vertical velocity of both objects are different
  - Angular
    - Initial angular velocity is zero
    - Initial angular velocity of both objects are the same
    - Initial angular velocity of both objects are different
  - Average (1D only)
    - ✓ Equal to displacement/time interval
    - ✓ Related to distance between dots on a motion diagram
  - Instantaneous
    - ✓ Equal to slope of position graph
    - ✓ Equal to derivative of position function
    - Constant
    - Not constant

- Acceleration
  - Average (1D only)
    - *Equal to change in velocity/time interval*
    - *Related to the change in distance between dots on a motion diagram*
  - Instantaneous
    - *Equal to derivative of velocity function*
    - *Equal to slope of velocity graph*
    - *Equal to curvature of position function*
  - Object is slowing
    - *Direction of acceleration and velocity are opposite*
  - Object is speeding up
    - *Direction of acceleration and velocity are parallel*
  - Constant
    - has magnitude = g
    - has magnitude = 0
    - has magnitude = g sinθ
    - has magnitude = v^2/R
  - Constant during intervals but changes suddenly in-between
  - Not constant

Answer:

Rationale:

Appendix F – Student Interviews

1. What do you believe is the purpose of the DALITE assignments?
    a. If the answer includes "learning" follow up with:
        - How do the assignments help you with your learning? Elaborate.
    b. If the answer does NOT include "learning" follow up with:
        - From your experience, what role has DALITE assignments played in your learning the content of this course, if any? Can you elaborate?
2. What <u>role has the in-class explanations</u> of DALITE assignments played in helping you learn, if any? With helping your understanding the concepts better?
    a. Follow up with:
        - What else could we do to make this part better?
3. What role has the <u>in-class tagging activities</u> played in your learning, if any?
    a. Follow up with:
        - Role of <u>the sorting tasks</u>?
        - Do you see similarities between these tasks and anything that is associated with DALITE?
        - Do you see similarities between DALITE questions and any other of the things you have been doing in this class?
4. What role has the <u>group work</u> with the in-class tagging activity played in your learning, if any?
5. Have you discussed your DALITE assignments other times in class? Elaborate if yes.
6. What would you suggest as ways to improve the usefulness of DALITE?

Additional questions:

7. Do you usually read the rationales in DALITE? Do you use these rationales to help you learn?

8. How similar is this process to the one you do when doing the clicker questions?

9. How long does it take to complete a DALITE question?

10. How do you use the “good” rationale? How useful is it to the way you use DALITE?

11. Do you have a problem with NOT seeing the “right” answer? i.e., waiting until class.

12. How difficult is it to write rationales? Has this changed over the course of the course?

13. Has the ability to read the rationales changed over time?

14. How different has this course been from other courses you’ve taken before in science?

L'équipe de recherche qui soumet ce rapport fût productive lors des trois dernières années ayant soumis plus de deux douzaines de publications, présentations à des conférences ainsi que plusieurs conférences invitées. En voici quelques-unes :
The research team submitting this report has had an active three years with over two dozen publications, conference presentations and invited talks between them. Below we list a few :

Publications:

Charles, E.S, Whittaker, C., Lasry, N., Dugdale, M., Lenton, K., Bhatnagar, S., Guillemette, J. (2014). Taking DALITE To the Next Level: What have we learned from a web-based peer instruction application. Polman,P.L., Kyza, E.A., O'Neill, D.K., Tabak, I., Penuel, W.R., Jurow, A.S., O'Connor, K., Lee, T., & D'Amico, L. (Eds.). *Learning and Becocming in Practice: ICLS2014 Conference Proceedings, Volume II*, pp.982-986. International Society of the Learning Sciences.

Lasry, N., Charles, E., & Whittaker, C. (2014). When teacher-centered instructors are assigned to student-centered classrooms. *Physical Review Special Topics-Physics Education Research*, 10(1), 010116.

Lasry, N., Guillemette, J., & Mazur, E. (2014). Two steps forward, one step back. *Nature Physics,* 10(6).

Lasry, N., Dugdale, M. et Charles, E. (2014). Zut! J'ai renversé ma pédagogie. *Pédagogie Collégiale*, 27, hiver.

Lasry, N., Dugdale, M. & Charles, E. (2014). Just in Time to Flip Your Classroom. *The Physics Teacher*, 52(1): 34-37.

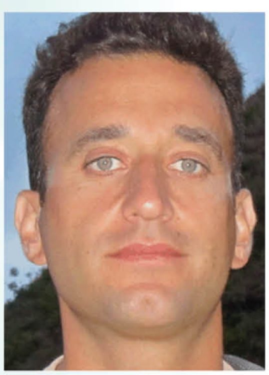
Charles, E., Lasry, N., et Whittaker, C. (2013). L'adoption d'environnements sociotechnologiques comme moteur de changement pédagogique. *Pédagogie Collégiale*, 26 (3).

Conférences:

Charles, E.S., Whittaker, C., & Lasry, N. (2014). Taking DALITE To the Next Level. Paper presented at *33rd annual AQPC symposium*: Laval, QC.

Charles, E., Whittaker, C., & Lasry, N. (2014). Active Learning Classroom Use in College Level Physics: The impact of technology design and adaptive orchestration. Symposium entitled: Space and Technologies for Learning in Schools, Museums and Workplaces: Recent Approaches in Design-Based Research. Paper presented at the *American Educational Research Association* (AERA): Philadelphia, PA.

Whittaker, C. & Charles, E.S. (2013). The Perfect Marriage - Technology and Pedagogy in a Next-Generation Active Learning Classroom. Paper presented at *32nd annual AQPC symposium*: Montreal, QC.



L'équipe de recherche PAREA - PA2011-06 : Research team for PAREA - PA2011-06
Dr. E.S. Charles, Dr. N. Lasry, Dr. K. Lenton, C. Whittaker, M. Dugdale, et S. Bhatnagar

Remerciement à nos adjoints de recherche : Acknowledgements to our research assistants
Dr. J. Guillemette, C. Zhang, et X. Wang