KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Science

Learning of physics and mathematics concepts in an integrated STEM curriculum

Stijn Ceuppens

Supervisors:
Prof. dr. Mieke De Cock
Prof. dr. Johan Deprez
Prof. dr. ir. Wim Dehaene

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Science (PhD): Physics

June 2019

Learning of physics and mathematics concepts in an integrated STEM curriculum

Stijn CEUPPENS

Examination committee:
Prof. dr. Hans Van Winckel, chair
Prof. dr. Mieke De Cock, supervisor
Prof. dr. Johan Deprez, supervisor
Prof. dr. ir. Wim Dehaene, supervisor
Prof. dr. Peter Lievens
Prof. dr. Wim Van Dooren
Dr. Jan Sermeus
Prof. dr. rer. nat. Gesche Pospiech
(Technische Universität Dresden)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Science (PhD): Physics

June 2019

© 2019 KU Leuven – Faculty of Science
Uitgegeven in eigen beheer, Stijn Ceuppens, Celestijnenlaan 200c box 2406, B-3001 Leuven (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

Preface

Een doctoraat komt er niet zomaar en zeker niet alleen. Het kan enkel tot stand komen dankzij de samenwerking en steun van vele mensen. Deze zou ik hier dan ook graag eerst willen bedanken want zonder hen zou dit werk niet zijn wat het vandaag is.

Allereerst en zonder twijfel is dat Mieke. Als mijn promotor heb je alles op de voet gevolgd en hebben we de voorbije jaren nauw samengewerkt. Zonder jouw inzichten, je kritische blik, je inzet om keer op keer mijn schrijfsels te finetunen, je enthousiasme, en je steun zou dit doctoraat er niet zijn. Je zorgde voor een goede mix van de nodige vrijheid en sturing, en stond telkens achter mij bij het maken van keuzes. En alsof het een fysische wetmatigheid is, kent elk doctoraat ups en downs, maar telkens heb je mij daar doorheen geholpen en stond je er wanneer dat nodig was, ook al had dat niet altijd met het onderzoek te maken. Dat niet elk gesprek over het onderzoek ging maar dat er ook aandacht was voor het leven daarbuiten is een onmiskenbare eigenschap van een goede promotor! Voor dat alles, voor nog veel meer, en om al deze jaren mijn gids te zijn, bedankt!

Johan en Wim, mijn twee copromotoren wens ik ook graag te bedanken. Johan, je zorgde voor de wiskundige balans in het verhaal door ook de wiskundige insteek te bewaken. Je feedback, je inzichten en je rustige to-the-point aanpak heb ik erg gewaardeerd. Dankzij jou zijn de fysica en de wiskunde weer wat meer geïntegreerd! Wim, als ingenieur en als grote probleemoplosser was het je al snel duidelijk dat drie promotoren wel eens een nodeloos complex probleem zou kunnen zijn. De oplossing was dan ook mooi en simpel, met Mieke en Johan als eerstelijnshulp, jezelf als tweedelijnshulp met een volledig vertrouwen in je collega's en in mezelf was mijn promotorenteam een succes. Bedankt voor je hulp en voor het vertrouwen! Ik kijk uit naar onze verdere samenwerking de komende jaren want misschien kan je deze fysicus wel wat engineering skills bijbrengen!

Dear members of the jury: Wim, Peter, Jan and Gesche, I wish to thank you for critically reviewing my dissertation and for providing me with valuable feedback and interesting questions. Wim, jou bedank ik graag extra voor je goede raad over statistiek, je waardevolle kritische blik en voor het reviewen van de testen. Hans, bedankt om de taak als voorzitter op jou te nemen, en om alles in goede banen te leiden. To the full examination committee: thank you for your time and effort, and for your interesting feedback which has helped me to lift my dissertation to a higher level.

Zonder project geen team en zonder team geen project. Ik wil dan ook graag het IWT bedanken om het SBO-project STEM@school mogelijk te maken. Dit project werd gedragen door vele collega's die elk op hun manier een unieke bijdrage hebben geleverd. Dat zijn uiteraard de drie professoren Mieke, Wim en Peter samen met Heidi, Jelle en Marie-Paule die elk hun sterke schouders onder STEM@school gezet hebben om het tot een succes te maken en om een aangename biotoop voor de PhD's te creëren. Ook bedankt aan de koepelmedewerkers Eddy, Emile, Marc, Hans, Didier en Luc voor jullie input, samenwerking en ondersteuning voor het project. Dan waren er nog zes anderen in het team. Zes doctoraatstudenten. Toch wel vrij zeldzaam want doctoraatstudenten werken vaak op een eigen project op een eigen eilandje, maar niet in dit project. Hier waren er zes unieke mensen die aan het begin van deze bijzondere reis aan elkaar beloofd hadden om dit allemaal samen tot een goed einde te brengen. Ik ben dan ook heel blij dat ik met dit doctoraat het zesde en laatste doctoraat van deze groep kan afronden. We hebben samen onze belofte gehouden! Dat zijn Annemie en Haydée — de twee van Antwerpen — Lieve, Leen en Jolien — die samen met mij de vier van Leuven vormden. Jullie zijn alle vijf fantastische mensen. Bedankt om samen de dalen en bergen te overwinnen, voor de vele gezellige momenten en zeker niet te vergeten voor de tweedaagse bijeenkomsten die voor hilarische herinneringen gezorgd hebben. Beste STEM@school-collega's, jullie hebben de Best Research and Project Award 2017 in Finland dubbel en dik verdiend!

Er zijn er drie die hier extra in de bloemetjes gezet mogen worden: Lieve, Leen en Jolien. We hebben samen lief en leed gedeeld, we hebben elkaar rechtgehouden en plezier gemaakt. Om maar enkele keywords op te noemen: koffies, Ierse pubs, memorabele time's up avondjes, wijn, Finland, cocktails, gestrand op de luchthaven, een mini-appartementje in Helsinki, kwakende kikkers in Wange, pizza's aan het kasteel, een BBQ die maar niet wilde branden, etc. Lieve, Leen en Jolien, bedankt om zoveel meer te zijn dan collega's!

En dan is er toch nog eentje die er op het einde is bijgekomen. Geen zevende doctoraatstudent, maar toch duidelijk eentje die snel part of the PhD-team was in Leuven. Marie-Paule, je was de kers op de taart in STEM@school en het is me dan ook een hele eer om samen met jou verder te werken. Bedankt, en

alvast bedankt, we maken er een dream team van en zetten de tradities verder!

Dear new MICAS-colleagues, thank you for welcoming me in the group and for not treating me like a MICAS-newby, yet. I'm looking forward to the many coffee breaks, lunches, MICAS-days and receptions for which you are famous!

Naast het team gaat mijn dank ook uit naar alle leerkrachten en medewerkers op de scholen die deelnamen aan STEM@school. Zonder jullie was er geen project, en zonder jullie bloed, zweet, tranen en niet-aflatende inzet in ons onderwijs zouden we onze jongeren niet kunnen voorbereiden om samen mee te bouwen aan een mooie toekomst voor iedereen. Het beroep van leerkracht wordt vaak onderschat, maar jullie zijn een cruciaal puzzelstuk in de maatschappij. Jullie verdienen dan ook het diepste respect. Bedankt voor jullie inzet!

De voorbije jaren heb ik het geluk gehad om naast het STEM@school-team ook verschillende andere mensen te leren kennen. Deze zou ik hier ook graag willen bedanken. Esther voor de ondersteuning en om mijn roommate te zijn gedurende die eerste jaren; Pieter, Laurens en Sofie om all-around toffe collega-doctoraatstudenten te zijn, en voor de leuke uitstapjes op de conferenties en summer schools in Denemarken, Polen of Ierland. Laurens, extra bedankt om mijn tweede roommate te zijn en voor de hulp met het onderzoek. Ook Aldine en Marjolein bedankt voor het uitwisselen van ideeën en voor de interessante besprekingen als PhD-collega's uit de andere disciplines.

Marie-Paule, Sofie en Hans, extra bedankt om mij te helpen voorbereiden op mijn verdediging. Sofie en Hans, jullie zijn de volgende PhD's bij Mieke. Het bureau nu officieel van jullie en ik wens jullie veel succes met jullie doctoraat!

Naast het onderzoek was er gelukkig ook plaats om als assistent mee te helpen in het onderwijs van de studenten in Leuven, iets dat ik er altijd met veel plezier heb bijgenomen en waar ik veel voldoening uit kon halen. De vlotte samenwerking met alle collega-assistenten maakte dit des te aangenamer, Veerle, Houria en Jozefien, bedankt daarvoor.

I would also like to thank all the wonderful people — in addition to those already named above — who joined me on the conferences and summer schools and turned them into fond memories: Jan, Spela, Maria, Mir, Marie-Annette, Lara, John, Magdalena and many others. Special thanks go out to Magdalena because it's great fun to keep in touch and to meet up!

Om zoiets als een doctoraat aan te pakken is ook een stevige thuisbasis nodig. En die was er! Ik ben dan ook erg dankbaar voor de kansen, mogelijkheden en ondersteuning die mijn ouders mij geboden hebben. Zonder hen was dit nooit gelukt. Mama en Papa, ik heb het jullie niet altijd gemakkelijk gemaakt. Een doctoraat is nu eenmaal geen job van 9 tot 5, het kruipt in je lijf en is pas

afgerond na de verdediging. En dat was er soms aan te merken. Bedankt om er altijd te zijn voor mij en ook om mij gewoon te laten doen terwijl ik 's nachts nog bezig was en jullie eigenlijk al lang hadden willen slapen. Ook mijn broer Tom met zijn gezin, Stephanie, Louie, Camiel en Rosalie, en mijn grootouders, Moemoe, Bomma en Bompa wil ik hier graag bedanken voor de steun en de nodige afleiding. Zeker Bomma en Bompa voor de regelmatige glaasjes wijn met een hapje en een babbel in het weekend!

Bij die thuisbasis horen ook goede vrienden. Bedankt om er te zijn voor mij, om voor de nodige afleiding te zorgen, voor de vele toffe momenten, de gezellige avonden, de etentjes, de uitstapjes, de reizen en zoveel meer! Lien en Robin, Hellen, Chrisje en Koenraad, Rens en Mathias, Philippe, bedankt. En extra bedankt aan Evi en Stijn, Britt en Dominik, Stephanie en Tom, en Kelly en Kevin, dat ik op jullie huis, huisdieren (katten, honden, konijn, kippen en vissen), of kindjes mocht passen wat voor een aangename afleiding zorgde. Regelmatig op reis gaan om de wereld te ontdekken en om te ontspannen is voor mij een must. Soms alleen, meestal met anderen. Voor al die mooie herinneringen, bedankt aan Jurgen, Evi, Stijn, Nina en Inne, Hellen, Kelly, Kevin, Emma en Loren, Vanessa, Nick, Elena en Victor, Dominik, Laurens, Philippe, Dirk en Wen!

Ook extra bedankt aan de wijn- en bierclub "Half-om-half": Mieke, Chris, Nicole, Koen, Ruben, Kevin, Jurgen en ook Rik, al hebben we van jou spijtig genoeg al afscheid moeten nemen. Dat er nog veel proefavonden mogen volgen!

In heel deze groep steekt er altijd iemand met kop en schouders bovenuit. Jurgen, je bent mijn beste vriend, je bent een absolute topkerel en je bent er altijd geweest voor mij. Ik kan oprecht zeggen dat ik hier vandaag niet had gestaan zonder jou. Bedankt voor alles!

Soms is zelfs al het bovenstaande nog niet genoeg en is wat extra hulp noodzakelijk, daarom ook bedankt aan Philippe en Johannes.

En de laatsten om te bedanken maar daarom alles behalve de minsten: ik heb altijd een sterke band gehad met onze honden thuis, en mijn Border Collie Floyd, van wie ik tijdens mijn doctoraat afscheid heb moeten nemen, verdient hier ook zeker zijn plaats. Je was fantastisch! En dan rest mij nog twee mensen te bedanken die op de een of andere manier niet meer in mijn leven zijn maar die toch mee hebben getimmerd aan het pad om tot hier te geraken: Joyce en Ellen.

Allemaal van harte bedankt!

Stijn Ceuppens
24 juni 2019

Abstract

Our highly technological society is confronted with multiple complex problems such as sustainable development, healthcare, ecological impact, and mobility. To solve these problems, increasing numbers of STEM (Science, Technology, Engineering, Mathematics) professionals are needed. Moreover, such a society requires STEM literacy for all citizens. Education therefore increasingly focuses on STEM education. In doing so, it faces some problems of its own, e.g., students' low perceived relevance of STEM, a significant number of students in STEM programs in secondary education who do not want to pursue a STEM program in higher education, the difficulty to achieve deep understanding of STEM topics, a decrease of students' performance in STEM fields, a need for a research-based STEM framework, a need for qualitative learning materials, etc. With these challenges in mind, the education research community is putting its efforts into studying integrated STEM (iSTEM), in which the integration between the various STEM disciplines is a proposed solution to address these issues.

In Flanders (Belgium), the STEM@school project aimed at contributing to this research by developing a framework, designing learning materials, and performing qualitative and quantitative research on the implementation thereof in Flemish secondary schools. As part of the STEM@school project, this dissertation cooperated with the development of the framework, studied 9th grade students' learning of discipline specific concepts in an integrated curriculum, and investigated their ability to link concepts between physics and mathematics. To study the link between physics and mathematics, we investigated students' representational fluency between graphs, formulas, and tables, as well as their understanding of linear function problems in kinematics and mathematics, and compared these between an iSTEM cohort and a control cohort. Furthermore, we assessed and compared students' conceptual understanding in kinematics.

The main conclusions from this research can be summarized as follows.

Concerning the specific topics of representational fluency and conceptual understanding in kinematics and mathematics for linear function problems, the iSTEM instructional approach from the STEM@school project resulted in a similar level of performance as the regular instructional approach since no significant difference was found between the two approaches. Furthermore, our studies show that students experience significantly more difficulties in kinematics compared to mathematics on isomorphic questions. We established that the type of linear function and the representational transition have a significant effect on students' performance, with negative values for the slope or the y intercept and transitions to formulas causing the most difficulties. Additionally, we found a significant asymmetry in the direction of the representational transitions containing a formula — which are the most difficult — for the direct linear relationship — which is the easiest. This unexpected finding was mostly confined to mathematics, which is — on average — the easiest context. Our results show that the slope is the most problematic concept in kinematics and that the y intercept is the most problematic in mathematics, though the latter is partly due to students' unfamiliarity with formal terminology in mathematics. Furthermore, we found a strong presence of interval/point confusions when determining the slope of a linear graph, which almost all occurred in kinematics and nearly none in mathematics. In addition, we identified three distinct categories of interval/point confusions in questions with formulas: numerical, algebraic, and unit-based. In mathematics, when asked about the y intercept given a graph, about a fifth of all students switched to an algebraic formula, which supports the idea that students often experience a pressure to use formulas to find and explain their answers, while they actually perform better on graphical questions in general. Overall, students rarely change representation to solve a question. In contrast to the similar accuracies in mathematics, in kinematics, students from the control cohort reason with steepness significantly more often and STEM students make significantly more use of the (correct) "ratio of differences" strategy and significantly less of the (incorrect) "ratio of coordinates" strategy compared to the control cohort. Concerning the conceptual understanding of kinematics concepts, our study does show a positive effect size of $r = 0.21$ for the iSTEM cohort, which was particularly the case for the concepts displacement and distance, velocity as a 1D vector, acceleration as a rate, and acceleration as a 1D vector. Furthermore, our results show that students' conceptual understanding in kinematics is very low in both cohorts, strongly implicating that further improvements to the learning goals and the curricula are necessary.

Specifically for our detailed topics, it can be concluded that the iSTEM approach succeeded in achieving similar or improved student performance and that further improvements are necessary.

Beknopte samenvatting

Onze hoogtechnologische samenleving ervaart meerdere complexe problemen zoals duurzame ontwikkeling, gezondheidszorg, ecologische impact en mobiliteit. Om deze op te lossen, is een toenemend aantal STEM (Science, Technology, Engineering, Mathematics) professionals noodzakelijk. Daarbij aansluitend is ook STEM-geletterdheid voor alle leden van onze samenleving belangrijk. Om aan deze noden te voldoen, focust het onderwijs steeds meer op STEM. Hierbij ervaart het zelf enkele moeilijkheden, zoals de laag gepercipieerde relevantie van STEM, het significante aantal leerlingen in STEM-richtingen in het secundair onderwijs dat er niet voor kiest om een STEM-richting te volgen in het hoger onderwijs, de moeilijkheid om conceptueel inzicht te verkrijgen in STEM-inhoud, de daling in de leerlingenprestaties in STEM-vakken, de noodzaak voor een onderzoeksgebaseerd STEM-didactiek, de noodzaak voor kwalitatief leermateriaal, etc. Met deze uitdagingen in het achterhoofd zet de onderzoeksgemeenschap in op geïntegreerde STEM (iSTEM), waarin de integratie tussen de STEM-disciplines centraal staat, als mogelijke oplossing.

In Vlaanderen (België), had het STEM@school project als doel om bij te dragen aan dit onderzoek door didactiek te ontwikkelen, leermateriaal te ontwerpen en kwalitatief en kwantitatief onderzoek uit te voeren implementatie daarvan in Vlaamse secundaire scholen. Als deel van dit project werd er in dit doctoraat, naast het ontwerpen en toepassen van de iSTEM-didactiek, gefocust op het leren van vakspecifieke concepten in een iSTEM curriculum, en wordt het linken van concept tussen fysica en wiskunde bij leerlingen uit het derde middelbaar onderzocht. Specifiek werd leerlingen hun representationele vlotheid tussen grafieken, formules en tabellen, alsook hun conceptueel inzicht in lineaire functies in kinematica en in wiskunde onderzocht. Dit werd vergeleken tussen iSTEM-leerlingen en reguliere leerlingen. Daarbovenop, werd ook hun conceptueel inzicht in kinematica onderzocht en vergeleken.

De hoofdzakelijke conclusies van dit onderzoek zijn als volgt. Betreffende de specifieke onderwerpen representationele vlotheid en conceptueel inzicht in

lineaire functies in kinematica en wiskunde, werd er vastgesteld dat de iSTEM-didactiek van het STEM@school-project resulteerde in een gelijkwaardig prestatieniveau als dat van reguliere leerlingen omdat er geen significante verschillen in hun prestaties werden gevonden. We toonden aan dat leerlingen significant meer moeilijkheden ondervonden in kinematica dan in wiskunde bij isomorfe vragen. Er werd vastgesteld dat het type lineaire functie en de representatieovergang een significant effect hadden op de leerlingenprestaties. Voor negatieve waarden van het y-intercept en voor de overgangen naar een formule werden de meeste moeilijkheden gevonden. Verder vonden we een significante asymmetrie in de richting van de representationele overgangen waarin een formule voorkwam — dewelke de moeilijkste waren — voor het rechtevenredige functietype — dewelke het eenvoudigste was. Dit onverwachte resultaat kwam hoofdzakelijk voor in wiskunde, wat gemiddeld gezien de eenvoudigste context was. De resultaten toonden aan dat de richtingscoëfficiënt het moeilijkste concept was in kinematica en dat het y-intercept het moeilijkste concept was binnen wiskunde, waarbij dit laatste grotendeels te wijten was aan de beperkte vertrouwdheid met formele terminologie. We vonden een groot aantal interval/punt-verwarringen bij het bepalen van de richtingscoëfficiënt van een lineaire curve in een grafiek, wat bijna uitsluitend voorkwam in kinematica en amper in wiskunde. Er werden drie verschillende categorieën van interval/punt-verwarring in vragen met formules onderscheiden: numeriek, algebraïsch, en eenheidsgebaseerd. Specifiek bij vragen naar het y-intercept van een grafiek in wiskunde, veranderde ongeveer een vijfde van de leerlingen de representatie naar een formule, wat het idee ondersteund dat leerlingen een druk ondervinden om formules te gebruiken in hun antwoorden. Dit terwijl ze gemiddeld beter presteerden met grafieken. In het algemeen veranderden leerlingen zelden van representatie om een vraag op te lossen. In tegenstelling tot de gelijkaardige accuraties in wiskunde, redeneerden de leerlingen uit de controlegroep in kinematica significant vaker met steilheid en maakten STEM-leerlingen significant vaker gebruik van de (correcte) "verhouding van verschillen" strategie en significant minder vaak van de "verhouding van coördinaten" strategie vergeleken met de controlegroep. Betreffende het conceptueel inzicht in kinematicaconcepten, resulteerde onze studie in een positieve effectgrootte van $r = 0,21$ voor de iSTEM-groep. Dit was vooral aanwezig bij vragen over de concepten verplaatsing en afstand, snelheid als 1D-vector, versnelling als verandering, en versnelling als 1D-vector. Onze resultaten tonen ook aan dat het conceptueel begrip in kinematica van de leerlingen in beide groepen zeer laag was, wat impliceert dat verdere verbeteringen aan de leerplannen en de curiccula noodzakelijk zijn.

Specifiek voor deze gedetailleerde onderzoeksonderwerpen, kan er besloten worden dat de iSTEM-didactiek kan resulteren in gelijkwaardige of zelfs verbeterde leerlingenprestaties en dat verdere verbeteringen noodzakelijk zijn.

List of abbreviations

General abbreviations:

aso Algemeen Secundair Onderwijs (General Secundary Education)

bso Beroepssecundair Onderwijs (Professionally oriented Secondary Education

CAD Computer Aided Design

CRT Criterion-Referenced Test

CUP Conceptual Understanding Programme

DeFT Design, Functions, Tasks

df Degrees of freedom

ERIC Education Resources Information Center

FCI Force Concept Inventory

FMCE Force and Motion Conceptual Evaluation

GEE Generalized Estimating Equations

ijk **representation** The vector representation using unit vectors $\hat{i}$, $\hat{j}$ and $\hat{k}$ in three dimensions.

iSTEM integrated STEM

KCT Kinematics Concept Test

kso Kunstsecundair Onderwijs (Arts Secundary Education)

LDR Light Dependent Resistor

MBT Mechanics Baseline Test

MER Mathematic Education Research

MERs Multiple External Representations

NRT Norm-Referenced Test

PER Physics Education Research

PISA Programme for International Student Assessment

R-FCI Representational variant of the Force Concept Inventory

RFS Representational Fluency Survey

RQ Research Question

SD Standard deviation

SE Standard error

SMEC Social and Societal Ethics Committee of KU Leuven

STEAM Science, Technology, Engineering, Arts and Mathematics

STEM Science, Technology, Engineering and Mathematics

TAKS Texas Assessment of Knowledge and Skills

tso Technisch Secundair Onderwijs (Technical Secondary Education)

TUG-K Test of Understanding Graphs in Kinematics

TUV Test of Understanding Vectors

Linear functions types:

NN Linear function with a negative y intercept and a negative slope

NP Linear function with a negative y intercept and a positive slope

ZN Linear function with a zero y intercept and a negative slope

ZP Linear function with a zero y intercept and a positive slope

PN Linear function with a positive y intercept and a negative slope

PP Linear function with a positive y intercept and a positive slope

Representational transitions:

TG Transition from a table to a graph

GT Transition from a graph to a table

FT Transition from a formula to a table

TF Transition from a table to a formula

FG Transition from a formula to a graph

GF Transition from a graph to a formula

Variables and other abbreviations in the analyses:

Comp. Comparison task

Det. Determine task

FType Function type

Qtype Question type

RepT Representational transition

Repr. Representation

slopesign The sign of the slope

Contents

Preface i

Abstract v

Beknopte samenvatting vii

List of abbreviations ix

Contents xiii

List of Figures xxi

List of Tables xxvii

1 Introduction 1

1.1 Introduction 1

1.1.1 The future 1

1.1.2 The past and the present 1

1.2 Challenges in STEM education 3

1.2.1 Challenge number one: increasing the relevance of STEM and the number of STEM graduates 4

1.2.2 Challenge number two: achieving deep understanding . 5

1.3 Two birds with one stone 7
1.3.1 Integrated STEM 7
1.3.2 The STEM@school project 9
1.3.2.1 Design of the framework and learning materials 10
1.3.2.2 Assessment of the new approach 12
1.4 The focus of this dissertation 14
1.4.1 This dissertation as part of the STEM@school project . 14
1.4.2 The focus 15
1.4.2.1 Physics and mathematics 16
1.4.2.2 Representational fluency 17
1.4.2.3 Linear functions 19
1.4.2.4 Kinematics 20
1.4.2.5 Summary 20
1.5 Research questions 21
1.6 Structure of the dissertation 22

2 iSTEM education: a theoretical framework 23

2.1 Introduction 24
2.1.1 Need for integrated STEM 24
2.1.2 Current challenges 24
2.1.3 Purpose and research questions of this study 25
2.1.4 Methodology 26
2.2 Results 27
2.2.1 Theoretical grounding 27
2.2.2 Instructional practices 28
2.3 Discussion and conclusion 36
2.3.1 A framework for teaching integrated STEM 36

2.3.2 Benefits . . . 37
2.3.3 Limitations . . . 38
2.3.4 Perspectives for future research . . . 39

3 iSTEM course modules 41

3.1 Summary of the iSTEM instructional approach developed by STEM@school . . . 41
3.2 Development and implementation of iSTEM course modules . 43
3.2.1 Design process of a course module . . . 43
3.2.2 The structure of a course module . . . 43
3.2.3 Implementation . . . 45
3.3 Course module: The museum security system . . . 46
3.4 Course module: The green wave . . . 51

4 Representational fluency 57

4.1 Introduction and Literature . . . 58
4.1.1 Physics and mathematics . . . 59
4.1.2 The influence of the representational format . . . 60
4.1.3 Representational fluency . . . 61
4.1.4 Literature on representational fluency, linear functions and kinematics . . . 62
4.1.5 Research questions . . . 64
4.2 Test design . . . 65
4.2.1 Representational transitions . . . 65
4.2.2 Linear function types . . . 66
4.2.3 Contexts . . . 66
4.2.4 Other design choices . . . 70
4.2.5 Test structure . . . 70

4.2.6 Expert revision 71
4.3 Test administration 72
4.3.1 Participants 72
4.3.2 Test procedure 73
4.4 Results 73
4.4.1 Validation 73
4.4.1.1 Internal consistency 75
4.4.1.2 Item difficulty index 75
4.4.1.3 Discriminatory power 76
4.4.2 Test results 77
4.4.2.1 Main effects and interaction effects 77
4.4.2.2 Distractor analysis 82
4.5 Discussion 84
4.5.1 Validation 84
4.5.2 Results 84
4.6 The effect of the iSTEM approach 88
4.7 How can these findings be used in teaching? 92
4.8 Future research 94
4.9 Conclusion 95

5 Linear function problems in kinematics and mathematics 99

5.1 Introduction 100
5.2 Literature 101
5.3 Research design 104
5.3.1 Research questions 104
5.3.2 Design choices 105
5.3.3 Test structure 106

5.4 Test administration . . . 107
5.4.1 Participants . . . 107
5.4.2 Test procedure . . . 109
5.4.3 Data analysis . . . 109
5.5 Validation and students' accuracy . . . 111
5.5.1 Validation . . . 111
5.5.1.1 Internal consistency . . . 113
5.5.1.2 Item difficulty index . . . 113
5.5.1.3 Discriminatory power . . . 114
5.5.2 Main effects and interaction effects . . . 115
5.6 Qualitative results . . . 119
5.6.1 Categorization scheme . . . 119
5.6.2 Interrater reliability . . . 123
5.6.3 Results . . . 124
5.6.3.1 Y intercept . . . 125
5.6.3.2 Slope . . . 126
5.7 Discussion . . . 127
5.7.1 Y intercept . . . 127
5.7.2 Slope . . . 129
5.7.2.1 General discussion . . . 129
5.7.2.2 Comparison to similar studies . . . 131
5.7.2.3 Interval/point confusions . . . 133
5.7.2.4 The sign of the slope . . . 136
5.8 The effect of the iSTEM approach . . . 141
5.8.1 Students' accuracy . . . 141
5.8.2 Qualitative results . . . 145
5.8.3 Conclusion . . . 147

5.9 Implications for teaching and future research 148
5.10 Conclusion 150

6 Conceptual understanding of 1D kinematics 153

6.1 Introduction 154
6.2 Integrated STEM 155
6.2.1 The standard curriculum 155
6.2.2 Integrated STEM and the STEM@school project 155
6.2.3 Kinematics course module 157
6.3 Literature on student difficulties in kinematics 158
6.4 Research questions 159
6.5 The modified KCT 160
6.5.1 Test structure 160
6.5.2 Test administration 164
6.6 Results and discussion 164
6.6.1 Student understanding of 1D kinematics concepts 165
6.6.2 Comparing iSTEM and the regular approach 167
6.6.3 Most notable student ideas 169
6.7 Conclusion 173

7 Conclusion 175

7.1 Summary of the scientific conclusions 176
7.1.1 Representational fluency in physics and mathematics . . 176
7.1.2 Conceptual understanding in physics and mathematics . 177
7.1.3 Conceptual understanding in 1D kinematics 178
7.2 Implications for instruction and policy 179
7.3 Critical reflection and discussion 181
7.4 Implications for future research 184

7.5 Final thoughts 189

A Appendix **191**

A.1 Kinematics concepts 191

A.2 Focus on misconceptions in the museum security system module 193

A.2.1 Introduction 193

A.2.2 Student difficulties in geometrical optics 194

A.2.2.1 Image formation by a plane mirror 194

A.2.2.2 Image formation by a convex lens 195

A.2.2.3 Indirect and direct observation of a real image formed by a convex lens 195

A.2.2.4 Ray diagrams 195

A.2.3 Student activities to tackle misconceptions 196

A.2.3.1 Image formation by a plane mirror 197

A.2.3.2 Image formation by a convex lens 199

A.2.3.3 Indirect and direct observation of a real image formed by a convex lens 204

A.2.4 Discussion and conclusion 205

A.3 Detailed distractor analysis for the representational fluency test 206

A.4 Representational fluency test 210

A.4.1 Physics 210

A.4.2 Mathematics 225

A.5 Conceptual understanding of linear functions 239

A.5.1 Kinematics 239

A.5.2 Mathematics 243

A.6 Modified Kinematics Concept Test 247

Bibliography **263**

List of Figures

1.1 Illustration to clarify the points of interest: ideal situation 21

1.2 Illustration to clarify the points of interest: non ideal situation . 21

2.1 Theoretical framework for instructional practices in integrated STEM 36

3.1 Sketch (top view) of the museum room 47

3.2 Example of a Master Plan to tackle the museum security challenge 48

3.3 (a) CAD render and scale model of a museum security system . 49

3.4 Illustration of part of the real-world problem of the Green Wave 52

3.5 Illustration of part of the scale model of the Green Wave 53

3.6 $x(t)$-graph of a model wit constant speed to get through three traffic lights. 54

3.7 $x(t)$-graph of a model wit constant acceleration to get through three traffic lights. 54

4.1 The six linear function types under study 66

4.2 Example items from the test: item P1 and P9 68

(a) Item P1: TG transition in physics 68

(b) Item P9: FT transition in physics 68

4.3 Example items from the test: item M21 and M9 69

(a) Item M21: GF transition in mathematics 69
(b) item M9: FT transition in mathematics 69
4.4 Interaction effect of representational transition and function type 79
4.5 Interaction effect of representational transition, function type and context (physics) 80
4.6 Interaction effect of representational transition, function type and context (mathematics) 80
4.7 Interaction effect of function type and gender 81
4.8 Interaction effect of function type, context and gender 81
4.9 Interaction effect of representational transition, function type, context and gender 82
4.10 Comparison between studies for the ZP function type 86
4.11 Comparison between studies for the PN function type 86
4.12 Comparison between studies for the PP function type 87
4.13 GEE results for the 2-way interaction term *RepT* * *Instruction* 90
4.14 GEE results for the 2-way interaction term *FType* * *Instruction* 90
4.15 GEE results for the 2-way interaction term *Context* * *Instruction* 91
4.16 GEE results for the 2-way interaction term *Gender* * *Instruction* 91
4.17 Illustrations for all-register synergy, semiotic chain and synergy of registers 93

5.1 Test structure including the slope sign, the concepts, the question types, and the contexts. 107
5.2 Illustrative examples from the test. 108
5.3 GEE results for the term *context* * *concept* 116
5.4 GEE results for the term *context* * *slopesign* 117
5.5 GEE results for the term *context* * *QType* 117
5.6 GEE results for the Slope in the term *context* * *QType* * *concept* 118
5.7 GEE results for the slope in the term *context* * *concept* * *slopesign* 119

5.8 Example of a categorization 122
5.9 Example of a categorization 123
5.10 Example of a student's response on question K10 (Case 1) . . . 133
5.11 Example of a student's response on question K12 and M12 (Case 2) 134
5.12 Example of a student's response on question K10 and M12 (case 2) 135
5.13 Example of a student's response on question K10 (Case 3) . . . 135
5.14 Example of a student's response on question K10 (Case 3) . . . 136
5.15 Example of a student's response on question K18 and M8 . . . 138
5.16 Example of a student's response on question K6 and M6 139
5.17 Example of a student's response on question K8 and M8 140
5.18 GEE results for the 2-way interaction term *context* * *instruction* 143
5.19 GEE results for the 2-way interaction term *QType* * *instruction* 143
5.20 GEE results for the 2-way interaction term *concept* * *instruction* 143
5.21 GEE results for the 2-way interaction term *slopesign* * *instruction* 143
5.22 GEE results for the 2-way interaction term *gender* * *instruction* 144
5.23 GEE results for the 3-way interaction term *context* * *gender* * *instruction* 144
5.24 GEE results for the 3-way interaction term *context* * *gender* * *instruction* 145

6.1 Score distribution for the full cohort 165
6.2 Comparison of the mean accuracy of each concept with those from the original KCT study 166
6.3 Comparison of the mean accuracy of each item with those from the original KCT study 166
6.4 Comparison of students' mean accuracy by concept between the iSTEM group and the control group 168

A.1 Illustration to clarify distance and displacement 192
A.2 Visual representation of the line-of-sight argument 194
A.3 Two methods to determine the image formed by a plane mirror 198
(a) The mathematical method 198
(b) The physical method 198
A.4 Ray diagram exercises eliciting the "too small mirror" error . . 199
(a) Exercise with a "too small mirror": example 1 199
(b) Exercise with a "too small mirror": example 2 199
A.5 Plane mirror exercise eliciting the line-of-sight argument 200
A.6 Plane mirror exercises in which to use reverse logic 200
(a) Exercise with reverse logic: example 1 200
(b) Exercise with reverse logic: example 2 200
A.7 Plane mirror exercises eliciting the "zooming capability" error . . 201
(a) Exercise with a "zooming capable mirror": example 1 . . . 201
(b) Exercise with a "zooming capable mirror": example 2 . . . 201
A.8 Standard problem for image formation by a convex lens 202
A.9 Possible outcomes for the exercise concerning the image formation by a convex lens 203
(a) Correct ray diagram without conceptual a conflict 203
(b) Incorrect ray diagram using the partial image logic 203
(c) Correct ray diagram with a dashed extended ray 203
(d) Correct ray diagram with the full light cone 203
A.10 Convex lens exercises 204
(a) Exercise eliciting the "too small lens" error 204
(b) Exercise eliciting the "asymmetric or partial lens" error . 204
A.11 Convex lens exercise using reverse logic 204

A.12 Indirect and direct observation of the real image formed by a convex lens 205

(a) Indirect observation by diffuse reflection 205

(b) Direct observation from behind the lens along the optical axis 205

(c) Determination of the observation area: step 1 205

(d) Determination of the observation area: step 2 205

(e) Determination of the observation area: step 3 205

List of Tables

2.1 Overview of the instructional practices from published papers . 29

2.2 Overview of the instructional categories present in each paper . 30

3.1 Summary of the main integrated topics from Science, Technology and Mathematics 47

3.2 Summary of the main integrated topics from Science, Technology and Mathematics 52

4.1 Validation quantities for the representational fluency test . . . 74

4.2 Interaction effect of representational transition and function type 78

4.3 Distractor analysis 82

4.4 Comparison of the GEE model effects between the model in the paper and the extended model including *Instruction* 89

5.1 Validation results for the conceptual understanding of linear functions test 112

5.2 Table of the significant two-way and three-way interaction effects 116

5.3 Bottom-up constructed categorization scheme. 120

5.4 Frequencies for the y intercept questions 125

5.5 Frequencies for the slope questions 126

5.6 Comparison of our results with other studies 132

5.7 Comparison of the GEE model effects between the model in the paper and the extended model including iSTEM 142

5.8 Frequency comparison of the strategies in the STEM and the control cohort by concept 146

5.9 Frequency comparison of the strategies in the STEM and the control cohort by context for the y intercept 146

5.10 Frequency comparison of the strategies in the STEM and the control cohort by context for the slope 147

5.11 Allocation of Transfer Score 149

6.1 Test structure of the modified KCT 161

6.2 Overview of possible misconceptions per concept 163

6.3 Overview of statistical quantities for the total and per concept accuracy of the iSTEM and control group 169

6.4 Distractor analysis for the iSTEM group and the control group 170

6.5 Percentages of students who believe that the sign of v and x must be the same in C2 items, and who believe that the sign of a and v must be the same in C6 items 171

A.1 Detailed distractor analysis for physics 208

A.2 Detailed distractor analysis for mathematics 209

Chapter 1

Introduction

1.1 Introduction

1.1.1 The future

A bright future requires a well-educated next generation and thus asks for high quality education.

That's it, the first line in this dissertation, the essence. It contains the goal, the means to achieve the goal, and the means to achieve the means to achieve the goal. Before continuing on how we can contribute to a bright future through education, it is insightful to look at the past and the present first.

1.1.2 The past and the present

What do Aristotle, Galileo Galilei and Leonardo da Vinci all have in common? They all embody the idea of the "uomo universale", the universal man, i.e. a person who is literate and skilled in many disciplines. To be fair, it was a lot easier to become anywhere close to an uomo universale in their lifetime because there was a lot less knowledge and skill to obtain. Considering the enormous increase in knowledge throughout the centuries since, the uomo universale is even more an ideal than an achievable reality today. Although some physicists and engineers might claim that everything is physics — from the smallest scale up to the largest scale, and everything in between — a family tree of science developed organically throughout these centuries, and many

individual disciplines and subdisciplines were born. This specialization has pushed the frontiers within each discipline ever further and many professionals have spent their entire career "within their field" contributing to this push and searching for answers to ever more specialized questions. Throughout the past few decades though, researchers have been exploring more and more connections between distinct disciplines, which is evidenced by the (slowly) increasing interdisciplinarity of science [1].

In education, this specialization into scientific disciplines is particularly reflected in secondary education. In secondary education in Flanders, students can enroll in ever more specialized educational programs at young ages (12, 14 and 16). Due to this early specialization, many of them therefore never receive any or very little education in e.g. physics or chemistry, and thus do not obtain a basic literacy in such disciplines. The students who opt to enroll in a rather broad educational program do encounter various disciplines, each in its dedicated course, e.g. mathematics, physics, economics, biology, social sciences, etc. These programs do allow the students to obtain some basic literacy in each of these fields and specialize further in higher education afterwards. The individual disciplines in these broad programs often stand on their own without many connections to other related disciplines, not unlike visiting an archipelago of disconnected islands of knowledge. So although these students can obtain some literacy in multiple disciplines, it is not necessarily interconnected. A shift towards more interdisciplinarity, similar to the one in science, is ongoing though. Also in education, interdisciplinarity is becoming more important and it strives towards a basic literacy for all.

Parallels can also be seen in the educational research community, in which researchers aim to gain more insight in teaching and learning to further optimize education. Since each specific discipline has its own intricacies, most research is either focused on a single discipline or is discipline independent. For a few decades now, more and more effort is going to research into closely related, yet separately taught, disciplines. This is especially true for STEM (Science, Technology, Engineering and Mathematics) education [2]. In secondary education, the courses related to STEM aim to cover such a vast body of knowledge through a comprehensive list of learning goals that the connection between these courses often takes a back seat to the discipline specific goals. Through linking of STEM courses, the many synergies that naturally exist between them can be employed and might benefit students' learning. This is more in line with the ideal of the uomo universale who has a network of knowledge and excels in combining different fields to create something new or to achieve a higher understanding and can thus apply this understanding to solve ever more complex problems.

However, current implementations of STEM education still face some challenges. In the next sections, we describe two of these challenges, discuss a proposed solution, and explain how this dissertation aims to contribute to that solution.

1.2 Challenges in STEM education

The content of this section is partly based on the following peer-reviewed publication:

> H. Knipprath, L. Thibaut, M. Buyse, S. Ceuppens, H. D. Loof, J. D. Meester, L. Goovaerts, A. Struyf, J. B. Pauw, F. Depaepe, J. Deprez, M. D. Cock, L. Hellinckx, G. Langie, K. Struyven, D. V. D. Velde, P. V. Petegem, and W. Dehaene, "Stem education in flanders: How stem@school aims to foster stem literacy and a positive attitude towards stem," *IEEE Instrumentation Measurement Magazine*, vol. 21, pp. 36–40, June 2018

In a society with an increasing demand for qualified STEM professionals [4], high quality education should include high quality STEM education as an essential part. STEM professionals are needed for tackling contemporary complex problems related to sustainable development, healthcare, ecological impact, mobility, etc. [5]. Moreover, in an information based and highly technological society, STEM literacy — i.e., the awareness of the nature of science, technology, engineering, and mathematics, and the familiarity with fundamental concepts from each discipline — is required for all citizens, not only for STEM professionals [6, 7]. Therefore, many governments have made this an important educational goal, e.g., in Flanders [8], in Europe [9], and in the USA [10]. This implies that secondary education should on the one hand motivate students enough to pursue a career in STEM, and on the other hand focus on basic STEM literacy for all. Although the latter is an important aspect, this dissertation and the project of which this dissertation is a part, will focus on the former, i.e. on motivating students to pursue a career in STEM. In doing so, we identify two specific challenges of interest in STEM education:

1. Current education does not succeed in motivating students enough to choose for STEM programs in education and for a STEM career thereafter. Directly related to this, is that students in wealthy countries have a low perceived relevance of STEM. The result is a shortage of STEM graduates in the job market in such countries. This challenge will be elaborated on in Section 1.2.1.

2. Students' performance on STEM topics in international comparison studies is decreasing and teachers face many difficulties to achieve deep understanding with their students. This necessitates continued efforts

to gain more insight into students' difficulties and into their learning and reasoning processes, which can then help teachers and researchers to further optimize STEM education. This challenge will be elaborated on in Section 1.2.2.

We will now discuss each of these in more detail.

1.2.1 Challenge number one: increasing the relevance of STEM and the number of STEM graduates

The first challenge essentially deals with quantity: we need more qualified STEM professionals. This demand in the job market is driven by the increasingly complex societal and technological challenges we face, e.g., the rapid iterations and evolution of technology, the need for sustainable energy, climate change, transportation, healthcare, aging of the population, recycling, food quality and security, etc. [5, 11, 12]. So, clearly the demand for STEM profiles is there, but what about the supply? Specifically for many developed areas, the share of STEM graduates already is or has been predicted to become insufficient in the near future, and the time to fill STEM job openings is longer than that of other sectors [7, 13–15]. A failure to attract new talent, to train them and to retain them will inevitably reduce a country's technological and economic competitiveness in the ever changing global playing field [2, 10, 16] and will delay solving our societal problems.

Furthermore, several studies have observed a decrease of students' positive attitudes towards science and mathematics with age, which is believed to be an important cause of the low number of STEM graduates in the job market. The percentage of students liking science and mathematics has been found to decline between the 4th and the 8th grade [17]. The recent STEM-monitor 2018 about Flemish STEM education provides some insights in the local situation. The percentage of students graduating secondary education with a STEM degree has reached an average of about 45% in 2016–2017. In the same year, the percentage of students with a STEM degree from secondary education who enrolled in a STEM program in higher education was only about 45%. In total, about one out of five students enrolling in higher education registered for a STEM program in 2014–2015, therefore ranking Flanders at the lower end compared to other European countries or areas. In 2016–2017 about 27% of degrees awarded in higher education were STEM degrees. This is less than 1 percentage point higher compared to eight years ago, although during this time, the number of students also grew with about 14%. This means that in absolute numbers, the number of STEM graduates increased, but only due to

an increase in student population, not due to educational changes or changes in students' attitudes and preferences.

So apparently there is sufficient initial interest in STEM, but during the transition from secondary education to higher education the "STEM pipeline" leaks, meaning that many of these STEM students do not pursue a STEM degree. This low level of interest and motivation for STEM has been observed in many highly developed countries and has been attributed to a lack of perceived relevance of STEM subjects by the students [18, 19]. The ROSE study [18] in particular, observed that students in wealthy countries have little interest in science or in becoming a scientist, do care about societal and environmental issues, but do not convincingly see STEM as a solution to those problems. Provided with this information, it is of interest to research how we can increase the relevance of STEM education and of STEM careers to students.

1.2.2 Challenge number two: achieving deep understanding

Students' performance in specific disciplines is regularly assessed in national and international studies. One of the more prominent ones is the Programme for International Student Assessment (PISA). The results from the latest PISA study [20] on Belgian students in 2015 show that their performance on science and mathematics is high compared to many other developed countries. Yet, at the same time, the study shows that average performance is decreasing, that the share of top performers is decreasing, and that the share of low performers is increasing. In light of this decline in student performance, it is necessary to invest in further optimization of education.

In this dissertation we will focus on two specific aspects, i.e., conceptual understanding, and the interplay between physics and mathematics. Researchers have been struggling to precisely define "conceptual understanding" [21]. Here we will consider it the deep understanding of a concept in such a way that the student can apply his/her understanding thereof in known and new contexts. This is very similar to McDermott's definition of functional understanding, which she preferred over the term conceptual understanding [22]. Many studies in discipline specific educational research have shown that conceptual understanding is difficult to achieve. One of the most important reasons for this is that students often enter education with very persistent beliefs and ideas, i.e. preconceptions, which might be incorrect [23–27]. Since students cannot be considered as tabula rasa [28] upon starting education, teachers are presented with the additional challenge of identifying and rectifying these beliefs, which is not a trivial task to accomplish [29]. Many student difficulties have been identified and instruments to study student understanding have

been developed. Some examples of which are: preconceptions in vector kinematics by Aguirre [24]; the Mechanics Baseline Test from Hestenes and Wells [30]; the Test of Understanding Graphs in Kinematics (TUG-K) from Beichner [31]; the Force Concept Inventory (FCI) from Hestenes, Wells and Swackhamer [26]; the Representational variant of the FCI, the (R-FCI) from Nieminen, Savinainen and Viiri [32]; student difficulties in optics from Goldberg and McDermott [33, 34]; the Kinematics Concept Test (KCT) from Lichtenberger *et al.* [35]; etc. Some examples from mathematics education research are: beliefs and conceptual understanding of the limit of a function [36]; Test of Understanding Vectors (TUV) from Barniol and Zavala [37]; reasoning about functions of two variables by Yerushalmy [38]; non-standard student conceptions about infinitesimals from Ely [39]; the Mathematical Literacy and Standardized Mathematical Assessments (using the Texas Assessment of Knowledge & Skills (TAKS) test for Grades 3 through 8) from Matteson [40]; the Baseline Study on Intuitions from Dreyfus and Eisenberg [41]; Adu-Gyamfi and Bossé's study on representations; etc. Furthermore, physics and mathematics have become inseparable throughout history [42] and the language used in physics can be considered a dialect of the mathematical language [43] due to the meaning added to it by physicists. Yet despite their strong connection, students' performance on isomorphic questions in these disciplines typically differs significantly, which illustrates their disconnect in the minds of the students [44–48].

The findings above show that there is still a need for further optimization of educational programs and instructional practices. The given examples illustrate this specifically for physics and mathematics. The results from such educational studies are an invaluable resource for further optimization of education and clearly show that new instructional approaches should be designed based on findings from education research on student learning. Researchers have developed and studied various innovative ideas such as concept cartoons [49], concepTests [50], peer instruction [51], the flipped or inverted classroom [52] and Conceptual Understanding Programme [53], all of which can result in positive effects — but dependent on the specific student profile since not every educational practice is beneficial for every student profile. Furthermore, such studies have usually been performed within a single discipline.

1.3 Two birds with one stone

The content of this section is also partly based on the following peer-reviewed publication:

> H. Knipprath, L. Thibaut, M. Buyse, S. Ceuppens, H. D. Loof, J. D. Meester, L. Goovaerts, A. Struyf, J. B. Pauw, F. Depaepe, J. Deprez, M. D. Cock, L. Hellinckx, G. Langie, K. Struyven, D. V. D. Velde, P. V. Petegem, and W. Dehaene, "Stem education in flanders: How stem@school aims to foster stem literacy and a positive attitude towards stem," *IEEE Instrumentation Measurement Magazine*, vol. 21, pp. 36–40, June 2018

Two challenges were identified before: to increase the number of STEM graduates, and to achieve deep understanding in STEM disciplines. These clearly highlight the need for a substantial reform of STEM education.

The next two sections will describe the proposed approach to address this need, and more specifically the approach chosen in the STEM@school project, of which this dissertation is a part.

1.3.1 Integrated STEM

The partial expression used as a title for Section 1.3 — "two birds with one stone" — is a cruel one, but it is an apt title nonetheless, because we just might solve two problems at once.

Concerning the two challenges defined above, the proposed solution to address them is to focus on the integration of STEM disciplines. By doing so, the hypothesis is that this will result in two effects. First, that it will increase the relevance of STEM for students. Second, that it will achieve a deeper understanding of STEM topics by students. From this, it is expected that more students will then graduate with a STEM degree and choose to pursue a STEM career.

This promising solution is called "integrated STEM" (iSTEM) or sometimes integrative STEM [2, 54], which has been defined as an educational approach in which teaching and learning of two or more of the STEM disciplines are integrated [2]. Other definitions however, say that the integration must include S, T, E, and M [7, 55]. So, instead of aiming for individual discipline specific pillars of knowledge and understanding in non-integrated STEM education — i.e. the traditional approach — the new goal is to construct a coherent network both in and between the STEM disciplines. The hypothesis is that this can improve transfer of understanding between disciplines and be beneficial to conceptual understanding within disciplines. This somewhat brings the idea of the uomo universale into the 21st century, but with the particular goal of him/her being an integrated one.

The approach of an iSTEM curriculum has its proponents and opponents. Typically, the discussion revolves around the relevance, the real life applicability, and the engagement in an integrated curriculum versus the historical and structural value of deep rigorous discipline specific knowledge and understanding [56, 57]. A meta-analysis from Hartzler on the effect of integrated instruction showed that students in integrated curricula consistently outperformed those in traditional curricula on standardized and program-developed assessments, and that integrated instruction can be successful at all grade levels [2, 58]. Another meta-analysis was performed by Becker and Park [59], which showed that many different implementations of iSTEM exist and that the implementation is highly dependent on the teachers' characteristics, thus resulting in a varied spectrum of iSTEM approaches in practice. After their analysis of different studies comparing a conventional and an iSTEM approach, they reported similar findings as those from Hartzler [58]: an iSTEM approach can have a positive effect on students' learning at all levels of education. Additionally, they found that the largest effect sizes were measured at the lower levels of education and with integration of all four STEM components, yet here we will focus our efforts on the leaky pipeline in secondary education. Such findings are encouraging and have stimulated interest in integrated STEM education from governments, educational organizations, researchers and individual schools leading to what Sanders described as "STEMmania" [2].

Mania or not, the implementation of an educational approach should be grounded in educational sciences. Integrated STEM specifically, has been found to be very much in line [60, 61] with the views held by social constructivism, i.e. the theory which considers learning a socially situated process in which knowledge is constructed through interactions with others and in which the learner is an active participant in the learning process [62]. This is in stark contrast to traditional instruction in which learners are considered individual passive receivers of knowledge. And as a respected physics education researcher once wrote [63]:

"A coherent conceptual framework is not typically an outcome of traditional instruction."

— Lillian C. McDermott, 1997.

In their effort to realize the transformation from traditional STEM to iSTEM, the education research community is confronted with some challenges. An important requirement is a research-based framework for iSTEM supported by evidence of its effectiveness. Therefore, learning materials and an accompanying instructional approach must be designed based on that framework.

These should then be used in trial implementations in education, and effect measurements must be performed on the relevant aspects to assess its effectiveness and its implementation conditions. Thereafter, and through various iterations, improvements can be made where necessary. The knowledge gained from such studies can then be used to educate teachers on the new approach and fully implement iSTEM throughout education. The full transition to iSTEM could then help solve the two challenges.

1.3.2 The STEM@school project

Considering the challenges described above, the STEM@school[1] project was set up in Flanders to kick-start and study iSTEM education. It is a collaborative effort of KU Leuven, Universiteit Antwerpen, and the largest two educational umbrella organizations in Flanders: Katholiek Onderwijs Vlaanderen and GO! Furthermore, the STEM@school project team also encouraged 10 schools to cooperate in the development of learning materials, and a total of 30 schools to pilot the implementation of the newly designed learning materials and instructional approach.

The inspiration for STEM@school is credited to Perkins' book [64] on "Making learning whole" in which he argues that students should learn to "play the whole game" — or more accurately "the junior version of the whole game" — from the start and specialize in specific components of the game afterwards and not the other way around. With the metaphor of children learning the complex rules and intricacies of baseball, Perkins suggests that children are better off learning about every player position in the game and how they all link together first, instead of specializing in each of them separately and then construct the network between them afterwards. Inspired by this idea and the challenges in iSTEM education, the STEM@school project aims to design, implement and assess research based teaching and learning materials for sustainable iSTEM projects for students in the 9th, 10th, 11th grade in Flanders. In doing so, the project strives to increase the relevance of STEM for students and to increase students' STEM literacy. Previously, we also described that there is an increasing awareness that a basic STEM literacy is needed for all. Although this is an important part of the global STEM effort, we do not directly address this in the STEM@school project, but will instead focus on the STEM literacy of students in secondary education who have already chosen to enroll in a STEM program. This focus is motivated by the observation that a large percentage of these students does not pursue STEM studies in higher

[1]The dedicatd website of the STEM@school project can be found at http://www.stematschool.be/en/.

education, which we discussed in Section 1.2.1. Essentially, the aim is to stop the "STEM-pipeline" from leaking.

The main activities in the project were the following: the design of an instructional framework and learning materials, and the assessment of the approach.

1.3.2.1 Design of the framework and learning materials

The STEM@school project team developed an instructional approach for iSTEM, inspired by existing frameworks and a review of existing literature. This work is described in detail in Chapter 2, and resulted in a framework based on five key principles (instructional practices) rooted in a social constructivist view on learning [61]:

Integration of STEM content Student activities and learning materials show that the different STEM disciplines are interwoven into a coherent whole and the strong connections between them are made explicit. In doing so it is important to maintain a sufficient investment into discipline specific learning [65]. This principle is essential in the development of iSTEM learning materials.

Problem-centered learning This falls in between project-based and problem-based learning. The students are challenged with authentic, (semi)open-ended, ill-structured, real-world problems in an engaging context at the start of the course module. The problem is treated in its totality, there is no detailed predetermined final solution but there is some guidance towards a certain type of solution, teachers offer guidance and the main goal is for students to develop their problem solving skills.

Inquiry-based learning Creating a learning environment and employing instructional practices which engage students in questioning, experiential learning and hands-on activities that allow them to discover new concepts and develop new understandings. Important aspects are questioning, reflecting, reasoning based on prior knowledge, experimenting, and demonstrating under the guidance of a teacher with the aim of developing critical thinking and constructing new knowledge.

Design-based learning The use of open-ended, hands-on design challenges that provide students with the opportunity to not only learn about engineering design processes and engineering practices, but also deepen their understanding of disciplinary core ideas. Students experience the

engineering design process and the need to learn and use new technologies to solve the challenge.

Cooperative learning Providing students with the opportunity to communicate and collaborate with each other to deepen their understanding. Learning takes place in a social context and among peers in which positive interdependence, face-to-face interaction, individual accountability, practicing interpersonal skills and reflecting on the group dynamics are encouraged and in which the teacher takes the role of a supporting coach.

Based on this framework, eight course modules were developed for the 9th, 10th and 11th grade curricula of Flemish educational programs with a strong focus on STEM. This was done by four teacher design teams supported by representatives from the umbrella organizations and by four PhD researchers from the project team, of which the author of this dissertation is one. The resulting course modules for the 9th and 10th grade can be freely accessed at the Scientix website (those for the 11th grade are still pending) and have the following topics:

9th grade course modules:

- "The green wave (autonomous driving car)", in which students are challenged to design and construct a scale model of a car which can drive autonomously through a predefined green wave of traffic lights [66].
- "A museum security system" in which students are challenged to design and construct an optics based security system for a museum [67].
- "A rehabilitation device", in which students are challenged to design and construct a rehabilitation device to assist a fellow student in his recovery after a traffic accident [68].

10th grade course modules:

- "A device to throw/catch (tennis) balls", in which students are challenged to design a ball throwing device and an accompanying device which can automatically manouvre towards the right location to catch the launched ball [69].
- "A passive house", in which students study various technologies to design a passive house, and construct a sun boiler with a control system to heat their scale model house [70].

11th grade course modules:

- "Drones", in which students study how drones work and then use them to map a geographical area and study the effects of flooding with imaging software.
- "Algae", in which students research algae and their properties, after which they study an application of choice in which algae are used to support a sustainable future.
- "Double stars", in which students study data collected from the Mercator telescope at La Palma and design their own analysis algorithms to deduce the kinematics and dynamics of a double star system in space.

1.3.2.2 Assessment of the new approach

After the development of the course modules was complete, teachers from up to 30 pilot schools were trained to teach these course modules to their students. During the implementation, there was a regular follow-up from the project team and a close feedback loop to address problems, make improvements, and to provide guidance and support. More detailed information about the development and implementation, as well as about the relevant course modules for this dissertation, is provided in Chapter 3.

During and after the pilot implementation, six STEM@school PhD researchers performed various qualitative and quantitative studies. This dissertation is one of them. Studies were performed on the design process and effectiveness of the course modules, on how teachers and schools can positively influence the implementation of iSTEM, and on STEM attitudes among teachers and students. To do so, multiple qualitative and quantitative assessment tools were developed, validated and used. These span a variety of types, ranging from paper-and-pencil test to digital multiple choice tests, free response tests, questionnaires, interviews, classroom observations, and focus groups. These assessments were done in the pilot schools and in a group of control schools in order to perform comparison studies and thus provide a broad assessment of the iSTEM instructional approach. These control schools were selected to be as comparable as possible to the pilot schools in terms of student population, soci-economic factors, location, school size, etc. We call the teaching in the control schools "the regular approach" to distinguish it from the STEM@school approach. In Flanders, schools must adhere to the mandatory learning goals which are defined for each individual discipline — and importantly, mostly independent from other disciplines. The learning goals for physics and mathematics are described in much detail and should be achieved in a limited amount of time. Although the content that should be covered is well defined, the way it is taught is the teachers' responsibility. Based on the

learning goals, publishers have developed discipline specific handbooks which include the required theory, exercises and class experiments. The practice many teachers base their lessons on are from these textbooks and they stick to it rather closely. In a typical physics or mathematics lesson, the teachers interact with students by questioning and students (collaboratively) solve problems and do experiments. Although teachers might frame the relevance of the learning content, illustrate its applicability and even connect with other disciplines, the achievement of the learning goals (content, attitudes and skills) remains the main goal for teachers and thus implicitly for their students. This means that the "regular approach" can be defined by the focus on the mandatory discipline specific learning goals, disconnected disciplines, dedicated discipline specific handbooks, and some variation due to teachers' personal styles. The main contrast with the STEM@school framework as briefly introduced above and which will be more elaborated upon in Chapters 2 and 3 is the following: in the STEM@school approach, students are challenged to solve an authentic real-world problem from the start which triggers the need for learning various new content from multiple disciplines while cooperating with each other, thus making the learning of the content relevant Furthermore, it is crucial to note here that in doing so, the STEM@school approach does not simply group and restructure the mandatory learning goals, but adds to them while employing possible synergies from the integration as much as possible which can compensate for the additional requirements. The additional goals aimed for in STEM@school are:

- Learning and applying STEM content;
- Integrating STEM content;
- Developing research and design skills;
- Thinking, reasoning, modeling and abstracting;
- Strategically applying and developing technology;
- Interpreting information and communicating about STEM;
- Cooperating like a STEM-team;
- Gain insight into the relevance of STEM.

The specific contribution and focus of this dissertation will be discussed in the next section. The other researchers and the titles of their studies in the STEM@school project are:

- Haydée De Loof: "Educating engaged and competent students for STEM: effects of integrated STEM education." [71]
- Jolien De Meester: "Designing iSTEM learning materials for secondary education." [72]
- Leen Goovaerts: "Development and assessment of iSTEM competencies." [73]
- Annemie Struyf: "How to engage students for STEM? Analysing the ingredients for a motivational cocktail in the learning environment." [74]
- Lieve Thibaut: "Implementing integrated STEM: teachers' attitudes, instructional practices and students' learning outcomes." [75]

Because most of these studies in the STEM@school project essentially compare the effects of the STEM@school implementation with a control group, it is necessary to perform a pretest to verify whether the "iSTEM cohort" and the "control cohort" are truly comparable and whether one of the cohorts is not biased in any way. If a significant difference between the cohorts is found, it needs to be taken into account in the analysis. A pretest which covers a broad spectrum of STEM topics (physics knowledge, physics application, mathematics knowledge, mathematics application, and technological concepts) was administered at the beginning of each school year by another researcher in the STEM@school project, i.e. Hadyée De Loof [71]. The results of her multilevel analysis showed that there was a significant difference between students from the few educational programs in the cohorts, but no significant difference when comparing within the same educational programs. Since the students in the target population of the dissertation presented here are all enrolled in the same educational program — i.e. "aso" education with the most hours of science and mathematics — there is no statistically significant bias which needs to be taken into account in the analysis, hence no additional pretest administration is performed in this dissertation.

1.4 The focus of this dissertation

1.4.1 This dissertation as part of the STEM@school project

The dissertation presented here contributes to the assessment of the framework at the project level by focusing on a few selected cognitive abilities in physics and mathematics. The studies it includes are in-depth probes which complement the broad scale research presented in the other dissertations. Each

of the choices made to narrow down the focus of this dissertation will be introduced in the next subsections and serve two purposes. The first one is to perform selective, yet highly detailed measurements which contribute to the assessment of the STEM@school approach. The second one is to contribute the our insight in discipline specific education research in physics and mathematics. Both of these are directly linked to the two challenges in STEM education which were identified above.

Since the STEM@school framework is strongly rooted in social constructivism, the assessment thereof should not be limited to measuring cognitive abilities but should include — if not focus on — assessing the learning processes and the instructional practices themselves as well as the social interaction between students. Although this was the case at the project level through the other dissertations, in this dissertation the focus lies on the assessment of a number of selected and very specific cognitive abilities to gain insight in whether students' performance is safeguarded or not when compared to a control group. Furthermore, it is important to clarify that each of these studies assess the effects of the full framework because it is not possible to disentangle any of the individual principles. Hence all measured effects should be viewed as results from the implementation of STEM@school designed course modules and instructional practices based on the full framework.

In the assessment studies in this dissertation, a few tests will be designed or an existing test will be modified. Each of these tests serves two purposes: to assess the mastery of a specific ability, and to compare this between the iSTEM cohort and the control cohort. This poses a conflict in the design of a test because we essentially aim for a combination of a Criterion-Referenced Test (CRT) which is in line with the first purpose, and a Norm-Referenced Test (NRT) which is in line with the second purpose. This test categorization is not necessarily a black and white situation as a test could serve both purposes within some boundaries. Overall we will prioritize the assessment of the mastery of a topic, i.e. our tests will tend to be more CRT-like, which is in line with the aim to perform in-depth cognitive measurements to compliment the other studies in the STEM@school project.

1.4.2 The focus

We hypothesize that the iSTEM approach based on the STEM@school framework can positively benefit the linking of concepts between disciplines and that the approach can improve student learning within disciplines as well. In this thesis we therefore focus on abilities and concepts in physics and mathematics, which are the two most strongly integrated disciplines in the

first course modules developed and implemented by STEM@school in the 9th grade. Specifically, we will study students' representational fluency and their conceptual understanding of linear functions in physics and mathematics, and their conceptual understanding of kinematics concepts. Each of these topics will now be introduced in this section and discussed in more depth in their respective chapters.

1.4.2.1 Physics and mathematics

It is argued by English [55] that current policy and curricula still have a limited focus on connecting core content knowledge and processes across disciplines. Often the focus is on one of the STEM disciplines, which might benefit performance in that particular discipline, but at the same time it might negatively affect that in another. This possible negative effect is not always measured though. The whole point of integration is that students will experience benefits in all the STEM disciplines and that iSTEM is more than the sum of its parts. Recent reviews such as the one from Honey *et al.* [7] conclude that, despite positive indications [58, 59], there is not yet any strongly convincing evidence that correlates integrated STEM education with student outcomes, which is a particular concern for discipline specific student outcomes [55]. So, more research on how to make STEM connections across disciplines more transparent and meaningful for students is needed [55], as well as more studies on the effects on students' learning in an integrated curriculum. The focus here is on physics and mathematics.

Physics and mathematics have always been strongly related and have benefited from each other and evolved together throughout history [76]. According to Karam [77], the way students see mathematics in physics is as a tool to facilitate calculations, or as a translator of physical thought tasked with representing it in an understandable way, or as a structure to structure physical thought similar to the skeleton of a body [76]. According to de Ataide and Greca [76], the role of mathematics in physics evolved through history from a descriptive role into a language and finally into a primary driver of new physical knowledge. But, as stated by Redish and Kuo, not only does mathematics drive physics, physics also drives mathematics, and physicists appear to use a different dialect of the "mathematical language" [43].

The close relationship between these two disciplines can be considered a burden or an opportunity. Whichever one it is, is likely dependent on the student, but one thing is for certain: their relationship is not always fully utilized to benefit understanding in both disciplines. This is evidenced by many studies which show the existence of a performance gap and/or a clear disconnect between

them, indicating compartmentalization of understanding in disciplines and a low transfer of knowledge between them [44–46, 46, 48, 78]. As argued, we hypothesize that an instructional approach in which integration of STEM-fields is a key principle, might reduce this performance gap and allow students to better relate concepts from those fields.

In Flanders, the mandatory learning goals for physics and mathematics in the 9th grade have a large potential for integration, but this potential is not fully exploited in the regular curriculum. In the STEM@school project, 9th grade course modules were designed based on the STEM@school framework and with a strong focus on a better integration of physics and mathematics. To achieve this, a better alignment and integration of the curriculum goals of both fields was developed, which implied multiple changes in the mathematics course and several ones in the physics course. Considering all of the above, physics and mathematics in the 9th grade STEM@school modules provide an interesting opportunity to assess the approach. Therefore we will study how students connect concepts between these fields, and also to what extent the approach supports the learning within these separate disciplines. Such studies in an integrated curriculum are of importance because we do not want to sacrifice the level of understanding within each discipline for the sake of integration [79].

1.4.2.2 Representational fluency

To communicate on science and mathematics concepts, the concepts are represented in one way or another: in words, in a graph, using a formula, a table, a diagram, a map, etc. In literature, these representations are referred to as the representational format, external representation or representation for short. Each of these representations conveys information about a concept and has specific advantages and disadvantages in a given situation, which makes them particularly useful in one situation but inappropriate in another. This can result in specific difficulties and can influence their success rate [78, 80, 81]. Using and interpreting these representations correctly is an explicit goal in many curricula, including those for mathematics and physics in Flanders [82], and the importance thereof is also internationally recognized [83].

There are various abilities linked to representations. These can be divided into two main groups: the first one is representational flexibility, which is the ability to select an appropriate representation for a given situation [84, 85]; and the second one is representational fluency, which is the ability to interpret, construct and use a representation, and to translate between representations [85–88]. This last aspect of representational fluency, i.e. translating between representations, is a point of interest in this dissertation,

and has been identified as a key ability to master learning and problem solving [89–91].

Studies on the role of representations in physics have revealed that the details in the representation used to portray a problem can affect students' solution strategy [81, 85, 92]. For example, an incorrect pictorial representation of the energy levels of an atom could bear strong resemblance to energy level diagram in the Bohr model which students are familiar with from their lessons and thus trigger them to select the most familiar looking answer without assessing whether the representation is correct or not [81, 92]. Such "cueing" is believed to at least partially explain some of the variation in student's performance in different representations [81, 85, 92]. Furthermore, details can be far more distracting for novices than for experts, resulting in important differences in representational abilities between novices and experts [93]. Furthermore, experts focus more on the underlying concept and less on the representation [93], and they are more flexible in choosing their initial representation when solving a problem and switch quicker to other suitable representations when required [94]. Other studies have identified students' difficulties with specific representations for specific topics, for example the representation of negative velocity in kinematics graphs [95], or the interpretation of ray diagrams for real image formation by a converging lens in optics [34].

Studies in mathematics have shown that students have a strong preference or feel a strong pressure to use algebraic formulas when given the choice of representation, and that this increases with the grade in which they are in [96]. Furthermore, the number and nature of errors is strongly dependent on the used representational format, representations including concrete function values are easier, and formulas are most difficult [78]. Other research has shown that the level of representational knowledge is predictive for student performance in problem solving, justification and explanation tasks [97]. Yet, others have shown that good performance on representation related tasks does not necessarily imply good mathematical understanding [98]. In summary, representational understanding (e.g. fluency) is certainly a part of conceptual understanding — particularly in mathematics [90] — but the former can only be considered an indicator of the latter.

To address the difficulties concerning representations reported in literature, integration of disciplines as part of the STEM@school framework might prove useful. Similar to the idea that integration of STEM disciplines might result in benefits for each discipline, the hypothesis is that "integration" of multiple representations can result in improved understanding of the represented concept. The word integration is not used though, researchers talk about using Multiple External Representations (MERs) instead. The use of MERs has been hypothesized to improve conceptual understanding, and has been

shown to do so in some studies [99–103]. Positive effects are not always guaranteed though [86, 104–106], as using MERs can be too overwhelming for some students, resulting in lower understanding. Hence it is important to appropriately scaffold the introduction of MERs. Specifically for more abstractly oriented students — such as those in our target population with the most hours of mathematics and physics — working with MERs could prove particularly beneficial. We therefore opted to include representational fluency as one of the topics of interest in this dissertation. A second argument to study representational fluency is that representations are an interesting topic for educational research by themselves because they are an integral part of learning in every discipline — and particularly in mathematics — as was illustrated by the literature referenced above.

1.4.2.3 Linear functions

In the discussion above, we have made clear that contexts (physics or mathematics) and representations are of interest for further investigation. In this dissertation we will study these topics for linear relations. Since linear relations are key in the 9th grade physics and mathematics curriculum, and because their most used representations are graphs, algebraic formulas and tables, it makes them a very suitable subject on which to study the effects of iSTEM.

The way the students in our target population typically receive education on linear relations is as follows. In the mathematics curriculum, there is a large focus on the understanding of linear relations in graphs, formulas and tables. Students learn to describe them as functions: $y(x) = ax + b$, in which y is the dependent variable, x is the independent variable, a is the slope and b is the y intercept. They learn to relate such formulas to their corresponding graphs, calculate the y and x entries for a table, and in general learn to relate the various representations with each other for linear functions. Besides studying the y intercept and the x intercept, a lot of attention goes to determining and understanding of the slope. Furthermore, the equation for a straight line through two points (x_1, y_1) and (x_2, y_2), i.e. $y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$, is also studied in the same year, and linked to determining the slope through the ratio $\frac{y_2 - y_1}{x_2 - x_1}$.

In physics, the students study 1D kinematics and more specifically, uniform linear motion. For this motion, the position x of the moving object as a function of time t is described by $x(t) = x_0 + vt$, in which x_0 is the initial position, and v is the (scalar component of the) velocity, which in the particular case of a uniform linear motion equals the average velocity. The average velocity is

introduced as a rate and is determined by $v_{avg} = \frac{\Delta x}{\Delta t}$. The emphasis in the 9th grade students' curriculum is on the graphical representation of the uniform linear motion, specifically $x(t)$- and $v(t)$-graphs. The representation by means of an algebraic formula is not mandatory.

Concerning linear functions, some specific studies have already been performed and many student difficulties have been identified. Relevant literature will be discussed later in this dissertation, some examples of which are: determining the slope in graphs [95], confusion between the slope and the height of a graph (known as the slope/height confusion) [107], comparison of slopes [47, 108], interpretation and use of formal and informal terminology for the y intercept [109], translating between representations of linear function [78, 98], understanding of linear graphs in different contexts [47], and categorizing students' interpretations of linear distance-time graphs [108].

1.4.2.4 Kinematics

The last topic we focus on is students' understanding of concepts in 1D kinematics. As already described above, students learn about the uniform linear motion in the first STEM@school course module. Therefore we will study the conceptual understanding of the relevant concepts from 1D kinematics in an iSTEM curriculum as well. These concepts are described in more detail in Appendix A.1 and are: distance, distance as area under a $v(t)$ curve, displacement, average speed, velocity as a rate, velocity as a 1D vector, acceleration as a rate, acceleration as a 1D vector. Concerning these concepts, many studies have already reported on student difficulties. Relevant literature will be discussed later in this dissertation, some examples of which are: the graphical representation of negative velocity [95], difficulties with the concept of velocity in 1D [110], the comparison of speeds [110], the general interpretation of kinematics graphs [111], the concept of slope in distance-time graphs [47], difficulties with the concept of acceleration in 1D [112]. In addition to assessing the iSTEM students and comparing with the control students, it will also be of interest to look at the overall understanding of the kinematics concepts because the mandatory learning goals related to kinematics are very basic and mainly serve to be applied in a later part of the curriculum. We therefore suspect that students' deep understanding of these concepts is rather limited.

1.4.2.5 Summary

As a summary, Fig. 1.1 provides an illustration of the main focus points of this dissertation. In this example two contexts (physics and mathematics),

and two representations (algebraic equations and graphs) are integrated. The concept under consideration is the rate of change, which is, e.g., the velocity in kinematics, and the slope in mathematics. Ideally, the concept is situated in the center where all contexts and representations overlap, indicating that the concept is understood in each of the individual contexts and in each of the representations. The main idea is that the more intersections are added to the figure, the better the students' understanding of the concept. Fig. 1.2 illustrates a less desirable outcome in which the rate of change is only understood in specific contexts and/or representations, thus resulting in an incomplete and fragmented network of understanding.

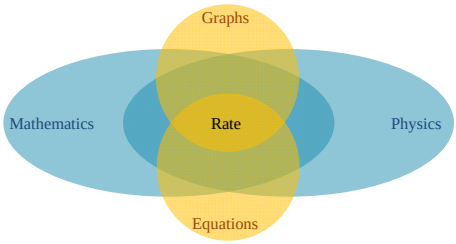


Figure 1.1: "Integration of contexts (blue) and representations (yellow) for the rate concept in which the concept is understood in all lobes and sections.

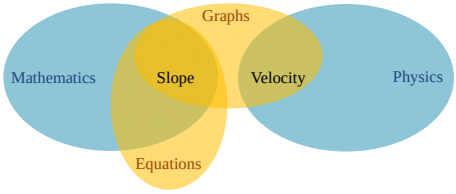


Figure 1.2: "Integration of contexts (blue) and representations (yellow) for the rate concept in which the concept is only understood in selected sections: in both representations in mathematics as "slope", and only in graphs in physics as "velocity", but not as velocity in physics equations. This means that slope and velocity are not linked and that understanding of velocity in equations is nonexistent.

1.5 Research questions

Considering the aforementioned challenges in STEM education, the proposed solution of iSTEM in the STEM@school project, and the focus points discussed

above, the research aims can now be presented. This dissertation, as part of a broader collection of dissertations in the STEM@school project, has the following research questions (RQ):

- RQ1: How does grade 9 students' representational fluency between graphs, formulas and tables for linear function problems compare between kinematics and mathematics and between the STEM@school instructional approach and the regular instructional approach?
- RQ2: How does grade 9 students' performance and conceptual understanding of linear function problems compare between kinematics and mathematics and between the STEM@school instructional approach and the regular instructional approach?
- RQ3: How does grade 9 students' conceptual understanding of kinematics concepts compare in and between the STEM@school instructional approach and the regular instructional approach?

1.6 Structure of the dissertation

In the dissertation presented here, we address the previously stated research questions. First, in Chapter 2 we discuss the founding for the STEM@school instructional approach in literature. Second, in Chapter 3 we elaborate on the practical implementation and conceptualization of the STEM@school approach in Flemish education, and describe two course modules which are relevant to the studies in this dissertation: the green wave and the museum security system. In the next three chapters we discuss our research related to the green wave course module in which kinematics and linear functions are key topics. Each of these chapters addresses one of the research questions defined above. In Chapter 4, we study representational fluency between graphs, algebraic formulas and tables in physics and mathematics on the topic of linear functions. To do so, we develop a dedicated representational fluency test and first discuss the results in general, thereafter we use the test to assess the effect of the iSTEM approach with the regular approach. A similar structure is found in Chapter 5, in which we study students' conceptual understanding of linear relations in kinematics and in mathematics. We also first develop a test and discuss the results in general and then use this test to assess the effect in an iSTEM cohort and in a cohort from the regular curriculum. In Chapter 6, we modify an existing test to study students' conceptual understanding of kinematics concepts and compare the results from the students in the iSTEM curriculum with those from students in the regular curriculum. Finally, Chapter 7 provides the conclusion, discusses the implications for instruction and provides a future outlook.

Chapter 2

iSTEM education: a theoretical framework

This chapter describes the theoretical grounding and the detailed characteristics of the five key principles used to design the STEM@school teaching/learning materials. These key principles for the iSTEM framework were selected based on a systematic review of existing literature. The content of this chapter was originally published in an international scientific journal with peer-review as:

> L. Thibaut, S. Ceuppens, H. De Loof, J. De Meester, L. Goovaerts, A. Struyf, J. Boeve-De Pauw, M. De Cock, W. Dehaene, J. Deprez, L. Hellinckx, H. Knipprath, G. Langie, K. Struyven, D. Van de Velde, P. Van Petegem, and F. Depaepe, "Integrated stem education: A systematic review of instructional practices in secondary education," *European Journal of STEM Education*, vol. 3, no. 1, p. 02, 2018

The development of the framework was a STEM@school team effort and the main credit for the paper goes to the lead author Lieve Thibaut. Partial contributions and revisions of the paper were done by all STEM@school team members. Its content is reproduced here and minor modifications have been added for further clarification.

2.1 Introduction

2.1.1 Need for integrated STEM

In recent years, the importance of providing students with a strong education in Science, Technology, Engineering and Mathematics (STEM) has been stressed. Qualified STEM professionals are needed to remain economically competitive in the global market and to fill contemporary demands such as ensuring sufficient and sustainable energy, efficient healthcare and well-considered technology development [5]. Moreover, all citizens, even non-STEM professionals, should have the skills and competences necessary to deal with the challenges of our information-based and highly technological society [113]. STEM-literacy, i.e. the awareness of the nature of science, technology, engineering, and mathematics and the familiarity with some of the fundamental concepts from each discipline, should be an educational priority for all students [7, 114].

A promising approach in this regard, is the use of an integrated STEM curriculum, which provides opportunities for "more relevant, less fragmented, and more stimulating experiences for learners" [115]. Real-world problems are not fragmented in isolated disciplines as they are taught in schools and to solve these problems people need skills that cut across the disciplines [79, 116]. Studies in a broad range of disciplines have shown that students involved in an integrated curriculum perform as well or even better than their peers in traditional instruction with separate disciplines [79, 117]. Moreover, the use of an integrated curriculum has been found to improve students' non-cognitive learning outcomes, such as interest in STEM [118, 119] and motivation towards STEM learning [120], which in turn could lead to increasing numbers of STEM graduates [7].

2.1.2 Current challenges

Despite the potential benefits and the increased focus on integrated STEM education, implementation of this new instructional strategy faces several challenges. First of all, implementing an integrated STEM approach in an educational system that has a very established segregated and discipline-based structure requires profound restructuring of the curriculum and lessons [121]. Moreover, integrated STEM education often requires numerous materials and resources for students such as construction tools (e.g., saws, measuring devices, and hammers), electronic materials (e.g., computers, design programs, robotics kits, and calculators) and other materials used in design (e.g., wood, styrofoam, glue, cardboard, or construction paper) [122]. Therefore, creating a school

culture and environment that supports an integrated STEM approach to teaching and learning can be costly and time-consuming [121, 123].

Furthermore, to effectively implement integrated STEM, teachers must have deep knowledge of the science, technology, engineering and mathematics content that they teach [124]. Additionally, they must also have specialized knowledge of how to teach STEM content to students, i.e., pedagogical content knowledge [125]. Moreover, many teachers report that they feel underprepared to use STEM applications with their students in the classroom [126]. Moreover, a study by El-Deghaidy and Mansour [126] showed that teachers do not have sufficient understanding of the T in STEM and that they may not have an adequate understanding of the nature of science and technology and the interactions between these two disciplines. Additionally, teachers' beliefs and views about teaching and learning, and their resistance or lack of motivation to change their beliefs and practice, may pose another challenge to the implementation of integrated STEM education [127].

In addition to the challenges of finding resources and insufficient teacher' knowledge, another major challenge for the implementation of integrated STEM education is the lack of consensus about how integrated STEM learning and teaching should be done. In recent years, many researchers have provided detailed information about the learning units for integrated STEM they designed. However, they often do not explain the instructional principles guiding their design [128, 129]. By contrast, other researchers have extensively described their ideas about instructional practices in integrated STEM. However, they often fail to provide a theoretical foundation for these practices chosen (e.g. Moore *et al.* [130], Sanders [2]) and different researchers seem to not always agree about the core instructional practices for teaching integrated STEM. Although several reviews about integrated STEM have been conducted [59, 131], they mostly focus on the effects of integration on students' learning outcomes. To our knowledge, no systematic review about instructional practices in integrated STEM education has been done.

2.1.3 Purpose and research questions of this study

This chapter aims at addressing the lack of consensus about how integrated STEM learning and teaching should be done. To do so, a systematic review of existing literature about learning theories for and instructional practices in integrated STEM was conducted. Moreover, based on the results of the systematic review, a framework for instructional practices in integrated STEM was constructed. By first conducting a systematic review of existing literature, a theoretical base for the instructional practices in the framework is provided.

In the absence of empirical evidence discerning which instructional elements are crucial for the successful implementation of integrated STEM, this is a valuable alternative to provide some form of evidence-based foundation. The specific research questions addressed in the systematic review are:

1. Which learning theories (behaviorism, cognitivism, social constructivism) form the basis of integrated STEM education?
2. Which instructional practices are used in integrated STEM in secondary education?

2.1.4 Methodology

A systematic review was conducted to identify, critically evaluate and summarize the findings of all relevant studies describing learning and teaching in integrated STEM [132]. Systematic reviews differ from narrative reviews, which tend to be mainly descriptive and usually focus on a subset of studies that were chosen based on availability or author selection and therefore often include an element of selection bias. By contrast, systematic reviews typically involve a detailed and comprehensive search strategy derived a priori, with the goal of reducing bias by identifying, appraising, and synthesizing all relevant studies on a particular topic [133]. For the current study, a systematic review was done using two databases: ERIC and Web of Science. The databases were browsed by using four different combinations of search terms: "Integrated STEM + secondary education" ($n = 184$), "Interdisciplinary STEM + secondary education" ($n = 86$), "Multidisciplinary STEM + secondary education" ($n = 14$), and "STEM integration + secondary education" ($n = 121$), leading to a total of 405 search hits. Overlapping search results were excluded and the data set was further reduced using four criteria.

Firstly, all selected articles had to be peer-reviewed journal articles or book chapters written in English, resulting in the exclusion of, for example, conference papers and dissertations. Secondly, the articles had to focus on the integration of at least three of the STEM disciplines. Articles studying the combination of only two disciplines, for example, interdisciplinary mathematics and science or technology integration in mathematics, were excluded. Thirdly, the articles had to provide a clear description of teaching practices for integrated STEM. Finally, all articles had to describe instructional practices for integrated STEM in secondary education. One rater checked whether the articles met the criteria and in case of doubt, papers were discussed with two other raters until consensus was reached. After applying the criteria only 15 articles remained in the sample. Given the small number of articles,

the "snowball approach" was used to retrieve additional publications [134]. Reference lists of all selected articles were inspected and eight additional relevant articles that met the inclusion criteria were added to the data set, resulting in a total of 23 articles.

To analyze the articles in the dataset, first a within-case analysis was conducted [135]. Each article was analyzed separately and summarized in a table, consisting of two categories: (1) theoretical grounding and (2) instructional practices. Secondly, a cross-case analysis was done [135]. The instructional practices extracted from all articles were rearranged and similar elements were grouped, leading to nine different categories. Next, the theoretical framework was composed by focusing on the learning theory and instructional categories named most frequently in the articles of the systematic review.

2.2 Results

2.2.1 Theoretical grounding

Learning theories are typically divided in three categories: behaviorism, cognitivism and social constructivism [136, 137]. Behaviorism is centered around the idea that one behavior leads to another and does not take into consideration the functions of the mind [138, 139]. According to behaviorism, knowledge is an organized accumulation of associations and skills that exists outside a person and can be gained through behavior modification [137, 139]. Therefore, learning is seen as a change in behavior that can be conditioned using positive and negative reinforcements such as reward and punishment [139].

By contrast, cognitivism focuses on students' internal mental structures and addresses the issues of how information is received, organized, stored, and retrieved by the mind [136]. In accordance with behaviorism, knowledge is considered to exist outside of the person. However, unlike behaviorism, cognitivism focuses on understanding how human memory works to acquire knowledge and promote learning, instead of focusing on behavior [139]. Therefore, learning is seen as the discrete changes between states of knowledge, rather than changes in behavior. Moreover, learning is described as a mental activity that entails internal coding and structuring by the student and the student is viewed as an active participant in the learning process [136].

Both behaviorism and cognitivism start from the underlying assumption that learning is an individual process [136]. By contrast, social constructivism states that learning is socially situated and knowledge is constructed through

interaction with others [62]. According to social constructivism, students learn by building personal interpretations of the world based on their experiences and interactions with the environment [136, 139]. Therefore, in accordance with cognitivism, the student is seen as an active participant in the learning process [136].

To answer the question: "Which learning theories form the basis of integrated STEM education?", the articles in the data set were examined for references to learning theories. Of the 23 papers, only seven mentioned an underlying learning theory for instructional practices in integrated STEM: Clark and Ernst [140]; El-Deghaidy *et al.* [141]; Guzey *et al.* [142]; Riskowski *et al.* [119]; Satchwell and Loepp [143]; Shahali *et al.* [144]; and Wang *et al.* [120]. Moreover, all seven articles referred to learning theories from the social constructivist category. As mentioned above, this category of learning theories states that knowledge cannot be transmitted, but rather is actively constructed by students based on their existing ideas and their experiences [145, 146]. Moreover, the social aspect of social constructivism dictates that learning is a shared, rather than an individual experience [147].

2.2.2 Instructional practices

Instructional practices were extracted from all papers and similar elements were grouped, as shown in Table 2.1. This led to nine categories of instructional practices: integration of STEM content, focus on problems, inquiry, design, teamwork, student-centered, hands-on, assessment and 21st century skills. A summary of the different instructional categories found in each article is shown in Table 2.2 and discussed more in detail below.

Table 2.1: Overview of the instructional practices from all papers, arranged in nine categories.

Category	Instructional practices (extracted from papers)
Integration of STEM content	• Multidisciplinary approach • Interdisciplinary approach • Content integration • Context integration • Integrated curriculum with equal attention to two or more disciplines • Curriculum integration with focus on content knowledge • Explicit assimilation / integration of concepts from more than one discipline • Integration of technology • Translation of representations from different STEM disciplines • Connections among learning goals, principles, concepts and skills across discipline specific domains • Infusing/merging of two or more STEM content areas
Focus on problems	• Problem-based learning • Problem-centered learning • Project-based learning • Defining, formulating, evaluating and solving problems • Meaningful/motivating/engaging content • Focusing on big ideas, concepts, themes • Open-ended, real-world, authentic problems
Inquiry	• Posing questions • Planning and carrying out investigations • Collecting, analyzing and interpreting data/information • Discovery learning • Inquiry-based instruction • Scientific inquiry • Authentic scientific practices/processes
Design	• Learning through design • Design-based learning • Developing and using models • Designing solutions • Engineering design • Design justification • Opportunities to learn from failure and to redesign based on that learning
Teamwork	• Collaborative learning • Cooperative learning • Communicating information • Teamwork • Working in small groups • Working with others • Interdependence in group work
Student-centered	• Student-centered pedagogies
Hands-on	• Hands-on learning • Hands-on activities • Effective use of manipulatives
Assessment	• Understanding student misconceptions & capabilities • Use assessment as a part of instruction • Performance as a part of instruction • Writing for reflection • Building on previous knowledge
21st century skills	• Development of 21st century skills [1]

Table 2.2: Overview of the instructional categories present in each paper.

Authors	Integration	Focus on problems	Inquiry	Design	Teamwork	Student-centered	Hands-on	Assessment	21st century skills
Asghar *et al.*, 2012 [127]	X	X	X		X				
Bryan *et al.*, 2015 [149]	X	X	X	X	X				X
Burrows *et al.*, 2014 [150]		X	X		X				
Bybee, 2010 [114]	X	X	X	X					
Clark and Ernst, 2007 [140]		X		X			X		
Debs and Kelley, 2015 [151]	X		X	X					
El-Deghaidy *et al.*, 2017 [141]	X	X	X		X				
Guzey *et al.*, 2016 [142]	X	X	X	X	X	X	X	X	
Guzey *et al.*, 2016 [152]		X	X	X	X				
Hernandez *et al.*, 2014 [153]		X	X	X	X				
James *et al.*, 2000 [154]	X	X	X		X				
Mathis *et al.*, 2017 [155]	X	X		X	X	X			
Pearson, 2017 [65]	X	X		X					
Riskowski *et al.*, 2009 [119]	X			X	X		X		
Roehrig *et al.*, 2012 [156]	X	X		X	X				
Satchwell and Loepp, 2002 [143]	X	X	X	X	X		X	X	
Shahali *et al.*, 2017 [144]	X	X	X	X	X				
Stohlmann *et al.*, 2012 [122]	X	X	X		X	X	X	X	
Stohlmann *et al.*, 2011 [157]	X	X	X	X	X	X	X		
Stump *et al.*, 2016 [158]	X	X	X	X	X				
Valtorta and Berland, 2015 [159]	X	X		X					
Wang *et al.*,2011 [120]	X	X	X						
Wells, 2016 [160]			X	X					

Integration of STEM content. The first category contains instructional practices that advocate making connections between the different STEM disciplines. Several approaches for achieving these connections exist and the terminology used for these approaches differs among the reviewed articles. For example, Wang *et al.* [120] distinguish between multidisciplinary and interdisciplinary approaches. According to them, in a *multidisciplinary approach*, subject-specific concepts and skills are learned separately in each discipline and students are expected to connect the content, taught in different classrooms, on their own. An *interdisciplinary approach*, on the other hand, starts with a real-world problem or issue and focuses on interdisciplinary content and skills (e.g., critical thinking and problem solving), rather than subject-specific content and skills. By contrast, Satchwell and Loepp [143], provide a different definition for interdisciplinarity. Moreover, they discriminate interdisciplinary approaches from integrated approaches, rather than multidisciplinary approaches. According to them, *interdisciplinary curricula* focus on instruction within one domain, while supporting the content with implicit connections to the other disciplines. *Integrated curricula*, on the other hand, explicitly assimilate concepts from more than one discipline and apply equal attention to two or more disciplines. A similar distinction is made by Roehrig [156] who differentiate between content and context integration. According to them, *content integration* focuses on the merging of the disciplines into a single curricular activity or unit to highlight "big ideas" from multiple content areas, while *context integration* focuses on the content of one discipline and uses contexts from others to make the content more relevant.

Although there is no consensus about the terminology, most of the reviewed articles do agree that making strong connections between the different STEM-disciplines is necessary for integrated STEM to work. Many papers (e.g., Satchwell and Loepp [143], Shahali *et al.* [144], Stump *et al.* [158]) emphasize the importance of applying equal attention to two or more STEM disciplines and/or explicitly assimilating concepts from different STEM-disciplines. Making integration explicit is critical, because students do not spontaneously integrate concepts across different representations and materials on their own [65]. Therefore, intentional and explicit support to help students build knowledge and skills across disciplines should be provided [65]. Nonetheless, although the importance of explicitly integrating large amounts of STEM content has been stressed, several papers (e.g., Guzey *et al.* [142], Pearson [65]) warn that integration should remain meaningful and purposeful and that more integration is not necessarily better. Moreover, students' knowledge in individual disciplines should be supported. Students require sufficient

understanding of the relevant concepts in the individual subjects to connect ideas across disciplines [65]. Therefore, integrated STEM education should also focus on learning goals and standards in the individual STEM subjects, so as not to inadvertently undermine student learning in those subjects [65].

Focus on problems. The second category entails the use of real-world problems tied to an engaging and motivating context. In the reviewed papers, different terms are used to designate this: problem-centered learning, problem-based learning and project-based learning. Although all of these approaches are student-centered, promote active learning and advocate the use of authentic real-world problems, specific differences between these instructional approaches exist [127]. In project-based learning, students are provided with the desired specifications for the end product and teachers serve as expert coaches that provide guidelines and suggestions for more effective ways to achieve this predetermined final product [127]. By contrast, in problem-based learning, there is no predetermined end product and students are required to identify and define the problem on their own. The goal of problem-based learning is for students to develop problem-solving skills by going through a realistic self-directed problem-solving process. Therefore, the teacher does not provide specific learning guidance, but rather (s)he serves as one possible resource that students can use to achieve their goal [127]. Finally, problem-centered learning holds the middle between project-based and problem-based learning regarding teacher guidance. Although the problem is more open-ended compared to project-based learning, problem-centered learning focuses mainly on applying and transferring knowledge to realistic contexts, while problem-solving skills are seen as an additional outcome [161,162]. Therefore, in contrast to problem-based learning, a carefully sequenced progression of problems and specific types of guidance are provided [161, 162].

Nonetheless, despite the different terminology, the same important aspects are emphasized in the different articles. First of all, instruction should start by presenting students with a problematic situation that serves as the organizing center and context for learning [114, 127]. By introducing the problem at the start of the learning unit, students can activate existing mental models early in the learning sequence and connect new information and experiences to their prior knowledge in a meaningful way [127]. Moreover, instruction should take place in a motivating and engaging context involving current events and/or contemporary issues. That way, students can link the knowledge and skills to be learned to their personal experiences and meaningful learning is encouraged [142]. Finally, the problems presented to the students should be authentic,

open-ended, ill-structured, real-world problems [143, 144, 150]. These are "messy" problems, often with an abundance of information, that resemble challenges encountered by engineers and scientists in the workplace and allow for multiple solution paths and answers [127].

Inquiry. A third category entails instructional practices referring to the use of inquiry. In inquiry-based learning, students engage in hands-on activities that allow them to discover new concepts and develop new understandings [143]. Thus, experiential learning is intentionally used to promote knowledge construction [160] and students are encouraged to test their existing ideas by taking things apart, making predictions, observing and recording their explanations [143]. Although inquiry-based learning originated in science education, where it usually entails that students have to engage in authentic science practices (e.g., planning and designing experiments and collecting data), it is not restricted to this domain, but also occurs in mathematical or technological contexts [143].

In the reviewed papers, several important aspects of inquiry-based learning are mentioned. First of all, questioning is an important part of inquiry-based learning, because it initiates all knowledge building [160]. Students are stimulated to question their current knowledge about a given topic and to identify which additional knowledge they require to move forward [158, 160]. Secondly, students should use this prior knowledge to generate new ideas, design and conduct investigations and discover new concepts. Moreover, they not only need to carry out the experiments, students also need to demonstrate their understanding of the concepts explored [143]. Finally, an appropriate amount of guidance should be provided to help students achieve the desired conceptual change [143,154]. Truly authentic inquiry experiences might be challenging for high school students, due to their lack of experience and knowledge, or because they have not reached the level of cognitive development required for abstract thought [163, 164]. Moreover, pure discovery learning without guidance might be ineffective because students may not come into contact with the to-be-learned content [165]. Therefore, teachers need to provide guidance by questioning students to help them discover flaws in their reasoning and/or research design, ultimately helping them to arrive at a solution [166].

Design. The fourth category refers to the use of technological or engineering design. By actively engaging students in engineering design challenges, they not only learn about engineering design processes and engineering practices, but also deepen their understanding of disciplinary core ideas [142,144,153]. Engineering design activities can strengthen students' knowledge of science, technology and mathematics, because they fill

the gap between factual content knowledge, abstract knowledge and application [119].

In the reviewed papers, several aspects of qualitative engineering design are mentioned. First of all, effective design challenges should be open-ended, authentic, hands-on, and multidisciplinary [144]. These are challenges that represent industry problems, allow students to explore or develop technologies and require them to work with incomplete information and to consider constraints, safety, risks, and alternative solutions [142]. Moreover, the engineering design process should entail different iterative phases, such a defining the engineering problem, designing engineering solutions, implementing a solution, testing the solution and evaluating and optimizing the solution [149, 160]. Furthermore, throughout the design process, students must manage risk and uncertainty, consider prior experience and learn from failure [142, 149]. In addition, a final requirement for qualitative engineering design is the incorporation of design justification [149, 160]. Students should be able to demonstrate what they have learned in the design process, justify their design decisions by means of this newly acquired knowledge and make recommendations about the design, based on the results of their tests [149,160]. This design justification serves as an effective mechanism for revealing student development and therefore facilitates conceptual change [160].

Cooperative learning. The fifth category entails the promotion of teamwork and collaboration with others. In the reviewed papers, two approaches to small-group learning are distinguished: collaborative and cooperative learning. Although the papers in the review not explicitly define both approaches, other researches have appointed their difference in teacher guidance as main discriminating factor. In collaborative learning, students structure their own group work, without receiving formal training in small-group social skills. The teacher does not actively monitor the groups and refers all questions back to them, because he wants students to resolve group conflicts on their own [167]. In cooperative learning on the other hand, the teacher moves from team to team, observes the interactions and intervenes when he feels it is appropriate. Moreover, training in small-group social skills is provided and the teacher encourages students to assess the functioning of the group in order to improve the levels of participation and performance [167].

Although almost all papers in the systematic review refer to the use of teamwork in integrated STEM, the description of how this teamwork should be organized or what the crucial aspects are, is much less elaborated compared to the previous categories. According to Guzey

et al. [142], it is important for students to have sufficient time and multiple opportunities to be involved in teamwork, such that they can improve their teamwork skills. Moreover, several authors (e.g., Bryan *et al.* [149], Roehrig *et al.* [156], Stohlmann *et al.* [157]) stress the importance of not only stimulating teamwork skills, but fostering communication skills as well. To do so, students need to be encouraged to communicate science concepts, mathematical and engineering thinking through reading, writing, listening and speaking [157]. Finally, positive interdependence between the group members should also be stimulated [127]. This means that students should work on tasks that they believe are only soluble if all group members contribute to the effort. Positive interdependence can be achieved through: (1) providing rewards for successful interdependence; (2) having activities in which resources are shared; or (3) providing a task that is too difficult for students to do individually [168].

Student-centered. The next category refers to the use of student-centered pedagogies and was only present in four of the reviewed articles. Guzey *et al.* [142] indicate that lessons and activities in an integrated STEM unit should be student-centered, because students develop better understanding and skills through active participation in learning activities. Nonetheless, specific guidelines for implementing student-centered learning are not discussed in the reviewed articles.

Hands-on. The instructional practices in the seventh category refer to the use of hands-on learning, hands-on activities and manipulatives. Through hands-on activities, students are less restricted and can actively experience learning. Moreover, relevance is conveyed through hands-on learning, because it allows students to observe the role of innovation in everyday life [140]. As with student-centered learning, no specific recommendations for implementing hands-on instruction are given in the reviewed papers.

Assessment. The eighth category deals with assessment and was only mentioned in three of the reviewed articles. According to Stohlmann *et al.* [157], assessment should be used as a part of instruction. Moreover, Satchwell and Loepp [143] dictate that assessments should present students with authentic tasks that require them to connect all key concepts studied in mathematics, science, and technology and should include a scoring rubric.

21st century skills. The final category comprises "21st century skills", referring to the knowledge, skills, and character traits that are deemed necessary to effectively function as citizens, workers, and leaders in the 21st century workplace [149]. They were only explicitly mentioned in

one of the reviewed articles, i.e. Bryan *et al.* [149]. Although several lists of 21st century skills exist, the ones named by Bryan *et al.* [149] include creativity and innovation, critical thinking, problem solving, communication and collaboration.

2.3 Discussion and conclusion

2.3.1 A framework for teaching integrated STEM

Based on the results of the systematic review, a theoretical framework for instructional practices in integrated STEM for secondary education was created. To do so, the five categories of instructional elements most commonly present in the papers of the systematic review were selected. These categories were deemed most essential for teaching integrated STEM. Moreover, when different instructional practices were named within the same category (e.g., cooperative learning versus collaborative learning), the strategy with the most emphasis on teachers' guidance was chosen. This approach was used, because the theoretical framework targets secondary students. Since their knowledge and experience is limited, they might need sufficient teacher assistance to achieve conceptual change [169]. The final framework consists of five distinctive but related key principles: integration of STEM content, problem-centered learning, inquiry-based learning, design-based learning and cooperative learning. All these principles are rooted in a social constructivist view on learning. An overview of the framework is given in Figure 2.1.

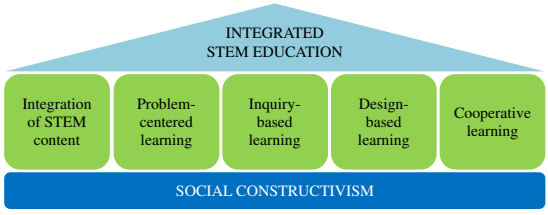


Figure 2.1: Theoretical framework for instructional practices in integrated STEM.

The first principle, integration of STEM content, refers to the explicit assimilation of learning goals, content and practices from different STEM disciplines. Due to the lack of consensus about terminology, no specific strategy (e.g., multidisciplinary, interdisciplinary or integrated learning) is proposed. Problem-centered learning, the second principle, indicates that

learning environments should involve students in authentic, open-ended, ill-structured, real-world problems to increase the meaningfulness of the content to be learned. The third principle, inquiry-based learning, refers to learning environments that engage students in questioning, experiential learning and hands-on activities that allow them to discover new concepts and develop new understandings. Design-based learning, the fourth principle, entails the use of open-ended, hands-on design challenges that provide students with the opportunity to not only learn about engineering design processes and engineering practices, but also deepen their understanding of disciplinary core ideas. Finally, the principle of cooperative learning indicates that students should get the opportunity to communicate and collaborate with each other to deepen their knowledge. All key principles are supported by a social constructivist view on learning, which dictates that knowledge is actively constructed by the students and that learning is a shared, rather than an individual experience.

Although the other instructional categories distinguished in the systematic review were not explicitly included in the framework, this does not mean they are considered to be unimportant. Rather, we believe that many of these categories are overarching aspects that are inextricably linked with one or more of the discerned key principles. For example, both problem-centered learning and inquiry-based learning are student-centered pedagogies and both inquiry-based and design-based learning advocate the use of hands-on learning. Moreover, the 21st century skills named by Bryan *et al.* [149] include creativity and innovation, critical thinking, problem solving, communication and collaboration. Hence, although the importance of developing these skills is not denied, they are already largely present in the other key principles, defying the need for a separate category. In addition, assessment is not explicitly incorporated in the framework, because it is an indispensable aspect of any instructional approach [170] and not specifically linked to integrated STEM education.

2.3.2 Benefits

The framework proposed in this paper is a valuable contribution to the challenge of implementing integrated STEM education and has several benefits. First of all, in contrast to many studies that reduce the instructional strategy for integrated STEM to either inquiry-based learning, problem-based learning or design-based learning [2, 118, 130], this framework allows for a more specific and detailed description of instructional practices through the formulation of five key principles. As can be seen by the description of the principles, they are complementary, but also partially overlapping. For example, teacher

guidance is an important aspect of both problem-centered and inquiry-based learning and the use of open-ended, authentic problems is advocated in both problem-centered and design-based learning. Nonetheless, all key principles have their own specific characteristics, making it worthwhile to include them all in the framework. Secondly, as opposed to other researchers who have described instructional practices for integrated STEM and failed to explain how the instructional elements were derived (e.g. Bryan *et al.* [149]), the current framework is the result of a systematic review of existing literature. In the absence of empirical research about which elements are crucial for the successful implementation of integrated STEM, this is a valuable alternative to provide some form of evidence-based foundation. Finally, the framework presented in this paper is widely applicable and can be used both in teaching practice and in educational research. While teachers and educators can use the five principles to develop curriculum materials for integrated STEM in a fine-grained manner, educational researchers can employ them to analyze instructional practices in more detailed and specific ways.

2.3.3 Limitations

Despite the benefits, the current study and resulting framework also have some limitations. Although a systematic review was done to reduce selection bias, it is still possible that other relevant articles exist that were not included in the data set. For example, we chose to focus solely on integrated STEM in secondary education. Therefore, other interesting articles about related subjects, such as integrated STEM in primary education [131] or design-based learning in higher education [171] were not taken into account. Moreover, the conclusions of the review are based on a small data set and many of the articles have one or several authors in common, which could partially explain why the same instructional categories were named in the different papers. Finally, although the framework can be used as a guideline for designing and implementing curriculum materials for integrated STEM, implementation of the five key principles is not straightforward. For example, creating meaningful connections between STEM disciplines is a challenging task. It requires an iterative act of matching and reorganizing learning goals between the different disciplines, searching an appropriate sequence for these goals and incorporating new learning goals. Therefore, teachers need to be willing to invest time and effort in the implementation of integration. Moreover, letting students conduct their own experiments or build a prototype requires extra time. Hence, teachers might struggle to incorporate new instructional practices, such as inquiry-based and design-based learning, since they usually need all their lessons to finish the traditional curriculum in time. Additionally, creating

an interdisciplinary learning approach requires cooperation between teachers of the different STEM disciplines, which demands sufficient human, financial, material and infrastructural resources.

2.3.4 Perspectives for future research

Although this paper provides an important step towards improving the implementation of integrated STEM education, the need for further research exists. Empirical research is needed to confirm the validity of a framework consisting of five key principles. Next, systematic examination of the effects of integrated STEM, implemented according to this framework, on students' cognitive and affective learning outcomes is required. Moreover, looking at each key principle's individual impact on these learning outcomes could provide information about the necessity of each principle and lead to a refinement of the framework. Finally, the influence of different factors (e.g., teachers' attitudes, school context) on the implementation of integrated STEM could be examined. Insight into these factors could help to improve the implementation of integrated STEM education, therefore ultimately contributing to students' increased motivation for STEM.

Chapter 3

iSTEM course modules

In Chapter 1 we introduced the STEM@school project and in Chapter 2 we described in detail the five key principles for the design of learning and teaching materials as well as for the instructional approach. In this chapter we illustrate the practical implications of our theoretical choices. First, we summarize the key principles from the previous chapter. Then, we elaborate on the development process and the implementation of the course materials in Section 3.2, and illustrate this with two examples. The first example is provided in Section 3.3 and is centered around the design of a museum security system based on geometrical optics. The design and implementation support for this course module was performed by the author of this dissertation in cooperation with a teacher design team consisting of in-service teachers of various STEM-backgrounds from four different schools. In the second example, described in Section 3.4, students are challenged to develop a working scale model car which can drive autonomously through a predefined green wave of traffic lights. The design and implementation support was performed by Jolien De Meester and her teacher design team. After the first year of implementation, the author of this dissertation updated this course module based on the feedback gathered during the first year.

3.1 Summary of the iSTEM instructional approach developed by STEM@school

The STEM@school project developed an instructional approach which was presented in detail in Chapter 2. The key principles of this approach are as

follows:

Integration of STEM content Student activities and learning materials show that the different STEM disciplines are interwoven into a coherent whole and the strong connections between them are made explicit. In doing so it is important to maintain a sufficient investment into discipline specific learning [65]. This principle is essential in the development of iSTEM learning materials.

Problem-centered learning This falls in between project-based and problem-based learning. The students are challenged with authentic, (semi)open-ended, ill-structured, real-world problems in an engaging context at the start of the course module. The problem is treated in its totality, there is no detailed predetermined final solution but there is some guidance towards a certain type of solution, teachers offer guidance and the main goal is for students to develop their problem solving skills.

Inquiry-based learning Creating a learning environment and employing instructional practices which engage students in questioning, experiential learning and hands-on activities that allow them to discover new concepts and develop new understandings. Important aspects are questioning, reflecting, reasoning based on prior knowledge, experimenting, and demonstrating under the guidance of a teacher with the aim of developing critical thinking and constructing new knowledge.

Design-based learning The use of open-ended, hands-on design challenges that provide students with the opportunity to not only learn about engineering design processes and engineering practices, but also deepen their understanding of disciplinary core ideas. Students experience the engineering design process and the need to learn and use new technologies to solve the challenge.

Cooperative learning Providing students with the opportunity to communicate and collaborate with each other to deepen their understanding. Learning takes place in a social context and among peers in which positive interdependence, face-to-face interaction, individual accountability, practicing interpersonal skills and reflecting on the group dynamics are encouraged and in which the teacher takes the role of a supporting coach.

These five principles are enveloped by the social constructivist approach and further enhanced by the inclusion of discipline specific educational research input where appropriate. This ensures that the specific intricacies, difficulties, student ideas and misconceptions of each discipline are taken into account. This

framework was used to design and implement the learning materials, which will be discussed in the next sections.

3.2 Development and implementation of iSTEM course modules

3.2.1 Design process of a course module

The course materials are designed by teacher design teams. Each team is composed of volunteering teachers from different STEM disciplines and from different schools, and is coached by academic researchers and supported by members of the two largest educational umbrella organizations in Flanders. Each team formulates an authentic challenge, combines different (mandatory) content elements to tackle the central challenge, and develops the different learning activities. This is an iterative process: a first proposal is formulated, this is reviewed by other teacher design teams, a revision is made using the feedback and a new review is performed. After two or three cycles the materials are made available in a digital format for the first trial implementation. All teachers implementing the materials are then given workshops and info sessions by both researchers and teachers from the design teams. The participation of in-service teachers in the development and in the workshops is deemed essential by the project team since this creates a buy-in [172] and bridges the distance between researchers and the implementing teachers. An online forum is also made available to facilitate and encourage a quick interaction channel for sharing experiences, supplementary ideas, feedback, correcting errors, and sharing tips for alternative tools, useful software and code as well as items for building prototypes and doing the experiments. During the implementation teachers from all schools are invited twice per trimester to share and discuss their experiences in "intervision meetings", which are semi-structured meetings in which feedback is gathered and ideas and experiences are exchanged. All feedback from the forum and from these meetings is then used to revise and fine tune the course materials by the next implementation year.

3.2.2 The structure of a course module

A course module is structured as a set of themed submodules. The main submodule challenges the students, helps them plan their solution and helps them identify their needs to learn more about various STEM topics which will be necessary to tackle the challenge. To meet these needs several submodules

are available, each exploring a specific theme or discipline, and designed to be consulted on an as-needed basis. This implies that students do not *have* to learn something, but that they will experience the *need* to learn something, which is key to show the relevance of the new learning content. This illustrates that the iSTEM approach is not a mere application of existing knowledge but rather the purposeful learning of new knowledge and skills within the context of an elaborate real-world problem.

To tackle the challenge, students will need to use certain technologies, which will require insight into certain science topics, which in turn creates the need to understand some mathematical concepts, ..., leading to the identification of a set of related needs to which the answers can be found in the various STEM disciplines or submodules. Essentially students will drill down from the real-world problem towards the different fundamental disciplines. Once new knowledge and skills are obtained, these can then be applied to solve the more complex challenge or a subproblem thereof. This "loop-process" of starting off with a complex problem, drilling down to gain more understanding and then applying this to the original problem, is an important mechanism in the iSTEM approach. To fully solve a (sub)problem, students must shift from one submodule to another repeatedly thus clearly identifying the different disciplines and avoiding a linear solution path. While doing so, teachers must explicitly make the links between these disciplines whenever appropriate to help students construct a coherent framework of knowledge.

The learning materials themselves are best described as annotated teaching and learning materials. They include explanatory texts, experiments, extra non-traditional exercises, reflection questions, exercises targeting misconceptions, and feedback links to the overarching challenge to provide as much relevance as possible to the learning content and to maintain an overview. All these materials are annotated with tips, suggestions and background information for teachers to assist them in using the materials as intended — which is a particular difficulty in educational research. These materials are provided digitally to the teachers thus allowing them to adopt them to their specific needs. The teachers are encouraged to actively engage with the material and make it their own while adhering to our guidelines and instructions for iSTEM. The motivation for this is to obtain a stronger sense of ownership with the teachers and to secure the continuation of the iSTEM approach in the future, after the STEM@school projects has ended [173].

3.2.3 Implementation

A total of 30 schools implemented the course modules developed in the STEM@school project. They did so voluntarily and committed to participate for the full duration of the project. Some of these schools were involved in the teacher design teams, the other schools only implemented the materials and provided feedback while doing so. Two types of schools were involved: the first type is called "aso" and its aim is to provide students with a broad education on various topics and prepare them for higher education. Every two years, the aso students can choose a more specific study program and specialize themselves more and more in e.g. science, languages, economics, etc., whilst still enjoying education on multiple other topics as well. The aso study programs which participated in the STEM@school project all have a strong focus on science and mathematics. The second type of secondary school involved in the project are tso schools, which are technically oriented schools aiming to prepare their students for higher education as well as for specific technical professions. The tso study programs which participated in the STEM@school project all had a strong engineering focus, these students usually aim to become industrial engineers. Although both types of schools are included in the project, the schools in the data sets in this specific dissertation are aso schools only. This choice was motivated by practical reasons such as the number of respondents and the limitation of the number of variables in the statistical models.

The teachers participating in the project — the ones in the teacher design teams and the ones implementing the courses — had various backgrounds in the STEM disciplines resulting in multi-disciplinary teams in which the teachers' experiences and knowledge would complement each others'. Most schools opted for an approach in which e.g. a section about a physics topic would be taught by the physics teacher, and a section about mathematics would be taught by the mathematics teacher. A minority of schools opted for co-teaching in which multiple teachers with various backgrounds are present in the classroom at the same time and teach together. Most teachers would have preferred to go with the co-teaching approach but practical constraints did not make that possible at every pilot school. The STEM@school project team welcomed both approaches and had a slight preference for co-teaching since it better illustrates the integration for the students and because teachers then also cooperate inside the classroom.

A course module is typically designed for a trimester or a semester. During each week, about four to five teaching hours of 50 minutes are allocated for teaching the iSTEM course modules. The aso schools, which are the relevant ones in this dissertation, had some choice in how they would divide the required teaching time. Typically two or three hours were allocated to teach a specific

STEM course to focus on the challenges in the course modules. The additionally required hours were allocated to specific disciplines. These hours were used to teach the learning content from those specific disciplines which was needed in the STEM course, often one or two hours in the physics course and one in the mathematics course. This time allocation did vary somewhat between schools as well as between course modules since some course modules include e.g. less physics and more biology.

During implementation, the project team performed visits and interviews, organized the inter-vision meetings discussed before, provided an online discussion forum for the teachers and the researchers for quick feedback and adjustments, organized professional development activities and workshops with and by teachers for teachers, set up a Facebook group, and finally also administered questionnaires and various types of tests and interviews for revisions and research purposes.

3.3 Course module: The museum security system

This section describes a course module for 9th grade students, which was designed by a teacher design team supervised by the author of this dissertation. The complete set of learning and teaching materials for this course module (in Dutch) are available for download free of charge from the Scientix project website:

> STEM@school, "STEM@school module 32 - A security system for a museum", 2018, available at Scientix via http://www.scientix.eu/resources/details?resourceId=20475 or at the STEM@school website: http://www.stematschool.be/nl/leermateriaal/3de-leerjaar.

This course module focuses on the integration of mathematics and physics content for students in the 9th grade. To relate these learning materials as much as possible to the regular physics and mathematics courses, the design team carefully took the (mandatory) curriculum goals into account during the development process. Given that a theme in the 9th grade physics course is geometrical optics and in the mathematics course is linear functions, a design challenge based on these topics was put forward. Students are challenged to design a museum security system for the situation shown in Fig. 3.1:

"*Design an invisible security system, so that for each attempt to theft an alarm sounds and a photo of the perpetrator is taken for identification.*"

A second challenge to assist in tackling the first one reads like:

"*Optimize and build a mini laboratory (a "Lightbox") to study the laws of optics.*"

Table 3.1 shows the topics from the different STEM fields in this module. Engineering is not explicitly mentioned since we consider this to be the overarching process of combining S, T and M to tackle the challenge.

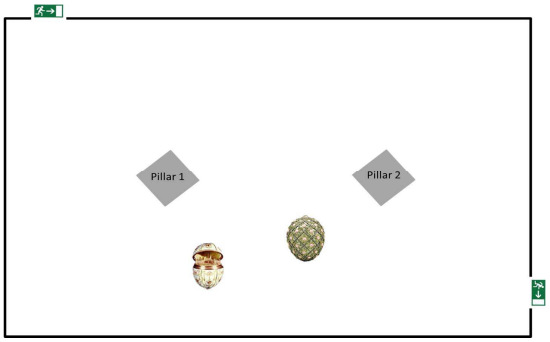


Figure 3.1: Sketch (top view) of the museum room containing two works of art and two pillars.

Table 3.1: Summary of the main integrated topics from Science, Technology and Mathematics.

Science (physics)	Technology	Mathematics
• The phenomenon light • Rectilinear propagation • Shadow • Reflection at a plane mirror • Refraction at a plane mirror • Convex lenses • Concave lenses • Camera system	• Lasers • Light sensors • Electronic components • Design and production of a Lightbox • Design and production of a scale model of the museum	• Equation of a straight line • Trigonometry • Unit circle • Similar triangles • Systems of equations

Confronted with the museum challenge students — guided by their teachers — design a solution ("The Master Plan") which creates a safe zone around the artworks by installing a laser that reflects on each of the pillars using plane mirrors and then illuminates a light sensor (LDR: Light Dependent Resistor). When the laser beam is interrupted, the alarm will sound and a camera system takes a photograph of the perpetrator. Given the location of the laser and the locations and angles of the pillars students must then determine the appropriate location of the light sensor. For differentiation purposes different groups can start with different values for the parameters and/or can be asked to find the

appropriate location of the laser or the correct angle of one of the mirrors instead.

An example of a possible solution by one of the participating classes is provided in Figure. 3.2, in which the students came up with a more elaborate Masterplan as the result of their brainstorm session. To solve the problem students design the entire solution and test it in practice on a scale model — which can also be designed and build by the students. A CAD rendering of a possible design and its real version are shown in Fig. 3.3. Such a 3D render was created by some classes — mainly in technically oriented schools (tso) — which was then produced in reality to experiment with.

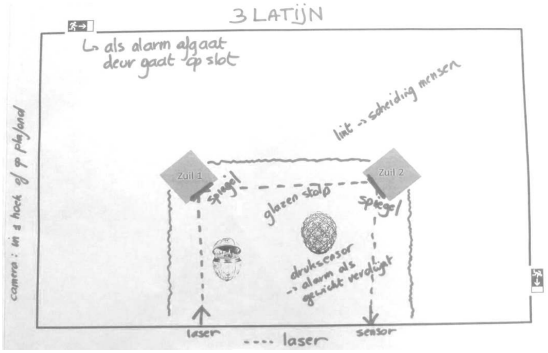
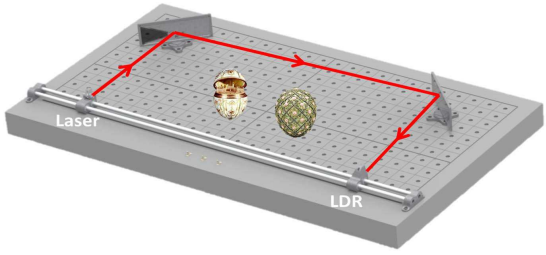
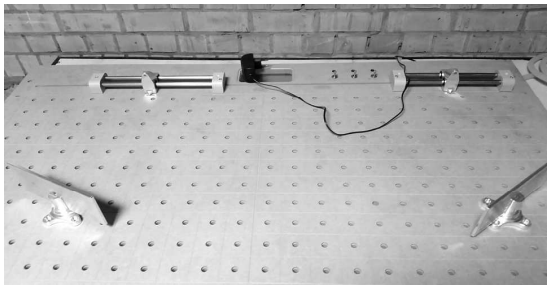


Figure 3.2: Example of a Master Plan with the following security measures: a large glass wall, pressure sensors beneath the works of art, a safe zone with a laser beam, two plane mirrors, a light sensor, a security tape to warn visitors where the safe zone is, a camera on the ceiling and automatic door locks which are activated when the alarm is triggered.

At the start and throughout the course, the students must formulate the main questions which help them identify their needs. A first important question which requires answering is “What is a laser beam?” This triggers a series of other questions such as “What is light?”, “How does light propagate”, “How is laser light different from ordinary light?”, ... This is a beautiful example of one of the “loop-processes” discussed above: students want to understand the more complex phenomenon of laser light, to do so they drill down to understand the properties and propagation of light and then work their way back to understand how laser light is different from natural light and how it is produced, thus understanding the phenomenon and answering the original question. Then, to describe a laser beam they need to learn how to describe a straight line in a mathematical way, followed by the need to understand what happens at a plane mirror. To calculate the point of incidence or the angle or



(a)



(b)

Figure 3.3: CAD render for the museum security system. The mirrors can be rotated and repositioned, and the laser and light sensor can rotate and slide along a rail system. (b) Real scale model based on the CAD render.

equation of the reflected ray the need arises to solve systems of equations and use geometry and similar triangles. Students can then calculate their way from the laser to the point of incidence on the wall where the light sensor should be placed. They learn about light sensors, build and/or simulate a scale model and test their calculations. The second part of the challenge requires them to understand a camera — which is again a "loop-process". To understand a camera they need to study lenses and to do so they need to understand refraction at a planar surface first. Combining these insights, they can then study a camera, and by making use of geometry and similar triangles they can derive a formula to calculate the field or angle of view of a camera and select the appropriate camera specifications to realize their solution to the challenge. When the challenge has been tackled, students can be challenged again buy reflecting on their solution and looking at it from the point of view of a thief, and try to find a way to circumvent their security system. After which they can again enhance the security measures. Critical reflection and the formulation of possible improvements then concludes the course module. Teachers are encouraged to have their students present their work through, for

example, a presentation, a demonstration, a poster or a project fair.

The five key principles from the iSTEM framework can be identified in this course module:

Integration of STEM content: The integration of the STEM content in this module is clear from Table 3.1 in which the focus for the science component was on geometrical optics in physics.

Problem-centered learning: The challenge to secure a space, and in particular a museum room, is an authentic challenge which is a truly open-ended, ill-structured problem to solve.

Inquiry-based learning: To tackle the security challenge, the students must learn to ask the right questions and find the appropriate answers through experimentation and inquiry under the guidance of a teacher. They develop and test their own Master Plan and use a self-designed Lightbox to study the optical phenomena they encounter, thus helping them to find answers to their questions and to learn from mistakes. The "loop processes" play an important role here.

Design-based learning: To design, construct and test their security system and their Lightbox, students use various technologies and learn how to use new ones, e.g.: 3D software, electrical components, laser, construction tools and methods, the use of proper materials, mathematical software, programming of microcontrollers, etc.

Cooperative learning: Students work in small groups to tackle the challenge and develop their Master Plan. They need to discuss and agree on a final design and work together as a team with shared responsibilities to realize their design.

In addition to the five key principles listed above, the findings from discipline based educational research were taken into account where applicable. In particular, the literature on eliciting and resolving misconceptions in geometrical optics was an important aspect in the design of this module. In the published paper included in Appendix A.2, we report how findings on particular student difficulties in optics informed our design of student activities to tackle these difficulties by using an inquiry based and cooperative learning approach.

During the first implementation of this course module, some teachers struggled somewhat to implement it as intended due to the considerable change in instructional approach and some practical hurdles which needed to be overcome. This illustrates that there is a learning curve which must be overcome

initially. After gaining experience during the first implementation, the second implementation went much smoother and was evaluated very positively by the teachers and the students. The many different designs across the classes and schools (ranging from simple yet functional cardboard models to precision machined or 3D printed models or classroom sized set-ups), as well as the additional side-challenges the teachers and students added themselves illustrate the success of this course module.

3.4 Course module: The green wave

This section describes a course module which was designed by a different teacher design team supervised by another academic collaborator of the STEM@school project, i.e. Jolien De Meester. The main author of this dissertation assisted in advising and reviewing the first version which was used in the first year of implementation, and revised the course module using the feedback gathered during the first implementation. The research described in Chapters 4, 5 and 6 is strongly connected to the content of this course module.

The learning and teaching materials for this course module (in Dutch) are available for download free of charge from the Scientix project website.

> STEM@school, "STEM@school module 31 - The green wave (autonomous driving car)", 2018, available at Scientix via http://www.scientix.eu/resources/details?resourceId=20474 or at the STEM@school website: http://www.stematschool.be/nl/leermateriaal/3de-leerjaar.

The green wave course module is designed for students in the 9th grade and focuses on the integration of kinematics in physics and of linear functions in mathematics. This module challenges the students to design and construct a working scale model of a car which can drive autonomously through a green wave of traffic lights with various distances and speed limits between them. Additionally it challenges the students to have the car pass through the wave as fast as possible while obeying the traffic rules:

"*Construct a programmable, autonomous car — similar to the Google Autonomous Car, but a scale model — and program it to drive through a Green Wave of traffic lights in one motion and as fast as possible.*"

Table 3.2 shows the topics from the Science, Technology and Mathematics courses which are integrated in this module. Engineering is once again considered to be the overarching process which tackles the challenge and employs the other disciplines to do so. A highlight of this course module is the focus on engineering-based modeling with the kinematics data which describes the motion of the car and has been suggested to help lift the engineering and mathematics components in integrated STEM [55].

Table 3.2: Summary of the main integrated topics from Science, Technology and Mathematics.

Science (physics)	Technology	Mathematics
• Position • Speed, velocity • Acceleration • Uniform linear motion • Uniformly accelerated linear motion	• Design, construction and coding of the scale model • Flowcode • Basic electronics • Microcontrollers	• Equation of a straight line • Linear function • Quadratic function • Basic vector properties

Confronted with this challenge, the students identify the relevant physical quantities such as position, distance and velocity, and scale them appropriately in order to achieve a miniature model which can be tested in the classroom. Fig. 3.4 illustrates a part of the real-world problem in which time and the speed limit are included, and Fig. 3.5 shows the matching scaled model.

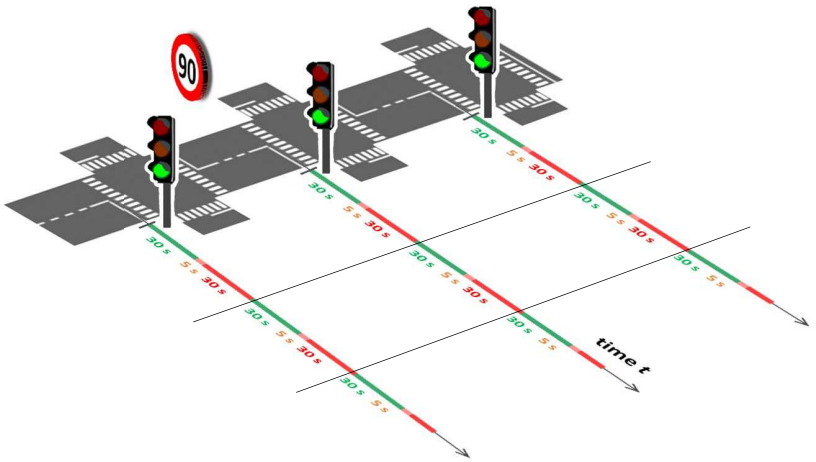


Figure 3.4: Illustration of a part of the real-world problem, including a speed limit and a time sequence for the traffic lights.

In parallel, students also start working on the design and production of their scale model car and learn how to code and program the electronics they need to make it work. They learn to work with easy to use 3D software, e.g. Google Sketchup [174], and work with existing electronic kits such as Arduino with programming in Flowcode. In doing so, they are confronted with the need to describe the motion of the car, thus needing linear equations and the equation

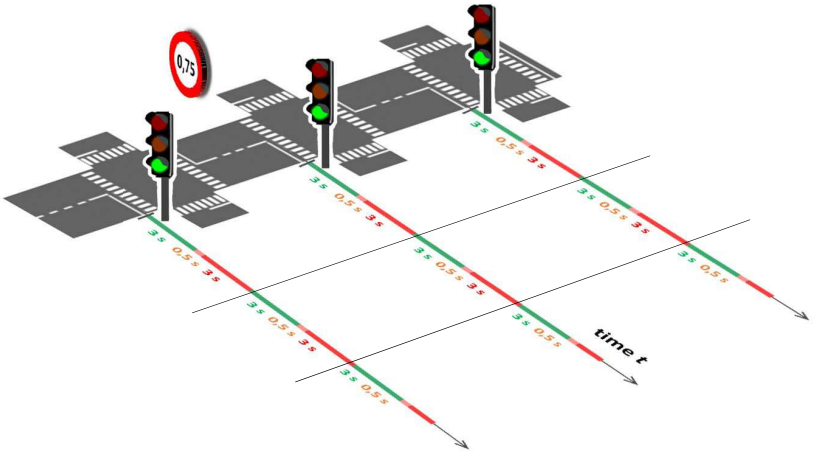


Figure 3.5: Illustration of a part of the scale model, with scaled speed and time.

of motion for a uniform linear motion. Once they start testing and graphing their test results they encounter deviations from the expected straight line graphs, which incentives them to study acceleration and include it in their model when traversing multiple speed zones. To do so, they must enhance their model and use quadratic equations and the equation for the uniformly accelerated rectilinear motion. In the simple model with constant speed they encounter several scenarios, some of which can offer a viable solution to traverse the green wave, and some which are unacceptable, as is shown in Fig. 3.6. In the model with constant acceleration, shown in Fig. 3.7, they find solutions which violate the speed limit, but also one which is a realistic solution. By changing the various parameters such as the number of traffic lights, the pattern of the traffic lights, the distances between the traffic lights, or the speed limits in various zones, many variations to the same basic system can be produced. Moreover, different questions can be asked, for example: one might be interested in optimizing the timing of the traffic lights or in optimizing the speed limits. The pilot schools had some freedom in which questions they aimed to address in the final part of this module, i.e. which variable they aimed to solve, so multiple variations on this were encountered.

When the main challenge had been tackled, some smaller ones can also be added if desired, such as finding a reliable method to assess the average speed of a car along the trajectory and fine perpetrators who where speeding. At the end of the course module, students can again present their work through presentations, posters, a home made movie, a stand on a fair, etc., and of course

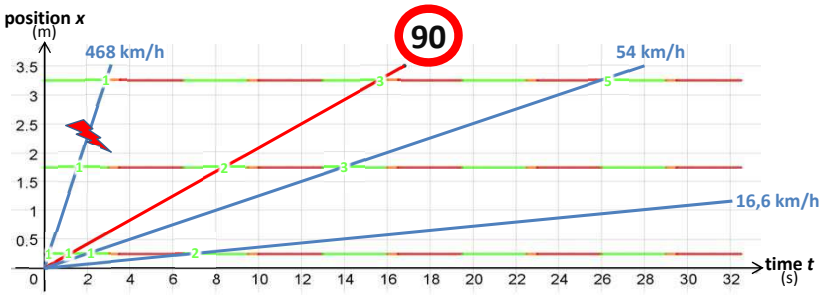


Figure 3.6: $x(t)$-graph of a model wit constant speed to get through three traffic lights. The first option is in violation of the speed limit, the second curve (in red) is the fastest within the speed limit.

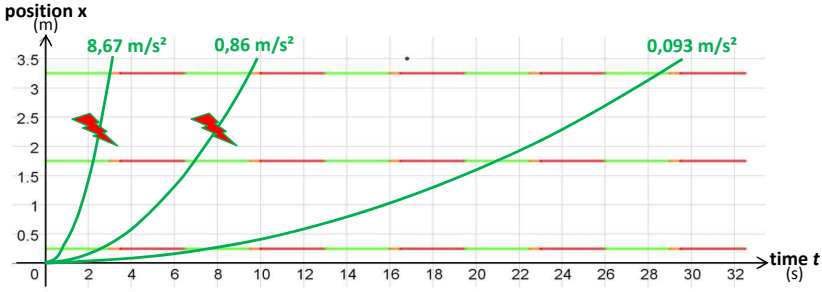


Figure 3.7: $x(t)$-graph of a model wit constant acceleration to get through three traffic lights. The first two curves are in violation of the speed limit at some point, the third is a possible option.

race their cars in a competition.

The five key principles from the iSTEM framework can be identified in this course module:

Integration of STEM content: The integration of the STEM content in this module is clear from Table 3.2 in which the focus for the science component was on kinematics in physics. Furthermore, this course module has a substantial technical side, since this is the module in which many students were confronted with microelectronics, programming, design and tooling for the first time, as this was the first course module in the curriculum.

Problem-centered learning: The challenge to design an efficient and autonomous transportation is very much in line with current efforts in research and industry and proved easy to relate to for students. It is by all means an authentic challenge, open-ended and ill-structured problem to tackle.

Inquiry-based learning: Similar to the previously described course module, the students must learn to ask the right questions and find the appropriate answers through experimentation and inquiry under the guidance of a teacher. They develop and test their own car design and perform various experiments to learn about new concepts relevant to their challenge.

Design-based learning: This course module is heavy on the engineering side because the end result must be a working scale model with "autonomous" capabilities. During this design process students encounter many technical challenges which allows the engineering process to really shine through in this course module.

Cooperative learning: Similar to the museum security module, students work in small groups to tackle the challenge and develop their solution. They need to discuss and agree on a design and tackle the technical difficulties as a team. Due to the many hands-on activities in this particular course module, students must come up with a fair distribution of the work and function as a real team.

This course module was the very first one to be implemented in the curriculum. Due to the very different instructional approach and the steep technical learning curve for students, and also for some teachers, the first implementation was not without difficulties. Revisions and the gained experience after the first implementation year resulted in a much improved second implementation. Also in this course module, the efforts from students and teachers from the various schools resulted in a wide spectrum of successful designs and used technologies (ranging from skateboard-like cars, to cardboard cars, store-bought toy cars, and very esthetically designed cars, and with a variety of electronics and motors). This leads us to conclude that the green wave module is a successful result of the iSTEM approach.

Chapter 4

Representational fluency in physics and mathematics: the case of linear functions

The course materials and instructional approach designed in the STEM@school project focus on the integration of the different STEM components in the context of authentic problems. In the module on the autonomous driving car, content from physics (1D-kinematics) and mathematics (linear functions) are combined to study and program the movement of the car. By showing how physics and mathematics come together, we not only hope to show the relevance of both fields but also aim at increasing student learning and at deepening their understanding in both topics, and in the mathematical description of physical concepts and phenomena. In this chapter, we study how fluent students are in translating between different representations of linear relationships in both physics and mathematics. Moreover, we look at whether the newly developed STEM@school curriculum has a possible impact on that. We therefore designed and validated a test that was administered with students in the STEM@school curriculum and in a regular curriculum. This chapter — excluding Section 4.6 is a reproduction of the following peer-reviewed and published paper:

S. Ceuppens, J. Deprez, W. Dehaene, and M. De Cock, "Design and validation of a test for representational fluency of 9th grade students in physics and mathematics: The case of linear functions," *Phys. Rev. Phys. Educ. Res.*, vol. 14, p. 020105, Aug 2018

The first author designed and administered the tests and analyzed the results. All authors contributed to the test design and reviewed the results from the analysis.

4.1 Introduction and Literature

Chapter 2 discussed the framework used in the STEM@school project and Chapter 3 provided examples of its implementation in Flemish secondary education. The green wave course module described therein challenges the students to design, build and test a scale model of a car which can autonomously drive through a predefined green wave of traffic lights and speed zones. This course module is strongly linked to the mandatory learning goals from the regular curriculum which center around kinematics and linear functions for physics and mathematics respectively. Because the kinematics content mainly covers the uniform linear motion, linear relationships are an important shared topic in both contexts and are therefore a very suitable subject to study. An important aspect in both kinematics and mathematics is the representation of a linear relationship, i.e. for example an algebraic formula or a graph. It is therefore interesting to study the link between kinematics and mathematics concerning the representation of linear functions. This is especially of interest in light of the integrated approach of the STEM@school project, because the integration between these might prove beneficial. This all relates back to our first research question which was formulated in Chapter 1: *How does grade 9 students' representational fluency between graphs, formulas and tables for linear function problems compare between kinematics and mathematics and between the STEM@school instructional approach and the regular instructional approach?* We aim to provide an answer to this research question in this chapter. To do so, we first present an overview of the literature on contexts, representational fluency and linear functions. Next, we formulate more specific research questions and design and validate a dedicated test to assess students' representational fluency. In practice, this test was administered to a cohort enrolled in the STEM@school curriculum and to a control cohort which was enrolled in the regular curriculum. A full analysis was then performed which showed that the instructional approach was not a significant variable in the statistical model. We therefore excluded this variable and also redid the analysis using the simpler model. In light of this dissertation we include a dedicated section — Section 4.6 — to discuss the effects of the iSTEM approach in addition to the results published in the paper. Thereafter, we provide suggestions for how our findings can be used in teaching and describe some implications for future research. Finally, a conclusion is formulated which summarizes all the findings from this chapter and answers the research questions.

4.1.1 Physics and mathematics

The motivation to study representational fluency specifically for the contexts of physics and mathematics is twofold.

First, mathematics and physics have always been deeply linked, although not always in the same way. As described by de Ataide and Greca [76] the role of mathematics in physics historically evolved in three stages: from a way to describe analogies between physical objects in the real world and structures in an ideal world to a mere language or a convenient tool to describe reality, and, finally, to a primary driver of new physical knowledge. This evolution came quite naturally with the ever-growing complexity of physical knowledge and the need for a consistent framework and unified theories to describe, to explain, to predict and to understand physical reality. Physics has become inseparable from mathematical modeling and mathematics is no longer simply being "applied" to physics as stressed by Kjeldsen and Lützen [42]. This view is further supported by Redish and Kuo [43] who noted, "we not only use math in doing physics, we use physics in doing math." So, the relation between these two fields is no longer a one-way street but has become a two-way interaction where mathematics can be a driver for new physics and vice versa. Mathematics has essentially become a language loaded with physical meaning plus reasoning to physics. According to Redish and Kuo [43] physicists appear to use a different dialect of the "mathematical language". This means that students must not simply learn mathematics, but also need to learn the specific dialect used in physics. They must learn to link the pure mathematical language to a physical context.

Second, good mathematical knowledge and skills are necessary for good problem-solving ability in physics, but are not sufficient on their own [43, 176, 177]. This indicates that there is more to it than just knowing mathematics and knowing physics. Linking these two fields is also part of the requirement for good problem-solving ability. This, however, is a persistent problem for students in all educational levels as is shown by several studies. Pollock, Thompson, and Mountcastle [44] compared questions related to the first law of thermodynamics and equivalent context-less questions in mathematics. They provided evidence that students' understanding was isolated in each field. Also Christensen and Thompson [45] studied slope and derivative in what they call "physics-less physics questions." These are questions that look like physics questions but are stripped of their physical context. They found that students have difficulties conceptualizing mathematical questions portrayed as physics questions indicating a lack of transfer of knowledge from mathematics to physics. Planinic *et al.* [47] compared students' understanding of slope and area under a graph in mathematics, physics and other contexts. They found that

mathematical questions were easier and that physics remains a more difficult context. Wemyss and van Kampen [108] studied students' interpretations of linear distance-time graphs and found that the ability to determine the slope of a y-x graph and having a qualitative understanding of distance-time graphs was insufficient to determine the speed on a distance-time graph.

The essential insight from the bulk of these studies is that good understanding in one field does not necessarily imply good understanding in the other field. Furthermore, the literature suggests that students' abilities in physics are not independent of their abilities in mathematics, nor are they independent of their abilities to link these two fields and to blend knowledge from both. It is clear then that the blending of these fields must be an area of focus in education and, consequently, requires various means for assessment and gathering data to improve insight therein. For this purpose we use a "three-pronged approach" where students' abilities are separately tested in a first context, in a — as isomorphic as possible — second context and then compared between contexts, which is in line with the approach in several other studies [44, 46, 47, 108].

4.1.2 The influence of the representational format

Besides the context in which a problem is formulated, another key aspect that influences the success rate and the solution strategy is the representational format in which the question is asked [80]. The representations used most often in scientific disciplines are formulas, graphs, tables, diagrams and verbal descriptions. Each of these can have many variants and each difference can influence the final solution and the strategy to obtain it. Various studies demonstrate these effects. Meltzer [80] studied students' responses on two of Newton's third-law questions, one in a verbal representation and the other in a diagrammatic representation, and found consistent differences in success rate and in the errors made. Additionally, he used questions in verbal, diagrammatic, mathematical or symbolic and graphical representations and also found evidence of an interaction effect between gender and representation. Similarly, Kohl and Finkelstein [81] concluded that the success rate in solving physics problems shows significant differences for near-isomorphic questions presented in a mathematical, pictorial, graphical and verbal representation. A second observation from their study is that allowing students to choose the representation in which they solve a given physics problem, can for some students increase and for others decrease the success rate. Acevedo Nistal, Van Dooren, and Verschaffel *et al.* [96] studied students' flexibility when choosing a representational format (formula or table) to solve linear function problems in mathematics in secondary education. They found that students have an overwhelming preference to use formulas which increases

with grade. Even when tables are expected to be more effective, students most often choose formulas and achieve greater accuracies when compared to using tables. A study from De Cock [85] also confirmed such effects by investigating the strategy students chose when presented with a question in either a verbal, pictorial or graphical representation in a physics context. By analyzing students' explanations in support of their answer, she was able to show that details in the used representational format influenced the chosen problem solving strategy. Lastly, De Bock, Van Dooren, and Verschaffel [78] conducted a study on the effect of representations in a more general context. Students had to model a textual description with either a graphical, tabular, or formulaic representation of the situation. Four different mathematical models were studied: proportional, inversely proportional, affine with negative slope and affine with positive slope. Their analysis identified a significant interaction effect between the mathematical model and representation, meaning that the representational format can have a facilitating effect for a particular model but an inhibiting effect for another. For both the proportional model and the affine model with negative slope, formulas resulted in significantly lower scores compared to tables and graphs. For the inversely proportional model though, results from formulas were significantly better compared to those from tables and graphs. For the affine model with positive slope, graphs resulted in significantly higher scores compared to tables and graphs. These studies demonstrate the important effect that using different representations can have. This is a strong argument to support students in learning different representations and the links and transformations between them.

4.1.3 Representational fluency

There are several representational abilities defined and categorized in literature. A recent insightful summary is given in Deliyianni *et al.* [178] as well as in Hill, Sharma, O'Byrne, and Airey [179]. In the work presented here we focus on representational fluency, defined as the ability to interpret and construct representations, and the ability to translate between them. This is in accordance with earlier work from De Cock [85], from Nistal *et al.* [84], from the more general work of Ainsworth, Bibby, and Wood *et al* [86], and from Even [87] on linking representations. Our definition of representational fluency is very specific and our focus here is on the ability to translate between representations. We are interested in how fluent students are in translating (identifying) the same concept — in this case a type of linear function — to (in) a different representation. This definition is closely related to interrepresentation flexibility which is the handling of between-representation transformations and combines both recognition and conversion competences, as

is defined by Deliyianni *et al.* [178]. The processes students use to perform these transitions are not part of this study, but we do reflect on this in Section 4.5.

Studies such as those of Ainsworth, Bibby, and Wood [89], Duval [90], and Kirsh [91] specifically identify the transition between representations as a key task in learning and problem solving; Duval even calls it the true challenge of mathematics education. Furthermore, Ainsworth, Bibby, and Wood focus heavily on transitions between multiple external representations (MERs) and proposes a functional taxonomy of MERs [180] which Ainsworth later incorporates into her DeFT (Design, Functions, Tasks) framework for learning with MERs [181]. She identifies three key functions: (i) complementary roles; (ii) constrain interpretation; (iii) construct deeper understanding. In this framework she makes clear that a multirepresentational environment — which should include representational fluency — can be of great benefit to the learning process.

4.1.4 Literature on representational fluency, linear functions and kinematics

In his study on fractions in mathematics, Niemi [97] showed that the level of representational knowledge — which he measured by the ability to identify correct alternative representations of a fraction given in a particular representation — is predictive for performance in problem solving, justification, and explanation tasks. When comparing novices and experts in physics several researchers have shown that there are important differences in representational abilities between these groups. A study by Chi, Feltovich, and Glaser [93] indicated that representational features are more distracting for novices than for experts, resulting in the inability to identify the underlying concept to solve the problem. They also found that when asked to categorize a selection of physics problems, novices sorted them by diagrammatic format and diagrammatic similarity whereas experts did so by the underlying physical concepts. Additionally Kohl and Finkelstein [94] learned that — when asked to solve a physics problem — experts are more flexible in the initially chosen representation to work with and are likely to switch quicker between representations than novices.

In their work, De Bock, Van Dooren, and Verschaffel [78] reported on two studies. In their second study they investigated linking graphical, tabular and symbolical (formulas) representations for four different mathematical models: proportional, inversely proportional, negative affine functions, and positive affine functions. Each question described a general context, included one of these models as either a graph, a table, or a formula and asked to identify

the matching option from the possible multiple-choice options. These choice options were presented as a graph, a formula or a table, but were always different from the representation used in the questions. They showed that the number and nature of errors strongly depend on the representational transition. Students performed best when switching or translating between graphs and tables (in both directions), and worst when switching or translating between formulas and graphs. They concluded that representations that contain concrete function values are easiest to transition from and to. Additionally they found students to have the most difficulty with models with decreasing functions. Their most important finding was the presence of a significant interaction effect between the representational transition and the type of function (model). They found that starting from a formula was the most difficult and that going to a table was the easiest for all function types, but that for affine functions with negative slope the accuracies were far lower when going to a graph or to a formula. In addition, for inversely proportional functions they found that when switching to or from a formula the transition is also far more difficult. They obtained the highest accuracies for proportional functions and affine functions with positive slope. In general they found that functions with a negative slope are the most difficult to tackle, similar to the results from other studies [47,108]. Research on student understanding of the y intercept of linear functions by Davis [109] showed that using terminology such as "starting point" for the y intercept in real-world contexts (e.g., a fall from a certain height) negatively impacted their performance on translating between symbolic, graphical, and tabular representations. This indicates the difficulties students have with transferring knowledge about the y intercept between representations and contexts similar to results from Schoenfeld *et al.* [182].

In the specific context of kinematics in physics such problems have been studied for a long time, with specific attention to position-time graphs in kinematics [31, 45–47, 95, 110, 112, 183, 184]. Of note is the TUG-K test from Beichner [31] which was specifically designed to assess student knowledge of graphs in kinematics, as well as the work by Planinic *et al.* [46, 47]. They also studied linear graphs in mathematics, kinematics and contexts other than physics and found that students in secondary education had more difficulty with the physics questions and specifically with negative velocity. University students, though, had closed the gap and achieved similar results on questions about slope, independent of context. This suggests they were able to successfully transfer their mathematical knowledge.

Representational fluency is often indirectly included in other tests that use various different representations, but it is rarely the goal of the test itself. One study we explicitly mention, is that of Hill *et al.* [179] that reports on the development and validation of the representational fluency survey (RFS)

for university physics students, which is a broad scale multiple-choice survey composed of previously validated items for other tests and newly written items in words, graphs, equations, and diagrams. In contrast with the very broad scale RFS test, our work focuses on linear relations and the differences between physics and mathematics specifically.

4.1.5 Research questions

The main goal of our study is to construct and validate a multiple choice test on representational fluency in linear function problems in mathematics and physics. In doing so, we adhere to the definition of representational fluency given above. However one could equally well use the term representational recognition for our study, which is the ability to recognize the same concept in multiple representations. We aim to set up a similarly structured test as that of De Bock, Van Dooren, and Verschaffel [78] but with a different selection of function types (models), a near-isomorphic comparison between physics and mathematics and a different age group. Their respondents had typically finished three years of nonuniversity higher education; our respondents are aged 14–15, still in secondary school and have only recently studied linear functions in mathematics and 1D kinematics in physics. We focus on these first order relations only. Variations are made through different sign combinations of y intercept and slope in mathematics and on initial position and velocity of uniform linear motion in physics. These types of linear relations are key topics in science education for our target students. With this test structure we aim to answer the following research questions:

1. How does student representational fluency between graphs, tables and algebraic formulas of linear functions differ between mathematics and physics (1D kinematics)?
2. How does student representational fluency between graphs, tables and algebraic formulas differ for different combinations of slope sign and y intercept sign in linear function problems in mathematics and physics (1D kinematics)?
3. What are the significant interaction effects between representational transition, context and linear function type (combination of slope and y intercept sign)?

When the test has been validated and these research questions have been answered, we can also use the test to compare students' representational fluency between the two different instructional approaches: iSTEM and the regular

teaching approach. We therefore use the test to answer the following research question — RQ1 in this dissertation:

5. How does grade 9 students' representational fluency between graphs, formulas and tables for linear function problems compare between kinematics and mathematics and between the STEM@school instructional approach and the regular instructional approach?

To answer these questions, we first describe the test design in Section 4.2, followed by a description of the test procedure in Section 4.3. The results are provided in Section 4.4 and are discussed in Section 4.5. Then, we focus on the effect of the instructional approach in a separate Section, i.e. Section 4.6. This is followed by a discussion about the implications for teaching and future research in Section 4.7 and 4.8 respectively. Finally, we summarize the main findings in Section 4.9.

4.2 Test design

For our purposes the test needs to incorporate the three aspects of interest: different representational transitions, different linear function types and different contexts. The full test is translated to English and available in Appendix A.4.

4.2.1 Representational transitions

Test items are formulated in one of the three representations: tables (T), graphs (G), and algebraic formulas (F). This choice is motivated by their abundant use by experts in both mathematics and physics, as well as by the many problems students experience when using graphs and the need they often feel [185] to use formulas to solve problems. The answer possibilities in the multiple-choice items are provided in either of the two remaining representations. This means students are forced to make the transition requested in the item thus allowing the test to measure representational fluency and not representational flexibility, which would assess the representational choices students make. This results in six possible representational transitions: table to graph (TG), graph to table (GT), formula to table (FT), table to formula (TF), formula to graph (FG), and graph to formula (GF).

4.2.2 Linear function types

Linear functions can be categorized in six different types when distinguishing between all combinations of negative or positive slope and negative, zero or positive y intercept as shown in a graphical representation in Fig. 4.1. To limit the number of questions in the test — and thus the time needed to solve them — and maximize the relevance of these possible combinations for kinematics, four out of six are selected to be studied here: negative y intercept with positive slope (NP), zero y intercept with positive slope (ZP), positive y intercept with positive slope (PP) and positive y intercept with negative slope (PN). For each multiple-choice question students are allowed to choose their answer from the full set of six possible combinations of y intercept and slope, thus also including the two remaining possible combinations (NN and ZN function types) as buffer distractors which are marked with dashed lines in Fig. 4.1.

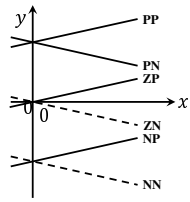


Figure 4.1: Graphical representation of the six possible answer options. Full lines are the function types under study, dashed lines are buffer distractors. Abbreviations denote the sign (N for negative, Z for zero and P for positive) of the y intercept and the slope respectively.

4.2.3 Contexts

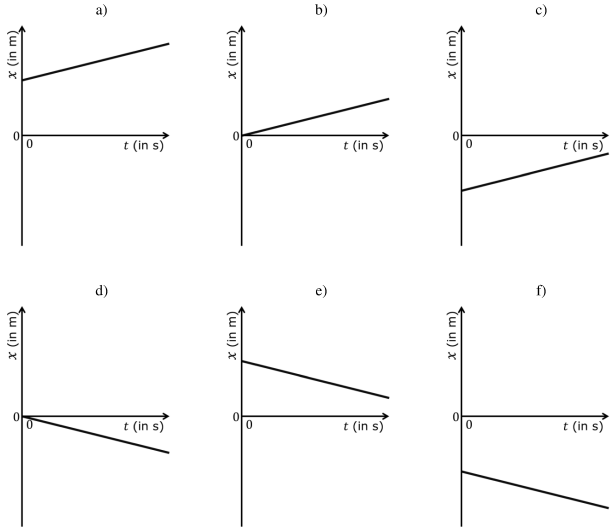
The test consists of two parts that are designed to be as isomorphic as possible. The first part is in a mathematics context, meaning questions provide information about the relation of y as a function of x and offer no additional context or interpretation of the dependent or independent variable. The second part is in a physics context (1D kinematics). The questions describe the uniform linear motion of a cyclist and provide information about the position x as a function of the time t. Depending on the representation, implicit or explicit information is provided about the sign of the initial position x_0 and the velocity v, where v is interpreted as the slope of the curve and can thus have a negative sign. The negative sign of velocity, particularly in graphs, has been shown to be problematic for many students [95] and is an extra point of interest in the test. Although the physics context and content in the physics questions might seem

very low to some, we expect that the differences with the mathematics questions are meaningful and significant enough for students and that they will trigger a different mindset in which to solve the questions and consequently make different errors. When comparing this to what Christensen and Thompson [45] labeled "physics-less physics questions", it becomes clear that our questions do not have the physics context stripped off, thus warranting the designation of true "physics items". Figs. 4.2 and 4.3 provide examples of four items from the test, including questions P9 and M9 — the physics (P) and the mathematics (M) version of question 9 — to illustrate the contextual differences.

P1 A cyclist is cycling at constant velocity on a straight road. The table describes the relationship between the position x (in meters) and the time t (in seconds):

t (in s)	x (in m)
0	-36
1	-30
2	-24
3	-18
4	-12

One of the graphs describes the same relationship between x and t. Which one?



(a)

P9 A cyclist is cycling at constant velocity on a straight road. The relationship between the position x (in meters) and the time t (in seconds) is given by $x = x_0 + vt$ with $x_0 < 0$ and $v > 0$.

One of the tables describes the same relationship between x and t. Which one?

a)

t (in s)	x (in m)
0	0
1	7
2	14
3	21
4	28

b)

t (in s)	x (in m)
0	42
1	49
2	56
3	63
4	70

c)

t (in s)	x (in m)
0	0
1	-7
2	-14
3	-21
4	-28

d)

t (in s)	x (in m)
0	-42
1	-35
2	-28
3	-21
4	-14

e)

t (in s)	x (in m)
0	42
1	35
2	28
3	21
4	14

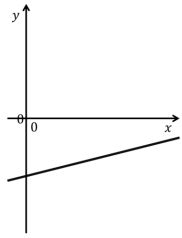
f)

t (in s)	x (in m)
0	-42
1	-49
2	-56
3	-63
4	-70

(b)

Figure 4.2: Example items from the test. Item P1: a physics question concerning the transition from table to graph. Item P9: a physics question concerning the formula to table transition.

M21 The variable y is dependent of the variable x according to a first degree equation. The graph describes this relationship between y and x:



One of the equations describes the same relationship between y and x. Which one?

a) $y = ax + b$ with $a > 0$ and $b > 0$.

b) $y = ax + b$ with $a < 0$ and $b = 0$.

c) $y = ax + b$ with $a > 0$ and $b < 0$.

d) $y = ax + b$ with $a > 0$ and $b = 0$.

e) $y = ax + b$ with $a < 0$ and $b < 0$.

f) $y = ax + b$ with $a < 0$ and $b > 0$.

(a)

M9 The variable y is dependent of the variable x according to a first degree equation $y = ax + b$ with $a > 0$ and $b < 0$.

One of the equations describes the same relationship y and x. Which one?

a)

x	y
0	-42
1	-49
2	-56
3	-63
4	-70

b)

x	y
0	0
1	7
2	14
3	21
4	28

c)

x	y
0	42
1	35
2	28
3	21
4	14

d)

x	y
0	42
1	49
2	56
3	63
4	70

e)

x	y
0	-42
1	-35
2	-28
3	-21
4	-14

f)

x	y
0	0
1	-7
2	-14
3	-21
4	-28

(b)

Figure 4.3: Example items from the test. Item M21: a mathematics questions concerning the transition from graph to formula. Item M9: a mathematics question concerning the formula to table transition.

4.2.4 Other design choices

Some additional design choices are made to keep the test consistently structured and to eliminate any superfluous distracting information that might result in an incorrect answer unrelated to the essential concept under investigation. These are as follows:

(i) All chosen linear function types must be clearly visually distinguishable from each other in the graphical representation.

(ii) Mainly the first and fourth quadrants in the graphical representation are shown to avoid the use of negative times in a kinematics context. A small section of the second and third quadrants are also shown. In mathematics, these small sections are used to make clear that the domain of the function value is $]-\infty,+\infty[$. In the kinematics context negative times are discarded and the domain for the time variable t is $[0,+\infty[$.

(iii) The root is not present for any of the graphs except for the directly proportional one to prevent possible confusion between the y intercept and the root.

(iv) The y intercept is explicitly present in both graphs and tables.

(v) In the tables, the values of the independent variable are always the same: 0, 1, 2, 3 and 4. The inclusion of value 0 facilitates the deduction of the y intercept and the use of equidistant steps facilitates deduction of the slope. A random rotation between multiples of 3, 4, 6 and 7 (positive and negative) is used for the slope and the velocity since these still result in small numbers, thus avoiding potential problems with large numbers, and are not the easiest number either, such as 1, 2, or 5.

(vi) All formulas are represented in the same general format accompanied by declarations about the signs of the slope and the y intercept. For physics the terms are ranked from lowest degree to highest degree, for mathematics from highest degree to lowest degree.

4.2.5 Test structure

All the previously listed choices result in a 48-item multiple choice test with a block of 24 mathematics questions and a block of 24 — as isomorphic as possible — physics questions. For each representational transition there are a total of 8 questions, and for each function type there are 12 questions. To prevent a possible effect from the order of the contexts — first physics then

mathematics or the other way around — about half of the students received the mathematics questions first, followed by the physics questions, the other half received the physics questions first and then the mathematics questions. Within each context block the items have a random order, but that order is the same in both context blocks. Each multiple-choice question offers six possible answers: the six possible sign combinations of y intercept and slope. As previously mentioned, these include the four function types under investigation and the two buffer distractors which are shown in Fig. 4.1 in dashed lines. While solving the total of 48 questions students might start to realize the two buffer distractors are never the correct answer. To prevent this as much as possible, the order of the six possible answers are randomized differently for each item, though the symmetry between the mathematics and the physics test is maintained to keep the two blocks as isomorphic as possible.

4.2.6 Expert revision

To validate the structure and set up of the test we consulted with a member of academic staff with expertise in this field. The feedback praised the very systematic structure of the test, allowing an in-depth analysis of all the variables involved. Advise on slight changes in wording were adopted in the final version. Some concern was expressed about the use of numbers in the tabular representations and the absence thereof in the graphical and formulaic representations. Our aim to avoid the use of numbers wherever possible — thus avoiding potential problems with numbers — was conceived positively, but the inevitable use of numbers in tables does create an asymmetry in the test structure which might benefit or disadvantage certain representational transitions. Also, the expert picked up on our choice to use equidistant numbers for the independent variable in the tables, which includes 0 and has a step size of 1 unit. We were urged to clearly motivate this choice. In addition to the main motivations already given in Section 4.2.4, this choice was also motivated by the time limit since students would likely require more time to deduce the signs of the y intercept and the slope when numbers are nonequidistant — which also automatically results in larger and likely more difficult numbers — and by our goal to question all 6 representational transitions for at least 4 function types and 2 contexts.

The last point of the feedback was about issues with negative velocity in 1D kinematics. In English there is the distinction between the magnitude “speed” and the vector quantity “velocity”, while in Dutch, there is only “snelheid” to describe both. In our test, we want to question students about “snelheid” in graphs with a negative slope, in formulas with a negative coefficient in the first order term and in tables with a negative difference between entries for the

position. This means we actually ask for the x component of the velocity and not for the speed. Strictly speaking, when using snelheid in 1D kinematics, we mean velocity as a vector $(v_x\vec{e_x})$. We therefore expect the student to not only give the length of the velocity vector, but also include information on the direction by assessing v_x, which can be negative. A specific example in which this is of importance is item P15, which is a physics questions with a table to formula transition for a PN function type. The table given in P15 contains positive but decreasing values for the position x similar to the one in option c) in M9 shown in Fig. 4.3. In our analysis, we consider the velocity in every PN question to be negative, which is important when interpreting and comparing our results.

In addition to the revision by the academic expert, we also asked mathematics and physics in-service teachers to provide feedback on the test items to assure content validity. Two in-service teachers in each field were addressed. Comments were very positive and the overall difficulty was perceived suitable for the target group. Small suggestions for better wording were adopted. Some remarks were made on the difference in difficulty between decreasing and increasing functions. In particular, the teachers pointed out that students are very likely to have more difficulties with the tables in which negative function values are provided, as well as with the graphs of the function type with negative intercept and positive slope — the ones where an increasing function is shown in the fourth quadrant. In summary, the teachers in both fields expected the most difficulties with the use of negative function values and concluded that the test has a high content validity. Note that for function types with negative intercept, all values in the tables are negative since the choice was made to not include the root — except for the ZP function type.

4.3 Test administration

4.3.1 Participants

The data used for the validation were gathered from administering the test in early 2017 in Flanders (Belgium) with students from the 9th grade aged 14–15, who had opted to study sciences in secondary education. These participants were from 13 different schools distributed across Flanders. All these schools provide the same type of education, called “aso education”. Their students receive education on a broad number of topics with selected specializations such as sciences, languages, economics, humanities, etc., to prepare them for higher education. These schools voluntarily agreed to have the test administered to students in a learning trajectory with a strong science component. Informed

consent from the students' parents was obtained for all 385 students who participated in this study (213 males and 172 females). All of these students had received instruction on the topic of linear functions in mathematics and on 1D kinematics in physics to achieve specific learning goals set by government standards. All gathered data were collected and treated anonymously to protect student privacy. Ethical approval for this study was granted by the Sociaal-Maatschappelijke Ethische Commissie (SMEC)–Social and Societal Ethics Committee of KU Leuven.

4.3.2 Test procedure

This test was administered digitally using the Qualtrics® online software platform for digital surveys. The digital administration facilitated the randomization of the test structure at the respondent level, meaning that the odds are very small that any two respondents received the exact same test. The randomization within a test is also beneficial to reduce possible learning effects due to the order of the questions and of the answers, e.g. no four TG questions in a row. Furthermore the digital nature of the administration allowed for forced responses that forces respondents to select an answer in order to proceed to the next question. This resulted in a full data set without missing values. When administering this test the following procedure was used: respondents were informed that they would be participating in a research project, that no educational credit could be earned through the participation and that the maximum time available to solve all 48 multiple-choice questions was 45 min, which — from our pilot test — proved to be adequate for solving the full test. Neither students nor their teachers knew about the contents of the test beforehand, thus excluding any possible influence or teaching to the test that might occur.

4.4 Results

4.4.1 Validation

In this section we present the results from the validation[1]. We provide a per-item overview of relevant validation quantities in Table 4.1. In this table the

[1] An older version of this test had been administered to 148 students in a pilot study in 2016 as a first validation which resulted in good values for all relevant validation quantities. Due to some changes made to the test for a second administration in 2017, and due to the larger number of respondents in this second administration, the updated test is validated again with the new data from 2017.

items are sorted by representational transition and by function type to facilitate easy interpretation. The actual item order in the test during administration was randomized.

Table 4.1: The relevant validation quantities for each question Q of the test: item difficulty index P, the standard deviation SD for P and the item discrimination index D. The items are sorted by representation transition ($RepT$), function type ($FType$) and context to simplify interpretation and comparison. Question numbers starting with "P" denote a physics context and those with "M" a mathematics context.

RepT	*FType*	*Q*	*P*	*SD*	*D*	*Q*	*P*	*SD*	*D*
TG	NP	P1	0.85	0.36	0.37	M1	0.86	0.35	0.43
	ZP	P2	0.90	0.30	0.30	M2	0.88	0.33	0.37
	PN	P3	0.85	0.35	0.35	M3	0.86	0.35	0.37
	PP	P4	0.86	0.35	0.37	M4	0.85	0.36	0.43
GT	NP	P5	0.85	0.36	0.39	M5	0.85	0.36	0.46
	ZP	P6	0.90	0.31	0.29	M6	0.88	0.32	0.39
	PN	P7	0.85	0.35	0.42	M7	0.88	0.32	0.37
	PP	P8	0.85	0.36	0.43	M8	0.82	0.38	0.44
FT	NP	P9	0.45	0.50	0.80	M9	0.52	0.50	0.67
	ZP	P10	0.67	0.47	0.66	M10	0.72	0.45	0.62
	PN	P11	0.45	0.50	0.54	M11	0.51	0.51	0.65
	PP	P12	0.60	0.49	0.73	M12	0.65	0.48	0.67
TF	NP	P13	0.45	0.50	0.60	M13	0.59	0.49	0.60
	ZP	P14	0.62	0.49	0.71	M14	0.59	0.49	0.74
	PN	P15	0.46	0.50	0.69	M15	0.51	0.50	0.75
	PP	P16	0.60	0.49	0.80	M16	0.65	0.48	0.67
FG	NP	P17	0.56	0.50	0.60	M17	0.58	0.49	0.67
	ZP	P18	0.67	0.47	0.62	M18	0.75	0.43	0.55
	PN	P19	0.51	0.50	0.67	M19	0.61	0.49	0.69
	PP	P20	0.69	0.46	0.70	M20	0.73	0.44	0.66
GF	NP	P21	0.54	0.50	0.59	M21	0.59	0.49	0.73
	ZP	P22	0.63	0.48	0.74	M22	0.64	0.48	0.72
	PN	P23	0.53	0.50	0.63	M23	0.56	0.50	0.72
	PP	P24	0.68	0.47	0.67	M24	0.73	0.45	0.65

4.4.1.1 Internal consistency

To assess the internal consistency of the full test, Cronbach's α is calculated which results in a value of $\alpha = 0.946$, indicating a high internal consistency since it surpasses the 0.9 threshold [186]. To check the effect of each question on the internal consistency, α is also calculated for the case in which each item i is omitted from the test. These per-item results show values between 0.944 and 0.946, meaning that all items are worth retaining since omitting any item would either keep the value for α or lower it.

4.4.1.2 Item difficulty index

For dichotomous items such as multiple-choice items, the item difficulty index P is the proportion of the number N_c of correct responses to the total number N of responses given to that item:

$$P = \frac{N_c}{N}.$$

A higher value means an item is less difficult and a lower value means an item is more difficult. This measure can thus be used to provide an overview of the difficulty distribution of all the test items. The SD has been calculated over the full sample of 385 respondents. We present the results and the standard deviation for each item in Table 4.1 and show that item difficulty indexes for this test range from 0.45 to 0.90 with fairly large standard deviation and an overall mean and standard deviation of 0.69 ± 0.24, respectively. The first eight questions in the table – the TG and GT transition – have very high item difficulty indexes close to 1, meaning these transitions are very easy for students.

Tests can be categorized into two different types: norm-referenced tests (NRTs) and criterion-referenced tests (CRTs) [187]. NRTs usually cover a large domain of learning tasks and aim to compare groups and individuals to groups which is why they need to provide good discrimination and average difficulty — depending on the number of options in multiple-choice tests. Ideally the item difficulty indexes (or average scores) are normally distributed with the bulk roughly between 0.4 and 0.6. CRTs aim to assess the mastery of a specific topic or an individual's ability in regard to a criterion and serve to diagnose students' difficulties and certify their competencies. The item difficulty index distribution is of lesser importance, but is generally strongly skewed to higher scores for instructed groups (more than 0.8) and strongly skewed to lower scores for uninstructed groups (less than 0.3) [188]. Furthermore, items which are found to be too easy or too difficult are not omitted from CRTs, whereas

they usually are from NRTs. Since our test is designed to assess a specific competency (representational fluency) for a specific topic, it is more CRT than NRT oriented. At the same time though, we also want to use it to compare the STEM cohort with the control cohort. As stated in Section 1.4, we do prioritize the CRT-aspect and can conclude that our results indicate that the item difficulty indexes we have found are arguably well distributed for our purposes.

4.4.1.3 Discriminatory power

A test should allow for discrimination between respondents with high ability and respondents with low ability. This is known as the discriminatory power of a test and can be studied at the item level as well as at the full test level. To describe the power of a single test item to discriminate between high- and low-achieving respondents, we calculate D, the item discrimination index [186]. To do so, all total test scores are ranked and an upper (U) group and a lower (L) group is selected. The U group contains the upper 27% of respondents and the L group contains the lower 27% of respondents in this ranking. Taking the difference of the item difficulty index PU for U and PL for L results in the item discrimination index D:

$$D = PU - PL.$$

Results of this calculation are provided in Table 4.1 and range between 0.29 and 0.80 with a mean and standard deviation of 0.58 ± 0.15 respectively.

Similarly to the results of the internal consistency and item difficulty index, we find that the item discrimination indexes of the first eight items in the table are significantly smaller, indicating low discriminatory power for all TG and GT transition items.

The second way to describe the discriminatory power of the test, but now only at the full test level is Ferguson's delta δ. It represents the ratio of the interperson differences to the maximum number possible:

$$\delta = \frac{(k+1)(n^2 - \Sigma_{i=1}^{k} f_i^2)}{kn^2},$$

with $k = 48$ the number of items in the test, $n = 385$ the sample size and f the frequency of a certain score i, with i from 1 to k. The resulting value for δ is 0.99 which easily satisfies the minimum required value of 0.9, meaning the test offers good discriminatory power [186].

4.4.2 Test results

4.4.2.1 Main effects and interaction effects

To analyze the data, we use the generalized estimating equations (GEE) approach [189, 190], which is a specific variant of logistic regression analysis. This approach is required since the study design includes repeated observations within the series of responses of each respondent, for each variable that we study. The GEE method – in contrast to regular logistic regression – includes adjustments for the correlations between the repeated dichotomous observations [191] which might otherwise result in incorrect conclusions. The analysis is performed with IBM ® SPSS ® Statistics version 24.0.0.0 software in which the binomial probability distribution and the logit link function are selected, which are the appropriate choices for our research design. Additionally the parameter estimation method selected is Fisher; type III error is selected for analysis type; Wald statistics is chosen for the chi-square statistics; and the Bonferroni correction is selected to adjust for multiple comparisons. We include four factors in the model: representational transition, function type, context, and gender. Gender is included as a subject variable since Meltzer [80] has already shown there can be an interaction effect between gender and representation. The full model is analyzed, i.e., all factorial terms were included: main, 2-way, 3-way, and 4-way effects.

Results show a significant main effect of the representation transition, with Wald $\chi^2(5) = 437.063$ and $p < 0.001$. The pairwise comparisons indicate that transitions between tables and graphs are by far the easiest to perform in either direction. The mean accuracy for TG is 0.87 ± 0.01 and for GT it is 0.86 ± 0.01. Transitions between formulas and tables are most difficult to perform with a mean accuracy for FT of 0.57 ± 0.02 and for TF of 0.56 ± 0.02. Finally the difficulty of translating between formulas and graphs falls in between with a mean accuracy for FG of 0.64 ± 0.02 and for GF of 0.61 ± 0.02. There is a noticeable symmetry for representational transitions (e.g. GF and FG). This symmetry will be discussed in more detail later. Pairwise comparison between the transitions also supports this assessment since these comparisons always show a significant difference with $p < 0.001$ — with an exception between FT and GF which is $p = 0.002$ — but none of the "symmetrical pairs" differ significantly. Overall, accuracy is higher when a graph is involved and lower when a formula is involved; and translation to a formula is the most difficult direction. The analysis also shows a significant main effect of function type with Wald $\chi^2(3) = 158.923$ and $p < 0.001$. The pairwise comparison indicates that the two function types with a negative value for either y intercept or slope - NP and PN - were most difficult with mean accuracies of 0.66 ± 0.01 and 0.66 ± 0.02

respectively. The function type with the highest mean accuracy is the directly proportional one with zero y intercept and positive slope (ZP), with a value of 0.76 ± 0.01. A close second is the combination of positive y intercept and positive slope (PP), which resulted in a mean accuracy of 0.74 ± 0.01. Overall, the use of negative signs for y intercept or slope tends to result in significantly lower accuracies by about 10%, and respondents achieve the highest accuracies for the directly proportional function type. Pairwise comparisons between the function types show that they all significantly differ with $p < 0.001$ except for the comparison between ZP and PP which do not differ significantly with $p = 0.067$. Furthermore, a significant main effect of the context is found, with Wald $\chi^2(1) = 5.831$ and $p = 0.016$. The pairwise comparison indicates that the mean accuracy in mathematics is significantly ($p = 0.015$) higher than that in physics with respective accuracies of 0.72 ± 0.02 and 0.69 ± 0.02. Lastly, gender has no significant main effect with Wald $\chi^2(1) = 1.960$ and $p = 0.161$ for which the pairwise comparison indicates the mean accuracy for males is 0.72 ± 0.02 and for females 0.69 ± 0.02 which is not significantly different with $p = 0.162$. The 95% confidence intervals on all of these values are within ± 0.03, except for gender where the values are within ± 0.04.

In addition to the main effects the GEE analysis also indicated four significant interaction effects. The most important one is the interaction between representational transition and function type with Wald $\chi^2(15) = 137.672$ and $p < 0.001$. These results are shown in Table 4.2 and in Fig. 4.4.

Table 4.2: Mean accuracies for the interaction effect of representational transition (TG, GT, FT, TF, FG, and GF) and function type (NP, ZP, PP and PN). The 95% confidence intervals for these numbers are always within ± 0.04 and standard errors are between 0.013 and 0.021. Students have greater difficulties with transitions including a formula (F) and/or function types with negative y intercept or slope — NP and PN. There is a significant ($p < 0.01$) asymmetry for the ZP function type between FT and TF transitions.

	TG	GT	FT	TF	FG	GF
NP	0.85	0.85	0.48	0.51	0.57	0.56
ZP	0.89	0.89	0.69	0.60	0.71	0.63
PP	0.86	0.84	0.62	0.63	0.71	0.71
PN	0.86	0.87	0.48	0.48	0.56	0.54

The table and the graph clearly show the dominant patterns of the main effects for representational transition and function type again. For transitions between tables and graphs all accuracies are very high and differences are very small. The more interesting part of the interaction between representational transition

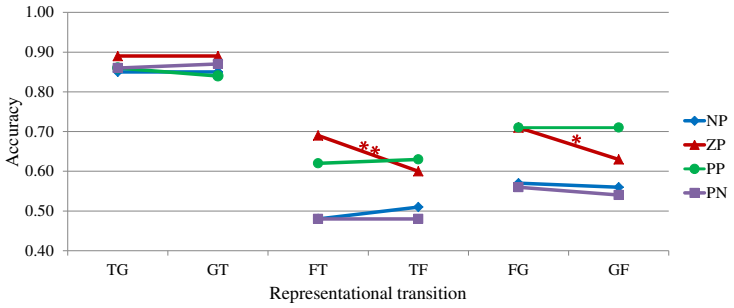


Figure 4.4: Data from the 2-way interaction effect of representational transition (TG, GT, FT, TF, FG and GF) and function type (NP, ZP, PP and PN). The data points connected by lines represent the pairs of transitions which are each other's opposite. The p value of the comparisons is indicated by: * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$. There are two significant asymmetries for ZP function type. The first between FT and TF transition with $p < 0.01$, and the second between FG and GF transitions with $p < 0.05$.

and function type is the significant asymmetry for the ZP function type between FT and TF transitions and between FG and GF transitions. There are two things to note here: first the asymmetry is consistent in the way that the accuracy is always the lowest when transitioning to a formula. Second, it is unexpected that this effect is so large for the ZP function type, which is the function type with the highest mean accuracy. A closer look at this effect for physics and mathematics separately reveals that most of it is confined in mathematics as is shown in Fig. 4.5 and Fig. 4.6 which essentially show the three-way effect of representational transition, function type, and context for which Wald $\chi^2(15) = 23.467$ and $p = 0.075$.

Another significant interaction effect revealed by the analysis is the two-way interaction of function type and gender for which Wald $\chi^2(3) = 10.089$ and $p = 0.018$, of which the results are shown in Fig. 4.7. The pattern from the main effect for function type is followed for both genders separately. The largest differences between them are for the NP and PN function types, which are also the two function types with the lowest mean accuracy. There are no interesting significant pairwise comparisons of note for this interaction.

The next significant interaction effect is the three-way interaction of function type, context and gender with Wald $\chi^2(3) = 8.207$ and $p = 0.042$, for which the results are shown in Fig. 4.8. There are no interesting significant pairwise comparisons of note for this interaction. The most interesting here, is that for the NP and PN function types there are slightly larger differences for

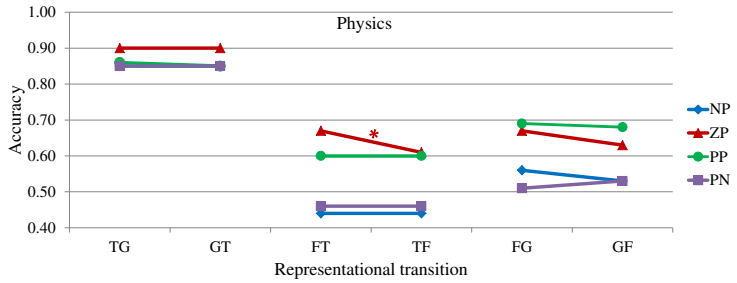


Figure 4.5: Data from the 3-way interaction effect of representational transition (TG, GT, FT, TF, FG and GF), function type (NP, ZP, PP and PN) and context for the context of physics. The data points connected by lines represent the pairs of transitions which are each other's opposite. The p value of the comparisons is indicated by: * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$. There is a significant ($p < 0.05$) asymmetry for the ZP function type between FT and TF transitions in the physics context.

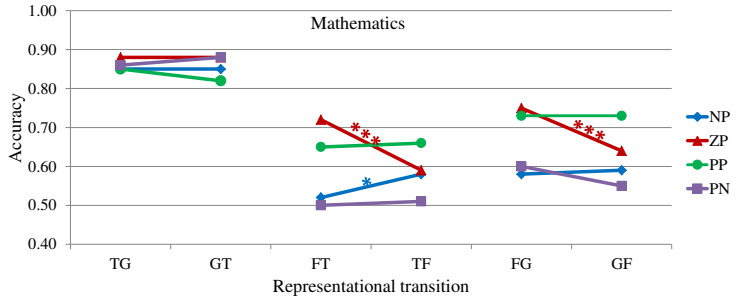


Figure 4.6: Data from the 3-way interaction effect of representational transition (TG, GT, FT, TF, FG and GF), function type (NP, ZP, PP and PN) and context for the context of mathematics. The data points connected by lines represent the pairs of transitions which are each other's opposite. The p value of the comparisons is indicated by: * for $p < 0.05$; ** for $p < 0.01$; *** for $p < 0.001$. In the mathematics context there are three significant asymmetries: for the ZP function type there is one between FT and TF transitions with $p < 0.001$ and one between FG and GF transitions with $p < 0.001$; for NP there is a significant asymmetry between FT and TF with $p < 0.05$.

males between physics and mathematics, and also that these function types correspond with the largest differences between female and male scores. For ZP and PP function types the scores are very similar; of note is that the accuracy in physics is slightly higher than in mathematics in one single case:

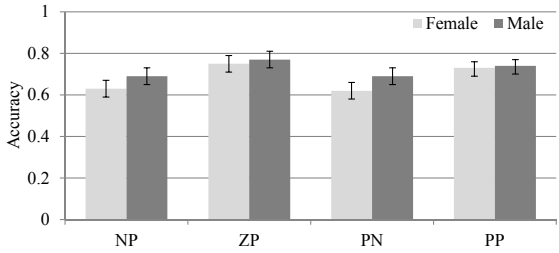


Figure 4.7: Data from the 2-way interaction effect of function type and gender. There are no significant differences between female and male for any function type.

for males solving ZP items.

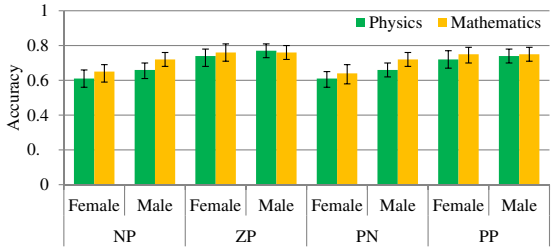


Figure 4.8: Data from the 3-way interaction effect of function type, context, and gender. There are no significant differences between physics and mathematics for any gender and function type combination.

The last significant interaction effect is the 4-way interaction of all factors: representational transition, function type, context and gender which has Wald $\chi^2(15) = 27.825$ and $p = 0.023$; this is shown in Fig. 4.9. The patterns of the main effects are still dominant here and the only significant difference between a representational transition and its opposite is with males in mathematics between the FG and GF transitions. There are other asymmetries as well - although not significant ones - they are mostly found in mathematics and almost exclusively for the ZP function type. In physics the largest asymmetry between transitions is again for the ZP function type, specifically with females. Furthermore the scores are spread out across a larger range for females as compared to males for both mathematics and physics.

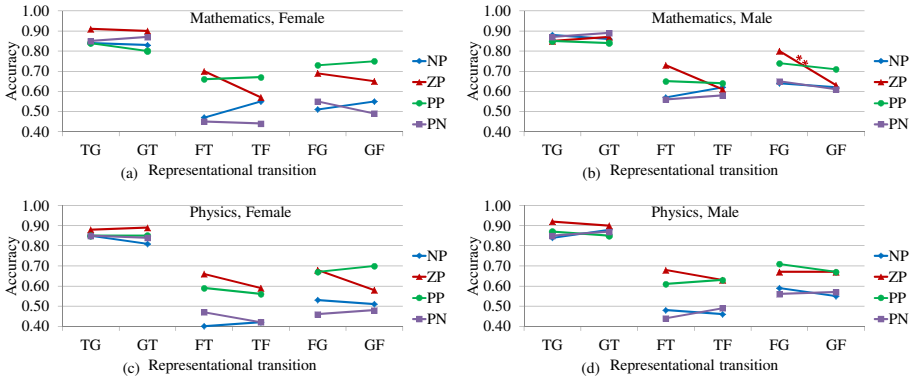


Figure 4.9: Data from the 4-way interaction effect of representational transition, function type, context and gender. (a) Mathematics and Female; (b) Mathematics and Male; (c) Physics and Female; (d) Physics and Male. The p value of the comparisons is indicated by: * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$. The only significant asymmetry is for the ZP function type between FG and GF transitions in the mathematics context with males.

4.4.2.2 Distractor analysis

Since the distractors for each question always contain the same six function types, the students' incorrect responses and their relation with representational transition and context can also be studied. Table 4.3 shows the frequencies for each of the distractors in the students' responses.

Table 4.3: Frequency (in percentage) of the chosen function type (NP, ZP, PP, PN, NN, or ZN) in the answer as function of the function type (NP, ZP, PP, and PN) in the question from all 385 respondents. Correct answers are in bold face. Three pairs of interest are marked in different shades of gray which indicate possible distinct errors which are further analyzed in the text.

	NP	ZP	PP	PN	NN	ZN
NP	**64.0**	5.2	6.1	9.5	9.5	5.7
ZP	3.5	**73.7**	8.4	4.4	3.5	6.5
PP	5.0	8.3	**72.7**	5.0	5.6	3.5
PN	9.9	6.0	9.3	**63.4**	6.1	5.4

None of the distractors has a prevalence higher than 10.0%, let alone higher than 16.7% which would be the prevalence if answers were chosen at random

in a test with 6 distractors. It is nonetheless insightful to have a closer look at the most frequent errors. Using Table 4.3, three pairs of question and answer function types stand out, which are marked in shades of gray. In the case of the darkest shade (error E1) students choose NP in a PN question, and vice versa, choose PN in a NP question. This error can be explained by either a "sign switching" error, in which the students switch positive and negative for both the intercept and the slope; or it can be explained by "concept switching," in which students switch intercept and slope. Analysis at the student level reveals that just shy of 30% of students who make this error, do so symmetrically: switching NP for PN and also switching PN for NP in the same context and for the same representational transition. This indicates some persistence to making a symmetrical E1 error.

The second pair, indicated by the middle shade of gray (error E2) in Table 4.3, identifies those who choose PP in a PN question and those who choose NN in an NP question, which can be explained by "sign switching" of the slope but at the same time correctly interpreting the sign of the intercept. Just over 20% of students who make such an error make it symmetrically: they choose PP in a PN question and also NN in an NP question. This also indicates at least some persistence to making a symmetrical E2 error, but less compared to the E1 error.

The last pair, indicated by the lightest shade of gray (error E3) in Table 4.3, shows switching of ZP and PP function types which suggests difficulty distinguishing between zero or positive intercept with a positive slope. Analysis at the student level shows that very few students make a symmetrical E3 error, most often students make this error only once.

In summary, we have identified the following three errors as the predominant errors made in the test, but since the prevalences in the total population are very low, there is no indication of any fundamental issues with teaching this subject matter:

- (Symmetrical) concept switching of intercept and slope.
- (Symmetrical) sign switching of the slope or of both slope and intercept.
- Mostly nonsymmetrical switching of ZP and PP function types.

More detailed data of the distractors as a function of context and representation transition are provided in Appendix A.3 in addition to a brief description of the most notable data points. Since these data are highly specific they are not discussed further in the main body of the paper.

4.5 Discussion

4.5.1 Validation

The validation results of the test show that the test is consistent, reliable, has a good distribution of item difficulty and good discriminatory power on a per item basis and for the whole test as well. This means this test is indeed appropriate to measure representational fluency for the specific use case of linear function types in physics and mathematics. In part, these good results are influenced by the relatively large number of questions (48) and their high similarity, which makes solving the test repetitive in nature. Furthermore, the very focused topic under study also helps in achieving these results. The TG and GT transitions stand out with their consistently high item difficulty indexes P and low item discrimination indexes D, signaling that these items might be too easy to be included in the test. Since these items are part of the research questions, they are not omitted or replaced by other items.

4.5.2 Results

All factors in the test design prove to have a significant main effect in the statistical model, thus supporting the inclusion of each one in the test design and their influence on the students' performance. One factor that is not in the test design but is included in the analysis, is gender; this factor proves to be nonsignificant as a main effect but is included in three out of four significant interaction effects, each time accompanied by the function type factor. This supports the inclusion of gender as a factor, which was inspired by the work of Meltzer [80] and again confirmed by our results. Teachers should be aware of this difference. In practical terms this implies that females could benefit from extra support for the most difficult function types (NP and PN) and/or the most difficult context (physics). Our research design does not allow us to specify how such extra support might differ between females and males, though. Although we do not find a significant gender gap in our test, other studies have shown gender gaps to exist and persist in various concept inventories and tests with a broader scope than ours [192–195]. Recent studies [196, 197] point to the use of a multidimensional context and the use of diagrams (phase diagrams, free body diagrams, multidimensional graphs, etc.) in the questions as factors which can create a bias in favor of male respondents, thus offering a possible explanation for the gender gap. Our results are complementary to these observations, since we find no significant gender gap in a test in which the questions are all restricted to a single dimension and in which no diagrams (other than 1D graphs) are present.

The main results are those concerned with representational fluency and the influence of function type and context. As was already clear from the validation, the TG and GT transitions are by far the ones with the highest accuracies, even so much that there is little distinction when comparing function types or contexts for these two transitions. Students are clearly very fluent at switching between tables and graphs — both ways — for all four function types in both contexts, which is in agreement with the results from De Bock, Van Dooren, and Verschaffel [78]. This can likely be explained by the direct visual correspondence between the numbers in the tables and the signs of the axes in the graphs which facilitates these transitions. In contrast, the transitions going from or to a formula are significantly more difficult for all function types and contexts, with the FT and TF transitions achieving the lowest mean accuracies, and FG and GF slightly higher ones. The cause of these lower accuracies is probably the higher abstraction level of formulas which is more difficult to relate with numbers. Contrary to our results, De Bock, Van Dooren, and Verschaffel. consistently found FG to be the single transition with the lowest mean accuracy for their selection of function types. A likely explanation for this difference is their inclusion of an inversely proportional function type and their exclusion of function types with negative y intercept, since our results show that function types with negative values for either y intercept or slope result in significantly lower mean accuracies, thus skewing these accuracies towards lower values, in particular for FT and TF transitions.

An important difference between our test and the one from De Bock, Van Dooren, and Verschaffel [78] is the abstract notation in the formula representation. We chose to make use of "greater than zero", "less than zero" and "equal to zero" instead of using numbers, which seems to be more difficult for students. Figures 4.10, 4.11 and 4.12 show the comparisons of the function types ZP, PN, and PP, which are investigated in each study. These figures show larger differences whenever formulas are involved, which suggests that our more abstract representation does negatively impact students' performance when compared to the use of numbers in formulas. Furthermore, they consistently show the largest differences between the two studies for the FT and TF transitions and smaller differences for FG and GF transitions. A plausible explanation for this is that the graphs mainly show the first quadrant which makes it easier to visually distinguish positive and negative slope, while in tables the use of negative numbers for the dependent variable (fourth quadrant) makes it more difficult to identify the sign of the rate of change. This in combination with the more abstract formulas might explain the observed differences. An additional issue with our more abstract notation

is that the "less than" symbol might be more intuitive[2] for students than the "greater than" symbol since we read and write from left to right and tend to sort numbers from smallest to largest. This means that students might find $a < 0$ and $0 < a$ easier to interpret than $0 > a$ and $a > 0$ for example. This is only speculative though, and could be an interesting topic for future research. There are only two cases in which our results are higher than those from De Bock *et al.*: the TG and FG transitions for the PN function type. In the TG situation, the difference is just within our 95% confidence interval, but in the FG case the difference is more significant. The latter case appears to be a unique situation in which the use of numbers actually negatively affects students' performance. The explanation for this is possibly linked to the negative slope of the PN function type in combination with the lack of numbers in our FG items.

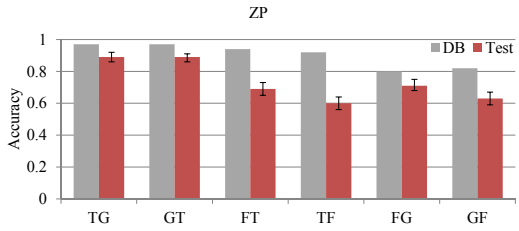


Figure 4.10: Comparison of the data from De Bock, Van Dooren, and Verschaffel [78] (DB) and our data (Test) for the ZP function type for each of the representational transitions.

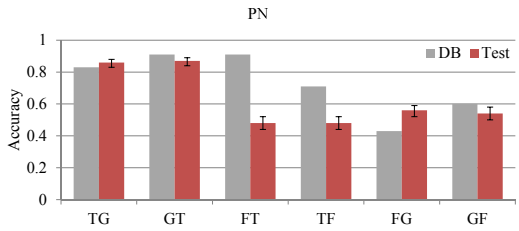


Figure 4.11: Comparison of the data from De Bock, Van Dooren, and Verschaffel *et al.* [78] (DB) and our data (Test) for the PN function type for each of the representational transitions.

[2]With *intuitive*, we mean in a way that does not involve analytical reasoning. An interesting analogon is that of a biological reflex which is a reaction to a stimulus that does not require processing by the brain or conscious thought.

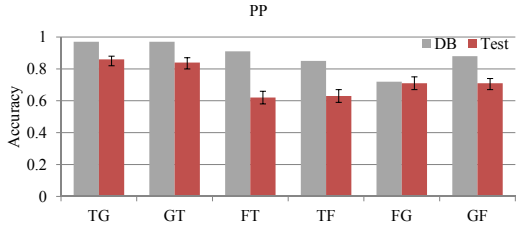


Figure 4.12: Comparison of the data from De Bock, Van Dooren, and Verschaffel [78] (DB) and our data (Test) for the PP function type for each of the representational transitions.

Ideally representational fluency between two representations is "bidirectional," meaning that both directions of the transition should be equally fluent. In other words, the part "the ability to translate between representations" of our definition of representational fluency is ideally a symmetric ability. When studying this symmetry, we found that the accuracy for the directly proportional function type (ZP) has significant asymmetries since the transition to a formula has significantly lower accuracies when compared to the transition starting from a formula. This effect is present in both contexts, but is more pronounced in the mathematics context. This means that the function type with the *highest* mean accuracy (ZP) has the largest asymmetry — which is an unexpected result. In relation to the representational transitions, the asymmetry is very pronounced with the most difficult transitions, i.e. FT/TF and FG/GF transitions. This means that the pairs of representational transitions with the *lowest* mean accuracy — FT/TF and FG/GF — have the highest asymmetry and that the pair of representational transitions with the *highest* mean accuracy — TG and GT — have the lowest asymmetry — which is more in line with what one would expect.

When comparing physics and mathematics our results indicate a significant difference of 0.03 between the mean accuracies for mathematics and physics, in favor of mathematics. This result is in line with that from Planinic *et al.* [47], who also found mathematics to be the easiest context when compared to physics. Their study was for different concepts though: graph slope and area under a curve. Furthermore in another study [46] of theirs — in which they studied both positive and negative slope of line graphs in physics and mathematics — they also found that questions were solved better in a mathematics context compared to a physics context and that questions with negative slope were more often answered incorrectly. Our data do not really allow us to identify why students experience more difficulties with physics compared to mathematics,

but comparing Figs. 4.5 and 4.6 does reveal that the largest discrepancies between the two contexts are with the NP questions with FT/TF transitions and also for the ZP questions with the FG transition.

Our findings provide accurate answers to research questions 1 and 2. The answer to research question 3 is that the most important significant interaction effect we find is the 2-way interaction between representational transition and function type; furthermore, there is a significant 2-way interaction between function type and gender; a significant 3-way interaction between function type, context, and gender; and a significant 4-way interaction between all factors involved: representation transition, function type, context, and gender.

4.6 The effect of the iSTEM approach

In this section, we discuss the effect of the instructional approach, i.e. iSTEM compared to the regular instruction, to answer the fourth research question in this chapter: *How does students' representational fluency between graphs, tables and algebraic formulas of linear functions in mathematics and physics (1D kinematics) compare between iSTEM students and students in the regular curriculum?*

To answer this question, we present the remaining part of the full analysis of the data by including the instructional approach as a variable in the GEE model. In doing so, we have the full group of 385 respondents which is composed of two distinct cohorts. The first cohort contains 195 iSTEM students with an average age of 14.9 year ($SD = 0.4$ year), of which 60 are female and 135 are male, from eleven different schools. The second cohort is the control group and consists of 190 students with an average age of 14.9 year ($SD = 0.4$ year), of which 112 are female and 78 are male, from nine different schools.

The test administration was performed after all students had received instruction about linear functions in mathematics and about 1D kinematics in physics. The main difference is that the iSTEM students used the Green wave course module from the STEM@school project, which was discussed in Chapter 3.

To compare these two cohorts, we extend the GEE model from the work above by including *Instruction*, which is either iSTEM or the regular instructional approach in the control group. This means that the extended model now includes all main effects and all interaction effects up to the 4-way effects. Table 4.4 shows a comparison of all relevant main effects and interaction effects between the previous GEE model and the extended GEE model.

Table 4.4: Significant main effects, significant interaction effects, and interaction effects of interest with Wald χ^2, the significance (p-value) up to three digits, and degrees of freedom (df) for the results from the paper and the extended model which includes the instructional approach (iSTEM or control group).

	Paper		Extension		
Term	χ^2	p	χ^2	p	df
RepT	437.063	0.000	418.376	0.000	5
FType	158.923	0.000	142.044	0.000	3
Context	5.831	0.016	6.525	0.011	1
Gender	1.960	0.161	3.215	0.073	1
Instruction	-	-	3.200	0.074	1
*RepT * FType*	137.672	0.000	117.784	0.000	15
*RepT * Instruction*	-	-	5.775	0.329	5
*FType * Gender*	10.089	0.018	8.584	0.035	3
*FType * Instruction*	-	-	6.843	0.077	3
*Context * Instruction*	-	-	0.002	0.967	1
*Gender * Instruction*	-	-	1.720	0.190	1
*ReptT * Gender * Instruction*	-	-	16.067	0.007	5
*FType * Context * Gender*	8.207	0.042	5.793	0.122	3

The GEE analysis shows a very similar structure with only a few small changes. *Instruction* is not significant as a main effect nor in any 2-way effect, but is found to be significant in $RepT * Gender * Instruction$. The one effect which changed the most is $FType * Context * Gender$ which is no longer significant. Since *Gender* is a factor in both these terms, it might explain these results. This is also supported by the imbalance between the gender ratio in the iSTEM cohort and in the control cohort: the former has a female/male ratio of 0.44 and the latter one of 1.4. Close inspection of the pairwise comparisons though, shows that this is not the case. The reason why these terms are significant is mainly because of the *RepT* and *FType* factors respectively. For example: for males in the iSTEM cohort, there is a significant difference between the TG transition and the GF transition. This is also true in the control cohort and also for females. An example of the other term is that there is a significant difference between the ZP and the PN function type for male students in physics. The same is found for females and also for males in mathematics but not for females in mathematics. A large number of such 3-way effects can be found in the results, but they do not merit much more discussion in light of the aims of

this dissertation.

The most important result in the extended GEE analysis is that, in addition to no significant main effect for *Instruction*, no significant 2-way effect was found including *Instruction*, which clearly shows that there was no significant effect from the new approach on students' representational fluency for the graphs, tables and algebraic formulas of linear functions in mathematics and 1D kinematics. The main effect for *Instruction* resulted in a mean accuracy for the iSTEM cohort of 0.69 ($SE = 0.22$) and for the control cohort of 0.74 ($SE = 0.18$). The results for the 2-way interaction effects with *Instruction* are shown in Figs. 4.13, 4.14, 4.15, and 4.16. Although the mean accuracies for the control cohort are consistently a few percent points higher than those for the iSTEM cohort, none of them are significant.

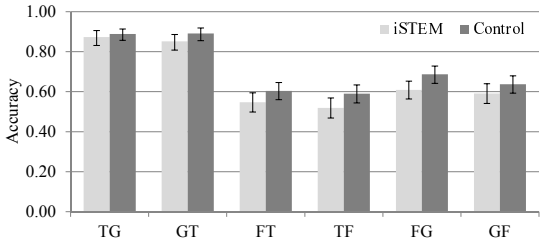


Figure 4.13: GEE results for the 2-way interaction term *RepT* * *Instruction*.

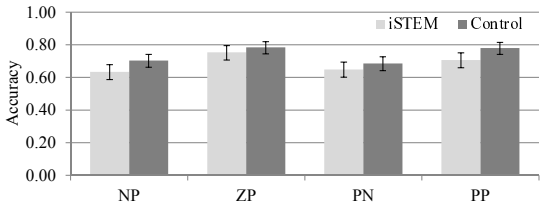


Figure 4.14: GEE results for the 2-way interaction term *FType* * *Instruction*.

Previously, we identified various significant asymmetries in the GEE model excluding the instructional approach. A close inspection of the GEE analysis can show whether these are specific to an instructional approach or not. The term *RepT* * *FType* showed significant asymmetries for the ZP function type between the FT and TF transition and between the FG and GF transition (Fig. 4.4). The results for the *RepT* * *FType* * *Instruction* term in the extended model show that there is only one significant ($p = 0.010$) asymmetry, which is for the ZP function type between the FG and GF transition in the

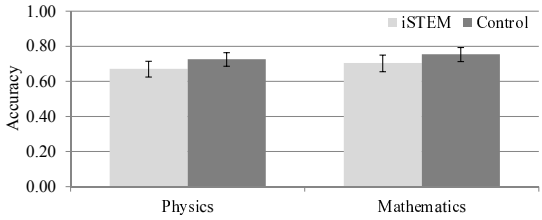


Figure 4.15: GEE results for the 2-way interaction term *Context* * *Instruction*.

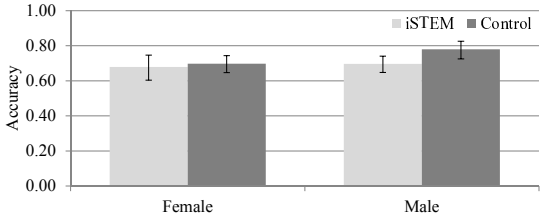


Figure 4.16: GEE results for the 2-way interaction term *Gender* * *Instruction*.

control cohort with mean accuracies of 0.75 and 0.63 respectively. This is not the case for the iSTEM cohort, for which we find $p = 1.000$ and mean accuracies of 0.68 and 0.63 for the FG and GF transition respectively.

The 3 way effect *ReptT* * *FType* * *Instruction* detailed this further in the first GEE model and identified one significant asymmetry in physics for the ZP function type between FT and TF transitions (Fig. 4.5), and three significant ones in mathematics (Fig. 4.6): one each for the NP and ZP function types between the FT and TF transitions, and one for the ZP function type between the FG and GF transition. Inspecting these again with the inclusion of the *Instruction* variable in the non significant *RepT* * *FType* * *Context* * *Instruction* term shows that there are no significant asymmetries in the iSTEM or the control cohort in the extended GEE model.

This means that, with the exception of the ZP function type between FG and GF transitions in the control cohort, none of these asymmetries are specific to the iSTEM or to the control cohort.

The extended GEE analysis results in a positive conclusion because it shows that students instructed with the iSTEM approach are able to achieve a of level representational fluency which is not significantly different from that of those instructed with the regular approach. This is especially noteworthy because

of the large required change in the mindset of both teachers and students in the iSTEM approach, and because of the fact that this data was gathered after only the second implementation thereof.

4.7 How can these findings be used in teaching?

Here, we link our theoretical findings to practice and teaching since it has often been acknowledged that there is a gap between educational research and practice [198],

The key idea is that the use of MERs can increase conceptual understanding [181]. Specifically, the sequence of learning with MERs, i.e.: the order in which a student learns to describe and comprehend the same concept in different representations. An optimal sequence would be scaffolded — since it could reduce the learner's cognitive load [199] — and start with the easiest representation and the easiest transitions and from there on increase in difficulty. This can be useful when the goal is to learn about the relation between representations. Which means learning how elements in one representation relate to elements in another representation — and not necessarily to learn about the transition between them — i.e., the exact method or procedure to do the transition, which is more complex. This is in line with Ainsworth's second design principle or use for learning with MERs [180], i.e.: "automating translation when MERs are used to constrain interpretation". Possible methods are automatic translation or dynamic linking [181] — i.e. the process where the manipulations performed by a learner in one representation are automatically translated by a third party (such as a teacher or a computer) to changes in another representation thus freeing the learner to learn about the relation between the representations instead of focusing on the more difficult task of translation [200,201]. For example, a computer applet could be designed in which a student can manipulate a linear curve in a graph by shifting it or changing the slope and immediately see the matching changes to a formulaic expression that represents the same linear relationship. Simple tasks such as trying to make one of the coefficients in the formula zero, positive, negative or a specific value can help students understand the relation between the representations and the meaning of certain elements in each representation. In the case of Ainsworth's third design principle or use for learning with MERs "scaffolding translation when MERs are used to develop deeper understanding," the stepwise increase in difficulty within a sequence would be complementary to the scaffolding of support by the teacher to gain deeper understanding. Thus, these design principles aim to reduce the cognitive load of the translation or

aim to construct a carefully scaffolded increase in difficulty. For both purposes, a well-structured sequence can be of great practical use to teachers.

According [3] to McGee and Martinez-Planell [202], a good sequence is a "synergy of registers" [90] and not a "semiotic chain" [203]. In a "synergy of registers" one can perform the transitions between at least two different representations, which is called a "two-register synergy." Preferably more registers are included such as with the "three-register synergy" shown in Fig. 4.17(a). In a semiotic chain — Fig. 4.17(b), one only learns to perform the transitions along the chain which might result in lower understanding. Thus it is advisable [202] to explicitly add missing transitions in a semiotic chain. Not every transition must be included though; sometimes a "base representation" exists. This is a representation that is used as a layover representation in the transition process. In Fig. 4.17(c) "Formula" is the "base representation" and serves as a layover between, e.g. "Diagram" and "Table". Heckler and Scaife [204] studied vector addition and subtraction in the arrow and ijk representation and presented empirical evidence for the existence of a natural hierarchy for teaching these operations in which the ijk format can be a useful base representation. A base representation can be present during the learning process and can remain so thereafter, but ideally an "all-register synergy" in which all transitions are present is achieved. In short, it is best to have as many "closed loops" as possible in a sequence diagram.

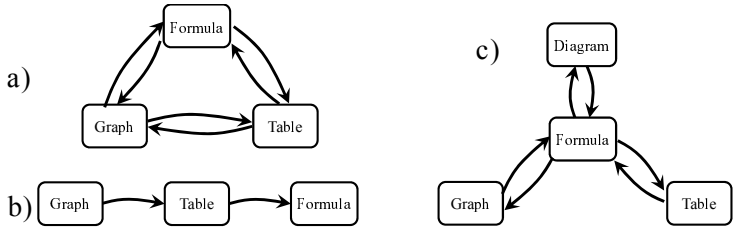


Figure 4.17: Hypothetical examples of (a) an all-register synergy; (b) a semiotic chain; (c) a synergy of registers with missing transitions and a central "base representation" (Formula). A formula can be any algebraic expression, and a diagram could be a free body diagram, or a phase diagram for example.

The average difficulties for the transitions in our data offer similar information as an all-register sequence diagram for tables, graphs, and formulas and suggest to start from graphs and tables when introducing linear relationships. In physics, one could think of plotting the $x(t)$ graph of a uniform linear motion using a motion sensor. The experiment will show a straight line graph that

[3] They — and some other authors — describe an external representation as a semiotic register, the two terms can be used interchangeably

can immediately be coupled to the corresponding table with the measurement points. Using different initial positions and or velocities, all different function types for linear relationships can then be identified (NN, NP, ZN, etc.). Next, the formula representation can be introduced by transitioning from graphs and tables to formulas by first matching the shape of the curve with the formula; and finally, the transitions between tables and formulas are treated. Furthermore, the general similarities for the representational transitions between the physics data and the mathematics data suggest that it would be best to do this in both contexts simultaneously. Ideally, all possible representations are then compared together at the same time so students can recognize the same function type in each representation. Such a summary exercise could be a perfect fit for the dynamic linking approach by creating a computer application which simulates or measures the motion of a moving object and represents the position and time data in various representations simultaneously. An important note here is the issue of acceleration when doing an experiment. A toy car starting from a standstill must first accelerate before it can achieve a nonzero constant velocity. This nonlinear behavior will show up in the measurement data. When doing such experiments it is highly recommended to use these phenomena as learning opportunities for students and teach them about acceleration and the limitations of a linear model when applied to reality. A useful comparison to make when doing such experiments is comparing a "rolling start" to a start from a standstill and see how the tables and graphs differ.

4.8 Future research

In this section we briefly discuss some possible future research directions as continuations and refinements of this study. First, the test structure can be easily applied to other representations, other transitions, other function types — or a different topic altogether — and other contexts. More subtle modifications to the current test can also be made, such as using different quadrants in the graphs and tables or including the roots in the graphs and tables which might open up new errors such as mistaking the root for the y intercept. Second, our study is more focused on the quantitative side, it would be insightful to complement this with qualitative research. A good suggestion would be to conduct think aloud interviews, which might provide a better understanding of the most frequently made errors. Third, the way we interpret the representational fluency in this multiple-choice test needs some caution. In each item, we show respondents a specific function type in a certain representation, such as a formula. At the same time and on the same page or screen the respondents are presented with six alternatives in another representation, such as tables. We assume that this example studies

the representational fluency of transitioning FT from a formula to a table but this might not be the way students actually solve the question. A student might identify the function type in the formula to a base representation (for example a graph), thus performing a FG transition and then transition each answer option to a graph to find the matching graph, thus performing the TG transition a maximum of six times. This example might actually be very realistic since our research — and that of other cited in this work — shows that graphs are the easiest representation of the three under study, which might indicate that it serves as a base representation when transitioning between other representations. Alternatively — and following the same example — a student might transition all six answer options to formulas, thus performing the TF transition 6 times and not the FT transition as intended. Furthermore, a student might make the FT transition once and then find the matching table from the six options, which is how we interpreted the results. Other constructs are of course also possible. To understand the exact process in a student's reasoning, the proposed think aloud interviews could provide more insight into this issue. This also reconnects with the remark made in the introduction and literature section which was that instead of using the term representational fluency, the term representational recognition might also be appropriate since students might actually recognize the identified mathematical concept of the question in the answer possibilities instead of performing an active conversion of the identified mathematical concept.

4.9 Conclusion

This study resulted in a validated 48-item multiple-choice test to study representational fluency in physics and mathematics on the topic of linear functions. The test has a high Cronbachs α of 0.946, a good distribution of item difficulty indexes and item discrimination indexes, as well as a very good Ferguson's δ of 0.99. Our results show significant main effects for representational transition, function type, and context. Students prove very adept in transitions between tables and graphs; they have the most difficulty with transitions which included a formula, and in particular for the directly proportional function type, the transition to a formula stands out from the analysis. Function types with negative values for either y intercept or slope also result in significantly lower mean accuracies indicating the difficulty students have with negative numbers in linear functions. When comparing contexts the analysis results in a significant difference between physics and mathematics. We found that the mean accuracies in physics are significantly lower compared to those in mathematics. The GEE analysis also shows four significant interaction effects, of which the most important one is between

representational transition and function type. The other significant interaction effects always include gender and function type. The main trend there is that the difference between male and female accuracies is larger for the most difficult function types and for the most difficult transitions. When comparing the symmetry of the representational transition between any two representations, the analysis clearly shows a significant asymmetry for the directly proportional function type. This unexpected result is most evident in the transitions between formulas and tables and between formulas and graphs in which the transition to a formula always has the lowest accuracy. Interestingly, the largest asymmetry is found for the — on average — *easiest* function type (ZP) and for the — on average — *most difficult* representational transitions, which are the ones containing a formula. Furthermore, this effect is most pronounced in the context of mathematics. The errors students make most in these particular cases are selecting the function type with positive y intercept and positive slope instead of zero y intercept and positive slope. The distractor analysis results in the identification of three distinct dominant errors: (i) (symmetrical) concept switching of intercept and slope; (ii) (symmetrical) sign switching of the slope or of both slope and intercept; (iii) (mostly nonsymmetrical) switching of directly proportional function type and the function type with positive y intercept and positive slope.

Overall, the analysis shows that students' knowledge is strongly dependent on representational transition, function type and interactions. Furthermore, our study confirms similar results about the difficulty of various representational transitions from De Bock, Van Dooren, and Verschaffel [78] and highlights the additional difficulties students experience with our more abstract representation in formulas compared to the use of numbers in formulas in their study. We emphasize that this study focuses on a specific definition of representational fluency and the ability to transition between representations. Our results provide quantitative insight about this, but do not allow us to determine how students perform the transition. To better understand these processes, additional and more detailed information on students' reasoning and their use of other representations is needed.

The last research question we aim to answer in this study concerns the effect of the instructional approach on students' representational fluency. The results show that there is no significant main or 2-way effect including the variable *Instruction*. The one significant 3-way effect including this variable, i.e. *RepT* * *Gender* * *Instruction*, did not result in any comparisons of interest. This means that, concerning students' representational fluency between tables — graphs and formulas of linear functions in physics and mathematics — the iSTEM approach resulted in student performance levels which do not significantly differ from those of the students in the regular

approach. Considering that the implementation of a new instructional approach takes some iterations before teachers and students are adapted to it, and that these measurements were done after only the second implementation, the achievement of a similar level of representational fluency is a positive outcome and provides a solid foundation from which to build upon in the future.

Chapter 5

Students' understanding and strategies when solving linear function problems in kinematics and in mathematics

While in the previous chapter we studied students' representational fluency between graphs, tables and formulas of linear relations in physics and in mathematics, this chapter studies more in detail students' understanding of the concepts slope and y intercept, both quantitatively and qualitatively. Again, answers from students in the STEM@school curriculum and from the regular curriculum are compared, although that section — Section 5.8 — is not included in the published paper. All other parts of this chapter are based on the following peer-reviewed paper which is reproduced here with only minor modifications:

> S. Ceuppens, L. Bollen, J. Deprez, W. Dehaene, and M. De Cock, "9th grade students' understanding and strategies when solving $x(t)$ problems in 1d kinematics and $y(x)$ problems in mathematics," *Phys. Rev. Phys. Educ. Res.*, vol. 15, p. 010101, Jan 2019

The first author designed and administered the tests and analyzed the results. All authors contributed to the test design and reviewed the results from the analysis.

5.1 Introduction

Linear functions are an essential part of any science curriculum but recent studies show that students' understanding thereof is subpar and that the context is an important factor [46, 47, 78, 98, 108, 109, 175, 184]. Therefore, and similar to Chapter 4, the subject of the study presented here is linear functions, which is an important topics in the mandatory curriculum and in the STEM@school curriculum, specifically the green wave course module described in Chapter 3. This time though, the focus is on 9th grade students' understanding and strategies when solving $x(t)$ problems in uniform linear motions in 1D kinematics and when solving as-isomorphic-as-possible $y(x)$ problems in mathematics. Specifically, we focus on the two essential concepts which define a linear relationship: the slope and the y intercept in mathematics and the velocity and the initial position of a uniform linear motion in kinematics. Additionally, we include the representation and the type of question on students' accuracy and strategies. We do this for both an iSTEM cohort and a control cohort with students enrolled in the regular curriculum.

In doing so, this chapter aims to provide an answer to the second research question as formulated in Chapter 1: *How does grade 9 students' performance and conceptual understanding of linear function problems compare between kinematics and mathematics and between the STEM@school instructional approach and the regular instructional approach?*

In order to do so we first provide an overview from literature about the relevant variables in our study. Next, formulate more specific research questions and describe the research design, the development of a test and the test administration. The test was administered in classes participating in the STEM@school curriculum and in classes following the regular curriculum. A full analysis was then performed, which showed that there is no significant effect from the instructional approach. Similar to the approach in our previous study, we then excluded this variable to simplify our statistical model and published a dedicated paper about the most effective model [205]. This paper is reproduced in this chapter and discusses the quantitative and qualitative results. Section 5.8 then focuses specifically on the effects of iSTEM approach in light of this dissertation. Thereafter, we formulate some suggestions for teaching and some implications for future research. Lastly, we summarize the results in the conclusion of this chapter.

5.2 Literature

Linear functions are of paramount importance to describe, understand and approximate many phenomena in physics. Students encounter them throughout their education and they are often the first step in learning about abstract relationships between dependent and independent variables. In most curricula, students start learning about linear functions in lower secondary education. At that time, they are likely to already have some deep-rooted beliefs and tend to operate in different belief systems: the real world, the physics world, and the mathematics world, which are not always well linked [27]. Multiple studies on this topic show that conceptual understanding of linearity is indeed not easily obtained and many difficulties have been reported across various contexts and across various representational formats such as graphs and formulas [45–47, 78, 81, 85, 92, 94, 96, 98, 107–112, 175, 183, 184, 206, 207]. The target groups in these studies range from secondary to higher education but the results are very similar which is suggestive of some obstinacy for these difficulties throughout education.

The slope of graphs of linear functions in physics has been extensively studied and the most common difficulties were grouped by Leinhardt *et al.* [107] into three distinct categories: slope/height confusion in which the height of the curve is mistaken for its slope; interval/point confusion in which a single point is considered when an interval is more appropriate, i.e. taking the ratio of coordinates instead of the ratio of intervals to determine the slope; or iconic interpretations in which the graph is viewed as a direct representation of reality. The latter was found to be the most frequently made error. Wemyss and van Kampen [108] categorized interpretations of numerical linear distance-time and $y(x)$ graphs from first-year university students enrolled in an algebra-based course. They found that both context and prior learning are likely of influence to explain poor performance and that the ability to determine the slope of an $y(x)$-graph and a correct qualitative understanding of distance-time graphs is insufficient for correctly determining the speed in a distance-time graph. Of note here is that they explicitly asked students to determine the slope or the speed *at a particular instant*. In their pretest data, they found that just under 20% of all students could determine the speed from an $x(t)$ graph, and that just over 50% could determine the slope in an $y(x)$-graph. Slope/height confusion and iconic interpretations were only present in a small minority of their data, whereas interval/point confusions occurred with 53% and 21% of their students in the kinematics and the context-free questions, respectively. When students were asked a similar question but with a graph in the context of water level versus time, their accuracy doubled compared to the distance versus time questions. Furthermore, almost twice the percentage of students used

interval reasoning compared to the percentage for the distance-time questions.

A follow-up study was performed by Bollen *et al.* [184]. They adapted the categorization and administered a similar test with first-year university students in calculus-based courses in the Basque country (Spain) and in Belgium. Their results showed similar difficulties, often related to dividing two coordinates to calculate the speed. Furthermore, they found that the context influences the success rate and that qualitative understanding of kinematics is important but not sufficient to determine the speed in $x(t)$ graphs.

The study by Planinic *et al.* [47] investigated students' understanding of linear graphs in mathematics, kinematics, and contexts other than physics. In their multiple choice test, students can choose between three comparative statements about the slope or the area under a graph, and must select the correct one. The results showed that mathematics is the easiest context for students and the other two are on par in difficulty. Additionally, determining the slope was found to be more difficult compared to determining the area under a graph — the other concept they studied. Concerning the y intercept, Davis [109] showed that students naturally and successfully start to use the informal terminology "starting point" when confronted with real-world contexts in multiple representations before the introduction of formal mathematics terminology. The study showed that a disconnect between formal and informal terminology in teaching activities results in a fragile understanding of the y-intercept, and strongly suggests to include explicit learning activities to connect these terminologies. After the introduction of formal mathematics terminology, students were confronted with abstract tasks (no real-world context) using the same representations and informal terminology which the researchers found to negatively impact students' performance. Adu-Gyamfi and Bossé [98] showed that even students who perform admirably at representation-related tasks concerning functions might still have limited mathematical understanding. This implies that testing students' understanding of functions should likely include, but certainly not be limited to, various representations. Also De Bock, Van Dooren, and Verschaffel [78] found that the representational format significantly influenced student performance as a main effect and as an interaction effect with the specific type of function (proportional, inverse proportional, affine with negative slope and affine with positive slope).

Ibrahim and Rebello [206] studied students' strategies when solving tasks concerning kinematics and work. They showed that, irrespective of the various representations offered in the questions' format (verbal, graphical and symbolic), students preferred manipulating equations and that they more often rely on equations to solve kinematics tasks whereas for tasks related to work they preferred a qualitative approach, thus indicating the dependency of the solution strategy on the context. Also Acevedo Nistal,

Van Dooren, and Verschaffel [96] came to the conclusion that their students in secondary education prefer to use formulas over tables or both to solve for the dependent or independent variable in linear function problems. They compared students' choices and performance in choice and no-choice conditions for the representation in which to solve the problem. Additionally, they found that the choice to use formulas, or the pressure students feel to use formulas, increases with grade. The representation in which a problem is presented and solved can strongly influence students' performance and the use of representations has been found to be very different between novices and experts [93]. Representational fluency in particular has been of interest to the research community, and notably the Representational Fluency Survey (RFS) [179] has been developed and applied [208] to assess this ability with university students in the context of physics. In light of this, we previously developed a multiple choice test [175] for students' representational fluency — specifically for the ability to translate between representations — of linear function problems in mathematics and kinematics. Our results showed that significant main effects of the representational transition, the function type (signs of slope and y intercept) and the context (kinematics and mathematics) exist. The study we present here is complementary to our previous one: now we focus more on qualitative insights and on the use of strategies and frequent errors.

All these findings point to a disconnect between contexts, a strong influence of and a disconnect between representations, and specific difficulties depending on the task and the concepts related to linear functions. This strongly suggests that students' understanding is compartmentalized to the specific situation in which they first learned about a specific topic. This is not unusual in the early stages of learning, but the referenced studies in higher education show that this compartmentalization continues to exist. To amend the situation, links should be actively constructed to connect these islands of understanding and achieve deeper understanding with the flexibility to apply knowledge and skills to new situations and achieve an efficient transfer thereof. To do so, it is important to understand students' difficulties that hamper the construction of these links. More than anything, the literature shows that these factors (contexts, representational format, function type, and task) should be included when assessing students' understanding.

5.3 Research design

5.3.1 Research questions

The broad goal of our study is to gain more insight in students' understanding of linear relations in kinematics ($x(t)$) and in mathematics ($y(x)$). Since effects from context, representational format, function type and task have been observed in literature, we include these factors in our research and ascertain their effects in a quantitative analysis followed by a qualitative analysis to gain as much insight as possible in students' strategies and errors. Our specific research questions are as follows:

1. How do the accuracy and strategy of grade 9 students compare between solving $x(t)$ problems in kinematics and solving isomorphic $y(x)$ problems in mathematics?
2. How do grade 9 students interpret negative velocity in $x(t)$ graphs and in algebraic formulas; and how does this compare to positive velocity and to negative and positive slope in the isomorphic $y(x)$ questions in mathematics?
3. How does the use of a graphs or algebraic formulas affect grade 9 students' accuracies and solution strategies when determining or comparing velocity or initial position in linear $x(t)$ problems in kinematics and slope or y intercept in linear $y(x)$ problems in mathematics?
4. How can we categorize grade 9 students' strategies when comparing and identifying velocity or initial position in linear $x(t)$ problems in kinematics and slope or y intercept in linear $y(x)$ problems in mathematics?

After investigating these questions, we use the test to assess the integrated STEM instructional approach and aim to answer the following additional research question:

5. How does grade 9 students' conceptual understanding and linking of linear functions compare between kinematics and mathematics and between the STEM@school instructional approach and the regular instructional approach?

To answer these questions we design and validate a test in the next sections. An English translation of the resulting test can be found in the Appendix A.5.

5.3.2 Design choices

A linear relation can be characterized by the combination of the signs of the y intercept and the slope. To limit the number of items, we select the two function types most familiar to our respondents: positive y intercept and positive slope, and positive y intercept and negative slope; each of these is represented by either a graph or an algebraic formula. The subject of each question is either of the two concepts: the y intercept or the slope. In addition, two kinds of tasks are used: determine the value of a concept or compare the value of a concept between two different situations. We combine the representations and the tasks into three different question types: determine via a graph, determine via a formula and compare via a graph. The other combinations are omitted because they are deemed less prone to errors and to limit the number of items. The contexts under investigation are kinematics and mathematics. In kinematics, students are presented with a description of one or two vehicles driving along a straight road with constant velocity, thus performing a 1D uniform linear motion. In mathematics, the questions concern a function $f(x)$ and no additional context is provided. To gain insight in student reasoning, we choose an open-ended question format including the explicit requirement for an explanation in some strategically selected questions. Finally, to obtain a consistently structured test and clearly defined questions, some additional design choices are made which are illustrated in the examples provided in Figure 5.2:

(i) In graphs, we focus on the first quadrant because we want to avoid issues which the use of other quadrants might trigger. We do show small sections of the other quadrants and we draw the curves in those quadrants depending on our students' familiarity with negative values for the respective variables. The result is that we include negative values for position, but exclude negative values for time. In mathematics we include negative values for all variables.

(ii) In comparison questions with graphs, each curve is drawn with a different thickness to better distinguish them. Furthermore, the two curves intersect in the first quadrant and each curve is labeled twice: once at the left and once at the right of the intersection point. These choices are motivated by the results from a pilot study which showed some confusing answers in which students referred to "the highest curve" or "the lowest curve" which makes interpretation of their explanations inconsistent.

(iii) In kinematics, units for the time t — always in seconds — and for position x — always in meters — are provided in the textual description of the situation and in the case of a graph they are also shown on the axis labels; coefficients in the formula representation do not include units.

(iv) Formulas in kinematics use the $x = x_0 + vt$ format in which the term with the lowest degree comes first, a format more often used in physics. In mathematics, formulas use the $f(x) = ax + b$ format in which the term with the highest degree comes first, which is more common in mathematics.

In addition, there is an important issue to highlight: a stationary object must accelerate before achieving a constant nonzero velocity. In the graphical kinematics items, we implicitly assume that the measurement starts at $t = 0$ and that the motion started before that moment, thus avoiding the acceleration issue.

5.3.3 Test structure

The design choices result in a 24-item test structured as shown in Figure 5.1. Figure 5.2 shows some examples. The test consists of two as isomorphic as possible context blocks: 12 kinematics items and 12 mathematics items. To avoid possible learning effects of solving a particular context block first, half of the respondents first complete the kinematics part and then the mathematics part, and the other half has the order reversed. The order of the items within each block is also randomized. Both linear function types under study have a positive y intercept but can have a positive or a negative slope. To distinguish between these we will simply refer to their slope sign. Each sign is questioned 6 times per context. Each item has one concept as its subject: the slope or the y intercept in mathematics and the velocity or initial position in kinematics. Furthermore, there are three different question types. In the first type, students must compare the slope or y intercept of two linear relations represented in a single graph. In the second, they must determine slope or y intercept given a graph. In the third question type, they must determine slope or y intercept given an algebraic formula. For the majority of the items the respondents are required to provide an explanation for their answer, using the additional instruction "*Explain your answer*". This was not done for all items to limit the test time. The items which require an explanation are marked in Figure 5.1 and are strategically chosen to gain as much understanding as possible, to avoid repetition of similar answers and to collect answers about the items for which we expected the most difficulties, e.g, the comparison of negative velocities in a graph. Since we expected fewer issues with the y intercept items, we requested an explanation more often with slope questions.

We designed this test to evaluate whether students can interpret concepts related to linear relations in different contexts. The test is not a concept test or an inventory [30] in the sense that it aims to measure students' conceptual

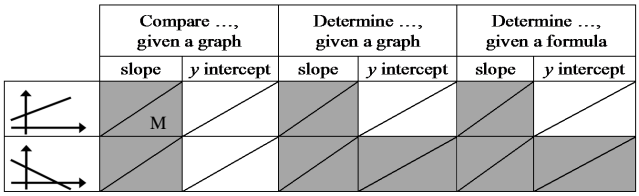


Figure 5.1: Test structure: two signs for the slope (positive and negative); two concepts (slope and y intercept); three question types using two tasks (compare or determine) and two representations (graph and formula); two contexts (kinematics (K); mathematics (M)). The items marked in gray require an explanation.

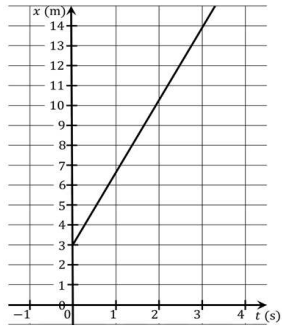
understanding in a particular domain, but rather aims at comparing students' understanding of particular isomorphic concepts between different domains like physics or mathematics. Its structure is in line with many other existing tests. It is most similar to the the ones from Bollen *et al.* [184] and Wemyss and van Kampen [108] which study a similar topic in isomorphic contexts and make use of written explanations for analysis. Also comparable are the test from Planinic *et al.* [46, 47] and Ivanjek *et al.* [48], who studied line graphs in various contexts using multiple choice tests and tests with open questions requiring an explanation. In addition, Bassok and Holyoak [209] compared constant acceleration kinematics and the isomorphic algebra question in a textual representation using think-aloud protocols. Furthermore, similar test structures have also been used for different topics by, e.g., Pollock *et al.* [44] when they compared the understanding of integrals in physics and mathematics; as well as by Barniol and Zavala [210] who studied the effect of context on students' understanding of vector concepts in their TUV-12 test and related studies [211], though the latter made use multiple-choice questions.

5.4 Test administration

5.4.1 Participants

The test was administered in 2017 in 17 classes distributed over 7 secondary schools in Flanders (Belgium) with one physics teacher and one mathematics teacher per school for these classes. A total of 253 students with ages 14–15 participated in the study. These grade 9 students were enrolled in programs with a strong science component. Our cohort consists of 117 females and 136 males. All of them had received instruction on linear relations in mathematics

K6
A train is moving along a straight track. The graph shows the position x of the train as function of the time t. The position is in meter, the time in seconds.



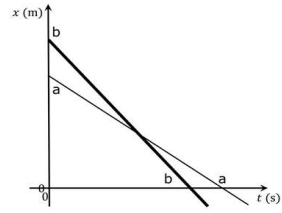
Determine the velocity of the train. Explain your answer.

K11
A car is moving along a straight road. The position x of the car as function of the time t is given by

$$x = 5 - 10t\,.$$

The position is in meter, the time in seconds.
Determine the initial position of the car. Explain your answer.

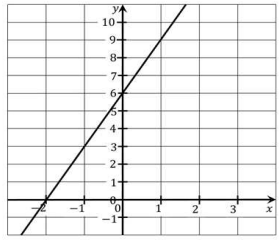
K4
Two cyclists a and b are riding along a straight bicycle path. The graph shows their position x as function of the time t. The position is in meter, the time in seconds.



Which cyclist has the largest velocity? Mark the correct answer and explain.

Cyclist a **Cyclist b** **Cyclists a and b have the same velocity**

M6
The figure shows the graph a function f.



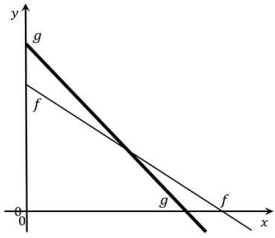
Determine the slope. Explain your answer.

M11
A function f is given by

$$f(x) = -8x + 2\,.$$

Determine the function value of 0. Explain your answer.

M4
The figure shows the graphs of two functions f and g.



Which of these functions has the largest slope? Mark the correct answer and explain.

f ***g*** ***f* and *g* have the same slope**

Figure 5.2: Three kinematics (K) items and the three isomorphic mathematics (M) examples illustrating the different factors in the test design.

and on basic 1D kinematics in physics in compliance with government imposed mandatory learning goals. Typically, instruction starts with examples from everyday life and simple experiments, followed by the introduction of the mathematical formula for average velocity, after which the graphical representation and interpretation in $x(t)$ graphs for uniform linear motions and other linear motions is discussed, and finally, the necessity to use vectors and the vector properties of velocity are made clear through examples. Additionally, experiments such as electric toy cars tracked by a motion sensor which outputs the measurement data in a graph are strongly encouraged. There is usually little to no focus on the algebraic formulation of a uniform linear motion in

the physics course at this point. Total time allocated for this is 6 teaching hours of 50 minutes each. In mathematics students learn about linear relations and linear functions in graphical representations, algebraic representations and tabular representations during 17 teaching hours of 50 minutes each. These lesson series start with the definition of the linear function, drawing linear graphs, translating between the various representations, discussing the meaning of the coefficients, studying the signs of the coefficients, of the dependent and independent variable, and solving standard exercises. Six schools participating in this study were part of a larger research project for which they had agreed to allow testing to take place within the classes with the required profile. One additional school was also contacted and voluntarily agreed to have classes with the requested profile to participate in the study. The locations of these schools are distributed across Flanders and all of them provide "aso education" in which students receive education on a broad number of topics to prepare them for higher education. All students in each class had to participate, but parents were given the possibility to have their child excluded from the test for, e.g., performance anxiety reasons, which only occurred a negligible number of times. Ethical approval for this study was granted by the Sociaal-Maatschappelijke Ethische Commissie (SMEC)–Social and the Societal Ethics Committee of KU Leuven.

5.4.2 Test procedure

The test is administered as a paper-and-pencil test to allow the respondents as much freedom as possible in the way they construct and explain their answers. The students are notified beforehand that a test will take place as part of a research project, that they do not need to prepare for it and that no credit can be earned. At the start, they are informed that the available time is 45 minutes, that it is crucial to explain their answer whenever this is required, and that they may use whatever method they prefer in their explanation, e.g., write a text, make a calculation, make a drawing, annotate the graph or formula, etc. Furthermore, if they are unable to answer a question they are asked to mark that question with a forward slash to indicate that they have read the question but do not know the answer.

5.4.3 Data analysis

In the quantitative analysis, coding of the data is dichotomous: unanswered and incorrect items are coded as incorrect (i.e., accuracy 0). An answer is considered correct if both the answer and the explanation (when required) are correct (i.e.,

accuracy 1). When the answer is correct but the explanation is not, the item is marked as incorrect. When the answer is correct but the required explanation is missing, the item is marked correct, giving the respondent the benefit of the doubt. For questions in which no explanation is required the answer is coded as 1 when correct and as 0 when incorrect. Item K4 (included in Figure 5.2) requires special attention. This item asks to compare two negative velocities in a graph, which can create a polarizing discussion in Dutch since there is no linguistic distinction between "velocity" and "speed", both are translated as "snelheid". This issue has also been discussed in our previous work [175]. In this study we choose to interpret snelheid as velocity in all questions, meaning that we expect students to take both the magnitude and the direction into account. Essentially this means that, since these are all 1D situations, we inquire about the scalar component v_x of velocity $\bar{v}$ whenever we use snelheid in this test. This choice means that we consider "Cyclist a" to be the correct answer for question K4, which is in line with the isomorphic question M4 (also included in Figure 5.2), in which answer "f"' is correct. Also in questions with an algebraic formula, the minus sign must be included in the answer. For example, given the algebraic expression $x = 3 - 12t$, i.e. question K12, the correct answer when asked to determine the velocity is -12 m/s.

Analysis is done using generalized estimating equations (GEE) [189, 190], which is an adaptation of logistic regression analysis and takes the repeated measurements of our dichotomous scoring into account [191]. The IBM ® SPSS ® Statistics version 24.0.0.0 software package is used. Settings in the GEE functionality are: binomial probability distribution with the logit link function, the Fisher parameter estimation method, the type III error analysis type, Wald statistics, and the Bonferroni correction to adjust for multiple comparisons. Our statistical model includes *context* (1D kinematics or mathematics), *QType* (compare given a graph, determine given a graph, determine given a formula), *concept* (y intercept, slope) and *slopesign* (+, −), *gender* (female, male). We include *gender* since this has been shown to have a possible (interaction) effect with the external representation [80], which is related to the question type in this case. All main effects, two-way effects and three-way effects are included in the model.

For the qualitative analysis, we use open coding to identify common strategies and recurring ideas across different students. This results in a first categorization scheme describing student explanations. To achieve a sound categorization, the scheme is — based on a subset of the data — refined by a second researcher and then independently used to analyze a new subset of the data by the first author and a colleague. The final categorization scheme is discussed in Section 5.6.

5.5 Validation and students' accuracy

In this section the validation of the test is discussed, followed by a quantitative analysis of the answers.

5.5.1 Validation

Here, we report on key validation parameters, which show that the test has good internal consistency, a wide range of item difficulty indexes and item discrimination indexes, and a good discriminatory power which makes the test well suited for our goals. An overview of these results on a per-item basis is provided in Table 5.1.

Table 5.1: Validation results sorted by *context* (indicated by "K" for kinematics and "M" for mathematics in the question number Q); by *Repr.* (*Representation*: Graph or Formula); by *task* (Compare Comp. or Determine Det.), by *slopesign* ("+" for positive, "−" for negative) and by *concept* (y intercept or slope). The statistical parameters are: item difficulty index P and its standard deviation SD; Cronbach's α when the item in question would be deleted; and item discrimination index D. Items marked with * require the respondent to provide an explanation for their answer.

Q	*Repr.*	*Task*	*Slope sign*	*Concept*	P	SD	α when deleted	D
K1	Graph	Comp.	+	y intercept	0.88	0.33	0.783	0.34
K2*	Graph	Comp.	+	Slope	0.92	0.28	0.790	0.18
K3	Graph	Comp.	−	y intercept	0.83	0.38	0.788	0.26
K4*	Graph	Comp.	−	Slope	0.13	0.34	0.795	0.12
K5	Graph	Det.	+	y intercept	0.74	0.44	0.783	0.46
K6*	Graph	Det.	+	Slope	0.48	0.50	0.794	0.32
K7*	Graph	Det.	−	y intercept	0.66	0.47	0.782	0.53
K8*	Graph	Det.	−	Slope	0.14	0.35	0.791	0.15
K9	Formula	Det.	+	y intercept	0.37	0.48	0.779	0.53
K10*	Formula	Det.	+	Slope	0.11	0.32	0.784	0.26
K11*	Formula	Det.	−	y intercept	0.35	0.48	0.781	0.51
K12*	Formula	Det.	−	Slope	0.05	0.21	0.791	0.07
M1	Graph	Comp.	+	y intercept	0.53	0.50	0.779	0.66
M2*	Graph	Comp.	+	Slope	0.85	0.36	0.786	0.28
M3	Graph	Comp.	−	y intercept	0.54	0.50	0.786	0.49
M4*	Graph	Comp.	−	Slope	0.41	0.49	0.791	0.44
M5	Graph	Det.	+	y intercept	0.39	0.49	0.776	0.66
M6*	Graph	Det.	+	Slope	0.70	0.46	0.780	0.54
M7*	Graph	Det.	−	y intercept	0.37	0.48	0.770	0.82
M8*	Graph	Det.	−	Slope	0.67	0.47	0.787	0.43
M9	Formula	Det.	+	y intercept	0.36	0.48	0.777	0.65
M10*	Formula	Det.	+	Slope	0.74	0.44	0.783	0.51
M11*	Formula	Det.	−	y intercept	0.36	0.48	0.772	0.74
M12*	Formula	Det.	−	Slope	0.71	0.45	0.780	0.62

5.5.1.1 Internal consistency

The internal consistency is assessed using Cronbach's α, which is $\alpha = 0.791$ and is an acceptable value just shy of 0.80, the criterion for a "fairly high" internal consistency [186]. Since omitting certain items from the test might increase this value, α was also calculated for each case in which an item i is omitted from the test. These values are shown in Table 5.1 and show that omitting a test item would only marginally increase the internal consistency by a maximum of 0.04 in the case of item K4.

5.5.1.2 Item difficulty index

Because of our dichotomous coding for the accuracy, the average accuracy of each item translates directly into the item difficulty index P:

$$P = \frac{N_c}{N},$$

in which N_c is the number of answers with accuracy 1, and N is the total number of answers, which is the number of participants: $N = 253$. A higher value for P is indicative of an easier item and a lower value indicates a more difficult question. Table 5.1 shows an item difficulty range between 0.05 and 0.92, and an average item difficulty of 0.51 ± 0.26.

The items with the highest average scores are K1, K2, K3, and M2, which are all "compare given a graph" questions, three out of four have a positive slope, and three out of four are kinematics questions. This illustrates the contrast between the contexts since the most difficult questions, as well as the majority of the easiest questions are found in kinematics, while the mathematics items do not have such extreme item difficulty indexes and are less broadly distributed. Furthermore, these observations highlight the important effect of the slope sign on the average accuracy and the difficulty students have with negative slopes in kinematics.

Items K4, K8, K10, and K12 stand out because they have very low item difficulty indexes accompanied with some of the lowest standard deviations, meaning that these items are very difficult for the large majority of students. The common factor for these four questions is that all are kinematics items in which the concept is the velocity. The sign of the velocity is negative for K4, K8, and K12; and the task in items K8, K10, and K12 is to determine the velocity. The counterparts of these questions in mathematics (M4, M8, M10, M12) show no such pattern and the counterparts which query about the y intercept (K3, K7, K9, K11) do not show any similarities either. Furthermore

it is clear from the lower item difficulty indexes of K9 and K11 that determining the initial position given a formula is far more difficult than when presented with a graph.

Of note here is the extreme contrast between item K2 and K4 which are analogous questions except for the sign of the velocity.

Another remarkable difference is clear when comparing the K5-K8 block with the isomorphic M5-M8 block, i.e., the items in which to determine the y intercept or slope given a graph. For these blocks, the item difficulty indexes for the kinematics items have the opposite structure of those in mathematics: it is easier to determine the initial position than the velocity in kinematics graphs compared to mathematics where it is easier to determine the slope than the y intercept. Part of this might be explained by the lower familiarity students have with the term "function value" in the wording "the function value of 0" in the mathematics items — compared to the more familiar "initial position" in kinematics. Although the wording is part of the official curriculum, many students made note of this lack of familiarity during the test or in their explanations. This is also clear in the items in which the y intercept has to be compared (K1, K3, M1, M3). For these items we did not require an explanation because we expected them to be easy and result in fewer interesting explanations. This proves to be correct in kinematics with accuracies for K1 and K3 above 0.80, but was not the case in mathematics (M1 and M3) which have an accuracy of 0.53 and 0.54 respectively. This difference between the isomorphic items indicates that students did not recognize the similarities between the contexts.

5.5.1.3 Discriminatory power

The discriminatory power of a test is a measure of its ability to discriminate between respondents with high and low ability. On a per-item basis, we can calculate the item discrimination index $D = PU - PL$, in which PU is the item difficulty index of the upper group of test scores, i.e., the 27% of highest scoring students; similarly, PL is the item difficulty index of the lower group, i.e., the group with the 27% lowest scoring students [186]. The higher the value of D for an item, the better that item discriminates. The results are shown in Table 5.1 and range between 0.07 and 0.82 with an average of 0.44 ± 0.20. Again the same four kinematics items (K4, K8, K10 and K12) stand out for their low score; which — combined with their lower item difficulty index — shows that the large majority of students has severe difficulties with these questions and that the score distribution for these four items is very narrow. Additionally, item K2 — a kinematics item to compare positive velocities in a graph — has

a very low item discrimination index in combination with a very high item difficulty index, indicating that this item is very easy for the large majority of students and that its score distribution is very narrow.

To assess the discriminatory power of the full test we use Ferguson's delta δ, which is the ratio of the interperson differences to the maximum number possible:

$$\delta = \frac{(k+1)(n^2 - \Sigma_{i=1}^{k} f_i^2)}{kn^2},$$

with $k = 24$ the number of items in the test, $n = 235$ the sample size and f the frequency of a certain score i, with i from 1 to k. This results in $\delta = 0.971$, which easily satisfies the minimum required value of 0.9, meaning that the test has good discriminatory power [186].

5.5.2 Main effects and interaction effects

In this section we discuss the results from the GEE analysis, the errors in this section are always standard errors. The GEE analysis shows four significant main effects. The first is *context*, for which we find Wald $\chi^2(1) = 30.756$ and $p < 0.001$ and a significant difference ($p < 0.001$) between a mean of 0.45 ± 0.02 for kinematics and 0.57 ± 0.02 for mathematics, which is in line with the results from our previous study [175]. The second significant main effect is that from *QType* with Wald $\chi^2(2) = 213.710$ and $p < 0.001$. The means are: 0.68 ± 0.02 for "compare given a graph", 0.52 ± 0.02 for "determine given a graph", 0.33 ± 0.02 for "determine given a formula", and each pairwise comparison is significant ($p < 0.001$). Thirdly, for *concept* the results are Wald $\chi^2(1) = 16.854$ and $p < 0.001$ with a significant ($p < 0.001$) difference between 0.55 ± 0.02 for y intercept and 0.46 ± 0.02 for slope. For *slopesign* this results in Wald $\chi^2(1) = 239.782$ and $p < 0.001$. Pairwise comparison of the slope sign shows a significant difference ($p < 0.001$) with a mean of 0.62 ± 0.02 for positive slope and 0.40 ± 0.02 for negative slope. Lastly, the variable *gender* — just like in our previous study [175] — does not result in a significant main effect: Wald $\chi^2(1) = 1.796$ with $p = 0.180$.

There are also many significant interaction effects, which are shown in Table 5.2.

First, the term "*context* * *concept*" for which the results are provided in Figure 5.3. This shows that, concerning the y intercept, students perform significantly ($p < 0.001$) better in kinematics than in mathematics, while the opposite is true for questions concerning the slope.

Table 5.2: Significant two-way and three-way interaction effects with Wald χ^2, degrees of freedom (df) and significance (p value) up to three digits.

Term	χ^2	df	p
context * *concept*	370.275	1	0.000
context * *slopesign*	83.070	1	0.000
context * *QType*	163.615	2	0.000
QType * *slopesign*	107.922	2	0.000
concept * *slopesign*	162.529	1	0.000
concept * *gender*	5.428	1	0.020
context * *QType* * *concept*	21.635	2	0.000
context * *concept* * *slopesign*	40.046	1	0.000
context * *slopesign* * *gender*	10.988	1	0.001
QType * *concept* * *slopesign*	86.085	2	0.000

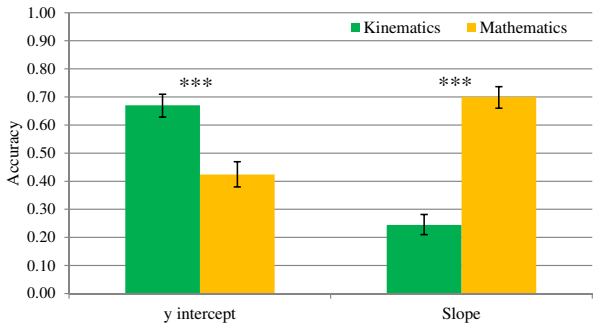


Figure 5.3: GEE results for the term *context* * *concept*. The p value of the comparisons is indicated by: * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$. Highly significant differences are found between the contexts for each of the concepts.

The results for the term "*context* * *slopesign*" are shown in Figure 5.4, which show that students' performance for questions with positive slope is very similar in kinematics and mathematics. For negative slope though, there is a significant ($p < 0.001$) difference between the two contexts which indicates that students have far more difficulties with negative velocity than they do with negative slope. The majority of this difference originates from questions in which the velocity or slope is the concept and the $\sim$ 50% drop between positive and negative scores in kinematics happens equally across all three questions types.

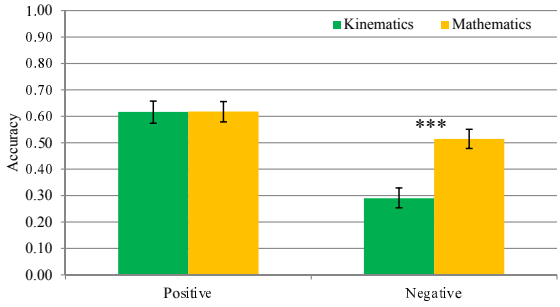


Figure 5.4: GEE results for the term *context* * *slopesign*. The p value of the comparisons is indicated by: * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$. A highly significant difference is found between the contexts for negative slopes.

Figure 5.5 shows the results for the term "*context* * *QType*", which illustrates that there are significant differences ($p < 0.001$) between the two contexts for "Determine via Formula" and "Compare via Graph", in which the former has the largest difference between the contexts. Furthermore, there are no significant differences among question types in mathematics, but very significant differences ($p < 0.001$) between all of them in kinematics.

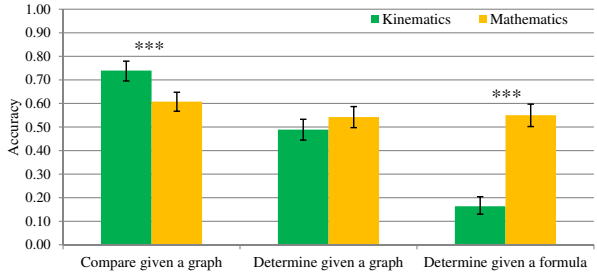


Figure 5.5: GEE results for the term *context* * *QType*. The p value of the comparisons is indicated by: * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$. Highly significant differences are found between the contexts for "Compare given a graph" and for "Determine given a formula" question types.

Since the slope causes many difficulties in kinematics, it is interesting to zoom in on the results for the slope in the term "*context* * *QType* * *concept*", which are shown in Figure 5.6. This graph shows significant ($p < 0.001$) differences

between contexts for the two question types including the "Determine" task. Furthermore, it shows no significant differences between question types in mathematics, but very significant ($p < 0.001$) ones in kinematics. The highest mean accuracies in kinematics are achieved in "Compare via Graph" because the very high accuracies for positive slope questions in this question type compensate the very low accuracies for negative slope questions in this question type, thus averaging out slightly above 50%. For the other two question types though, the mean accuracies in kinematics are very low for both the negative and positive slopes resulting in significantly lower mean accuracies overall.

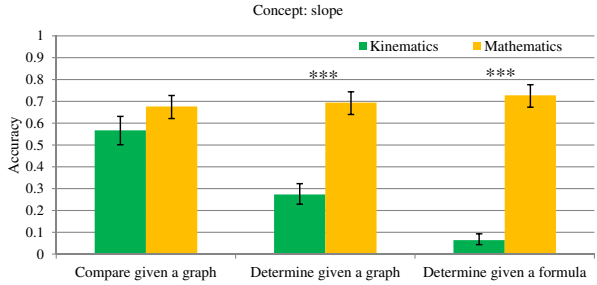


Figure 5.6: GEE results for the Slope in the term *context* * *QType* * *concept*. The p value of the comparisons is indicated by: * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$. Highly significant differences are found between the contexts for "Determine given a graph" and for "Determine given a formula" question types.

Figure 5.7 shows the significant difference ($p < 0.001$) between physics and mathematics and positive and negative slope for questions concerning the slope. Most importantly, it shows that negative slope is far more difficult in kinematics compared to mathematics.

These results support those from the validation, which is that students have the most difficulties with items in kinematics, concerning slope, with negative velocity in all three question types, i.e., items K4, K8 and K12. In addition item K10, in which students must determine the positive slope when given a formula, also poses increased difficulties for students. In mathematics the main difficulty is determining the y intercept, in both graphs and algebraic formulas. This is possibly explained by the relative unfamiliarity with the terminology "function value" in the wording "the function value of 0" to indicate the y intercept in the mathematics items. Furthermore, there is a larger focus on the x intercept than there is on the y intercept in mathematics lessons. In kinematics the initial position of a motion is of greater importance since in this context the choice of the coordinate system and its origin is often intelligently

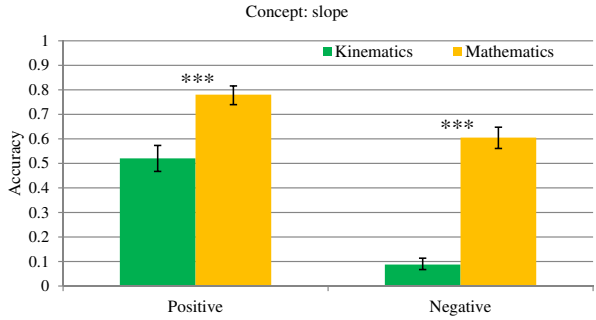


Figure 5.7: GEE results for the slope in the term *context* * *concept* * *slopesign*. The p value of the comparisons is indicated by: * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$. Highly significant differences are found for the slope between the contexts for both slope signs.

chosen so that the initial position is in the origin. Furthermore, terminology such as "initial position" is more intuitively clear for students compared to "the function value of 0" as was also the case in Davis' study [109].

5.6 Qualitative results

For the qualitative analysis, a categorization scheme was constructed bottom-up from the raw data. First, a pilot study with 181 students and a first version of the test was performed and a first version of the scheme was constructed. A second researcher made refinements to the test to make the graphs and the wording clearer and to create a better distinction between the two contexts, but the structure of the test remained. The scheme was optimized by grouping some categories and creating a better distinction between others. That optimized scheme was then used to categorize a subset ($n = 42$) from the new data from the refined test by a third researcher which resulted in some additional refinements. The final scheme was then used to categorize all the new data ($n = 253$) by the second researcher. In this section we describe the scheme, validate its reliability, and discuss trends in the data.

5.6.1 Categorization scheme

The final categorization scheme is shown in Table 5.3. It is split by concept (y intercept or initial position and slope or velocity). The first seven categories

are shared between both concepts followed by two and four concept specific categories. Answers usually fall in only a single category, but a non-negligible number of questions had answers which made use of multiple strategies in which case they were put into all suitable categories.

Table 5.3: Bottom-up constructed categorization scheme.

y intercept / initial position	slope / velocity
I1: Location in an equation	S1: Location in an equation
I2: Identification of the x intercept	S2: Identification of the x intercept
I3: Construct an equation	S3: Construct an equation
I4: Construct a graph	S4: Construct a graph
I5: Construct a table	S5: Construct a table
I6: Intersection with vertical axis	S6: Reasoning with "steepness"
I7: Calculate a specific function value	S7: Drawing a triangle on a graph
	S8: Ratio of differences
	S9: Ratio of coordinates
I99: Other	S99: Other

I1/S1: Location in an equation A coefficient in the equation is marked by underlining it or encircling it. Alternatively, the student explicitly writes that, e.g., "6" is the answer because it is located before the "x".

I2/S2: Identification of the x intercept Instead of identifying the y intercept or slope they identify the x intercept (graphical or symbolic).

I3/S3: Construct an equation Change of representation. When presented with a graph, an equation was constructed.

I4/S4: Construct a graph Change of representation. When presented with an equation, a graph was constructed. This need not be a detailed one, a simple sketch suffices.

I5/S5: Construct a table Change of representation. A table is constructed with two or more data points.

I6: Intersection with vertical axis When a graph is given or constructed, the y intercept is marked by, e.g., encircling it or drawing an arrow towards it. Often this is accompanied by written explanations such as "This is where the line/graph/cyclist starts/begins."

I7: Calculate a specific function value When an equation is given or constructed, the value of the dependent variable is calculated for a specific value of the independent variable, often with $x = 0$ or $x = 1$.

S6: Reasoning with "steepness" Qualitative reasoning concerning the steepness of the graph. Often wording similar to "Graph a is steeper than graph b" is used.

S7: Drawing a triangle on a graph Drawing a triangle with a vertical and a horizontal side under or above the line of a graph to indicate the step change in both directions. Often these triangles are accompanied by the change in value along each axis with notations such as "+1" and "-2". The size of the triangle does not matter.

S8: Ratio of differences Calculation of the slope using $\frac{\Delta y}{\Delta x}$ or $\frac{y_2-y_1}{x_2-x_1}$, or written explanations such as "In $1s$ the cyclist traveled $2m$;" "If x increases with 1, then y increases with 3;" "If you go one unit to the right, then you go 3 units up." A specific case in this category is when the ratio of the y-value of the y intercept over the x-value of the x intercept is calculated, which is essentially the ratio of the differences from the intercepts with respect to the origin. More qualitative explanations also fit in this category, such as: "In the same time, cyclist b traveled more meters;" "Cyclist b starts later, but still overtakes cyclist a".

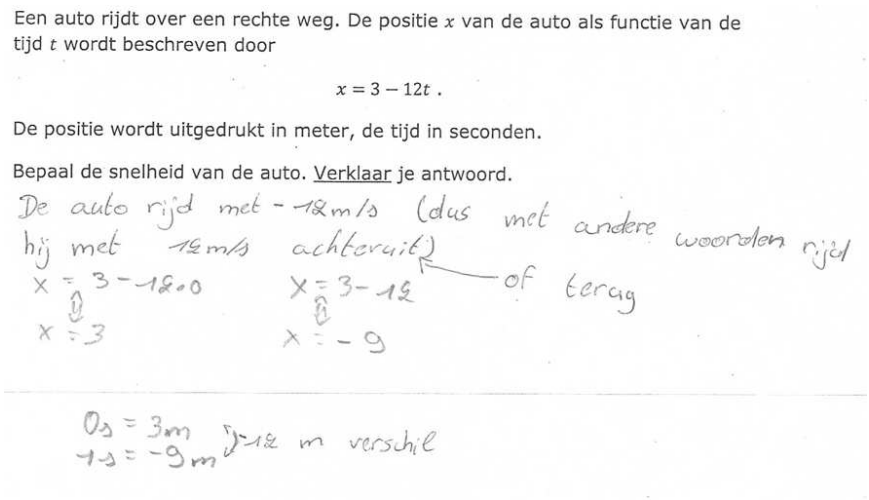
S9: Ratio of coordinates The ratio $\frac{y}{x}$ of some coordinate (x, y) is calculated. This includes cases in which the correct formula is written (e.g. $v = \frac{\Delta x}{\Delta t}$, but actually $v = \frac{x}{t}$ is calculated, signaling the disregard or the misunderstanding of Δ in the formula. Often $\frac{f(1)}{1}$ or $f(x = 1)$ is calculated. Also written explanations such as "The y-value of function f is larger than that of function g for the same x-value" with comparison questions are in this category.

I99/S99: Other All answers which are less frequent (a few percent) *and* do not fit in any of the other categories. E.g., confusion of y intercept and slope; calculations with the coefficients of a formula such as $a * b$ or $a + b$ in $f(x) = ax + b$; identifying the slope as $6x$ in $f(x) = 6x + 2$; in the case of a y intercept question, providing coordinates such as $(0, 2)$ given the equation $f(x) = 6x + 2$, which is always considered incorrect since the questions ask for the function value or the initial position; or in case of a y intercept question, providing a coordinate constructed of the x intercept and the y intercept such as $(-\frac{1}{3}, 2)$ for $f(x) = 6x + 2$.

Categories I3/S3, I4/S4, and I5/S5 all describe a change in representation, these categories are always accompanied by another strategy to fully answer

the question. We deliberately keep these categories separate so we can compare their prevalences. Categories S7 and S8 are both essentially a ratio of differences, but our data strongly suggest that it is useful to distinguish the two.

We illustrate the categorization using Figure 5.8. In this figure (item K12), the students are asked to determine the velocity given a formula. Since the answer ($-12m/s$) and the explanation are both correct, the student receives an accuracy of 1. The explanation is categorized as S8 because the student calculates the value of x for two different values of t, i.e. $t = 0$ and $t = 1$, then calculates the differences of position and time taking into account the sign. Thereafter he/she calculates the correct ratio to find the velocity. This example also illustrates a doubt about the interpretation of the minus sign, i.e. whether the car is reversing or turning back. A second example is provided in Figure 5.9 (item M4). The student has drawn two triangles with a stepsize of $+1$ in the x direction and measured the corresponding step in the y direction including while accounting for the negative sign. Since $-0.5 > -1.2$ the correct answer is f.


Een auto rijdt over een rechte weg. De positie x van de auto als functie van de tijd t wordt beschreven door

$$x = 3 - 12t\ .$$

De positie wordt uitgedrukt in meter, de tijd in seconden.

Bepaal de snelheid van de auto. Verklaar je antwoord.

De auto rijd met -12m/s (dus met andere woorden rijd hij met 12m/s achteruit)

$x = 3 - 12 \cdot 0$ $\quad$ $x = 3 - 12$ $\quad$ of terug

$x = 3$ $\quad$ $x = -9$

$0_s = 3m$

$1_s = -9m$ } 12 m verschil

Figure 5.8: Example of a categorization for an answer to item K12. The matching category is S8 and the answer is correct. The text in Dutch reads: "The car is going $-12m/s$ (so, in other words is driving $12m/s$ backwards or back" and "$-12m$ difference".

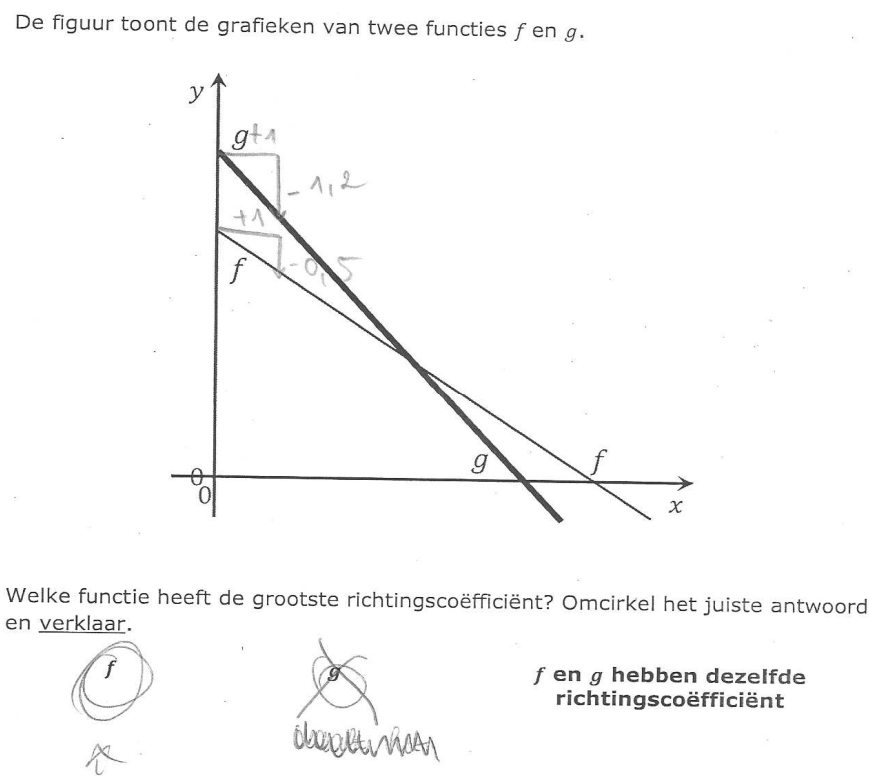


Figure 5.9: Example of a categorization for an answer to item M4 which asks to determine the largest slope. The matching category is S7 and the answer is correct.

5.6.2 Interrater reliability

To validate the classification scheme, the inter-rater reliability is calculated using Cohen's kappa κ. This quantity provides a measure for the agreement between two raters of the same data by calculating the ratio of the observed agreement excluding the probability of agreement by chance over the potential agreement excluding the probability of agreement by change. The coefficient can be calculated as follows:

$$\kappa = \frac{p_0 - p_e}{1 - p_e},$$

in which p_0 is the observed agreement between the raters, i.e., the percentage of agreement; and p_e is the probability of agreement by chance. Due to the fact

that we consider strategies, students can employ multiple strategies in the same question meaning that an answer is not always uniquely categorized. Therefore, we cannot directly apply the expression to calculate p_e:

$$p_e = \frac{1}{N^2} \sum_k n_{k1} n_{k2},$$

in which N is the number of items, k is the category, and n_{k1} and n_{k2} are the number of times category k was identified by rater 1 and 2 respectively. First, we did calculate an estimate for p_e. In doing so it is important to note that the number of categorizations per rater is different because e.g. the first rater might identify two categories in a particular answer whereas the second rater might only categorize one. Therefore, we used the smallest number (either n_1 or n_2) to obtain the most conservative estimate, which resulted in a value of 0.111 for p_e and a corresponding interrater reliability of $\kappa = 0.779$. But due to the possibility to assign multiple strategies to an answer, this is not really correct. We therefore opted to use $p_e = 0.5$ to calculate an even more conservative estimate which resulted in $\kappa = 0.607$. We also calculated the agreement per item and using those values we additionally calculated κ on a per item basis, in which we considered p_e equal to the inverse of the number of possible categories, which is 1/10 for slope items and 1/8 for y intercept items. These results showed values often much higher than the κ for the full test and indicate good agreement [212, 213] between the categorization of the second and third researcher for a subset of the data ($n = 42$).

5.6.3 Results

In this section we describe students' use of the strategies in the categorization scheme. Since not all questions in the test require an explanation and since not every answer contained an explanation (required or not), we express frequencies relative to the number of answered questions for which at least one strategy was discerned by the researcher. This means that unanswered questions are not included. With this definition, the maximum frequency of 100% for a certain category means that every answered question was provided with an explanation and that this category was used in all those explanations. To clarify this a bit more, let us assume that the data contains 1000 questions which were answered *and* for which an explanation was provided which is categorized in *at least* one category. Since some explanations contain multiple categories, the total number of items in all categories is higher than 1000, but 1000 is the maximum prevalence for each single category, i.e., this prevalence would result in a frequency of 100%. If 200 out of 1000 cases with *at least* one category are categorized in category X, then the frequency for X is 20%. A discussion

about the highlights will be provided and the main discussion will be done in Section 5.7.

5.6.3.1 Y intercept

The results for the y intercept questions are shown in Table 5.4. Answers without any explanation are omitted from this analysis. When an explanation is provided but not requested, the answer is included in the data. This results in far more data than anticipated than from the four questions which explicitly require an explanation for the y intercept (shown in Figure 5.1). The resulting data set contains 1600 uses of *at least* one strategy across the *twelve* questions about the y intercept, 723 in kinematics and 877 in mathematics. The gathered data structure is largely in line with the full test structure in Figure 5.1. The data have a determine/compare ratio close to 2/1, a graph/formula ratio close to 2/1, and a positive/negative slope ratio close to 1/1. Because the inclusion of all items accompanied by an explanation, about 13% of the kinematics questions in the data set are comparison items. This means that the frequencies provided are predominantly from questions with a "determine" task, that half of them use a graph and the other half a formula, and that they are almost exclusively from questions with a negative slope.

Table 5.4: Frequencies (in %) for the y intercept questions. f is across both contexts, f_K is in the kinematics context and f_M is in the mathematics context. The percentages are relative to 1600, 723 and 877 respectively.

Category	f	f_K	f_M
I1: Location in an equation	4	5	4
I2: Identification of the x intercept	27	10	42
I3: Construct an equation	13	4	20
I4: Construct a graph	3	5	1
I5: Construct a table	1	1	2
I6: Intersection with vertical axis	24	40	12
I7: Calculate a specific function value	24	18	29
I99: Other	17	25	11

The most common strategy — and simultaneously the most common error — in general, is identification of the x intercept (I2), with a frequency of 27%. A very large difference between the contexts is found: only 10% in kinematics and 42% in mathematics. The change of representation also shows interesting results. Constructing an equation (I3) when a graph is provided also shows

a stark difference between kinematics (4%) and mathematics (20%) with an overall frequency of 13%. Furthermore, when construction of a graph (I4) is used, it mainly happens in kinematics. Tables are rarely used to determine the y intercept. The most used strategy in kinematics is the determination of the intersection with the vertical axis (I6) of a graph with a frequency of 40%, which is in stark contrast with that for mathematics, which is only 12%.

The calculation of a specific function value (I7) is also a frequently used method, especially in mathematics with 29%, compared to 18% in kinematics. A total of 83% of the answers in I7 are correct, meaning students indeed calculate $f(0)$ in the majority of these cases.

5.6.3.2 Slope

The results for the slope questions are shown in Table 5.5. Answers without any explanation are omitted from this analysis. The data set contains 2501 uses of *at least* one strategy across the twelve questions about slope, 1202 in kinematics and 1299 in mathematics. As shown in Figure 5.1, all twelve slope questions require an explanation. The data structure is in line with the structure of Figure 5.1. This means that the data have a determine/compare ratio close to 2/1, a graph/formula ratio close to 2/1, and a positive/negative slope ratio close to 1/1.

Table 5.5: Frequencies (in %) for the slope questions. f is across both contexts, f_K is in the kinematics context and f_M is in the mathematics context. The percentages are relative to 2501, 1202, and 1299, respectively.

Category	f	f_K	f_M
S1: Location in an equation	16	2	29
S2: Identification of the x intercept	2	2	3
S3: Construct an equation	5	2	7
S4: Construct a graph	2	2	2
S5: Construct a table	2	2	1
S6: Reasoning with “steepness”	19	15	23
S7: Drawing a triangle on a graph	13	6	20
S8: Ratio of differences	31	42	22
S9: Ratio of coordinates	10	20	1
S99: Other	13	20	6

The most used strategy for slope is calculating the ratio of differences (S8) at 31% overall and with a substantial difference between the contexts with 42% in kinematics and 22% in mathematics. The second most used strategy is more qualitative: reasoning with steepness (S6) at 19%, which is more often used in mathematics (23%) than in kinematics (15%). Triangles are drawn on a graph (S7) much more often in mathematics (20%) compared to kinematics (6%) whereas calculating the ratio of coordinates (S9) is far more frequent in kinematics (20%) than in mathematics (1%). Another stark contrast between contexts is the determination of the slope through the location of the coefficient in an equation (S1) which achieves an overall frequency of 16% with 29% in mathematics and a mere 2% in kinematics. The representational transitions are not often used, only the construction of an equation (S3) in mathematics is notable with a frequency of 7% compared to 2% in kinematics.

5.7 Discussion

The results in Section 5.5 and Section 5.6 reveal significant main effects and interaction effects, and substantial differences in strategy between the two contexts. In this section, we discuss these results as well as selected examples from student explanations.

5.7.1 Y intercept

The most difficult kinematics questions for this concept are K9 and K11 which both require determining the initial position when a formula is given. Very similar results are found for the mathematics questions M9 and M11, which illustrates the difficulty students have with interpreting an abstract algebraic expression. We assume that experts are more likely to determine the y intercept by identifying the coefficient of the zeroth order term (I1), but to our knowledge there is not yet any study available to support this claim. Our data show that students make very little use of this method and instead attempt a calculation to do so, or determine the x intercept (I2) instead. Furthermore, the change in the order of the terms in an equation between kinematics and mathematics (design choice (iv)) does not have a noticeable effect, though particularly because strategy I1 is infrequently used.

Concerning representational transitions (I3, I4, I5), we find that many students construct an equation (I3) in mathematics but do this far less in kinematics. This increased use of algebraic expressions in mathematics is in line with the conclusions from Acevedo Nistal, Van Dooren, and Verschaffel [96] who

studied students' flexibility in choosing a representation to solve linear function problems. They found that students have a very strong preference or feel very strong pressure to use formulas, which increases with the grade they are in. The categorization scheme could be simplified here by grouping I4 and I5 into a single category due to the low frequencies, but category I3 should definitely be kept separate.

The most common strategy — and simultaneously the most common error — in general, was identification of the root, i.e., (I2) confusing the y intercept with the x intercept with far more errors in the mathematics context. As mentioned before, the large number of errors in mathematics can probably be attributed to a low familiarity with the term "function value of 0" which was used in the item formulation. The low familiarity was indicated by the students, though it is part of their official curriculum. This compares well with the GEE results in Figure 5.3 which show a significant difference in average accuracy between kinematics and physics for the y intercept. These results can be linked to the conclusions from Davis [109] which also indicated that students have more difficulties with using the official terminology whereas they were more successful with informal terminology such as "the starting point" in physics questions. The use of terminology such as "starting point" or "initial position" might also have a negative effect though. A small number of students wrote replies similar to "the initial position is always 0 m" in the kinematics questions. These explanations were categorized in the "Other" category (I99) and did not occur in the mathematics questions. In many standard kinematics exercises, students are taught to intelligently choose the reference system such that the initial position is in the origin, which simplifies the equations and the calculation. Quotes such as the one above show that some student believe that the point 0 m or the origin in a position - time graph is always the "initial position", almost as if that is its name. This implies that students at this early stage of learning kinematics have difficulties understanding the concept of a reference frame. They seem unaware that shifting a reference frame changes the graph and the equation, but still describes the same physics and the same reality. Such difficulties show that this might be an opportune time — or perhaps already too late — to introduce the first notions of relativity.

Category I99 contains another notable error, although also with low frequency: the use of the coordinate notation, e.g. (x, y) or (t, x), which is most often a combination of the x component of the x- and the y component of the y intercept. This occurs in both contexts and most often in questions with a graph with a negative slope. This use of coordinate notation relates to the "Cartesian connection" which is the realization that each point on a line represents an ordered pair that satisfies the equation representing the line [109]. Davis [109] states that the y intercept can be used to promote the Cartesian

connection since it can be directly identified in the equation — which we categorized as strategy I1 and assume to be the preferred method by experts. Our results show that a minority (less than 10%) use the Cartesian connection to answer a y intercept or initial position question, but do so incorrectly since we explicitly inquire about the function value or the position, which is a single number, not a coordinate.

In general, the most interesting results concerning the y intercept are found in mathematics since here the accuracies are far lower than in kinematics, students' use of strategies shows more variation and use of informal terminology has a bigger impact. Furthermore, no specific cases are of note for the y intercept questions, the categorization scheme — when not accounting for the representation transitions — can be used very well to classify the most common errors and strategies. Finally, the data suggests that it might prove useful for mathematics teachers to use examples from kinematics to explain the meaning of the y intercept.

5.7.2 Slope

In this section we first discuss the GEE results and the most common strategies and errors. Next, we zoom in on the interval/point confusions by discussing three different cases and compare between kinematics and mathematics. Afterwards, we have a similar discussion about how students treat the sign of the slope in kinematics and mathematics.

5.7.2.1 General discussion

Questions concerning slope are found to be particularly difficult in the kinematics context with a negative slope and the most difficult when presented with a formula. In "determine via formula" questions about the slope in kinematics, the average accuracies are extremely low compared to their isomorphic version in mathematics. Furthermore, negative slope questions result in rather low average accuracies in mathematics, but extremely low average accuracies in kinematics. A large factor of influence in these results is the inclusion of the possible minus sign for the one-dimensional velocity as a criterion for a correct answer. The results show that the large majority of students does not include the minus sign in kinematics, i.e., they consider the magnitude and omit the direction. This pattern is not present in the mathematics questions, so when isomorphic equations and graphs are used, students are far more likely to include the minus sign.

Part of the large overall difference between the two contexts can be explained by the more expert-like use of strategies in mathematics by determining the slope from an equation through locating the coefficient in the expression (S1). This strategy achieved a 29% frequency in mathematics compared to a mere 2% in kinematics. This shows that students do not see the link between the slope of linear functions and of velocity in $x(t)$ expressions, which indicates a weak link between contexts and a compartmentalization of understanding within contexts.

By far the most preferred method in kinematics is calculating the ratio of differences (S8) (mainly in questions with a graph). The method is not always applied correctly though. After calculating the ratio of differences, students often omit the minus sign (when present) from the result. From students' explanations it is clear that this is triggered by the use of the formula $v = \frac{\Delta x}{\Delta t}$ in kinematics and $a = \frac{y_2 - y_1}{x_2 - x_1}$ in mathematics. Although this is essentially the same formula, there are a few differences in students' use. In mathematics, the formula is usually used correctly. In physics students usually write down the correct formula but often calculate $v = \left|\frac{\Delta x}{\Delta t}\right|$. The second most common error in kinematics — also often after writing down the correct formula — is misinterpretation of the Δ and calculation of the ratio $\frac{x}{t}$ of coordinates (x, t) for some point on the graph (S9). This S9 strategy has a frequency of 20% in kinematics but only 1% in mathematics. In mathematics, a lot more students resort to drawing a triangle on a graph (S7), which in almost all cases has a set horizontal length of 1 and a vertical length proportional with the slope — usually also including the correct sign in mathematics. This difference between the two contexts for S7 confirms the useful distinction between calculating the ratio of differences (S8) and drawing a triangle on a graph (S7) in the categorization scheme, despite that both essentially calculate a ratio of deltas. The answers in category S8 are correct in 62% of these cases with errors mainly due to omissions of the minus sign or miscalculations. In contrast, the answers in S9 — a frequent strategy in kinematics questions — are always incorrect and contain various interesting errors which will be discussed in Section 5.7.2.3. This distribution of students' strategies between S7, S8 and S9 — which are all predominantly used in questions with graphs — and the frequent use of the expert-like strategy S1 in mathematics are the most striking results for the slope in this data.

Concerning the representation transitions (S3, S4, S5), the results show very low frequencies. Just like with the y intercept, the construction of an equation (S3) is the dominant one and it is more frequently used in mathematics than in kinematics. A straightforward simplification of the categorization scheme is to take S3, S4 and S5 together in the same category called "Change in representation", or at least group S4 and S5 into a single category called

"Change in representation other than equation".

5.7.2.2 Comparison to similar studies

We can compare our results with the categorization into slope/height, interval/point, and iconic representations from Leinhardt *et al.* [107]. First, we find a negligible number of slope/height confusions in both contexts, hence this is not a separate category in our scheme and we categorize these answers into category S99: Other. Second, we do find a substantial number of interval/point confusions — essentially S9 in our scheme — which almost all occur in kinematics questions. Last, we find only a small number of iconic interpretations which are almost always in addition to another strategy and are also categorized in the "Other" category (S99) due to their low occurrence.

We can also compare our results about the ratio of differences and the ratio of coordinates to those from Wemyss and van Kampen [108] and to those from the follow-up study from Bollen *et al.* [184]. The key differences between our study and the others are that we ask to determine the velocity, whereas they ask to determine the *instantaneous speed*, and that our respondents are 9th grade students in the Flemish educational system whereas their students are first-year university students in the Irish, Flemish or Basque higher educational system. Table 5.6 provides an overview including the relevant data which can be compared, specific details about this are provided in the footnote [1]. In general,

[1]The questions selected in our test are all those with the "determine via graph" question type in which the concept is the slope/velocity. This includes all kinematics and mathematics questions and both positive and negative slope. For detailed comparison between the data from the other two studies we refer to the study from Bollen *et al.* [184]. Note that KU Leuven and UPV/EHU data is from undergraduate students in a calculus-based course and that from W & vK and DCU are from students in an algebra-based course at DCU. The selected data from the two referenced studies is from Table II and Table III in each of them. This includes all questions with distance-time graphs (kinematics) in which the respondents are asked to determine the *speed* at a specific *instant*, and all $y(x)$-graphs (mathematics, which they call context-free) in which the respondents are asked to determine the slope at a specific *point*. In our data, questions without an explanation are not included. In the referenced studies though, there is a category "No answer" in the mathematics data. For a fair comparison, we subtract the cases in "No answer" from the total and recalculate the frequency. In their kinematics questions there is a category "Other or no answer" or "No response or incoherent or other", which complicates the comparison. In the study from Wemyss and van Kampen we see that in the mathematics table the total prevalences in "Other" and "No answer" are very similar. In line with this we subtract half of the "Other or no answer" prevalence from the total n in the kinematics table. We did the same with the data from the Bollen *et al.* This seems appropriate for the DCU data since the numbers in "Other" and "No answer" are very similar, but less so for KU Leuven and UPV/EHU. For these last two universities, we calculate that the maximum deviation is only about 3%, which would occur if e.g. "No response or incoherent or other" would actually all be "No response", which is unlikely. There are two useful categories to compare. The first are the ones in which some form of a ratio

we find that the results from our students in Flemish secondary education are similarly structured (lowest and highest frequencies, largest differences, ratios of frequencies) as those from the Flemish undergraduate students at KU Leuven, which is that they use the ratio of differences about twice as often as the ratio of coordinates in kinematics, and that they use the ratio of differences far more often than the ratio of coordinates in mathematics. This shows that the effects from the educational system last from secondary education to higher education. Furthermore, all cohorts show large context gaps with kinematics being the most difficult, indicating that students have difficulties linking their understanding of linear function in mathematics to a similar topic in kinematics.

Table 5.6: Comparison between our results (Flanders) and those from Wemyss and van Kampen [108] (W & vK) and those from Bollen *et al.* [184] (KU Leuven in Belgium, UPV/EHU in Basque country in Spain, DCU in Ireland). The categories are: $\Delta y/\Delta x$ in which in some way a ratio of differences is used (combined $S7 + S8$); and y/x in which in some way a ratio of coordinates is used ($S9$). Frequencies f_K in kinematics and f_M in mathematics are expressed in percentages with respect to the total number n of answered questions.

			Category	
Study	$f_{Context}$	n	$\Delta y/\Delta x$	y/x
Flanders	f_K	433	63	27
	f_M	455	86	1
W & vK	f_K	403	22	56
	f_M	355	57	25
KU Leuven	f_K	202	50	22
	f_M	202	81	5
UPV/EHU	f_K	121	46	30
	f_M	224	74	9
DCU	f_K	480	18	66
	f_M	355	57	25

of differences is used — which are S7 and S8 in our scheme — and their category $\Delta y/\Delta x$ and $\Delta x/\Delta t$. The second one is the one in which some form of a ratio of coordinates is used, meaning S9 in our scheme and y/x and x/t in theirs.

5.7.2.3 Interval/point confusions

In questions with graphs, our data contain many examples of the well documented interval/point confusion; in questions with an algebraic representation though, this — to our knowledge — has not yet been well studied. In questions with a "determine via formula" task, different and kinematics-specific errors occur in which students exhibit interval/point confusion often combined with the use of dimensional arguments. In addition, many of the examples in this section illustrate that different contexts often result in different strategies and errors. Three cases are discussed:

Case 1: Numerical interval/point confusion A value for one of the variables is entered into the equation and the matching value for the other variable is calculated. The ratio $\frac{x}{t}$ is then calculated. This is illustrated in Figure 5.10 in which the student chose $x = 2$, calculated the corresponding value for t and then calculated the ratio of both. Note that this student correctly wrote down the formula for velocity but failed to interpret it correctly.

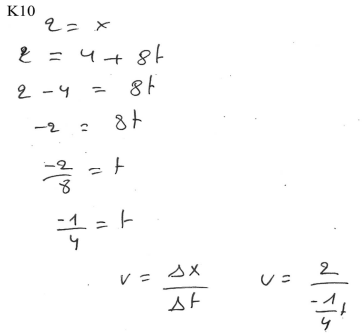


Figure 5.10: Example of Case 1 with question K10 in which the equation $x = 4 + 8t$ was provided: a numerical interval/point confusion and additionally the Δ in the formula for v is written but not taken into account.

Case 2: Algebraic interval/point confusion The student seems to look for the ratio of x over t, which by some is literally written as $\frac{distance}{time}$, while actually $\frac{position}{time}$ is calculated. The focus here is to find "an x" and "a t" or "a distance (position)" and "a time" because the formula requires them. In contrast with case 1, the first step here is to find algebraic expressions for x and t instead of finding numerical values. This confusion is illustrated in the answer to K12 in Figure 5.11, where the student also omits the deltas from the formula for velocity and tries to find an algebraic expression for x, for which the

student uses the equation of motion. After calculating the ratio, the student is stuck and does not know how to continue to end up with a numerical value. In contrast, in the isomorphic question M12 (also shown in Figure 5.11), the student manages to solve the question with an (assumed) expert-like strategy by identifying the slope through its location in the equation (S1). Another

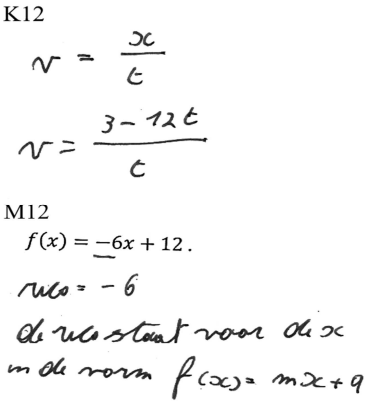


Figure 5.11: Example of Case 2 with question K12 in which the equation $x = 3 - 12t$ was provided: an algebraic interval/point confusion in which the equation of motion is substituted in the formula for velocity. Additionally the isomorphic question M12 from the same student is shown, in which strategy S1 is used: the location of -6 is underlined in the question and the explanation translates as: "the slope is located before the x in the form $f(x) = mx + q$."

illustration of an algebraic interval/point confusion is given by the answer to K10 in Figure 5.12, in which the student first makes clear that the equation of motion provides the expression for x, and then manipulates the equation to find an expression for t. Next, the correct formula for v is written, and the previous two expressions are incorrectly entered in that formula like an interval/point confusion. Finally, the student is stuck. Interestingly, in the comparable question M12 — in which only the sign of the slope/velocity is different — the student applies the (assumed) expert-like strategy S1 in which the slope is identified through its location in the equation.

Case 3: Unit-based interval/point confusion Here, the student tries to find a ratio of something with units in meters over something with units in seconds to achieve meter per second which they know is a proper unit for velocity. Recall that we opted to not include the units in an equation in kinematics, as stated in design choice (iii). Figure 5.13 shows an example in which the student tries to find an expression in units of meters and an expression in units of seconds and then takes the ratio. For the numerator the

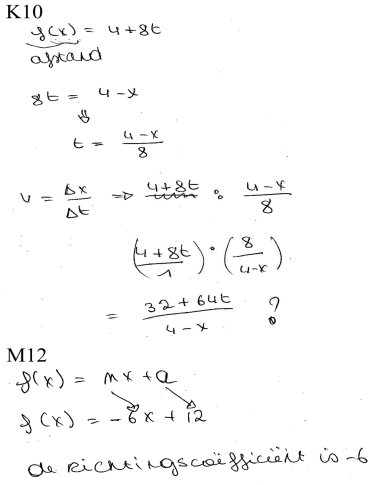


Figure 5.12: Example of Case 2 with question K10 in which equation $x = 4+8t$ was provided: an algebraic interval/point confusion in which the equation of motion is used for x and a manipulated version for t. The comparable question M12 from the same student is solved using strategy S1 by identifying the slope through its location in the equation, made clear by comparing with a standard form $f(x) = mx + q$. The written explanation translates as: "the slope is -6".

equation of motion is chosen since this results in a value expressed in meters. Also, the student incorrectly indicates that the term $8t$ has units of seconds and chooses the full term (coefficient and variable in this particular case) for the denominator. An incorrect calculation then results in $4\frac{m}{s}$.

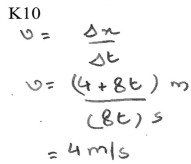


Figure 5.13: Example of Case 3 with question K10 in which equation $x = 4+8t$ was provided: a unit-based interval/point confusion in which the equation of motion and a term in the equation are thought to be expressed in units of meters and units of seconds respectively.

Figure 5.14 shows another example of question K10 in which the student identifies the suspected units of the coefficients and takes the ratio to achieve a

value with units $\frac{m}{s}$. The term $8t$ is incorrectly thought to be expressed in units of seconds and the other term is correctly thought to be expressed in units of meters.

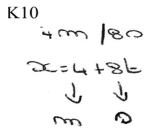


Figure 5.14: Example of Case 3 with question K10 in which equation $x = 4+8t$ was provided: a unit-based interval/point confusion in which the coefficients from the equation were each linked with a unit and then the appropriate ratio was taken to calculate the velocity.

These cases and the examples illustrating them, also provide evidence of poor algebraic manipulation skills, as well as poor understanding of units in algebraic expressions, of the Δ symbol and of the equation of motion. This highlights the weak link between students' understanding in mathematics and in kinematics. The quality of the link between mathematics and kinematics is likely low in Case 1 and Case 2, since these are basically mathematical approaches without any sign of proper understanding in physics other than a formula learned by heart. Yet at the same time the kinematics context does cause difficulties that do not arise in the isomorphic mathematical questions. In Case 3, the students take more of a physics approach since they take the units into account — which is promoted by many teachers — but here they fail to connect the units to the correct parts in the algebraic expression, which highlights their difficulties with algebraic expressions in contexts other than mathematics.

5.7.2.4 The sign of the slope

In questions with a negative slope, and mainly in kinematics questions, students sometimes made statements related to the motion of the cyclist such as: "the cyclist is returning" or "is riding backwards" or "is slowing down". The first two statements indicate that these students include a sense of direction in their interpretation, but some have difficulty to understand and/or express moving with or opposite the direction of the position axis. There was even a respondent who repeatedly wrote "riding backwards or back" thus illustrating doubts about the correct interpretation. The "slowing down" statement illustrates the difficulty students have with interpreting $x(t)$-graphs with constant velocity by confusing them with $v(t)$- or $v(x)$-graphs. Our results show that the sign of the slope in mathematics is no problem for the majority of students, but when

confronted with a negative sign for the velocity in kinematics they frequently omit the minus sign. Recall that Dutch speaking students only have the single word "snelheid", whereas English speaking students have velocity and speed.

This difference in student strategy between mathematics and kinematics concerning the negative slope is well illustrated with some examples from the test. Figure 5.15 shows isomorphic questions K8 and M8 solved by the same student, both solved by using a triangle (S7). In kinematics the minus sign is omitted, but is taken into account in mathematics. Also note that in kinematics the students used the intercepts while in mathematics a horizontal step size of $+1$ was used.

Figure 5.16 also shows two isomorphic questions solved by the same student, but in this case with positive slope. The kinematics question is solved incorrectly and shows an interval/point confusion, while the mathematics question is approached correctly by drawing a triangle (S7) with a horizontal step size of $+1$. In this case M6 was eventually also incorrect because the ratio was inverted.

Figure 5.17 shows isomorphic questions K8 and M8 solved by the same student who answered the questions in Figure 5.16. This student incorrectly solved the kinematics question (interval/point confusion) and omitted the minus sign, but answered the mathematics question by drawing a triangle (S7) with a horizontal step size of $+1$ and includes the minus sign. Again, the mathematics question was also incorrect because this student consistently used the inverted ratio for the slope.

Another question which challenged students' reasoning with negative velocities is K4 in which two negative velocities are compared graphically. As stated before, the average scores were very low for K4, but they were also at the low end for M4. The reason why both were not solved very well is because when asking to compare two negative numbers e.g. -5 and -2, many students reasoned that -5 was the largest because $5 > 2$, thus not taking into account the minus sign. This occurred in kinematics and in mathematics. An additional problem for kinematics — as made clear in the discussions above — is that students tend to consider speed instead of velocity, which increased the number of incorrect answers significantly. These two difficulties fully explain the difference between the average accuracies for K4 and M4. No such issues occurred in questions K2 and M2 which are the isomorphic versions with a positive slope.

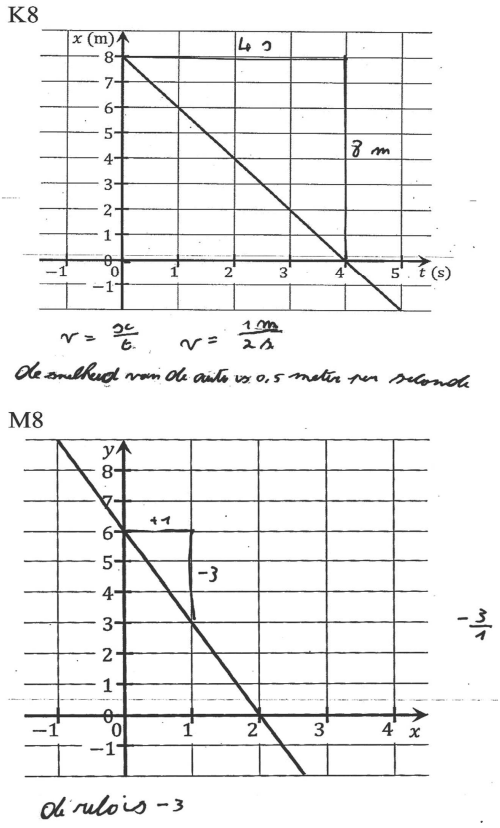


Figure 5.15: Example of isomorphic questions K8 and M8 solved by the same student. In both cases strategy S7 is used, but in kinematics the minus sign has not been taken into account in contrast to M8. Translated wording in K8 and M8 are "The speed/velocity of the car is 0.5 meters per second." and "The slope is -3".

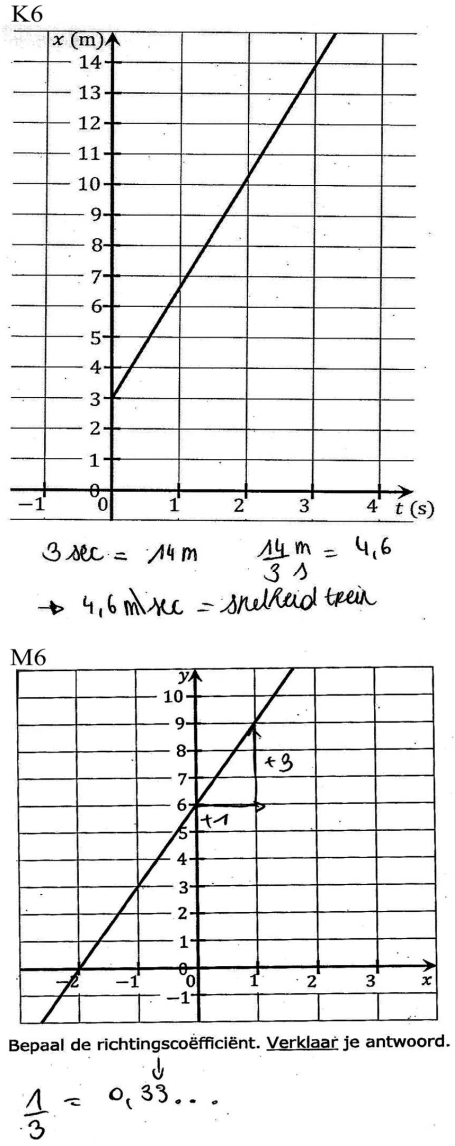


Figure 5.16: Example of isomorphic questions K6 and M6 (with positive slopes) solved by the same student. In K6, the interval/point confusion occurs resulting in an incorrect answer, but in M6 a triangle is drawn (S7) with a horizontal step size of +1 resulting in a correct approach but an incorrect answer because the ratio was inverted. The text in K6 translates as “speed/velocity train”.

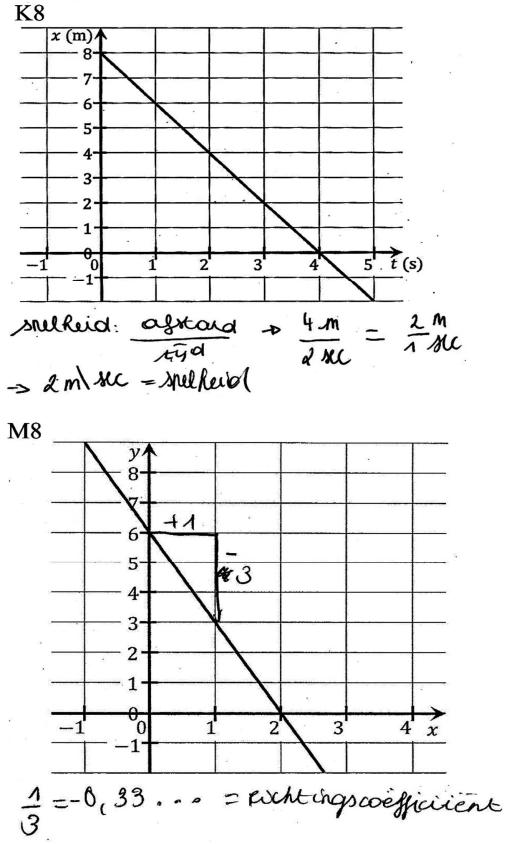


Figure 5.17: Example of isomorphic questions K8 and M8 (with negative slopes) solved by the same student who answered the questions in Figure 5.16. In K8, the interval/point confusion occurs and the minus sign is omitted. In M8 a triangle with horizontal step size of +1 is drawn and the minus sign is included. In K8, the text translates as "*velocity* : $\frac{distance}{time}$.

5.8 The effect of the iSTEM approach

5.8.1 Students' accuracy

In the previous sections of this chapter, we have studied students' performance and students' conceptual understanding of linear function problems in kinematics and mathematics. The GEE model allowed to study the relevant aspects independent of the instructional approach. In this section we expand the previous GEE model with an additional factor to include the instruction, i.e. *instruction*, and study the effects thereof. In doing so, we aim to answer the fifth research question in this chapter: *How does grade 9 students' performance and conceptual understanding of linear function problems compare between kinematics and mathematics and between the STEM@school instructional approach and the regular instructional approach?*

To answer this question, the full group of 253 respondents from the 9th grade was divided into two cohorts. The first cohort contains 121 iSTEM students with an average age of 14.4 year ($SD = 0.6$ year), of which 44 are female and 60 are male, from five different schools. The second cohort is the control group and consists of 132 students with an average age of 14.4 year ($SD = 0.4$ year), of which 72 are female and 60 are male, from five different schools.

Similarly to the representational fluency test, this test was administered after all students had received instruction about linear functions and about 1D kinematics in physics. The students in the control cohort followed the regular curriculum and those in the iSTEM cohort followed the integrated STEM approach from the STEM@school project and specifically the green wave module which we discussed in Chapter 3.

To compare the two cohorts, we extend the GEE model with an *instruction* variable which represents the instructional approach, i.e. either the iSTEM approach or the regular one in the control cohort. This extended model includes all main effects and all interaction effects up to the 4-way interactions. All significant effects and effects of interest including *instruction* are shown and compared to the previous GEE model in Table 5.7.

The extended GEE model shows a near identical structure to the previous one, and most importantly shows that *instruction* is not a significant main effect with Wald $\chi^2(1) = 0.777$ and $p = 0.378$. The extended GEE model results in a mean accuracy and a standard error for the iSTEM cohort and the control cohort of 0.50 ($SE = 0.21$) and 0.53 ($SE = 0.22$) respectively. This means that test scores of students in the iSTEM cohort do not differ significantly from the the scores of students in the control cohort.

Table 5.7: Significant main effects, interaction effects, and interaction effects of interest with Wald χ^2, the significance (p-value) up to three digits, and degrees of freedom (df) for the results from the basic model and results from the extended model which includes the instructional approach (iSTEM or control group).

	Basic model		Extended model		
Term	χ^2	p	χ^2	p	df
context	30.756	0.000	21.837	0.000	1
QType	213.710	0.000	168.623	0.000	2
concept	16.854	0.000	15.654	0.000	1
slopesign	239.782	0.000	196.699	0.000	1
gender	1.796	0.180	2.249	0.134	1
instruction	-	-	0.777	0.378	1
context * *QType*	163.615	0.000	170.457	0.000	2
context * *concept*	370.275	0.000	332.174	0.000	1
context * *slopesign*	83.070	0.000	71.697	0.000	1
context * *instruction*	-	-	0.024	0.876	1
QType * *slopesign*	107.922	0.000	104.269	0.000	2
QType * *instruction*	-	-	1.845	0.398	2
concept * *slopesign*	162.529	0.000	165.119	0.000	1
concept * *gender*	5.428	0.020	5.181	0.023	1
concept * *instruction*	-	-	0.094	0.759	1
slopesign * *instruction*	-	-	1.581	0.209	1
gender * *instruction*	-	-	1.485	0.223	1
context * *QType* * *concept*	21.635	0.000	24.640	0.000	2
context * *QType* * *slopesign*	5.475	0.065	13.025	0.001	2
context * *concept* * *slopesign*	40.046	0.000	43.343	0.000	1
context * *slopesign* * *gender*	10.988	0.001	6.353	0.012	1
context * *gender* * *instruction*	-	-	5.539	0.019	1
QType * *concept* * *slopesign*	86.085	0.000	90.230	0.000	2
concept * *slopesign* * *gender* * *instruction*			14.372	0.000	1

The *instruction* factor is not included in any significant 2-way interaction effect. To illustrate this, the results for the 2-way interaction effects including *instruction* are shown in Figs. 5.18, 5.19, 5.20, 5.21, and 5.22, thus further supporting the conclusions that the instructional approaches do not result in significantly different performance on the test.

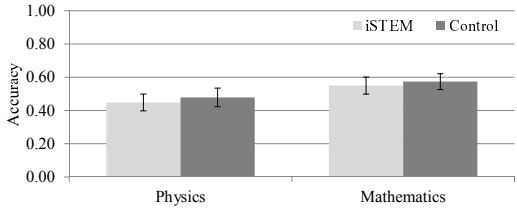


Figure 5.18: GEE results for the 2-way interaction term *context* * *instruction*

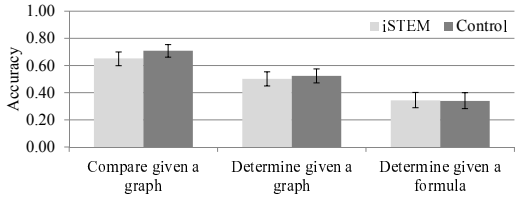


Figure 5.19: GEE results for the 2-way interaction term *QType* * *instruction*

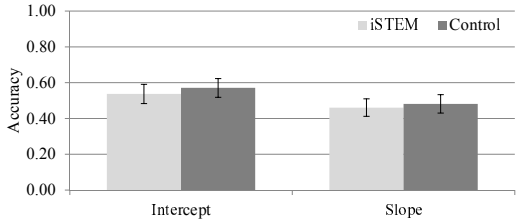


Figure 5.20: GEE results for the 2-way interaction term *concept* * *instruction*

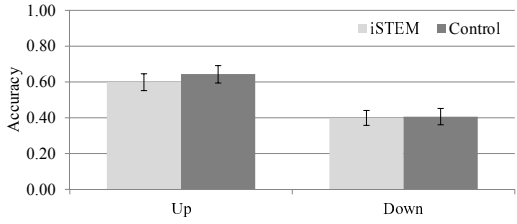


Figure 5.21: GEE results for the 2-way interaction term *slopesign* * *instruction*

The variable instruction does appear in a significant 3-way effect: *context*

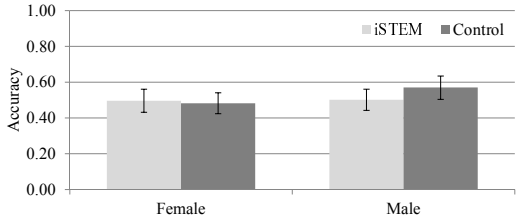


Figure 5.22: GEE results for the 2-way interaction term *gender* * *instruction*

* *gender* * *instruction* with Wald $\chi^2(1) = 5.539$ and $p = 0.019$. Further inspection of this term reveals two significant comparisons which are illustrated in Figs. 5.23 and 5.24. Fig. 5.23 shows a significant difference ($p = 0.043$) between the genders for kinematics in the control group in the advantage of males. Fig. 5.24 shows that females in the control group perform significantly ($p = 0.001$) better on mathematics compared to kinematics. Both of these observations illustrate the differences present between the students of the control group. No significant differences between the instructional approach were present though.

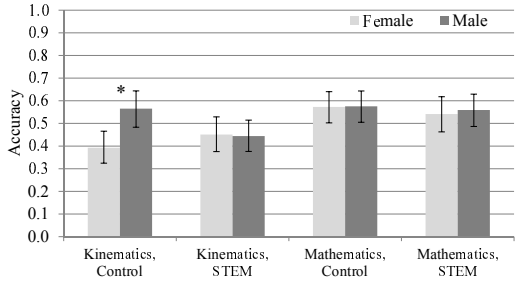


Figure 5.23: GEE results for the 3-way interaction term *context* * *gender* * *instruction*, structured for gender comparison.

Furthermore, there is one new significant 4-way interaction effect: *concept* * *slopesign* * *gender* * *instruction* with Wald $\chi^2(1) = 14.372$ and $p = 0.000$. Studying the detailed GEE output more closely reveals that there are no significant pairwise comparisons between the two instructional approaches within this term. It does confirm earlier results, in particular the significant differences between positive and negative slope for both genders and both contexts.

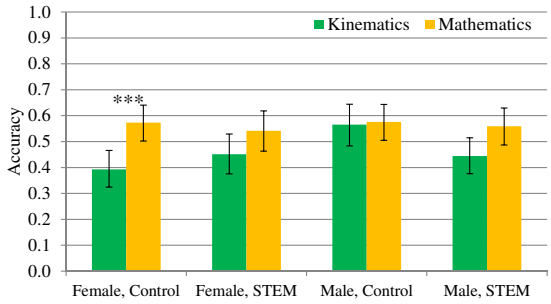


Figure 5.24: GEE results for the 3-way interaction term *context* * *gender* * *instruction* structured for context comparison.

5.8.2 Qualitative results

A comparison of students' strategies between the two cohorts is shown in Table 5.8. This table reflects a very similar use of strategies in both cohorts. There is one exception, concerning the y intercept items and category I2 (identification of the x intercept). This incorrect strategy was used significantly ($\chi^2(1) = 20.356, p < 0.001$) more often by the students in the control cohort (28%) compared to those in the STEM cohort (19%). This difference is partly compensated (11% and 19%) by category I99 (Other). Category I99 contains many different low-prevalence strategies of which the most prominent one is providing a coordinate consisting of the first component of the x and the second component of the y intercept combined. This means that students from both cohorts tend to use the x intercept in their answer and that STEM students combine this more often with the y intercept into a single coordinate. The actual difference concerning the use of the x intercept is therefore smaller than it appears from these numbers.

Tables 5.9 and 5.10 provide more detail about the use of strategies within the two contexts. First, they reflect the effect of the context which confirms the result of the GEE analysis. Second, Table 5.9 again shows that the control cohort employs incorrect strategy I2 significantly ($p < 0.05$) more often compared to the STEM cohort in both kinematics and mathematics separately.

Table 5.10 shows the results for the slope items and shows significant differences at the $p < 0.05$ level in kinematics: the STEM students use the (correct) method S8 (ratio of differences) significantly more often compared to the control students, and the control students use the incorrect method S9 (ratio of coordinates) significantly more frequently. On the other hand — and also for kinematics items — the students from the control cohort use strategy S6

Table 5.8: Frequencies (in %) for the STEM and the control cohort by concept. The percentages are relative to the total number n of strategies employed per concept.

y intercept			Slope		
Category	STEM $n = 875$	Control $n = 945$	Category	STEM $n = 1344$	Control $n = 1450$
I1	4	3	S1	13	15
I2	19	28	S2	2	2
I3	12	11	S3	5	3
I4	3	2	S4	2	1
I5	1	1	S5	1	1
I6	21	22	S6	15	19
I7	21	21	S7	12	11
			S8	30	27
			S9	8	10
I99	19	11	S99	12	11

Table 5.9: Frequencies (in %) for the STEM and the control cohort by context for the y intercept. The percentages are relative to the total number n of strategies employed per context and cohort.

Concept: y intercept				
	Kinematics		Mathematics	
Category	STEM $n = 395$	Control $n = 377$	STEM $n = 480$	Control $n = 568$
I1	5	3	3	4
I2	6	13	31	38
I3	5	3	19	15
I4	5	4	1	1
I5	1	1	1	1
I6	34	40	10	10
I7	16	17	24	24
I99	28	18	12	7

(reasoning with steepness) significantly more frequently ($p < 0.05$). When looking at mathematics, the results are much more similar between the two cohorts.

Table 5.10: Frequencies (in %) for the STEM and the control cohort by context for the slope. The percentages are relative to the total number n of strategies employed per context and cohort.

Concept: slope				
	Kinematics		Mathematics	
Category	STEM $n = 650$	Control $n = 681$	STEM $n = 694$	Control $n = 769$
S1	1	1	23	27
S2	1	2	3	2
S3	2	2	8	5
S4	3	2	2	1
S5	2	2	1	1
S6	11	16	19	21
S7	5	5	18	17
S8	41	35	19	20
S9	15	20	0	1
I99	19	16	5	6

5.8.3 Conclusion

Overall, the comparison between the iSTEM instructional approach and the regular approach in the control cohort did not result in any significant differences in accuracy. This is again a positive result since this shows that the integrated teaching approach managed to achieve a similar level of performance and understanding as the regular approach after only the second implementation, this despite a substantially different instructional approach for both students and teachers.

When looking at the use of strategies, we found that control students include the x intercept in their answers to y intercept items significantly more frequently, but that this is partly offset by the STEM students' higher tendency to use a combination of x and y intercept as a single coordinate in the I99 category. Furthermore, this difference was consistent in both kinematics and mathematics. Lastly, we found that control students reason with "steepness" significantly more frequently in kinematics; and more importantly and also specifically in the kinematics context, that STEM students make significantly more use of the ratio of differences (S8) and significantly less use of the incorrect ratio of coordinates (S9) compared to the control cohort.

5.9 Implications for teaching and future research

The validation shows that teachers can use the test for groups or individuals [186, 214] to assess students' understanding, to identify difficulties, and to uncover solution strategies and confusions. This can be used for targeted remediation. Moreover, the test can also be used as a pre-test and post-test for intervention studies similar to how Hill *et al.* [215] used the Force Motion Concept Evaluation [216] and the Representational Fluency Survey [179] to measure learning gains after an intervention. To suit other research questions, the test could of course be modified, in which case some interesting changes can be made: e.g. require an explanation in every question, include negative values for some variables in the graphs, drop some items in favor of items with other sign combinations or other quadrants, expand the test to include other linear graphs from 1D kinematics ($v(t)$, $v(x)$, $a(t)$, $a(x)$, ...), include other contexts, etc.

Our comparison with other studies shows substantial differences between different cohorts depending on age group, the educational system and the type of course (algebra- or calculus-based). Because of our fairly large number of respondents, distributed over multiple schools and classes across Flanders, the results are representative for students in a science oriented curriculum in the early stages of learning about kinematics and linear functions. Additional studies with different age groups and curriculum programs would be welcome, as well as longitudinal studies which follow students' progression concerning this topic.

To improve students' understanding and their linking between contexts, we suggest introducing isomorphic quantities such as slope and velocity in parallel. Comparing isomorphic situations from different contexts could be a good approach, e.g. show a graph and the matching algebraic expression in mathematics and in kinematics side by side and annotate them while highlighting the similarities. The STEM@school project already aimed at using such an approach but considering the implementation fidelity further improvements can still be made. Additionally, we suggest discussing examples in different reference systems to illustrate the effect on equations and graphs. We also observed difficulties with the Δ symbol, which is why we advise to explicitly write the incremental steps in calculations to clarify its meaning in both contexts, e.g.: $v = \frac{\Delta x}{\Delta t} = \frac{x_2 - x_1}{t_2 - t_1}$, and elaborating on the concept of a negative slope and its meaning. Furthermore, the significant main and interaction effects in our study strongly imply that students could benefit from an increased focus on linking different contexts; therefore, this issue also requires appropriate attention during teacher education.

Due to the isomorphic structure of our test and the direct mapping between e.g. item M1 and item K1, the test can also be considered a transfer of learning test. Transfer of learning is often defined as the ability to use and apply skills and knowledge in a different context from the one in which they were learned [217–219]. To quantify transfer of learning is not a straightforward task, but Britton *et al.* [220] have made progress by introducing the Transfer Rating [220], which was subsequently improved by Roberts *et al.* [219] when they introduced the Transfer Index. Based upon correlating performance between matching questions in two contexts, a Transfer Score can be ascribed to each matching set using the system presented in Table 5.11. For a test with n pairs of matching items, the Transfer Index is calculated as follows:

$$Transfer\ Index = \frac{\sum_{i=1}^{n} Transfer\ Score}{2n} \times 100,$$

resulting in a value between 0 and 100, with higher values indicating higher transfer of learning.

Table 5.11: Allocation of Transfer Score [219].

	Accuracies			
Item accuracy in context A	1	0	1	0
Item accuracy in context B	1	1	0	0
Transfer Score from A to B	2	1	0	0

Calculating the Transfer Index from mathematics to kinematics for our data we find a value of 37.63 ± 17.45. Although it is always possible to calculate this index, a consistent interpretation requires an appropriate research design. Our design was chosen to study how well students in the Flemish education system — which uses mandatory learning goals — perform in two related contexts without controlling for the specific implementation in the classroom, essentially considering the classroom a black box governed by the educational system. This means that our data are not fully suited for interpretation in terms of transfer of learning since we cannot verify one of the important conditions to assess transfer between contexts, which in our study would be that learning took place in mathematics *before* it took place in kinematics. Though it is highly likely to be the case, we did not include this information in our original design. Furthermore, as discussed in Roberts *et al.* [219], to verify the causal

relationship of improvements in a context due to transfer of learning from another context, all other variables which can influence the performance should be checked and included, which is not possible with our design. For illustrative purposes, we also calculated the Transfer Index from kinematics to mathematics which results in 41.58 ± 21.21 which is actually higher than that for mathematics to kinematics. Furthermore, we found a Spearman's rank correlation coefficient ρ of 0.356 between kinematics and mathematics items. These results show that the interpretation of the Transfer Index is not always straightforward and should be done with appropriate care. With a matching research design and detailed control of classroom activities though, our test can certainly be used in combination with this index to study transfer of learning, which would be an interesting subject for future research. Additionally, it would also be interesting to calculate the Transfer Index between mathematics and kinematics for, e.g., the items concerning graphs with those concerning formulas separately, which could provide more insight in the effect of representations on transfer.

Since we did not find a significant effect from the instructional approach, no specific suggestions can be formulated that would differ from the suggestions for the regular instructional approach and the above is applicable for both.

5.10 Conclusion

In conclusion, we studied student understanding of the concepts y intercept and slope in linear function problems with graphs and algebraic expressions in isomorphic kinematics and mathematics questions. Test validation resulted in good values for internal consistency, item difficulty indexes, item discrimination indexes and Ferguson's delta. Our results show that students' difficulties are concentrated in physics, algebraic expressions and questions with negative slope. Results from a GEE analysis of the accuracy show that context (kinematics or mathematics), concept (y intercept or slope), slope sign (positive or negative), and question type (determine via graph, determine via formula or compare via graph) are all highly significant factors, as well as almost all interaction effects (up to three way interactions). The study also resulted in a bottom up constructed categorization scheme for students' explanations which proved a reliable tool and showed that there are only a few strategies/errors for questions concerning the y intercept, and a more detailed scheme for slope questions. The scheme included representational transitions, but we found that such strategies were not frequently used in general with only one exception, namely the transition from a graph to an algebraic expression in mathematics questions concerning the y intercept. The main error for the y intercept was the determination of the x intercept instead and students very rarely used the

location in an algebraic expression to identify the y intercept. Additionally, we found that students had poor understanding of formal terminology concerning the y intercept since the wording "function value of 0" caused some difficulties. More variation was found in students' explanations for the slope questions. For questions with an algebraic expression we found little use of identification of the velocity or the slope through the location in the expression in kinematics but far more in mathematics. Slope/height confusion and iconic interpretations were infrequent in our results, but interval/point confusions were seen often. In graphical questions we confirmed presence of the interval/point confusion in physics, but found almost none in mathematics. In questions with an algebraic expression we also found a high frequency of interval/point confusions in kinematics but almost none in mathematics and discussed three notable cases: numerical interval/point confusion in which the focus lies on finding the numbers for x and t; algebraic interval/point confusion in which the focus lies on finding expressions for x and t; and unit-based interval/point confusion in which the focus lies on finding expressions or values with the requested units for x and t. These were often interwoven with additional errors resulting from poor algebraic manipulation skills, poor understanding of units, poor understanding of the Δ symbol and poor understanding of the equation of motion. Additionally, isomorphic examples illustrated that there is a weak link between kinematics and physics and that students are unable to successfully transfer their mathematical understanding to kinematics. Finally, we highlighted the sign issue of velocity in kinematics, which for our respondents was a particular issue due to linguistic difficulties. Isomorphic questions again showed that students consider the minus sign in mathematical questions, but ignore the minus sign in kinematics questions, and tend to reason with speed rather than velocity.

Our last research question in this section aimed to compare the iSTEM and the regular teaching approach. The results show that there are no significant differences between these approaches, neither in the main effect nor in any of the interaction effects with the other independent variables. We can conclude that the iSTEM approach achieved a similar level of performance and conceptual understanding as the regular teaching approach. Furthermore, for items asking to determine the velocity, we found that STEM students use the ratio of differences significantly more, and the ratio of coordinates significantly less frequently compared to the control students; and that the latter reason with steepness significantly more frequently.

Chapter 6

Students' conceptual understanding of 1D kinematics in an iSTEM curriculum

In this Chapter, we aim to provide an answer to research question three as formulated in Chapter 1: *How does grade 9 students' conceptual understanding of kinematics concepts compare in and between the STEM@school instructional approach and the regular instructional approach?* To do so, we use an existing conceptual understanding test and modify it to match our students' learning goals.

The work described in this chapter will be submitted to an international peer-reviewed journal with the following title and co-authors:

> S. Ceuppens, J. Deprez, W. Dehaene, and M. De Cock, "Students' conceptual understanding of kinematics in an integrated stem curriculum," *European Journal of STEM Education*

The first author administered the tests and analyzed the results. All authors contributed to the test design and reviewed the results from the analysis.

6.1 Introduction

Demand for STEM (Science, Technology, Engineering, Mathematics) graduates is predicted to increase in the future in countries all over the world [4] as well as in Flanders (Belgium) [8], and already outgrows or will outgrow supply in the job market in many countries [13–15]. One of the problems is that a sufficient number of students opt for STEM studies in secondary education but many of them do not choose to pursue a STEM degree in higher education. This phenomenon has been attributed to a lack of perceived relevance of STEM subjects by the students [18] and consequently a decline in the interest and motivation for STEM has been observed in highly developed countries [54]. To increase the quantity and quality of graduates with a STEM oriented degree, many educational institutions, teachers and researchers are searching for appropriate ways to adapt the curricula and educational approaches. Often, these are structured around integrated approaches and problem centered learning in which students are challenged with a more complex problem requiring elements from the multiple STEM disciplines to solve it. The STEM@school project in Flanders (Belgium) is an example of such efforts [3,61,222]. It is a research project which — amongst other goals — aims to increase students' motivation to follow an integrated STEM (iSTEM) program in secondary education and in the long run increase the throughput to STEM programs in higher education. In the STEM@school curriculum, students are confronted with complex and open-ended design challenges, which they take on in small teams during the period of a few months. A crucial aspect is that students learn new subject matter from different STEM disciplines because they need this knowledge and these skills to solve the problem, which we hypothesize will increase the perceived relevance of the subject matter. This relevance and the path towards it, are key in the STEM@school approach and are a clear step away from the traditional approach of individually solving small standard exercises after learning the subject matter in one distinct discipline without any links to other STEM disciplines. New educational approaches can be assessed by comparing various aspects of the new approach and the reference approach, such as students' attitudes, acquired skills, critical thinking, motivation, teacher attitudes etc. In this study, we focus on assessing and comparing the conceptual understanding of key concepts in the physics curriculum, more specifically, in one dimensional kinematics.

First, in Section 6.2, we provide a description of the regular curriculum in Flanders, followed by a summary of our integrated STEM approach, after which we elaborate on the STEM course module based on that approach. In Section 6.3 we summarize literature about student difficulties in kinematics. Section 6.4 then lists the research questions for this study. Next, we explain

how we make use of an existing test to answer these questions, and discuss the administration of a modified test in Section 6.5. The results are provided and discussed in Section 6.6 and a conclusion is formulated in Section 6.7.

6.2 Integrated STEM

6.2.1 The standard curriculum

In Flanders, the standard curriculum is based on mandatory learning goals set by the Flemish government as was more elaborately explained in Section 1.3.2. At the start of the curriculum, the students are introduced to 1D kinematics with a main focus on uniform linear motion, velocity, interpretation of graphs, and vector properties of velocity. The main reason for this rather limited and introductory treatment of velocity and motion is to be able to use the concept of velocity in other topics such as optics and energy. We hypothesize that this does not promote good conceptual understanding in kinematics.

6.2.2 Integrated STEM and the STEM@school project

The STEM@school project [3,61,222] in Flanders (Belgium) is a collaboration of two universities (KU Leuven and Universiteit Antwerpen), counselors of the two largest educational umbrella organizations in Flanders, and in-service teachers from thirty secondary schools. The main goals of the project are to develop, implement and validate research based learning materials for integrated STEM projects for grade 9 up to grade 12. With these learning materials and an accompanying educational framework we aim to increase STEM literacy and show students the relevance of STEM in everyday life. Through a systematic review of existing literature we developed this framework consisting of five related but distinct key principles rooted in a social constructivist view on learning. These define our iSTEM approach which is explained in full in the work by Thibaut *et al.* [61].

Integration of STEM content Student activities and learning materials show that the different STEM disciplines are interwoven into a coherent whole and the strong connections between them are made explicit. In doing so it is important to maintain a sufficient investment into discipline specific learning [65]. This principle is essential in the development of iSTEM learning materials.

Problem-centered learning This falls in between project-based and problem-based learning. The students are challenged with authentic, (semi)open-ended, ill-structured, real-world problems in an engaging context at the start of the course module. The problem is treated in its totality, there is no detailed predetermined final solution but there is some guidance towards a certain type of solution, teachers offer guidance and the main goal is for students to develop their problem solving skills.

Inquiry-based learning Creating a learning environment and employing instructional practices which engage students in questioning, experiential learning and hands-on activities that allow them to discover new concepts and develop new understandings. Important aspects are questioning, reflecting, reasoning based on prior knowledge, experimenting, and demonstrating under the guidance of a teacher with the aim of developing critical thinking and constructing new knowledge.

Design-based learning The use of open-ended, hands-on design challenges that provide students with the opportunity to not only learn about engineering design processes and engineering practices, but also deepen their understanding of disciplinary core ideas. Students experience the engineering design process and the need to learn and use new technologies to solve the challenge.

Cooperative learning Providing students with the opportunity to communicate and collaborate with each other to deepen their understanding. Learning takes place in a social context and among peers in which positive interdependence, face-to-face interaction, individual accountability, practicing interpersonal skills and reflecting on the group dynamics are encouraged and in which the teacher takes the role of a supporting coach.

After completing the framework, several teacher design teams were set up, and based on this framework, they developed various iSTEM course modules under the guidance of members from the academic staff. These course modules were then implemented in multiple schools, during which several support and feedback activities were organized, more details of which can be found in the work by Knipprath *et al.* [3]. Each course module is designed to be taught over a period of about three months, with about four to five hours a week of teaching time in the classroom. This teaching time was allocated both within the various courses such as physics, mathematics, engineering, and within the flexible teaching time — a few hours a week which can be allocated to a topic of choice by each school individually. Because of this construction, teachers from these STEM disciplines worked together and aligned their lessons. Where possible, co-teaching was implemented, meaning that teachers with different

specializations taught together in the same classroom. Our approach made clear that an integrated educational approach not only means that courses are integrated at the student level, but also that the teachers must integrate and align their teaching. Before, during and after teaching each course module, the teacher teams were provided with support by the STEM@school team through workshops, an online support forum, school visits and regular semi-structured "intervision meetings" guided by STEM@school project members, in which the teacher teams from various schools came together so they could exchange experiences and learn from each other.

6.2.3 Kinematics course module

The relevant STEM@school course module for our study here, is the *Green Wave* unit for grade 9 students which is freely available for downloading in Dutch from the Scientix project website [66]. The challenge for students is to design and build a functioning scale model of a car that can drive autonomously through a predefined green wave of traffic lights along a straight one dimensional road. At the start, students receive the basic information about the real life distances, the speed limits for each road section and the time intervals of the traffic lights. First, students design their own toy car with an appropriate driving mechanism and electronics. Second, they study the green wave and analyze the best way(s) for the car to drive through the green wave while taking many relevant conditions and limitations into account. While doing so, students discover the need to learn new skills and acquire new knowledge in the various STEM disciplines, thus making all the subject matter relevant to the challenge. Some of the main activities include learning about the relevant physical quantities such as position, velocity and acceleration, their relationship and how they can be represented in graphs and formulas. The main goal is then to model the motion of the car and learn how to program the necessary instructions so the car performs the desired motion with the required accuracy and precision, thus solving the challenge. It is important to note that a large portion of the time allocated for this course module was spent on many other learning goals unrelated to kinematics in addition to the mandatory learning goals for kinematics. One of the important aspects in this module, for the study presented here, is the focus on the integration between physics and mathematics. In the regular curriculum, kinematics and mathematics are taught separately and related learning content is not aligned in time, e.g. linear functions in mathematics and uniform linear motion in kinematics. In the iSTEM curriculum these learning contents were tightly integrated so that students would encounter the need to mathematically describe a motion with a constant velocity and gain insight therein to be able to solve the green wave

challenge. Additionally, extra care was put into tackling misconceptions by explicitly notifying teachers where these could occur and suggesting additional exercises and activities to elicit and resolve them. Key to this is linking the content matter to the real world aspect of experiments with the functioning miniature car driving along a down scaled trajectory with traffic lights. By using distance sensors, the motion of the car was tracked and graphed in real time so students could recalibrate and reprogram their control chips and interpret not only the velocity but also the acceleration at the start of the trajectory and when changing speed zones.

6.3 Literature on student difficulties in kinematics

Kinematics is a challenging topic. Many researchers have reported on student difficulties and multiple tests have been developed to assess them. Trowbridge and McDermott studied student understanding of 1D velocity [110] and 1D acceleration [112], and they identified issues with comparing speed, determining relative velocity, distinguishing velocity and acceleration, and taking into account the correct time interval. Goldberg and Anderson [95] as well as McDermott, Rosenquist and van Zee [183] found student difficulties with graphical representations of negative velocities. The latter authors also encountered student difficulties with interpreting the shape of a graph as the shape of the path, identifying relevant features in kinematics graphs, relating graphs to each other, interpreting the area under a graph and connecting graphs to real world motions. Leinhardt *et al.* categorized the most common difficulties with graphs in three groups: slope/height confusions, interval/point confusions, and iconic representations [107]. Beichner's Test of Understanding Graphs in Kinematics (TUG-K) [111] provides a valuable assessment tool for many of the previously mentioned issues and administration thereof also confirmed the many difficulties reported earlier. This test was recently updated by Zavala *et al* [223]. Some of these issues are also covered in the Force and Motion Conceptual Evaluation (FMCE) from Thornton and Sokoloff [216] and the difficulties with the area under a curve in physics problems have also been studied by Nguyen and Rebello [224]. Due to the persistent nature of these difficulties with graphs in kinematics, they also remain a focus in more recent studies [45–47, 108, 175, 184, 205]. Aguirre and Rankin [25] studied conceptions about vector kinematics and found that many college students still hold on to intuitive beliefs about the vector aspects of kinematic quantities, specifically concerning the independence of orthogonal velocities. They observed that about half of their respondents believe that a velocity in one direction is influenced by a velocity in an orthogonal direction, illustrating that they do not grasp the formal vector treatment of composition

of orthogonal component velocities. Hestenes and Wells developed the Force Concept Inventory (FCI) [26] and the Mechanics Baseline Test (MBT) [30] to study students' preconceptions and assess students' understanding after formal education respectively. Additionally, in our previous studies we have investigated students' representational fluency of linear functions in 1D-kinematics (as well as in mathematics) [175] and students' conceptual understanding of linear functions and solution strategies in kinematics and mathematics [205]. In the former, we identified increased difficulties with the slope and specifically with a negative slope in kinematics, as well as significantly lower performance with representational transitions containing formulas. In the latter, we also found increased difficulties with negative slope in kinematics as well as with comparing and determining positive and negative slopes. Moreover, we confirmed the high prevalence of interval/point confusions in graphs and identified three distinct types of interval/point confusions in formulas, i.e.e numerical, algebraic and unit-based. A recommended overview of many of these issues for the whole of mechanics is also provided by McDermott [63].

6.4 Research questions

Provided the above, we aim to answer the following three research questions:

(1) To what degree does an introduction on 1D kinematics based on a limited set of learning goals result in conceptual understanding of displacement, distance, average speed, velocity as a rate, velocity as a 1D vector, distance as area under a $v(t)$ curve, acceleration as a rate and acceleration as a 1D vector?

(2) How does students' conceptual understanding of selected 1D kinematics concepts compare between students in the regular curriculum and students in the iSTEM curriculum?

(3) Which are the most notable student ideas in the regular curriculum and in the iSTEM curriculum and how do they compare?

To answer these research questions, we need to assess students' conceptual understanding in kinematics. To do so, we make use of the recently developed and thoroughly validated Kinematics Concept Test (KCT) from Lichtenberger *et al.* [35], from which we select the relevant items which match our students' curriculum. The multiple choice items in this test are linked to specific concepts and the distractors are based on known misconceptions thus making the KCT an ideal assessment tool for our study.

6.5 The modified KCT

6.5.1 Test structure

To assess students' conceptual understanding in kinematics we make use of a modified version of the recently developed kinematics concept test (KCT) from Lichtenberger *et al.* [35] Their 49-item multiple-choice test includes items from other tests such as the Force Concept Inventory (FCI) [26] and the Test of Understanding Graphs in Kinematics (TUG-K) [111] and is complemented by new items. Furthermore, the distractors for each item are based on known preconceptions and misconceptions. Depending on the item, the number of the relevant misconceptions varies, thus the number of distractors is not always the same and can vary between three and seven. The original KCT covers seven kinematics concepts: C1 - velocity as rate, C2 - velocity as one-dimensional vector, C3 - addition of velocities in two dimensions, C4 - displacement as area under the $v(t)$ curve, C5 - acceleration as rate, C6 - acceleration as one-dimensional vector, and C7 - velocity change as area under the $a(t)$ curve. Furthermore, it makes use of three representations: pictures, tables and graphs. The inclusion or exclusion of other concepts was guided by their target respondents, which were Swiss high school students aiming to enter university.

Our target respondents are grade 9 students in Flemish secondary education who have only just been introduced to kinematics for the first time with a main focus on displacement and distance, average velocity, uniform linear motion and the graphical representation of motion. Related topics such as instantaneous velocity and the vector properties of velocity, as well as concepts concerning acceleration usually receive very little to no attention at this early stage. As discussed before, the students in our integrated STEM cohort were more exposed to all these concepts, albeit in a problem-centered learning approach in which they were necessary to solve the problem. Because of the differences with the Swiss Gymnasium students we created a modified KCT in Dutch, of which an English translation is available as in Appendix A.6. Note that we translated the original Swiss German version from Lichtenberger *et al.* to Dutch and that our test in Appendix A.6 is a retranslation to English which differs somewhat from that published by Lichtenberger *et al.* to reflect our Dutch translation as much as possible. For easy identification and comparison of the items we continue to use the original item numbering and concept numbering from Lichtenberger *et al.* throughout this paper. The structure of our modified KCT is illustrated in Table 6.1. Note that in concept C4, we replaced the word "displacement", which was used in Lichtenberger *et al.*, with the word "distance", which is a more accurate terminology when considering

the wording in the questions. The items are categorized into seven categories and into three representations. These kinematics concepts are described in more detail in the Appendix A.1. An overview of the known misconceptions — originally developed by Lichtenberger *et al.* but not included in their paper — linked to each of the concepts in the test is shown in Table 6.2.

Table 6.1: Overview of the structure of the modified test by category and representation. Items in parentheses are linked to two concepts.

	Representation		
Concept	Picture	Table	Graph
C0: Displacement and distance	01, 02	-	-
C00: Average speed	001, 002	-	-
C1: Velocity as rate	1, 25	3, 27	10, 11, 16, 20 26, 33, (37)
C2: Velocity as 1D vector	5, 41	-	21, (37), 43
C4: Distance as area under a $v(t)$ curve	-	-	30, 35, 38, 49
C5: Acceleration as rate	-	-	6, (23), 32
C6: Acceleration as 1D vector	9, 36	-	(23), 24, 48

After personal correspondence with the authors of the KCT, we were able to verify that our modifications did not result in any significant changes to the factor structure or the reliability of the test. The following modifications were made. 1) We deleted concept C3 (items 4, 8, 13, 18, 45) and concept C7 (items 22, 40, 44, 46) which are not part of our students' curriculum (iSTEM and regular). 2) We deleted items 2, 12, 17, 31, 34, and 42 to reduce the number of items, especially highly similar ones, since these items have high correlations to items within the same concept category. 3) We reduced the number of items in concepts C5 and C6. These concept categories are not mandatory in the regular curriculum and are only touched upon in the STEM curriculum, thus we want to see how students would perform intuitively[1] in these categories. Therefore, we included some items from these categories — mainly focusing on graphs and on vector aspects in pictures. In category C5, we only kept graphical items: one qualitative item (6), one quantitative item (32), and one item linked to two concepts (23). In category C6, we kept the picture items (9 and 36), two qualitative questions (24 and 48), and also one item linked to two concepts (23).

[1]With *intuitive(ly)*, we mean in a way that does not involve analytical reasoning. An interesting analogon is that of a biological reflex which is a reaction to a stimulus that does not require processing by the brain or conscious thought. Yet in contrast to the biological analogon, intuition in mathematics [41, 225] — and in our view similarly for other sciences — can be trained by increasing experience until things become self-evident, i.e. intuitive.

4) Items 7 and 37 are both part of C1 and C2, so we opted to only keep item 37 since it was the most relevant for our students and has no parabolic sections in the graphs. 5) The suggested changes [35] by Lichtenberger *et al.* to items 6, 26 and 49 were taken into account. These modifications resulted in the retention of 26 items. In addition to these, we included four newly developed items to cover categories C0: displacement and distance, and C00: average speed, which our students do encounter in their curriculum, resulting in a 30-item test.

A pilot study with this first version of the test was performed as a paper and pencil test in 2015–2016 with 145 students, with an average age of 14.78 year ($SD = 0.35$ year) from four different schools. Written annotations and remarks from students on that test, in combination with their achieved accuracies, allowed us to make improvements to some questions. These include small modifications to pictures; the use of integers instead of decimal numbers in item 3; using a step size of 1 m/s in the graphs in items 32 and 35; the use of different line styles instead of colors to identify different curves on the same graph in items 6, 26 and 49; identify the different curves on the same graph more clearly by annotating them with "A", "B" or "C" both left and right of the intersection in items 6 and 26; add grid lines to some graphs; shift point 2 in the graph of item 11 downward on the slope and use letters instead of numbers to mark the points; change the curve in items 16 and 20 into the graph of a piecewise linear function; change curve c in item 26 into a straight line; delete superfluous information; and improve wording of the questions and the answer options. For some items, a small number of students selected none of the answer options available and wrote their own answer which could be linked to misconceptions in Table 6.2. In these cases we included those answers as distractors in the new version because other students might have the same misconceptions. This happened with items 1, 21, 25 and 43. Furthermore, we have also changed the order of the answer options in some items to improve clarity. After this revision the test was reviewed by three members of the academic staff with expertise in physics and/or mathematics, and then digitized for digital administration.

Table 6.2: Overview of the possible misconceptions in each concept category. Codes such as "M5C" identify a specific misconception and are included for easy reference. Every distractor in the test is linked to one or more of the items below.

Possible misconceptions in the distractors for each concept
C0: Displacement and distance
• M0A: Displacement equals distance: $\Delta x = d$.
• M0X: Other.
C00: Average speed
• M00A: Average speed equals the arithmetic mean: $v_{avg} = \frac{v_1+v_2}{2}$.
• M00X: Other.
C1: Velocity as rate
• M1A: Velocity equals position: $v = x$.
• M1B: Velocity equals position over time: $v = \frac{x}{t}$.
• M1C: Velocity equals difference in time over difference in position: $v = \frac{\Delta t}{\Delta x}$.
• M1D: Velocity only depends on the difference in position Δx.
• M1E: Velocity only depends on the difference in time Δt.
• M1F: Objects at the same position always have the same velocity.
• M1X: Other.
C2: Velocity as 1D vector
• M2A: Position vector $\vec{x}$ and velocity vector $\vec{v}$ point in the same direction.
• M2X: Other.
C4: Distance as area under a $v(t)$ curve
• M4X: Other.
C5: Acceleration as rate
• M5A: Acceleration equals position: $a = x$.
• M5B: Acceleration equals velocity: $a = v$.
• M5C: Acceleration equals velocity over time: $a = \frac{v}{t}$.
• M5D: Acceleration equals difference in time over difference in velocity: $a = \frac{\Delta t}{\Delta v}$.
• M5E: Acceleration only depends on the difference in velocity Δv.
• M5F: Acceleration only depends on the difference in time Δt.
• M5G: Objects at the same position always have the same acceleration.
• M5H: Objects with the same velocity always have the same acceleration.
• M5X: Other
C6: Acceleration as 1D vector
• M6A: Acceleration vector $\vec{a}$ and velocity vector $\vec{v}$ point in the same direction.
• M6X: Other

6.5.2 Test administration

In the final administration, 339 students from the 9th grade participated.

There are two cohorts, the first cohort contains 165 iSTEM students with an average age of 14.9 year ($SD = 0.5$ year), of which 56 are female and 109 are male, from nine different schools. The second cohort is the control group and consists of 174 students with an average age of 14.9 year ($SD = 0.3$ year), of which 109 are female and 65 are male, from five different schools. The nine schools in the iSTEM group are all schools in which the STEM@school team considers the curriculum implemented as intended. This assessment of high implementation fidelity was made by using all the information from interactions with the teachers and the schools during visits, interviews and intervision meetings in combination with questionnaires after each course module was taught. All participants from all schools in the cohort are enrolled in what is called "aso education" in Flanders, which aims to prepare students for higher education and covers a broad spectrum of predefined courses regulated by government learning goals. Within this type of education, all students in our cohort opted for the program with a strong focus on science. All schools voluntarily participated, informed consent from the students' parents was obtained. At the time of administration, the students had received education on concepts C0 through C2. Those in the iSTEM cohort had also received introductory education on concepts C4, C5 and C6 while those in the regular science curriculum had not because these are not mandatory. Teachers and students were notified of the test administration in advance, but were not informed about the content of the test. The ethical approval for our study was granted by the Sociaal-Maatschappelijke Ethische Commissie (SMEC) Social and Societal Ethics Committee of KU Leuven.

For the test administration, we used the Qualtrics ® software platform for digital surveys. The 30 items were randomized for each student individually but the order of the answer options was kept identical for each item. Students were shown one item at a time. We used the forced response setting which forces students to choose a single answer before allowing them to continue to the next item. Returning to a previously answered item was not allowed. The total time available to complete the full test was 45 minutes.

6.6 Results and discussion

In this section, the results are presented and discussed. The score for each question was determined by awarding one point for selecting the correct answer

option, and zero points for selecting an incorrect option. The accuracy is the percentage score and ranges from 0 to 1.

6.6.1 Student understanding of 1D kinematics concepts

Our first research question is: To what degree does an introduction on 1D kinematics based on a limited set of learning goals result in conceptual understanding of displacement, distance, average speed, velocity as a rate, velocity as a 1D vector, distance as area under a $v(t)$ curve, acceleration as a rate and acceleration as a 1D vector? To answer this questions, we first provide an overview of the score distribution in Figure 6.1 to illustrate the cohort's overall performance. The distribution is strongly skewed and has a mean score — for the full cohort — of $M = 8.50$ ($SD = 4.55$) which is an accuracy of $M = 0.28$ ($SD = 0.15$), and has a long tail towards higher scores. The median score is 7, or 23.3%. Only 36 out of the 339 students, i.e. 10.6%, obtained scores equal to or higher than 15 out of 30. The most frequent score is 6, which was achieved by 42 students (12, 4% of the full cohort). Five students achieved the lowest score of 6.7%, which means that only two questions were answered correctly.

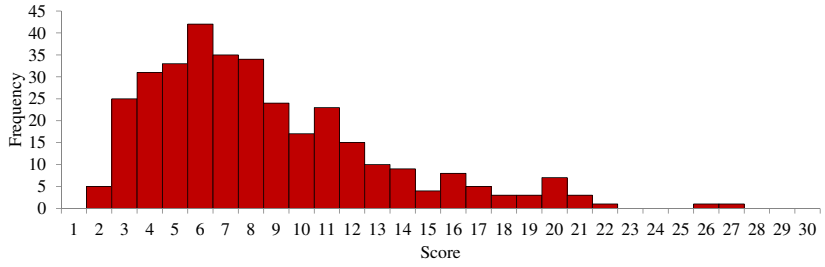


Figure 6.1: Histogram showing the distribution of the scores from all 339 students.

Second, Figure 6.2 shows the achieved accuracies for each concept compared to the values from the original KCT paper. The accuracy of a concept is equal to the mean accuracy of all items related to that concept.

Third, Figure 6.3 shows our per item mean accuracies in comparison to the results from Lichtenberger *et al.* which are provided as reference values. The mean accuracy for the 30-item test is $M = 0.28$ ($SD = 0.15$) and the median accuracy is 0.23 (7 out of 30); for the 26 KCT items the mean accuracy is $M = 0.25$ ($SD = 0.17$) and the median accuracy is also 0.23 (6 out of 26). For the original 26 KCT results, we can estimate $M = 0.56$, by calculating

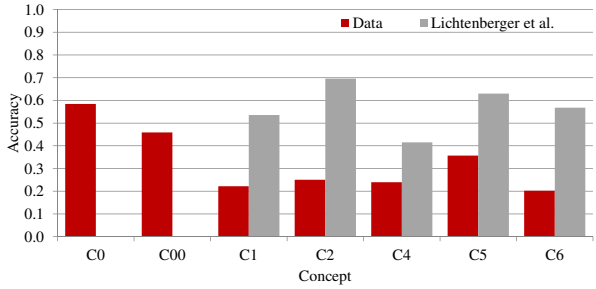


Figure 6.2: Plot of the mean accuracy of each concept. Items linked to two concepts have been included in both means. The data from our administration is shown in red ($n = 339$). The data from Lichtenberger *et al.* is shown in grey ($n = 343$) and represents the accuracies from the original 26 items included in the KCT.

the mean of the average scores of these items instead of averaging over the participants, which provides the correct value for M.

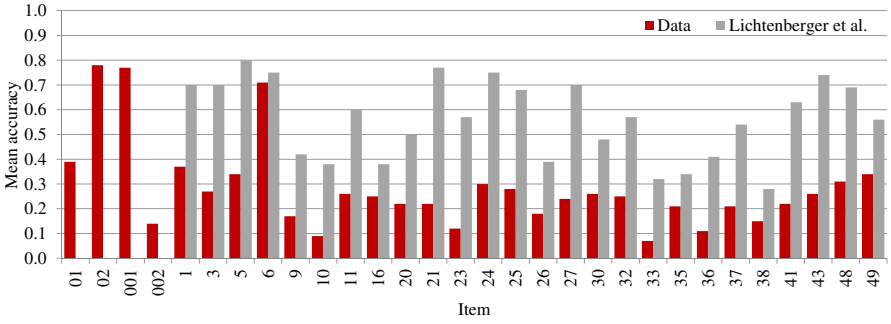


Figure 6.3: Plot of the mean accuracy for each item. Our data is shown in red ($n = 339$) and the data from Lichtenberger *et al.* in grey ($n = 343$). Only our data is available for the first four items since these are not part of the KCT.

These results show a clear trend, i.e. the overall performance is very poor. Some items do have a higher accuracy, such as 02 and 001, which we suspect are because they can be solved very intuitively and because they also have fewer distractors. Also item 6 scores relatively high and close to that in the original KCT data. This is unexpected since this item concerns acceleration, to which the students had only limited or no educational exposure. We can only speculate as to why this accuracy is much higher compared to the other original KCT items. A cautious hypothesis is that acceleration as a rate — particularly

positive acceleration — is a rather intuitive concept to many students because they can directly experience acceleration in everyday life through the force exerted on them when e.g. a car accelerates. We do not have any means in this study to verify this hypothesis though. In any case, it would be interesting to investigation this curiosity further in a new study. Concerning negative acceleration, literature already shows that it is a known challenge for students [226]. Since C5 only contains three items (6, 23 and 32), it also becomes clear why C5 has a somewhat higher mean accuracy. This also explains the somewhat unexpected result that the mean accuracy for acceleration as a rate (C5) is higher than that for velocity as a rate (C2). For all the other items, we find a stark contrast between the results from the 26 KCT items and those from our items. On a per item and also on a per concept basis, the results are slightly below half of those from the KCT validation. Although all performance differences between the data sets are very large, those for C2 and C6 are notable because they indicate low understanding of the vector properties of velocity and acceleration; as well as those for C1 and C5 which indicate very low understanding of rate.

The answer to the first research question is clear: the results for the full cohort show that instruction based on the limited set of mandatory learning goals in kinematics does not result in proper understanding of the kinematics concepts included in this test. This begs the question: is it useful to briefly introduce these concepts without immediately targeting deeper understanding and relying on in-depth instruction at a later time? Given these results, we argue that this brief introduction is likely to be of little benefit to students. Specifically for the iSTEM group though, it is important to look at the broader effect since the iSTEM approach targets additional benefits such as problem solving skills, critical thinking, integration, transfer of learning, etc. In the next research question we will investigate whether there are any gains in conceptual understanding in the iSTEM group compared to the control group. The answer to research question one does imply though, that conceptual understanding must be safeguarded, no matter what the educational approach is.

6.6.2 Comparing iSTEM and the regular approach

In the second research question, we look at how students' conceptual understanding of selected 1D kinematics concepts compares between students in the regular curriculum and students in the iSTEM curriculum. To answer this question we compare student's performance between the two cohorts on the full test and for each concept category individually. First, Figure 6.4 shows the comparison of the mean accuracies. This figure illustrates that the

STEM cohort outperforms the control cohort on the full test as well as on each individual category.

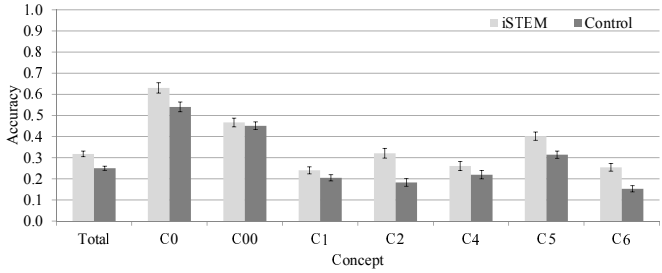


Figure 6.4: Comparison of students' mean accuracy for each concept between the iSTEM group (red, $n = 165$) and the control group (blue, $n = 174$). Items linked to two concepts have been included in both means, e.g. item (37) is included in concept both C1 and C2. The error bars represent the SE for each mean accuracy. The mean accuracy from the STEM students is higher in each comparison.

To determine whether the differences between the cohorts are significant, we use a Mann-Whitney U test because the accuracies of the concept categories are not normally distributed due to the low number of items per category. Table 6.3 shows the results of the analysis. The effect size r was calculated [227] using the z score by $r = \frac{z}{\sqrt{n_1+n_2}}$.

The results for the full test show that the iSTEM cohort performs significantly better than the control cohort with $U = 10838.5$, $z = -3.91$, $p < 0.001$ and a small effect size of $r = 0.21$. Furthermore, the iSTEM group outperforms the control group for each individual concept, and does so significantly for four out of seven concepts with small effects sizes, i.e. for C0, C2, C5, and C6. The largest effect size ($r = 0.26$) is found for concept C2 "Velocity as a 1D vector", and a close second ($r = 0.23$) is for concept C6 "Acceleration as a 1D vector". This shows that most progress for the iSTEM students compared to the control group is found for the vector concepts. The smallest effect size ($d = 0.03$) is found for C00 "Average speed", which can be explained by the polarizing nature of the items 001 and 002. The first is very intuitive and accuracies are high for this question, the second is not intuitive and requires true understanding of the definition of average velocity, which many students struggle with. This is true for both groups which we believe is why the effect size is so small.

Table 6.3: Overview of statistical quantities for the total and per concept accuracy the iSTEM and control group: number of respondents n, mean M, median Mdn, z value, U value from the Mann-Whitney U test, 2-tailed significance (p value), and effect size r (in which < 0.3 is small, $0.3 \leq r < 0.5$ is medium and > 0.5 is a large effect size [228]). The iSTEM group scores higher in each category and significant differences between the groups are found for Total, C0, C2, C5 and C6.

Concept	Cohort	n	M	Mdn	z	U	p (2-tailed)	r
Total	iSTEM	165	0.32	0.23	-3.91	10838.5	< 0.001	0.21
	Control	174	0.25	0.27				
C0	iSTEM	165	0.63	0.50	-2.70	12216.0	0.007	0.15
	Control	174	0.54	0.50				
C00	iSTEM	165	0.47	0.50	-0.52	13991.0	0.620	0.03
	Control	174	0.45	0.50				
C1	iSTEM	165	0.24	0.18	-1.27	13228.0	0.205	0.07
	Control	174	0.21	0.18				
C2	iSTEM	165	0.32	0.20	-4.84	10174.0	< 0.001	0.26
	Control	174	0.18	0.20				
C4	iSTEM	165	0.26	0.25	-1.43	13142.5	0.154	0.08
	Control	174	0.22	0.25				
C5	iSTEM	165	0.40	0.33	-3.39	11607.0	< 0.001	0.18
	Control	174	0.31	0.33				
C6	iSTEM	165	0.26	0.00	-4.32	10671.0	< 0.001	0.23
	Control	174	0.15	0.20				

6.6.3 Most notable student ideas

Our last research question asks for the most notable differences in distractor choices between students in the regular curriculum and students in the iSTEM curriculum. To answer this question, we perform a distractor analysis, the results of which are shown in Table 6.4. While interpreting this table it is useful to remember that the distractors are based on known misconceptions [35] and that their prevalence can be considered significant when it exceeds 10% [229–232], which is the case for many of them. We will highlight the most notable differences between the iSTEM cohort and the control cohort.

Table 6.4: Distractor analysis for the iSTEM group ($n = 165$) and the control group ($n = 174$). All frequencies are in percentages. When an item has less than six answer options, the non-existent ones are marked with '-'. Correct answers are marked in bold face. It is important to note here, that our answer options are sometimes in a different order than those from the original KCT, hence A, B, C, D, E and F are sometimes different from those in the study from Lichtenberger *et al.* [35].

	iSTEM						Control					
Item	A	B	C	D	E	F	A	B	C	D	E	F
01	**46.7**	48.5	4.8	-	-	-	**32.2**	66.1	1.7	-	-	-
02	12.7	**79.4**	7.9	-	-	-	17.2	**75.9**	6.9	-	-	-
001	7.9	**76.4**	10.3	5.5	-	-	10.3	**78.2**	9.2	2.3	-	-
002	**17.0**	68.5	8.5	6.1	-	-	**12.1**	74.1	6.9	6.9	-	-
1	4.8	4.8	6.7	35.8	**38.8**	9.1	4.6	2.9	6.3	36.2	**35.1**	14.9
3	5.5	3.6	45.5	**28.5**	17.0	-	3.4	7.5	46.0	**25.9**	17.2	-
5	1.2	4.2	42.4	**44.2**	7.9	-	4.6	5.2	56.3	**24.1**	9.8	-
6	4.2	1.8	**73.9**	9.7	10.3	-	1.1	1.7	**67.8**	13.2	16.1	-
9	33.3	**21.8**	12.7	13.3	18.8	-	40.8	**13.2**	19.0	6.3	20.7	-
10	8.5	**10.3**	21.2	3.0	57.0	-	8.0	**8.6**	16.1	5.2	62.1	-
11	5.5	33.3	4.2	**21.2**	35.8	-	7.5	23.6	4.6	**31.0**	33.3	-
16	1.2	**28.5**	28.5	6.7	33.3	1.8	0.6	**21.3**	33.3	5.2	36.8	2.9
20	7.9	19.4	20.0	**24.2**	26.1	2.4	9.2	12.1	26.4	**20.1**	29.9	2.3
21	10.9	8.5	2.4	32.7	**26.1**	19.4	6.9	4.6	3.4	42.0	**17.8**	25.3
23	**14.5**	15.2	41.8	12.1	16.4	-	**8.6**	9.8	48.3	11.5	21.8	-
24	23.0	7.9	7.3	**39.4**	22.4	-	33.3	9.2	7.5	**21.3**	28.7	-
25	4.2	9.1	4.2	43.0	**30.3**	9.1	3.4	9.2	5.2	44.8	**25.3**	12.1
26	**18.8**	8.5	20.6	10.9	41.2	-	**16.7**	8.6	22.4	12.1	40.2	-
27	9.1	7.9	46.7	**27.3**	9.1	-	10.3	6.9	51.1	**21.8**	9.8	-
30	**26.1**	20.0	30.9	17.0	6.1	-	**26.4**	13.8	31.0	23.0	5.7	-
32	3.0	**32.1**	19.4	10.3	32.7	2.4	4.0	**17.8**	17.8	16.1	38.5	5.7
33	30.3	**8.5**	5.5	6.1	49.7	-	25.9	**5.2**	2.3	10.9	55.7	-
35	38.2	9.1	8.5	14.5	**21.8**	7.9	45.4	5.2	8.6	16.7	**19.5**	4.6
36	15.2	49.1	**10.3**	12.7	12.7	-	20.7	43.7	**11.5**	13.8	10.3	-
37	10.3	7.9	41.8	**27.9**	12.1	-	20.7	17.8	35.6	**14.4**	11.5	-
38	7.9	8.5	38.2	15.8	**17.0**	12.7	5.7	5.2	46.0	13.2	**13.2**	16.7
41	0.6	17.0	**30.3**	45.5	6.7	-	3.4	14.9	**14.9**	55.7	10.9	-
43	7.3	3.0	5.5	24.8	**32.1**	27.3	10.9	7.5	4.0	21.3	**20.1**	36.2
48	27.9	3.0	10.9	**41.2**	17.0	-	47.1	6.3	8.6	**21.8**	16.1	-
49	7.3	**39.4**	11.5	18.2	23.6	-	10.3	**28.7**	13.2	21.8	25.9	-

The results for item 01, combined with those of 02, show that students in the control cohort have more difficulty to discern between the distance and the displacement than the iSTEM students.

The items in C2 and C6, which all concern the vector properties of velocity or acceleration, are of particular interest because these concepts result in the largest effect sizes between the two cohorts. All items in C2 or C6 require assessment of the sign of the velocity or the acceleration respectively. The results show that a large percentage of students believes that the sign of v and x must be the same, and that the sign of a and v must be the same, i.e. that the respective quantities have the same direction. To highlight this, Table 6.5 shows the percentages of students in each cohort who have such a misconception (M2A or M6A) for the items in these categories. All these numbers are high to very high and in all but one case (item 36) higher for the control group.

Table 6.5: Percentage of students in each cohort who believe that the sign of v and x must be the same (misconception M2A) in C2 items, and who believe that the sign of a and v must be the same (misconception M6A) in C6 items.

Item	Question type	iSTEM	Control
21	Graph, $v > 0$	32.7	35.6
43	Graph, $v < 0$	62.4	75.9
37	Graph, $x(t)$ to $v(t)$	52.1	56.3
5	Picture, $v > 0$, horizontal	42.4	56.3
41	Picture, $v < 0$, horizontal	45.5	55.7
48	Graph, $a > 0$	47.9	69.5
24	Graph, $a < 0$	52.7	69.5
23	Graph, $v(t)$ to $a(t)$	53.9	59.8
9	Picture, $a < 0$, horizontal	33.3	40.8
36	Picture, $a < 0$, vertical	49.1	43.7

When viewing Table 6.5, it is important to recall that the vector properties of velocity are an explicit learning goal in the curriculum, and that there is a large focus on graphs as well. Acceleration is introduced in the iSTEM cohort, but receives little to no attention in the control cohort. This is well illustrated here since larger percentages of the control cohort have these particular misconceptions compared to the iSTEM cohort, except for item 36. A notable contrast is found between positive and negative velocity in a graph when comparing items 21 and 43, since students in both cohort have the M2A misconception about twice as often in item 43 ($v < 0$) compared to item 21

($v > 0$).

The results (Table 6.4) for C2 items with a graph (21 and 43) show that many students consider the sign of the velocity to be determined by both the sign of x and the sign of the slope, i.e. they believe that the velocity is negative if and only if both the curve and its slope are negative or zero (option F in item 43, or similarly option D in item 21).

The comparable acceleration items 24 (option A) and 48 (option A) show similar results in which both the location of the $v(t)$ curve with respect to the horizontal axis and the slope of that curve are taken into account to determine the sign of the acceleration, i.e. they must all be positive (in 48) or all be negative (in 24). The prevalence of this distractor (option A) is the main reason why the difference between the two cohorts is so large for these items. Besides these sign problems, items 37 and 23 show that many students do not understand the relation between an $x(t)$ graph and a $v(t)$ graph or between a $v(t)$ graph and an $a(t)$ graph. The majority selected the same shape (option C for both items) in both pairs of graphs. Although the iSTEM students score higher on average for both items, more of them (41.8% versus 35.6%) do make this mistake for the acceleration item (37), but less of them (41.8% versus 48.3%) for the velocity item (23).

Item 11 (C1: Velocity as rate) is of interest since it is one of the very few at which the iSTEM students clearly perform worse compared to the control students (21.2 versus 31.0). The main difference here is the prevalence of option B, which shows that an additional 10 percentage points of iSTEM students believe that the highest point on the $x(t)$ graph represents the highest velocity, which is essentially a slope/height confusion [107, 205] in graphs and is not uncommon for students at this point in their education.

Concerning rate, item 32 is of interest because the iSTEM group outperformed the control group by a large margin (32.1 versus 17.8). The surplus of incorrect answers from the control students is distributed across all distractors so not a single one really stands out. The main error, for both groups, is option E which, similar to issues in the velocity as a rate category, is a slope/height confusion, but this time in a $v(t)$ graph.

The last item to highlight is item 49 (C4: Distance as area under a $v(t)$ curve) which also shows a substantially higher accuracy for the iSTEM group (39.4 versus 28.7). Again the results show that the control students did not select a particular distractor notably more often than another one when compared to those from the iSTEM cohort. For both cohorts, the most common error was option E which states that there is not enough information available to determine the answer, meaning that about a quarter of all students simply

does not believe it is possible at all to determine the distance traveled from a $v(t)$ graph. About an additional fifth of the students believe that all runners covered the same distance. These rather poor results for the determination of the distance from a $v(t)$ graph are in line with the conclusions from Planinic *et al.* [47] and from Nguyen and Rebello [224]. In many curricula, including in the Flemish one, more time is spent on interpreting the slope of a curve than on interpreting the area under the curve. There is also a mismatch with mathematics education here because, although slope is introduced rather early, derivatives are introduced rather late and integrals are introduced after kinematics education. Though it is clear that these mathematical concepts must be introduced within a more elaborate and consistent mathematical framework, the late timing thereof does appear to hamper a more complete understanding of topics such as kinematics.

6.7 Conclusion

In this study we used a modified version of the KCT to assess grade 9 students' conceptual understanding of kinematics concepts. All respondents were enrolled in the Flemish educational system — based on mandatory learning goals — and opted for programs with a strong focus on science. In the 9th grade, the kinematics part of the curriculum is very basic and is only taught at that point because some of the kinematics concepts are used in other parts of the curriculum, e.g. optics. The full cohort was divided into two groups, the first one is representative for Flemish education and is used as a control group, and the second one used an integrated STEM approach. Overall conceptual understanding was assessed and the two cohorts were compared. Several conclusions can be made based on the results.

First, the KCT proved to be a useful tool to assess conceptual understanding of kinematics concepts. It provides rich data for analysis and was a successful tool to compare groups with a different teaching approach.

Second, we found severe underperformance and a very low mean accuracy of $M = 0.28 \pm 0.15$ for the full cohort which shows that a deep conceptual understanding of the studied kinematics concepts was not achieved. These results show that introducing a concept superficially and only to be applied somewhere else later, results in low conceptual understanding. Based on these findings we suggest changes must be made, either to the learning goals, the structure of the curriculum and the instructional approach.

Third, the iSTEM group significantly outperformed the control group on the total test which resulted in a small effect size of $r = 0.21$. Furthermore,

the iSTEM cohort outperformed the control cohort for each concept category individually, and we found significant differences in favor of the iSTEM cohort for the following concepts: Displacement and distance, Velocity as 1D vector, Acceleration as rate, and Acceleration as 1D vector. The largest effect sizes per concept were found for the velocity and acceleration as a 1D vector, $r = 0.26$ and $d = 0.23$ respectively, meaning that the largest gains from the iSTEM approach in conceptual understanding were found in understanding the vector properties of velocity and acceleration. Although gains were realized through the iSTEM approach, overall accuracy was still sub par and further improvements are necessary.

Fourth, some observations stood out when analyzing the distractors. We found that large percentages of each cohort believe that the sign of the velocity is the same as the sign of the position, and that the sign of the acceleration is the same as the sign of the velocity. Furthermore, large subgroups of those believe that — in items with a graph — the sign of the velocity is determined by both the sign of the position and the sign of the slope of the $x(t)$ graph — which must be the same sign or zero. The conclusions are similar for acceleration: a large number of students believes that the sign of the acceleration is determined by the sign of the velocity and by the sign of the slope in a $v(t)$ graph. In most cases, there were few differences between the two cohorts. Slope/height confusions in $x(t)$ did occur slightly more often with STEM students compared to control students, and in $v(t)$graphs, the frequency of the slope/height confusion was very similar.

This study has shown that basic education on kinematics does not result in deep conceptual understanding and leaves many misconceptions unresolved. The results do show that integrated STEM education as implemented by the STEM@school project can somewhat improve students' conceptual understanding in kinematics with a small effect size. Despite these gains, we also find bad performance in the STEM group. From this, it is clear that more fundamental changes to the structure of the curriculum are necessary. The time spent on teaching these concepts should be used more appropriately. A restructuring to a dedicated section about kinematics before continuing with a different topic should be a first step. Furthermore, alternative approaches such as teaching kinematics using vectors from the start, i.e. teaching in 2D or 3D, while learning about vectors in mathematics in parallel is an approach worth investigating further.

Chapter 7

Conclusion

In Chapter 1, we highlighted two challenges in STEM education. The first one was the need to attract and retain more students in STEM programs to satisfy the current and future demand for STEM graduates in order to help solve the many challenges faced by our society. The second challenge was to further optimize learning in STEM disciplines and was motivated by an observed decrease in students' performance and by the difficulties faced by teachers to achieve deep understanding with their students, e.g. conceptual understanding of a specific topic or the linking of understanding between two fields. The dissertation presented here contributed indirectly to a solution for the first challenge through the collaboration on the construction of a framework for iSTEM, and through the design and implementation of iSTEM course modules as part of the broader STEM@school project which aimed to design, implement and study integrated STEM in secondary education in Flanders. Furthermore, this dissertation contributed to a solution for the second challenge by studying conceptual understanding in physics and mathematics, as well as the interplay between these two contexts. These studies added insight into students' difficulties and into their learning and reasoning processes, which is crucial in order to optimize learning. Specifically, we studied representational fluency in the context of linear relations in kinematics and mathematics, conceptual understanding of linear functions in kinematics and mathematics, and conceptual understanding of 1D kinematics. This chapter summarizes the main findings from the previous chapters in Section 7.1, describes implications for instruction in Section 7.2, offers critical reflection in Section 7.3, provides implications for future research in Section 7.4, and concludes this dissertation with some final thoughts.

7.1 Summary of the scientific conclusions

This dissertation encompasses three studies which are presented in Chapters 4, 5, and 6. In this section, we summarize the main scientific findings which provide an answer to our aims and research questions as formulated in Chapter 1:

- RQ1: How does grade 9 students' representational fluency between graphs, formulas and tables for linear function problems compare between kinematics and mathematics and between the STEM@school instructional approach and the regular instructional approach?
- RQ2: How does grade 9 students' performance and conceptual understanding of linear function problems compare between kinematics and mathematics and between the STEM@school instructional approach and the regular instructional approach?
- RQ3: How does grade 9 students' conceptual understanding of kinematics concepts compare in and between the STEM@school instructional approach and the regular instructional approach?

7.1.1 Representational fluency in physics and mathematics

To answer RQ1, we formulated subquestions in Chapter 4 which aimed to study the effect of the context, of the function type, and of their interactions on representational fluency, and also an additional section which aimed to study the effect of the iSTEM approach. A test was designed, validated, and used to assess representational fluency by means of a GEE analysis. We found that several factors had a significant effect on students' representational fluency. The first one was the transition between representations showing that those between graphs and tables in both directions are the easiest, and transitions including a formula are significantly more difficult, especially when translating to a formula. The second one was the function type which showed that negative signs significantly impacted students' performance in a negative way. The third one was the context, which showed that questions asked in a kinematics context, resulted in a decrease in performance. Significant interaction effects were also found, of which the most important one was the interaction between the representational transition and the function type. Additionally, we found that some transitions had an asymmetric difficulty, the most important of which concerned the translations between a formula and a table and between a formula and a graph for the directly proportional function type in mathematics. This

was unexpected since, although it concerned the most difficult representational transitions, it also concerned the easiest function type and the easiest context.

The distractor analysis identified three distinct errors: switching between concepts, sign switching of the slope or of both the slope and the y intercept, and switching of the direct proportional function type and the function type with positive y intercept and positive slope.

Lastly, this representational fluency test was used to compare the iSTEM approach with the regular instructional approach. The results showed no significant differences.

7.1.2 Conceptual understanding in physics and mathematics

RQ2 was divided into more specific questions in Chapter 5 targeting students' accuracy in kinematics and mathematics, difficulties with negative values for velocity and slope, the effect of the representation (graphs or formulas), the categorization of students' solution strategies, and the effect of the iSTEM approach on the results. To answer these questions, we designed a test, validated it and used it to study students' conceptual understanding of linear functions through GEE analysis and a qualitative analysis of students' responses. The results showed that the context, the question type, and the slope are all highly significant variables. Most difficulties were concentrated in kinematics, in algebraic expressions and in questions with a negative slope. Many interaction effects were also found to be significant, in particular those including the context.

The qualitative analysis resulted in a categorization scheme with a set of categories for solution strategies for the y intercept and another one for the slope. Results showed that students rarely change representations, except in mathematics items asking to assess the y intercept when given a graph. In that case, about one out of three students transitioned to a formula, thus supporting the idea that students often experience a pressure to use formulas due to a strong focus of education on the use thereof [96]. The most frequent error concerning the y intercept was the determination of the x intercept in mathematics which is likely due to the unfamiliarity with the formal terminology. In slope items in mathematics, students more often used the location of a coefficient in an equation to identify the slope, which occurred rarely in kinematics which showed that students had limited insight in the equation to describe uniform motion. Furthermore, our results showed very few slope/height confusions and iconic interpretations [107]. Interval/point confusions [107] though, occurred frequently. We confirmed their presence in graphical questions in physics, but found almost none in mathematics.

In questions with an algebraic formula we found a high occurrence of interval/point confusions and we further categorized them into three types: numerical, algebraic, and unit-based. The presence of interval/point confusions with algebraic formulas in kinematics and in mathematics, including the categorization into three types was — to our knowledge — not yet reported in literature. Results also showed that multiple errors and strategies often occurred simultaneously due to poor algebraic manipulation skills, poor understanding of units, or poor understanding of the Δ-symbol. The qualitative analysis confirmed the findings from the GEE analysis that there is a weak link between kinematics and mathematics. Lastly, we found considerable difficulties with negative sign for the slope and for the y intercept, and especially in kinematics where students were more inclined to use speed rather than velocity.

This test was also used to compare the implementation of the iSTEM approach with the regular teaching approach. The GEE analysis showed no significant effect. In the qualitative analysis for items asking to determine the velocity, we found that STEM students use the ratio of differences significantly more, and the ratio of coordinates significantly less compared to the control students; and that the latter reason with steepness significantly more frequently.

7.1.3 Conceptual understanding in 1D kinematics

In Chapter 6 we divided RQ3 into three subquestions. These concerned: the effects of a basic introduction based on a limited set of learning goals for kinematics on conceptual understanding, the comparison of students' performance in the iSTEM cohort and the control cohort, and an analysis of students' errors. To answer these questions, we modified the Kinematics Concept Test [35] and administered it in both cohorts. Results showed a very low mean accuracy in both cohorts and consequently that deep conceptual understanding was not achieved. When comparing the iSTEM cohort with the control cohort, we found that the iSTEM cohort significantly outperformed the control cohort with a small effect size of $r = 0.21$. More specifically, the iSTEM cohort significantly outperformed the control cohort for the categories "displacement and distance", "velocity as 1D vector", "acceleration as rate", and "acceleration as 1D vector". The largest gains — although still small — were found in the vector concept categories. Moreover, the study confirmed the occurrence of known misconceptions, of which the most prominent were: the belief that the sign of the velocity must be the same as the sign of the position [31,110,183] and that the sign of the acceleration must be the same as the sign of the velocity [31,183]; the belief that in graphical questions the sign of the velocity is determined by both the sign of the position and the sign of the slope of the $x(t)$ graph [31,183]; and the belief that the sign of the acceleration

is determined by the sign of the velocity and by the sign of the slope in a $v(t)$ graph [35].

The poor results from this study clearly show that the current curriculum does not promote conceptual understanding of kinematics concepts, and that the iSTEM approach was able to slightly improve the results but that further improvements are necessary.

7.2 Implications for instruction and policy

Although it has often been acknowledged that there is a gap between educational research and practice [198] the application of the studies reported here is rather straightforward.

The tests which we developed or modified have proven successful and can be used in education to assess the corresponding abilities. It is not necessarily the exact test which is the most relevant, but rather the test structure and the corresponding GEE analysis, distractor or error analysis, and categorization. As suggested in their respective chapters, many modifications can be made to study additional representations, contexts, function types etc. depending on the needs of the user. Specifically for the representational fluency test, we suggested an additional use case: to investigate which sequence of learning representation is best suited for a scaffolding approach to teach the various representations of a concept. This could provide valuable insights to optimize teaching strategies and to achieve a synergy of registers or at least to establish a base representation from which to expand to other representations at a later time.

The GEE results from the two tests on the topic of linear functions identified the important variables which influence students' performance. These results can be used in instruction to identify where the highest gains can be achieved and can thus be used to make improvements to instruction in a very detailed way.

Furthermore, the most frequent errors and strategies identified and quantified in the categorization scheme constructed in Chapter 5 provided valuable insights which can be directly targeted in practice. Although we have not studied ways to address incorrect strategies, we do recommend that teachers use questions such as the ones in the test to assess to what extent they are present with their students, and to direct more care and to perform additional exercises to this type of confusion when needed.

The (modified) kinematics concept test is more broadly applicable and can be used to assess students' conceptual understanding before, during, or after instruction as was intended by the authors of the original KCT. The most significant differences between the instructional approaches were found for the vector concepts velocity and acceleration. Though it is difficult to verify the precise cause for these results, we believe that the combination of the iSTEM framework and consequently the application of the vector concepts in real-world situations have contributed to those improvements. For these reasons we recommend the iSTEM approach for use in education and also want to suggest that introducing kinematic quantities as vectors in 1D or 2D from the start might be a useful approach in education, which merits further investigation.

In general, after implementing the iSTEM approach we have shown that students can achieve the same or a higher level of performance compared to students in the control cohort. This conclusion is valid only for the specific topics which we studied but can be considered indicative for other topics. Our very detailed findings from cognitive tests in combination with the findings from the other studies performed in the STEM@school project indicate that the iSTEM framework can be a successful alternative to current educational practices which — in addition to the learning goals — also strive to add to them.

In addition to the detailed implications described above, our experiences in the STEM@school project allow us to make more general suggestions for policy and implementation of STEM education. Specifically we suggest:

- To align the learning content from the STEM disciplines where possible in such a way that similar concepts are introduced and explicitly linked in all relevant STEM disciplines around the same time in the curriculum. This suggestion serves to make the most from the links which naturally exist between the various disciplines.
- To construct overarching STEM learning goals which connect those from the individual STEM disciplines and stimulate a direct integration of the learning content. With this suggestion, the integration becomes an explicit and an achievable goal for both teachers and students.
- To pursue integration of as many disciplines as possible, also outside of STEM, while respecting the internal logic of an individual discipline and while keeping the quality of learning within each discipline high. This suggestion must ensure that the individual disciplines benefit from the integration, or are at least not negatively affected by it. This does not mean that each discipline should be integrated in each learning activity, the aim should be to integrate whenever it makes sense to do so.

- To not shy away from introducing more difficult concepts in an iSTEM project but to employ the real-world challenge and the integration to facilitate the learning thereof. This suggestion is made because a concept which is traditionally considered difficult within the structure of a discipline, might be more easily understandable within the context of an integrated STEM project. Or in other words: teach students to "play the whole game" [64].

Considering our results in relation to educational policy, we strongly suggest to restructure the curriculum in such a way that each block of learning content becomes as relevant and as integrated as possible within its discipline and in connection with other disciplines. In doing so, it is crucial to maintain dedicated time and an in-depth focus for achieving conceptual understanding of the topic under study.

7.3 Critical reflection and discussion

Our studies have some limitations, which should be considered and can also stimulate future research.

The first ones are the geographical and educational characteristics of our cohorts. All the data was gathered from students in Flanders who are enrolled in the Flemish educational system. The detailed results could therefore differ significantly with other regions and educational systems. Nonetheless, we do expect that the larger trends in the data will also be present elsewhere, e.g. the most difficult representational transitions or the most typically made errors when determining the slope. In addition, our cohorts were limited to students who had opted to enroll in educational programs focused on science and mathematics, since we deliberately focused on the "leaky STEM pipeline". Therefore our research does not directly contribute to the need for a basic STEM literacy for all, but could serve as a starting point for future research since abilities such as conceptual understanding and representational fluency are useful for all learners, not just for future STEM-professionals.

Directly related to the reflection above are the characteristics of the students in relation to the iSTEM educational approach. In the study about representational fluency we discussed that the use of MERs can positively or negatively impact students' understanding. Positively because it can facilitate the construction of a coherent framework of understanding, and negatively because for some students the cognitive load of using MERs is too high which impairs their learning. This effect is generally true for other

instructional practices as well, as the effect thereof always depends on the students' characteristics. Consequently, this is also true for each of the five key principles of the iSTEM approach. This implies that, although we aim for the iSTEM approach to improve students' learning, it might have an adverse effect on some students. Such possible negative effects were not taken into account. What we do know is that the target population in our studies consists of students who chose to study sciences and are thus more likely to have a higher level of abstraction abilities. Although only speculative, this might relate positively with the five key principles from the iSTEM approach.

Concerning the five key principles, it is important to stress that in each of the studies every effect observed between the two cohorts can be attributed to the combination of all five key principles. Although the integration was highlighted most often, it is the full approach which was evaluated because the effects of the individual principles can not be untangled and were not studied separately. Furthermore, other variables were either comparable or randomized thus making the attribution to the different approach valid.

Another critical reflection is that it takes time to truly change the mindset of both teachers and students, meaning that the true effects are best measured after several implementations [172]. Although the second iteration of the iSTEM approach — which already had some revisions to the course materials — is the one after we collected our data, one would ideally perform a new assessment after the fifth or tenth implementation cycle to study the long-term effects.

Next, one of the most difficult aspects in our studies — and in most educational studies — is the assessment of the implementation fidelity. This is the case for both the iSTEM cohort and the control cohort. In both cohorts, the teachers adhere to the learning goals and to the guidelines for teaching, but every teacher does so in his or her own way. Some aspects might get stressed more while others might get stressed less. This is in addition to their freedom to personalize their own teaching. This is especially the case for the iSTEM teachers because due to the ill-defined and open-ended nature of the iSTEM challenges, each school has in effect a different implementation based on the same course materials and solution paths. The project team from STEM@school followed up on the implementation as much as practically possible through various means which were described in section 3.2.3. This is imperfect and partly based on self reported information from the teachers, but did provide us with sufficient confidence to assess the implementation and include the schools with the best implementation in the comparisons between the two approaches. Still, this remains an aspect which could be improved. A possible suggestion for future research is to use less schools and monitor them more closely. The downside in that case would be a smaller population to analyze.

Related to this, is the issue of time on task when comparing the control cohort with the iSTEM cohort. At first sight, it might appear that the iSTEM students spent more time with e.g. linear functions than the students in the control cohort because the STEM students had more dedicated STEM hours in total each week. This is however not the case since during the STEM hours those students also learned additional learning content and skills which are not part of the regular curriculum. The amount of actual time spent on the same learning content therefore does not vary more between the two cohorts as it does between schools within cohorts. This assessment was made from the same observations as the ones for the implementation fidelity, and can thus be improved upon, possibly through e.g. teacher diaries in which the teachers take note of their students' activities.

In each of the three research studies students' intuition was included at some point as a possible — though in this dissertation unproven — explanation for the observed results. In representational fluency we suggested that students might find the "less than" symbol more intuitive than the "greater than" symbol as a possible explanation for the discrepancy between their performance on algebraic questions with negative signs. In the study on students' understanding and strategies of solving linear functions, we suggested that the "initial position" is more intuitive than the "function value of 0" as an explaining factor based on the work of Davis [109]. In the study on conceptual understanding in kinematics, we suggested that positive acceleration — compared to negative acceleration — is likely more intuitive for students, and that average velocity is intuitively seen as the arithmetic average. Considering the cases listed above, it is clear that intuition might positively effect students' performance, e.g. reasoning about a positive acceleration although it was not studied in depth. In contrast, intuition might negatively effect students' performance, e.g. using the arithmetic average for the average of two velocities. The positive or negative effect of intuition can be an important factor when explaining results but it is difficult to detect. Furthermore, a positive effect might be mistaken for true understanding and a negative one might be mistaken for a preconception or a misconception. The role of intuition is therefore not easy to identify when interpreting the results and might bias the conclusions concerning conceptual understanding.

A final critical reflection on our work concerns the validation of the kinematics concept test. The original KCT was thoroughly validated in Swiss German [35]. After translation to Dutch and a reduction of the number of items, the test should ideally have been validated again. A full validation was not performed though, but four collaborators in or close to the STEM@school project and with an academic background in physics or mathematics did assess the modified KCT items. In addition, and as noted in the Chapter 6, a paper-and-pencil

pilot test was performed the year prior to the final administration. From these results it was clear that some minor modifications had to be made and that — although students had understood the questions — they experienced a lot of difficulty to answer them correctly. This high difficulty makes it less ideal to compare groups, but the test was mainly intended as a criterion-referenced test, which serves to assess how well a student performs on predefined criteria, which in this case is conceptual understanding in kinematics. The alternative was to make the test items easier on average while maintaining their ability to assess conceptual understanding, which is unlikely to be possible since the items had been ratified in multiple previous studies — usually with an older student population — and most importantly because the distractors are known misconceptions.

7.4 Implications for future research

Since we have focused on a few detailed topics of interest and on the effects of the instructional approach on students' abilities concerning those topics, there is plenty of room for future research. In general, a dissertation ends with more questions than it started with. This one is not any different. In this section, we will first discuss the opportunities concerning the studies reported on in this dissertation and afterwards formulate more general suggestions.

The results from our three studies are valuable and contribute to our insight in the matter, but to obtain a deeper understanding of our observations, we propose that future research focuses on student interviews. Methods such as think-aloud interviews could provide more insight in the thinking process of students when solving the tests in our studies. This is an important part of truly understanding why students make errors and how they find the correct results. The answers provided by students are often only the final result of their thought process. This means that valuable information for researchers and teachers is lost. The open-ended nature of our questions in the conceptual understanding of linear functions test in combination with the request to explain the answer did provide a level of insight to a certain degree, but still does not reflect the full thought process.

Specifically for the representational fluency, it would be of interest to know how students solve such questions. As explained in that study, it is possible that in a table to formula questions, a student actually translates the formulas to the table and not the other way around. It is also possible that an intermediate "base representation" is used and that a student first translates the table to a graph and then to a formula which is then compared to the formulas in

the answer options. Once again, interviews, and in particular think-aloud interviews could provide valuable information in such cases.

An aspect which we touched upon, but did not include in our research design is "transfer", specifically the transfer of knowledge from physics to mathematics or from mathematics to physics. We did use our data from the conceptual understanding of linear functions test and calculated the corresponding transfer index [219] for both transfer directions. These results showed that there is a certain ambiguity to this measure of transfer. It highlights how difficult it is to truly assess transfer of knowledge and that more work needs to be done. Our suggestions for these topics are to perform an assessment including think-aloud interviews under very strict conditions. Important factors to verify first are that the knowledge must be present in the first context and not in the second context before learning that knowledge in the second context. After the assessment, the researcher must assess whether the knowledge is or is not present in the second context. Ideally, also the knowledge in the first context is assessed again since this might also have changed.

The poor results of the kinematics concept test for both cohorts was a notable finding. Although the direct cause could not be identified with our data, the diversity of schools, gender, teachers, and the medium effect size of the iSTEM instructional approach versus the regular instructional approach, indicated that any of these was unlikely the main cause. We highly suspected that the limited introduction on 1D kinematics is an important factor. The cause for such an introduction is the limited set of mandatory learning goals, which are mainly in the curriculum because the concept velocity is needed in optics and dynamics thereafter. To address this, we suggested that a better approach should be pursued in the educational system. The small positive effect size from the iSTEM approach provided an indication of the direction in which useful changes might be found. The main difference is that the students in iSTEM also received limited education on acceleration and that they applied the physical quantities (position, velocity and acceleration) in an integrated approach following the five key principles from STEM@school which aimed to make these variables and their vector aspects more relevant by requiring them to be introduced to solve the problem. Our results showed that the iSTEM students achieved the most significant positive effects on precisely those categories which describe the vector properties of velocity and acceleration. Thus, these results can provide a small indication that additional benefits can be found in such an approach. This is in contrast with the regular 1D approach in which the focus is on position and speed, in which the vector properties are not stressed as much, and in which they are not applied in integrated challenges. Extrapolating from these insights, it might prove useful to teach kinematics using 1D vectors from the start. Or perhaps even start with 2D vectors to stress the vector aspects even

more. The possible benefits of such an approach would be a very interesting topic for future research, especially in combination with the parallel dilemma of introducing vectors in mathematics and the fine tuning of the instruction thereof between physics and mathematics.

A recurring topic in our studies was the influence of language on the understanding of very well defined scientific concepts. This is also related to the vector properties discussed in the previous paragraph. The particular example in this dissertation is the difficulty of expressing the difference between velocity and speed in Dutch. Additionally, we also encountered difficulties with the use of formal and more informal terminology in case of the y intercept and the initial position. These issues suggest that the interplay and the influence of a specific language and the specific wording used in classrooms are an important factor in learning, which warrants further research efforts.

As a final suggestion for future research related to our detailed topics, and in line with our first point in the critical reflection, we highly recommend to perform replication studies with different populations and educational systems. Additionally, our collected data is very rich in information and has most likely not yet been exhausted, hence further analysis could result in additional findings. E.g. we have mainly looked at group level comparisons, but comparisons at the student level could also contain interesting results.

From our experience with designing, implementing and studying integrated STEM, we can also formulate some broader areas of interest for future research related to STEM:

- Is it better to integrate known learning content or new learning content? In the STEM@school project we focused on the integration of as much new learning content as possible. Furthermore, the introduction of this new learning content was driven by the necessity to do so, because without it, the challenge of course module could not be solved. In contrast, many STEM projects focus on integrating known learning content and are only applying existing knowledge. This is an important distinction and future research could help to determine which approach produces the most desirable results.

- Does integrated STEM typically benefit specific disciplines or pairs of disciplines (e.g. physics and mathematics), and if so, which ones? Such insights are important because an essential boundary condition of integrated STEM should be that all individual disciplines must benefit from it, or at least do not experience any negative effects. Whereas integration between physics and mathematics is rather obvious, different combinations of (sub)disciplines should be studied as well.

- The balance between discipline specific learning time and integrated learning time. Integration is one of the important key principles of the iSTEM approach, yet we also strongly feel that each individual discipline should continue to receive dedicated teaching time in any iSTEM curriculum to address the discipline specific intricacies and to construct the internal structure of each discipline. But what is the right balance?

- Research into other aspects of students' learning in iSTEM education, such as procedural knowledge and metacognition. These aspects might differ significantly — and hopefully in a positive way — in the context of an integrated STEM project.

- As was referenced in the introduction, there is still a great need for high quality teaching and learning materials for iSTEM programs. Developers of such materials would welcome research-based guidelines to format them. As explained in Chapter 3, every STEM@school course module was composed by several themed submodules of which one was the main submodule which challenges the students. The other submodules can be consulted on an as needed basis, but alternative formats might prove equally or more useful. Furthermore, the most suitable format might be age dependent. In addition to the format, it is important to develop a collection of suitable challenges or to continuously develop new ones. In doing so, developers and teachers should make sure that the projects are not implemented the same way every year because sequential iterations are likely to result in smaller and smaller levels of improvement and also because the teachers might become too familiar with the course module. The latter might seem strange, but it is not only important to keep up the challenge for the students, but also for the teachers to keep them at the top of their game. As a general rule, it seems useful to switch to a different challenge once the current one tends to result in highly similar solutions and when the module has become too "standard" for the teacher.

- The target group in the STEM@school project were students in the 9th, 10th, 11th and 12th grade who were enrolled in an educational program with a strong focus on science and mathematics. But as discussed in Chapter 1, there is a need for STEM literacy for all students. Therefore, it would be useful to research ways in which all students can achieve at least a basic STEM literacy. A possible way to implement this could be to have students from different educational programs work together on a larger challenge in which each addresses the fields which they opted to study. This way the students can focus on their own specialty and interest while still connecting with the STEM disciplines, experiencing

them on a basic level, and learning how they are needed to solve certain aspects of the challenge.

- The long term effects of integrated STEM education on students, on teachers, on schools, on the number of STEM professionals in the job market, and on their performance and well-being in their employment.
- More research on integrated STEM is needed, but the next evolutionary step is already here: STEAM, which includes "A" for the Arts (visual arts, music, languages, ...). The integration of the Arts component has the potential to connect almost all disciplines and could extend the potential benefits from STEM to the whole curriculum.

7.5 Final thoughts

In this dissertation, we have studied various aspects of the integrated STEM instructional approach and have contributed to discipline specific educational research for physics and mathematics. These efforts have increased our insights in ways to further improve education and will hopefully also result in more and better STEM graduates to help society tackle its many challenges.

Perhaps this is the point where this dissertation truly ends and we can come full circle by asking ourselves why we are doing all of this in the first place. The answer, it seems, is twofold. First, humans have a thirst for knowledge. We want to know. We want to understand. It is the natural curiosity for the unknown that drives us to know more. And second, because we all have the desire to create a better future. The arrow of time points to the future and we all want it to be better than the past. We want to advance, and we will not settle for less than what we already have. To do so, it is important to realize our place in our habitat and also our influence on it. To obtain a bright future for ourselves and for the next generations, it is paramount to understand that we are borrowing our habitat from them, and hence should put appropriate effort in taking care of it. Science can help us understand and can help us fix what we have broken, so we can guarantee that bright future for all.

With this, I write my last words in this dissertation. I will dedicate the very last ones to a final quote from a STEM enthusiast, a physicist, an entrepreneur, an explorer, a humanitarian, an *uomo universale* of our time.

"It's important to teach problem solving or teach to the problem, not to the tools. So, this would be like... let's say, you are trying to teach people about how engines work. A more traditional approach would be to say: we're going to teach them all about screwdrivers and wrenches, and you can have a course on screwdrivers and a course on wrenches and all these things. And it is a very difficult way to do it. A much better way would be like: here's the engine, now let's take it apart. How are we going to take it apart? Oh, you need a screwdriver — that's what the screwdriver is for. You need a wrench — that's what the wrench is for... And then, a very important thing happens. Which is that the relevance of the tools becomes apparent."

— Elon Musk.

Appendix A

Appendix

A.1 Kinematics concepts

Grade 9 students in science oriented programs in Flemish education receive instruction on basic kinematics concepts. This appendix serves to describe the relevant ones for this dissertation. Although it is not part of the mandatory curriculum, we do highlight the detailed vector aspects of some of these quantities because many student difficulties are directly related to them.

Distance The distance d is a scalar quantity which equals the length of the chosen path between two points, i.e. it is a measure of the distance covered *along* that path between those two points. For the same two points, a different path can have a different length, thus the distance along each path can be different. In a flat plane, the shortest distance between two points is a straight line, any other path connecting those points results in a longer distance between these two points. The distance is always a positive quantity.

Displacement Our grade 9 students learn that the displacement Δx along a straight line between two points x_1 and x_2 is a measure for the change in position, i.e. the difference between x_2 and x_1: $\Delta x = x_2 - x_1$, and that it can be negative. What they do not learn though, is the vector aspect of the displacement. In three dimensions, it is a vector quantity $\Delta\bar{r} = \bar{r}_2 - \bar{r}_1$ with starting point $\bar{r}_1$ and end point $\bar{r}_2$. Its magnitude equals the shortest distance between the two points, and its scalar components can be negative. In the 9th grade, our students' education is limited to

the scalar component of the 1D displacement. The difference with the distance d is often made clear by a 1D motion of a person who returns at some point and covers a particular section twice. Fig. A.1 illustrates the difference between distance and displacement.

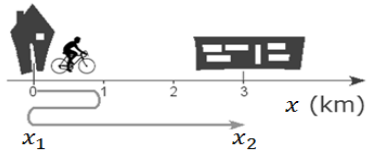


Figure A.1: Typical illustration to clarify the difference between distance d and displacement Δx in 1D. This cyclist has covered a distance of 5 km and has a displacement of 3 km between points x_1 and x_2.

Velocity In general, the velocity $\bar{v}$ is a vector quantity and is a measure for the rate of change of the position in a given reference frame with respect to the time, i.e. $\bar{v} = \frac{d\bar{r}(t)}{dt}$, which is more accurately known as the instantaneous velocity. The speed is the magnitude of the velocity, and is always positive. Our grade 9 students have not studied derivatives yet so they start with learning the "average velocity in 1D" as: $v_{avg} = \frac{\Delta x}{\Delta t}$, where Δx is the displacement and Δt the time interval. This value can be negative because Δx can be negative. What they actually learn though, is the scalar component $v_{avg,x}$ of the average velocity $\bar{v}_{avg}$. But they do not realize this since their vector knowledge is very limited or non existent. They do not yet understand what a scalar component, a vector component or a unit vector is. This is made even more confusing because of a linguistic issue: there is no distinction between speed and velocity in Dutch, they are both translated as "snelheid". In English, it is possible to distinguish between the average speed $\frac{d}{\Delta t}$ in which d is the distance and is always positive, and the (scalar component of the) average velocity $\frac{\Delta x}{\Delta t}$ with Δx the displacement which can be negative. This distinction is not so easily made in Dutch and is often ignored in textbooks. A consistent set of 1D definitions in Dutch would be:

- the average snelheid: $\bar{v}_{avg} = v_{avg,x}\bar{e}_x = \frac{\Delta \bar{x}}{\Delta t} = \frac{\Delta x}{\Delta t}\bar{e}_x$,
- the magnitude of the average snelheid: $|\bar{v}_{avg}| = \frac{|\Delta \bar{x}|}{\Delta t} = \frac{|\Delta x|}{\Delta t} \geq 0$,
- the average magnitude of the snelheid: $|\bar{v}|_{avg} = \frac{d}{\Delta t} \geq 0$,
- the scalar component of the average snelheid: $v_{avg,x} = \frac{\Delta x}{\Delta t} \in \mathbb{R}$.

Instead, only one terminology is used: "the average snelheid" and it is interpreted as $v_{avg} = \frac{\Delta x}{\Delta t}$. I will continue to use this flawed terminology

because this is the way students use it, but it should be corrected in education nonetheless. In an $x(t)$-graph, the average velocity can be determined by calculating the slope. From a $v(t)$-graph, the displacement can be derived by calculating the oriented area under the curve, and by calculating the non oriented area under the curve, the distance can be found.

Acceleration In general, the acceleration $\bar{a}$ is a vector quantity and is a measure for the rate of change of the velocity in a given reference frame with respect to the time, i.e. $\bar{a}_{avg} = \frac{d\bar{v}(t)}{dt}$, which is more accurately known as the instantaneous acceleration. In line with the average velocity, the "average acceleration in 1D" is defined as $a_{avg} = \frac{\Delta v}{\Delta t}$ which again actually means the scalar component $a_{avg,x}$ of the average acceleration $a_{avg,x}\bar{e}_x$ in 1D. Learning about acceleration is not mandatory, but some teachers might briefly touch this concept. The STEM@school curriculum does include understanding of acceleration as a rate, and acceleration as a vector as supplemental learning goals, though these are not treated in depth. In $a(t)$-graphs, the average acceleration can be determined by calculating the slope.

A.2 Focus on misconceptions in the museum security system module

In the museum security system module described in Chapter 3, particular attention was given to tackling misconceptions in geometrical optics. In this appendix we describe and illustrate how we included insights from literature and how we contributed to it. This work has been published as a peer-reviewed publication:

S. Ceuppens, J. Deprez, W. Dehaene, and M. De Cock, "Tackling misconceptions in geometrical optics," *Physics Education*, vol. 53, no. 4, p. 045020, 2018

The work reported on here was developed by the first author in close cooperation with the last author. The second and third author supervised the process and reviewed the work. Except for the abstract, the paper is reproduced here in full with minor textual enhancements.

A.2.1 Introduction

Optical phenomena and optics-based technology are everywhere and their applications are only increasing. Therefore, it is not surprising to find

geometrical optics in most science curricula and a topic of interest in physics education research (PER). Studies from PER have repeatedly shown that students have difficulties to understand concepts underlying optical phenomena both pre-instruction and post-instruction [34, 234–243]. The preconceptions students construct before entering formal optics education are so entrenched in their reasoning that classical optics teaching does not always allow to change their ideas to more scientific ones. In order to have students change their ideas, they should be mentally activated in targeted student activities (e.g. [244]). The Flemish STEM@school project aims to develop innovative learning materials on integrated STEM (Science, Technology, Engineering and Mathematics). One of the underlying principles in the design of these materials is taking into account findings from Discipline Based Educational Research. In this paper, we report how findings on particular student difficulties in optics informed the design of student activities to tackle misconceptions, using an inquiry based and cooperative learning approach.

A.2.2 Student difficulties in geometrical optics

A.2.2.1 Image formation by a plane mirror

Many students have difficulty explaining the image location of an object placed in front of a plane mirror. They state the image is formed on the mirror's surface or use a "line-of-sight" argument to incorrectly state that the image of an object is located at the intersection of the mirror plane and the line-of-sight connecting the observer and the object, or at the point symmetrically behind the mirror but still on the line-of-sight [234, 235], as is illustrated in Figure A.2.

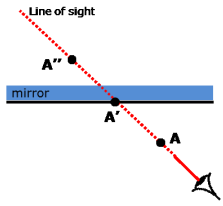


Figure A.2: Visual representation of the line-of-sight argument where an observer incorrectly believes that the image of object A is located at the symmetrical point A" along the line-of-sight or at the intersection of the line-of-sight and the mirror plane.

Also, predicting whether the image of an object is visible from a given point of view proves to be difficult for both students and adults [236]. This relates to

the effect of the point of view and the size of the mirror on the field of view. Many people incorrectly believe that plane mirrors have "zooming capabilities": by moving closer you see less of yourself and by moving further away, you see more [234, 235, 242]. Additionally, Bianchi and Savardi have shown that about half of their adult respondents are strongly biased towards thinking that the field of view of a mirror is the orthogonal space in front or beyond it for eccentric viewpoints [236].

A.2.2.2 Image formation by a convex lens

Goldberg and McDermott [34] reported significant student difficulties with real image formation by a converging lens when students were presented with a simple setup of a light bulb, a converging lens and a screen and asked what would happen if half of the lens were covered. Before instruction, almost all respondents replied that half of the image would vanish. After instruction, this incorrect idea persisted with 55% to 75% of the students, depending on the instruction received.

A.2.2.3 Indirect and direct observation of a real image formed by a convex lens

Goldberg and McDermott [34] also reported that many students — when given a simple setup of a luminous object and a converging lens — were unable to identify the full conditions to observe the image. Students think that a screen is necessary to see the image — i.e. indirect observation — or that the observer's eye should be at the location of the screen, effectively replacing the screen — i.e. direct observation. Most students did not know that the image could be observed directly by looking from anywhere in a well-defined space along the optical axis behind the lens.

A.2.2.4 Ray diagrams

Constructing a ray diagram is a powerful geometrical algorithm to depict phenomena that can be explained by geometrical optics principles in which a light ray is the key concept allowing to explain the rectilinear propagation of light and image formation. However, research has shown that students do not always recognize these rays as useful geometrical tools but tend to see them as physical entities [34, 245, 246]. A particular problem is that the rays/lines in a ray diagram do not always have a matching equivalent in reality which can create confusion about the interpretation of a diagram and its relation with

what happens in reality. In a diagram with a plane mirror for instance, the rays can be extended behind the mirror to determine the image location, but there are no rays behind the mirror in reality since it is a virtual image (see Section A.2.3.1). Also the principal rays used to determine the location of an image formed by a converging lens are a geometrical tool since even if the matching ray in reality were blocked, the principal ray in the diagram can still be used to determine the image location (see Section A.2.3.2). Similarly with a diverging lens, the extensions of the rays help to determine the image location. To distinguish between the rays that exist in reality and those that do not have a matching equivalent, the former are drawn as full lines and the latter are drawn as dashed lines. This distinction is important and must be explicitly made by the teacher, ideally also illustrated with an experimental setup to show the presence and absence of light rays in which a laser can serve as a light ray. In addition, an insightful ray diagram contains only a sufficient number of well chosen rays and students tend to forget that the infinite number of other possible rays also contribute to the image formation. Lastly, the conditions to observe an image are often disconnected from ray diagrams. Students usually understand that a screen at the image location allows for indirect observation, but they often fail to understand how the image can be observed directly. These issues show that ray diagrams can cause student difficulties, especially when translating a real system into a ray diagram and vice versa. However, when used appropriately, ray diagrams offer opportunities to elicit conceptual conflicts and to resolve misconceptions by applying them consistently in carefully designed exercises.

A.2.3 Student activities to tackle misconceptions

To tackle the described difficulties, we designed student activities using the "elicit, confront and resolve" approach [244] combined with ideas from Conceptual Understanding Program (CUP) [53], which are direct applications of inquiry and cooperative learning. We focus particularly on:

- The use of ray diagrams: in all exercises we scaffold students to use ray diagrams.
- The use of observer's eyes, which is twofold: Firstly, they indicate the locations (along a line) from which a particular image point can be observed. Secondly, in diagrams containing lenses it makes clear that light rays don't end in the image point but continue until they are absorbed by an object or an eye. In our diagrams we always draw a single eye for each single ray. It is important to note that because of this, students might think that a single ray is sufficient to determine the location of the image

point, which it is not. It is a very different task to determine the location of an image point or to determine the location from which an observation can be made. The former requires a minimum of two rays in the case of image formation by a lens and when using the physical method with plane mirrors (discussed in Section A.2.3.1) and the latter only requires a single one — using backwards selection of a path of light [245] — after the image point has been determined.

- The implementation of activating student activities: First, students learn the basics of ray diagrams; afterwards, they receive provocative exercises which confront them with apparently conflicting situations that challenge their understanding. In these exercises, students first solve a problem individually, then discuss their solutions in small groups and finally discuss with the whole class. When consensus is reached the teacher can confirm the answer or indicate potential conflicts in the students' reasoning and guide them to the correct answer through inquiry. Students are actively engaged and responsible for their own learning with minimal guidance from the teacher and they learn with and from their peers. By applying the same concepts and reasoning, reflecting on them and then generalizing them in a variety of such exercises, students can construct a coherent conceptual framework applicable across all contexts instead of memorizing specific solutions to specific problems and ending up with an incoherent conceptual framework [244].

To support these activities, teachers are provided with a list of tips and possible pitfalls, e.g.: draw arrows on the rays, draw the reflected or refracted rays through until they reach an observer's eye, draw regular rays in addition to principal rays, etc., which they can then make explicit to students. In the following sections we describe selected examples in which we apply this approach.

A.2.3.1 Image formation by a plane mirror

To determine the image formed by a plane mirror, students are taught two methods. Figure A.3 (a) shows the "*mathematical method*" in which symmetric distances are measured with respect to the mirror's plane. This does not provide much insight in the phenomenon of reflection though, hence students are also taught the "*physical method*": using ray diagrams and the law of reflection as shown in Figure A.3 (b). Students are instructed to primarily use the latter. Note that a grid is absent in Figure A.3, since the presence of a grid might stimulate the use of the mathematical method and the absence of a grid might stimulate the use of physical method. Additionally, in Figure A.3 (b),

each observer eye receives only a single ray since the situation depicted here is that of a point object. Teachers should be aware that students might interpret this in such a way that they conclude that only a single ray is necessary to determine the image point, which is incorrect since at least two are necessary in this method.

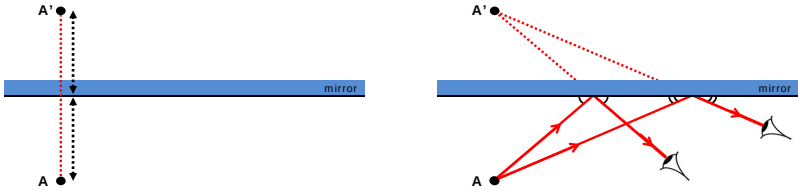


Figure A.3: (a) Mathematical method: using equal distance with respect to the mirror plane to find image point A'. (b) Physical method: using the law of reflection through a ray diagram with equal angle of incidence and angle of reflection for two randomly chosen rays starting from object point A and ending in "observer eyes". The intersection of the extended reflection rays identifies image point A'.

Students then work on problems evoking the earlier mentioned difficulties, using the described approach.

- In Figure A.4 (a) and A.4 (b) students are asked to determine the image of objects which are not completely in the orthogonal space in front of the mirror or for which the mirror is "too small". Using ray diagrams can help students understand that also the points outside the orthogonal space have an image and that determining the image of an object is not the same as determining the field of view of the mirror.
- The problem in Figure A.5 confronts students with the "line-of-sight" argument [34, 234, 247] by locating the image point for two different observers. By explicitly asking to use ray diagrams, they realize the image is independent of the observers' locations.
- In the "reverse engineering problems" (Figures A.6 (a) and A.6 (b)) students must apply known principles to non-standard situations: in these examples using the physical method in reverse to find the mirror's plane instead of the image. These two examples clarify the misconception that light rays are not emitted from the eyes in order to see [243], but that starting from the eyes can be a useful solution strategy. The exercise in Figure A.6 (b) includes a catch — there is no realistic solution because one of the observers would be behind the mirror — which requires students

to reflect on their solution and interpret the ray diagram back to the real world situation.

- The exercises in Figures A.7 (a) and A.7 (b) show two examples which can help tackle the idea of a mirror's "zooming capacities" and help to correctly determine the minimum mirror size to see the full image of a person in a mirror. Students are provided with multiple versions of this exercise in which e.g. the distance between the person and the mirror varies. Additionally, we encourage to use an interactive applet from the Wolfram Demonstrations Project [248] or similar, which allows students to dynamically study the effect of mirror size and the distance to the mirror. Also, students carry out a simple experiment in which they use a plane mirror and a marker and indicate which part of their face they can see in the mirror and compare this for different object distances.

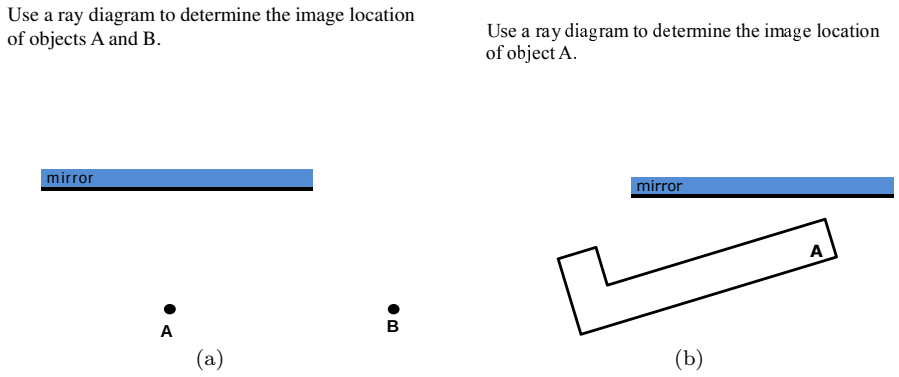


Figure A.4: Ray diagram exercise on reflection at plane mirrors, abbreviated instructions are shown. Elicited issues: (a), (b) Some points are outside the perpendicular space of the mirror, the mirror is "too small".

A.2.3.2 Image formation by a convex lens

To tackle the problem with the partially covered lens, Goldberg and McDermott [34] advise teacher demonstrations to clarify the effect of covering parts of the lens. Based on their work, we developed a more active tutorial in which students cooperate in groups of four. Students start with a typical situation with an object and a converging lens as shown in Figure A.8.

- Students individually determine the image location. Each must do this differently: one has to use all principal rays, the other three each have to

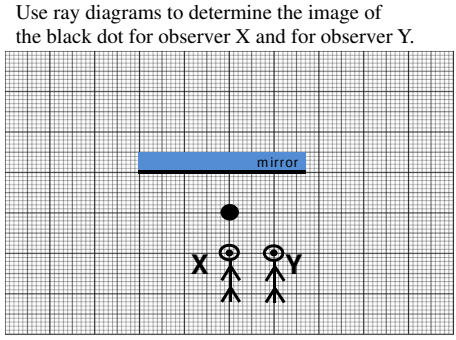


Figure A.5: Ray diagram exercise on reflection at plane mirrors, abbreviated instructions are shown. Elicited issues: Line-of-sight argument, the image is observer dependent.

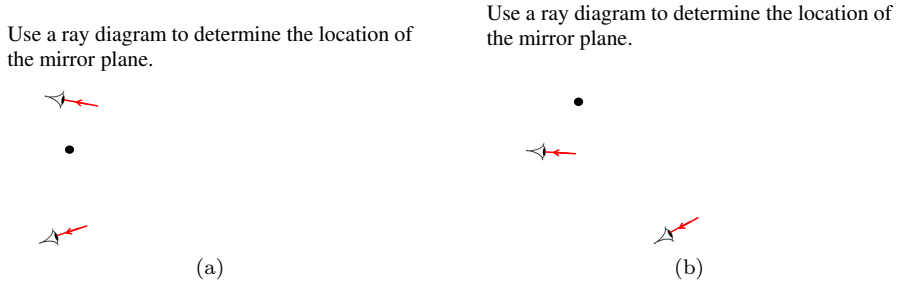


Figure A.6: Ray diagram exercise on reflection at plane mirrors, abbreviated instructions are shown. Elicited issues: (a), (b) Reverse application of known principles, light must be emitted by the eyes to see an object/image.

use a different pair of principal rays. They must also draw at least one "regular" ray in a different color.

- Next, they draw a predetermined opaque block on their diagram — which "blocks" the principal ray through F_2 — and predict the effect on the image using a ray diagram. The situation in Figure A.9 (a) shows what a student using the principal rays through F_1 and O is likely to find, which is unlikely to cause any concern since no rays are "blocked". Figure A.9 (b) shows a possible situation of a modified diagram of a student using the principal rays going through F_1 and F_2 and intuitively — and incorrectly — believes that a partial image will be formed and tries to justify this by drawing the rays starting from the highest non-blocked object point. Figure A.9 (c) shows a correct diagram in which

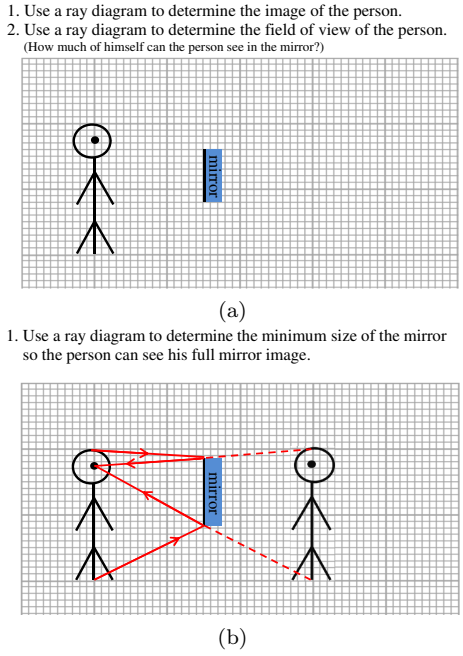


Figure A.7: Ray diagram exercise on reflection at plane mirrors, abbreviated instructions are shown. Elicited issues: zooming capability of a mirror, dependency of object distance and mirror size; (a) Exercise in which the mirror is not large enough to see the full image; (b) Ray diagram to determine the minimum required mirror size to see the full image.

the principal rays through O and F_2 are used and in which the latter one can still be used to determine the location despite the dashed part of the ray not existing in reality, i.e. the dashed part does not match a ray which exists in reality but is simply a geometrical tool to determine the image location. This diagram shows that the image is fully formed but still doesn't illustrate the entire effect on the image. Figure A.9 (d) shows a similar diagram as in Figure A.9 (c) and additionally illustrates that a section of the light cone is blocked which decreases the amount of light contributing to the brightness of the image point, thus resulting in a fully formed but less bright image. This diagram — which students are unlikely to come up with themselves — is the one which should be understood at the end of this exercise, since it shows that the principal ray through F_2 can still be used and explains the effect on the image V'.

- Students compare approaches and discuss till consensus. Each group

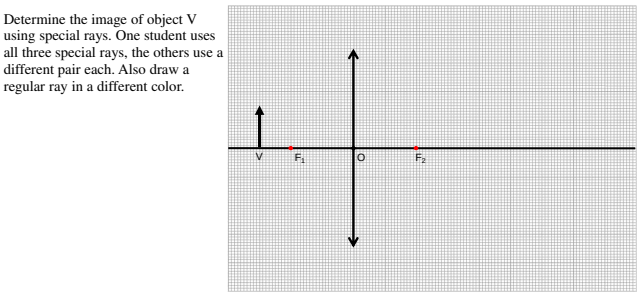


Figure A.8: Standard situation with object (V) and a convex lens with focal points F_1 and F_2 for which students must determine the image of V.

draws its final diagram and compares with other groups. After a class consensus the teacher reveals the correct answer and discusses the incorrect ones. Throughout the tutorial we encourage to use of real optical artifacts (a luminous object, a converging lens and a screen) to better link the ray diagrams with the real world system.

- Besides "blocked" principal rays, other issues can challenge students' understanding, e.g. not passing through the drawn lens. In Figure A.10 (a) the lens seems to be "too small" for the object and two principal rays will not intersect the lens.

- The exercise in Figure A.10 (b) relates to the completeness of the image. It shows an asymmetric or partial lens which might confuse students since regular exercises have lenses with their optical center in their geometrical center. This aims to clarify that even asymmetric/partial lenses produce a full image if there are rays from all object points reaching a part of the lens and that the image can always be determined through consistent application of ray diagrams.

- Figure A.11 shows a "reverse engineering" problem in which students must determine the lens location and focal points by using the same ray diagram techniques as in regular exercises, but reverse the logic. We also ask students students to calculate the results using the lens formula, thus showing that both ray diagrams and formulas give the same results when applied and interpreted correctly.

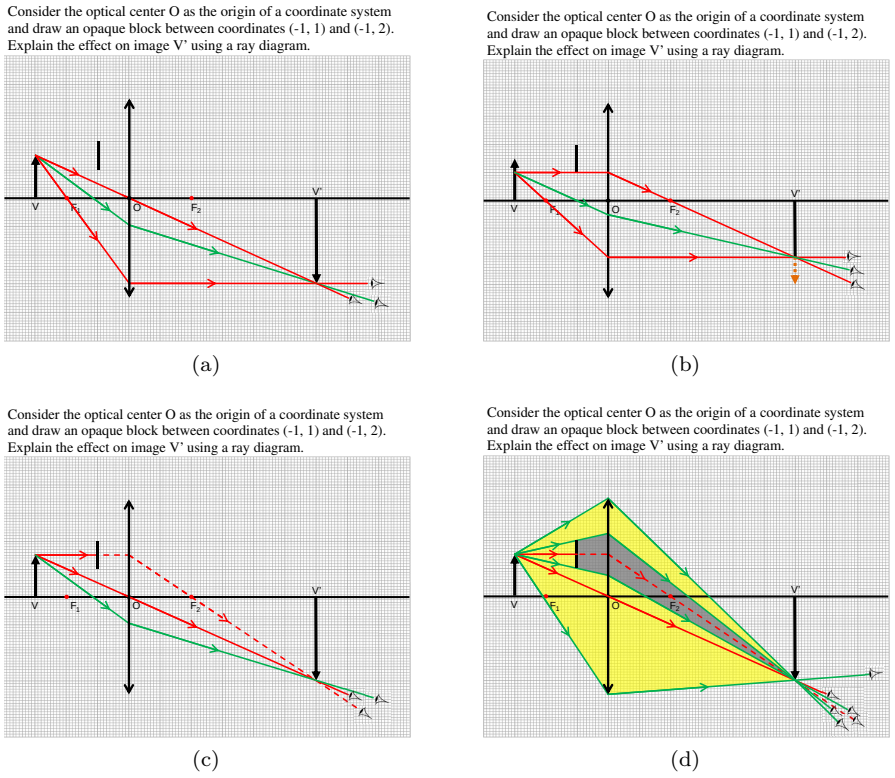


Figure A.9: Possible outcomes with principal rays in red and regular rays in green. (a) Correct diagram without any "blocked" rays, which does not create a conceptual conflict. (b) Incorrect diagram which attempts to explain the formation of a partial image by drawing two principals rays from the highest non-blocked point of the object thus creating the idea that the top (orange) part of the image is not formed. (c) Correct diagram to determine the image location. Though the principal ray through F_2 is blocked in reality the dashed extension is still used to correctly determine the image location. (d) Correct ray diagram using the same principal rays as in (c) to determine the image location. In addition this diagram also illustrates the effect of the opaque block: the light cone which determines the brightness of the image point is shown and part of the cone is blocked, thus resulting in a less bright image point.

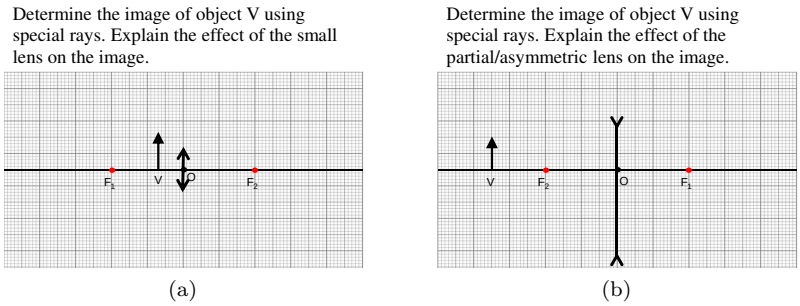


Figure A.10: (a) A "too small" lens with two principal rays not intersecting the drawn lens; (b) Asymmetric or partial (convex) lens.

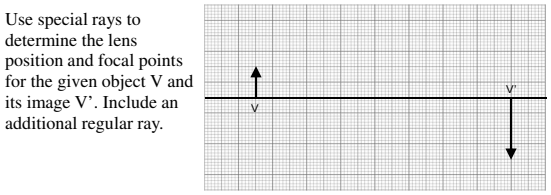


Figure A.11: Reverse application of known principles to find the lens position and focal points.

A.2.3.3 Indirect and direct observation of a real image formed by a convex lens

Also to distinguish between direct and indirect observation and study their conditions, we use the same approach.

- First students figure out how both observers can indirectly see the image using a screen and motivate this with a ray diagram. The goal is to end up with a diagram similar to Figure A.12 (a).
- In a second step, they determine for which of the given observers the image is directly visible — see Figure A.12 (b) — and identify the area from which the full image is always directly observable. The conclusion should be that there is an area along the optical axis behind the lens which can be determined by drawing the rays that start at the end points of the object and pass through the end points of the lens as shown in Figure A.12 (c) and in Figure A.12 (d). Note that these are regular rays, which aides in countering the idea that only principal rays contribute to the image formation. Another important point is that the rays don't

stop at the image point but continue and can reach the eye of an observer. Next, the overlapping area which allows for the direct observation of the full image is identified.

- Finally, the diagrams for indirect and direct observations are compared and discussed.

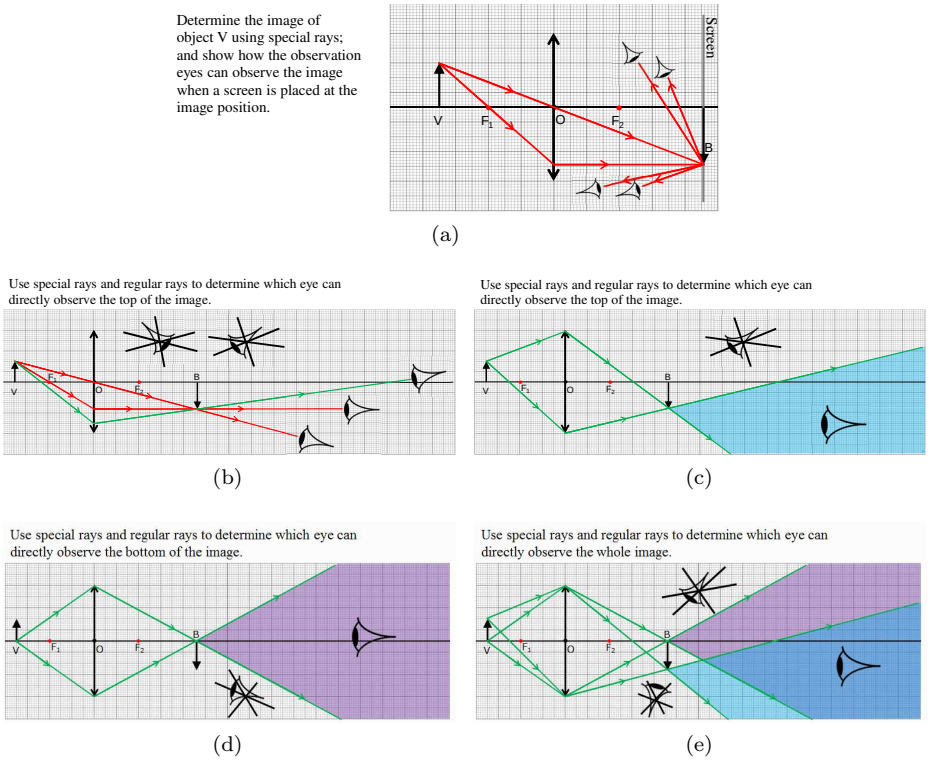


Figure A.12: (a) Indirect observation by diffuse reflection; (b) Direct observation from behind the lens and along the optical axis; (c), (d) and (e) Steps to determine the area from which the whole object can be observed directly.

A.2.4 Discussion and conclusion

A total of 30 secondary schools from across Flanders voluntarily implemented these activities during two consecutive years (2015 – 2017) in the 9th grade with a typical student age of 14 to 15 and a typical boy/girl ratio of 1.5.

All of the students participated in the STEM@school project, meaning they opted for a STEM-curriculum with a strong science component. Before, in between and after teaching the curriculum, the teachers — who range from new to experienced — were invited to semi-structured meetings with their colleagues from the other STEM-disciplines and from other participating schools to discuss the implementation, share their experiences and learn from each other. After the implementation, online questionnaires were administered to each team of teachers — one per school — to gather additional feedback for future improvements to the STEM-curriculum. The teachers' feedback and assessment of their students is based on the day-to-day interaction with their students and their own assessment tools (not-standardized across the various schools). This teacher feedback showed that they much appreciated the focus on misconceptions and that many found the exercises an improvement compared to classical textbooks. When queried about misconception retention after the course was taught, teachers' opinions were that low performing students likely still retain some misconceptions but that others have less or almost none as far as the teachers are aware. Furthermore, teachers reported increased students' enthusiasm and engagement for cooperation. These observations are in line with our goal to improve students' conceptual understanding through CUP and exercises with ray diagrams to provoke a conceptual conflict and elicit, confront and resolve misconceptions. Based on these positive experiences, we advise to include this approach in any new course design.

A.3 Detailed distractor analysis for the representational fluency test

Table A.1 and A.2 provide a highly detailed overview of the distractor data from the representational fluency test in Chapter 4 as a function of the context and the representation transition. The table clearly shows that most of the erroneous selections made by students are in the physics context and mainly with the transitions which include a formula as was already clear from the validation and the discussion in the body of the paper. We will highlight the most notable data points in the table.

The highest frequency of an incorrect answer is found in physics: 20.3% of students chose PP in a PN item with a TF transition — whereas only 14% chose PP in the isomorphic mathematics item. This means that the single most frequent error students made was when they had to identify a negative velocity in a table in physics and chose a positive velocity instead. This issue is likely explained by the tendency to identify velocity as speed in Flemish — thus ignoring the minus sign — and/or the difficulty students experience

with identifying a series of increasing numbers with negative signs. Since this erroneous choice was made substantially more often in physics, the issue with velocity and speed likely contributed to this difference between the contexts.

The second highest frequency in the table is that 18.7% of students chose ZP in a PP question in physics with the FT transition — for the isomorphic mathematical question this is 14.3%. This means that these students, when confronted with a formula with all positive coefficients, identify this as a directly proportional function ZP.

The third highest number is 18.4% in ZP-TF questions in mathematics for which students chose PP — in physics the frequency is 12.5%. These students — when confronted with a table including the point $(0, 0)$ — still chose a formula with a positive y intercept. Comparing this with the frequency of only 4.9% under the TG transition, the difficulty here is most likely not from interpreting the table, but from translating to a formula. A possible explanation is that students might consider a ZP table as the stereotypical table for a linear function and at the same time consider the PP formula as the stereotypical formula for a linear function, which could cause a mismatch.

Table A.1: Frequency (in percentages) of the chosen function type in the answer (A) as function of the function type and the representational transition in the question (Q) from all 385 respondents for all questions with a physics context. Correct answers are in bold face and all incorrect choices with frequency greater than 10.0% have colored cells.

		Physics					
Q	A	TG	GT	FT	TF	FG	GF
	NP	**84.7**	**84.9**	**44.7**	**44.7**	**56.1**	**53.5**
	ZP	2.3	2.6	11.4	6.8	12.2	3.6
	PP	2.2	1.8	8.1	9.1	7.3	7.5
NP	PN	2.6	2.3	14.0	15.8	14.5	14.0
	NN	4.4	4.9	11.9	15.3	3.9	16.9
	ZN	3.9	3.4	9.9	8.3	6.0	4.4
	NP	1.8	1.8	5.7	5.5	4.2	4.2
	ZP	**90.4**	**89.6**	**67.0**	**61.6**	**67.3**	**63.1**
	PP	3.1	2.1	7.3	12.5	10.6	11.2
ZP	PN	1.8	2.9	6.8	4.4	4.7	4.4
	NN	1.6	1.0	6.0	6.0	2.1	4.9
	ZN	1.3	2.6	7.3	10.1	11.2	12.2
	NP	1.0	2.1	4.7	8.8	6.5	7.0
	ZP	6.0	6.0	18.7	5.7	10.1	5.5
	PP	**86.2**	**85.2**	**59.7**	**60.3**	**69.1**	**68.1**
PP	PN	2.1	2.6	8.1	9.6	6.2	5.2
	NN	2.3	1.8	4.4	10.1	5.7	9.1
	ZN	2.3	2.3	4.4	5.5	2.3	5.2
	NP	3.4	3.4	12.5	10.6	14.5	11.4
	ZP	3.1	3.9	10.1	9.1	10.4	5.2
	PP	3.6	1.6	9.1	20.3	10.1	15.1
PN	PN	**85.5**	**85.5**	**45.5**	**46.0**	**51.4**	**53.2**
	NN	2.1	3.1	11.4	6.2	6.5	9.1
	ZN	2.3	2.6	11.4	7.8	7.0	6.0

Table A.2: Frequency (in percentages) of the chosen function type in the answer (A) as function of the function type and the representational transition in the question (Q) from all 385 respondents for all questions with a mathematics context. Correct answers are in bold face and all incorrect choices with frequency greater than 10.0% have colored cells.

		Mathematics					
Q	A	TG	GT	FT	TF	FG	GF
	NP	**86.2**	**84.9**	**52.5**	**58.7**	**58.4**	**59.2**
	ZP	1.6	2.3	3.9	4.7	6.0	4.4
	PP	0.8	2.3	11.4	8.3	8.3	6.0
NP	PN	2.3	2.9	11.2	11.9	13.2	9.6
	NN	5.7	4.7	11.7	11.9	6.5	15.8
	ZN	3.4	2.9	9.4	4.4	7.5	4.9
	NP	1.0	1.3	3.4	4.4	4.2	4.7
	ZP	**87.5**	**88.3**	**71.7**	**59.2**	**75.1**	**63.6**
	PP	4.9	5.7	8.3	18.4	7.0	9.6
ZP	PN	2.6	1.3	8.3	4.9	5.2	5.5
	NN	2.3	1.8	1.8	6.0	2.9	5.7
	ZN	1.6	1.6	6.5	7.0	5.7	10.9
	NP	2.3	2.1	5.5	7.0	6.2	6.5
	ZP	4.4	8.8	14.3	6.8	7.8	5.2
	PP	**84.9**	**82.3**	**65.2**	**65.5**	**73.2**	**72.7**
PP	PN	3.4	3.4	6.5	4.4	3.9	4.2
	NN	2.1	2.3	4.7	11.2	6.0	7.5
	ZN	2.9	1.0	3.9	5.2	2.9	3.9
	NP	1.6	2.1	13.8	15.1	14.3	15.8
	ZP	3.6	1.6	8.1	6.8	7.0	3.6
	PP	5.2	3.9	9.6	14.0	7.8	11.4
PN	PN	**86.2**	**88.3**	**50.9**	**51.4**	**60.5**	**55.8**
	NN	0.8	0.5	10.4	8.1	5.5	9.1
	ZN	2.6	3.6	7.3	4.7	4.9	4.2

A.4 Representational fluency test

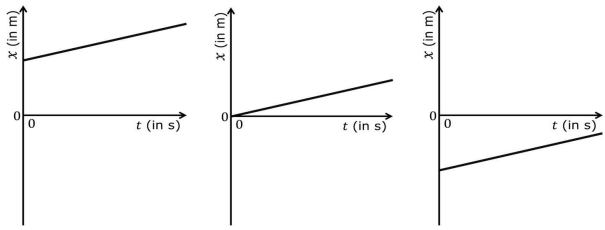
This Appendix contains the full 48-item representational fluency test which is discussed and used in Chapter 4. First we provide the 24-item kinematics (physics) part, and then the as-isomorphic-as-possible 24-item mathematics part. The questions are structured in line with the order in the tables presented in Chapter 4.

A.4.1 Physics

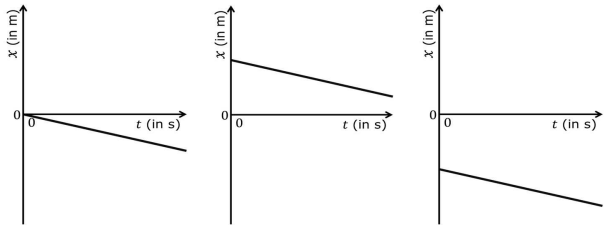
P1: A cyclist is cycling at constant velocity on a straight road. The table describes the relationship between the position x (in meters) and the time t (in seconds):

t (in s)	x (in m)
0	-36
1	-30
2	-24
3	-18
4	-12

One of the graphs describes the same relationship between x and t. Which one?



(a) (b) (c)

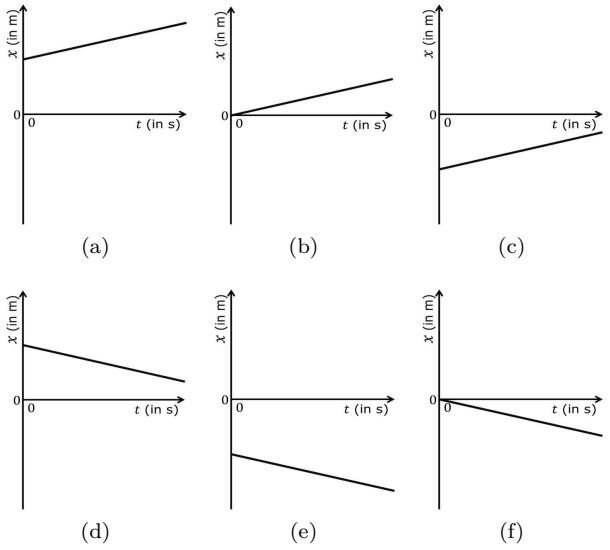


(d) (e) (f)

P2: A cyclist is cycling at constant velocity on a straight road. The table describes the relationship between the position x (in meters) and the time t (in seconds):

t (in s)	x (in m)
0	0
1	7
2	14
3	21
4	28

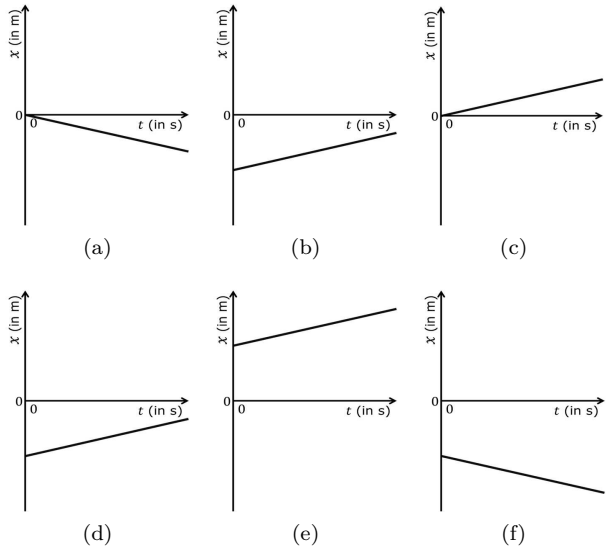
One of the graphs describes the same relationship between x and t. Which one?



P3: A cyclist is cycling at constant velocity on a straight road. The table describes the relationship between the position x (in meters) and the time t (in seconds):

t (in s)	x (in m)
0	24
1	20
2	16
3	12
4	8

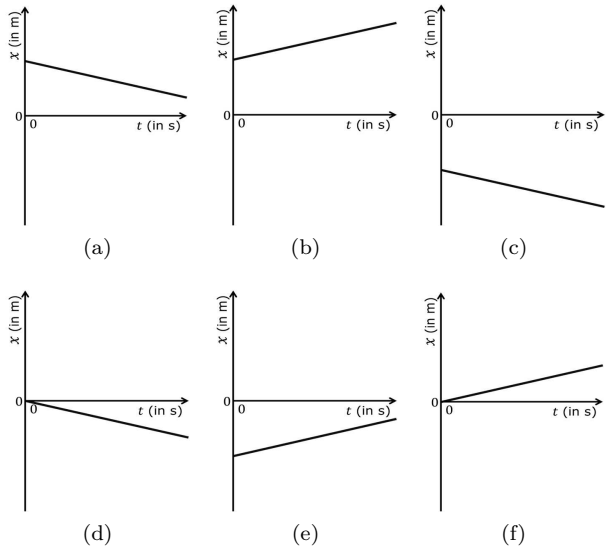
One of the graphs describes the same relationship between x and t. Which one?



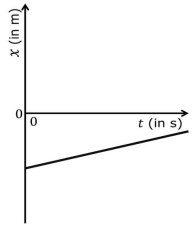
P4: A cyclist is cycling at constant velocity on a straight road. The table describes the relationship between the position x (in meters) and the time t (in seconds):

t (in s)	x (in m)
0	18
1	21
2	24
3	27
4	30

One of the graphs describes the same relationship between x and t. Which one?



P5: A cyclist is cycling at constant velocity on a straight road. The graph describes the relationship between the position x (in meters) and the time t (in seconds):



One of the tables describes the same relationship between x and t. Which one?

t (in s)	x (in m)
0	-18
1	-21
2	-24
3	-27
4	-30

(a)

t (in s)	x (in m)
0	18
1	21
2	24
3	27
4	30

(b)

t (in s)	x (in m)
0	-18
1	-15
2	-12
3	-9
4	-6

(c)

t (in s)	x (in m)
0	18
1	15
2	12
3	9
4	6

(d)

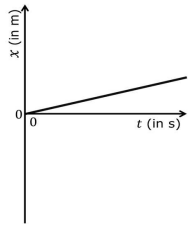
t (in s)	x (in m)
0	0
1	3
2	6
3	9
4	12

(e)

t (in s)	x (in m)
0	0
1	-3
2	-6
3	-9
4	-12

(f)

P6: A cyclist is cycling at constant velocity on a straight road. The graph describes the relationship between the position x (in meters) and the time t (in seconds):



One of the tables describes the same relationship between x and t. Which one?

t (in s)	x (in m)
0	-24
1	-20
2	-16
3	-12
4	-8

(a)

t (in s)	x (in m)
0	24
1	28
2	32
3	36
4	40

(b)

t (in s)	x (in m)
0	-24
1	-28
2	-32
3	-36
4	-40

(c)

t (in s)	x (in m)
0	24
1	20
2	16
3	12
4	8

(d)

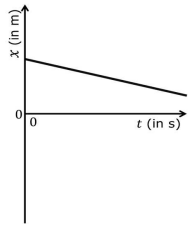
t (in s)	x (in m)
0	0
1	4
2	8
3	12
4	16

(e)

t (in s)	x (in m)
0	0
1	-4
2	-8
3	-12
4	-16

(f)

P7: A cyclist is cycling at constant velocity on a straight road. The graph describes the relationship between the position x (in meters) and the time t (in seconds):



One of the tables describes the same relationship between x and t. Which one?

t (in s)	x (in m)
0	36
1	30
2	24
3	18
4	12

(a)

t (in s)	x (in m)
0	-36
1	-42
2	-48
3	-54
4	-60

(b)

t (in s)	x (in m)
0	0
1	6
2	12
3	18
4	24

(c)

t (in s)	x (in m)
0	-36
1	-30
2	-24
3	-18
4	-12

(d)

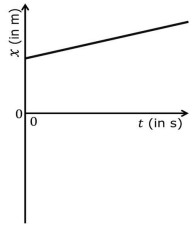
t (in s)	x (in m)
0	36
1	42
2	48
3	54
4	60

(e)

t (in s)	x (in m)
0	0
1	-6
2	-12
3	-18
4	-24

(f)

P8: A cyclist is cycling at constant velocity on a straight road. The graph describes the relationship between the position x (in meters) and the time t (in seconds):



One of the tables describes the same relationship between x and t. Which one?

t (in s)	x (in m)
0	0
1	7
2	14
3	21
4	28

(a)

t (in s)	x (in m)
0	42
1	35
2	28
3	21
4	14

(b)

t (in s)	x (in m)
0	-42
1	-35
2	-28
3	-21
4	-14

(c)

t (in s)	x (in m)
0	0
1	-7
2	-14
3	-21
4	-28

(d)

t (in s)	x (in m)
0	-42
1	-49
2	-56
3	-63
4	-70

(e)

t (in s)	x (in m)
0	42
1	49
2	56
3	63
4	70

(f)

P9: A cyclist is cycling at constant velocity on a straight road. The relationship between the position x (in meters) and the time t (in seconds) is given by $x = x_0 + vt$ with $x_0 < 0$ and $v > 0$. One of the tables describes the same relationship between x and t. Which one?

t (in s)	x (in m)
0	0
1	7
2	14
3	21
4	28

(a)

t (in s)	x (in m)
0	42
1	49
2	56
3	63
4	70

(b)

t (in s)	x (in m)
0	0
1	-7
2	-14
3	-21
4	-28

(c)

t (in s)	x (in m)
0	-42
1	-35
2	-28
3	-21
4	-14

(d)

t (in s)	x (in m)
0	42
1	35
2	28
3	21
4	14

(e)

t (in s)	x (in m)
0	-42
1	-49
2	-56
3	-63
4	-70

(f)

P10: A cyclist is cycling at constant velocity on a straight road. The relationship between the position x (in meters) and the time t (in seconds) is given by $x = x_0 + vt$ with $x_0 = 0$ and $v > 0$. One of the tables describes the same relationship between x and t. Which one?

t (in s)	x (in m)
0	18
1	15
2	12
3	9
4	6

(a)

t (in s)	x (in m)
0	18
1	21
2	24
3	27
4	30

(b)

t (in s)	x (in m)
0	-18
1	-21
2	-24
3	-27
4	-30

(c)

t (in s)	x (in m)
0	0
1	3
2	6
3	9
4	12

(d)

t (in s)	x (in m)
0	-18
1	-15
2	-12
3	-9
4	-6

(e)

t (in s)	x (in m)
0	0
1	-3
2	-6
3	-9
4	-12

(f)

P11: A cyclist is cycling at constant velocity on a straight road. The relationship between the position x (in meters) and the time t (in seconds) is given by $x = x_0 + vt$ with $x_0 > 0$ and $v < 0$. One of the tables describes the same relationship between x and t. Which one?

t (in s)	x (in m)
0	24
1	20
2	16
3	12
4	8

(a)

t (in s)	x (in m)
0	24
1	28
2	32
3	36
4	40

(b)

t (in s)	x (in m)
0	0
1	4
2	8
3	12
4	16

(c)

t (in s)	x (in m)
0	0
1	-4
2	-8
3	-12
4	-16

(d)

t (in s)	x (in m)
0	-24
1	-20
2	-16
3	-12
4	-8

(e)

t (in s)	x (in m)
0	-24
1	-28
2	-32
3	-36
4	-40

(f)

P12: A cyclist is cycling at constant velocity on a straight road. The relationship between the position x (in meters) and the time t (in seconds) is given by $x = x_0 + vt$ with $x_0 > 0$ and $v > 0$. One of the tables describes the same relationship between x and t. Which one?

t (in s)	x (in m)
0	0
1	-6
2	-12
3	-18
4	-24

(a)

t (in s)	x (in m)
0	-36
1	-30
2	-24
3	-18
4	-12

(b)

t (in s)	x (in m)
0	0
1	6
2	12
3	18
4	24

(c)

t (in s)	x (in m)
0	36
1	30
2	24
3	18
4	12

(d)

t (in s)	x (in m)
0	36
1	42
2	48
3	54
4	60

(e)

t (in s)	x (in m)
0	-36
1	-42
2	-48
3	-54
4	-60

(f)

P13: A cyclist is cycling at constant velocity on a straight road. The table describes the relationship between the position x (in meters) and the time t (in seconds):

t (in s)	x (in m)
0	-18
1	-15
2	-12
3	-9
4	-6

One of the equations describes the same relationship between x and t. Which one?

a) $x = x_0 + vt$ with $x_0 > 0$ and $v > 0$.

b) $x = x_0 + vt$ with $x_0 < 0$ and $v < 0$.

c) $x = x_0 + vt$ with $x_0 = 0$ and $v > 0$.

d) $x = x_0 + vt$ with $x_0 > 0$ and $v < 0$.

e) $x = x_0 + vt$ with $x_0 = 0$ and $v < 0$.

f) $x = x_0 + vt$ with $x_0 < 0$ and $v > 0$.

P14: A cyclist is cycling at constant velocity on a straight road. The table describes the relationship between the position x (in meters) and the time t (in seconds):

t (in s)	x (in m)
0	0
1	6
2	12
3	18
4	24

One of the equations describes the same relationship between x and t. Which one?

a) $x = x_0 + vt$ with $x_0 < 0$ and $v < 0$.

b) $x = x_0 + vt$ with $x_0 = 0$ and $v > 0$.

c) $x = x_0 + vt$ with $x_0 > 0$ and $v < 0$.

d) $x = x_0 + vt$ with $x_0 > 0$ and $v > 0$.

e) $x = x_0 + vt$ with $x_0 = 0$ and $v < 0$.

f) $x = x_0 + vt$ with $x_0 < 0$ and $v > 0$.

P15: A cyclist is cycling at constant velocity on a straight road. The table describes the relationship between the position x (in meters) and the time t (in seconds):

t (in s)	x (in m)
0	42
1	35
2	28
3	21
4	14

One of the equations describes the same relationship between x and t. Which one?

a) $x = x_0 + vt$ with $x_0 > 0$ and $v > 0$.

b) $x = x_0 + vt$ with $x_0 = 0$ and $v > 0$.

c) $x = x_0 + vt$ with $x_0 < 0$ and $v > 0$.

d) $x = x_0 + vt$ with $x_0 = 0$ and $v < 0$.

e) $x = x_0 + vt$ with $x_0 > 0$ and $v < 0$.

f) $x = x_0 + vt$ with $x_0 < 0$ and $v < 0$.

P16: A cyclist is cycling at constant velocity on a straight road. The table describes the relationship between the position x (in meters) and the time t (in seconds):

t (in s)	x (in m)
0	24
1	28
2	32
3	36
4	40

One of the equations describes the same relationship between x and t. Which one?

a) $x = x_0 + vt$ with $x_0 > 0$ and $v < 0$.

b) $x = x_0 + vt$ with $x_0 = 0$ and $v < 0$.

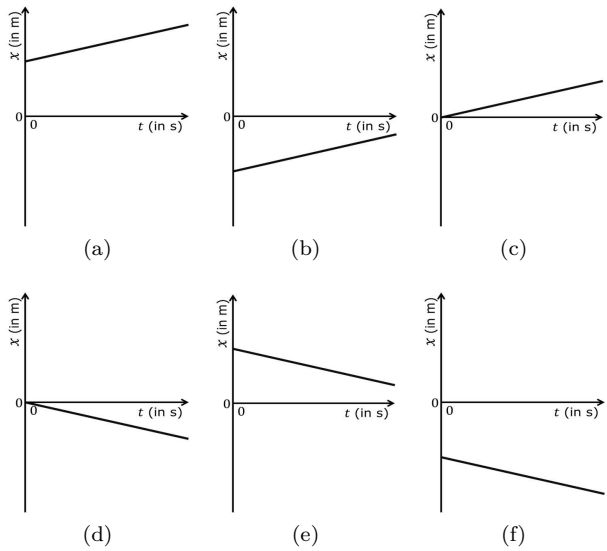
c) $x = x_0 + vt$ with $x_0 > 0$ and $v > 0$.

d) $x = x_0 + vt$ with $x_0 < 0$ and $v > 0$.

e) $x = x_0 + vt$ with $x_0 < 0$ and $v < 0$.

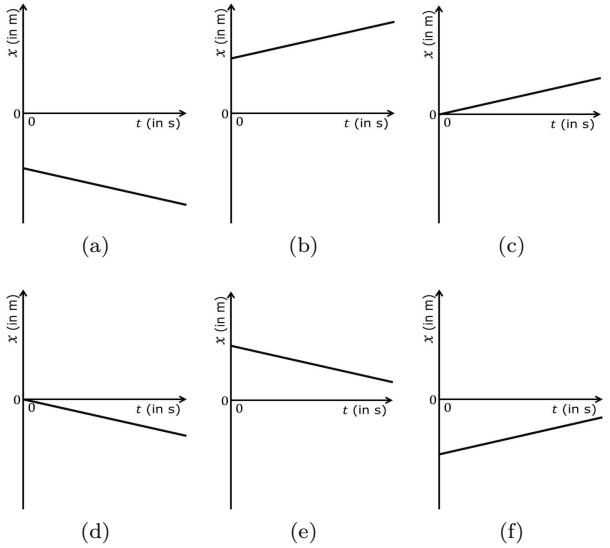
f) $x = x_0 + vt$ with $x_0 = 0$ and $v > 0$.

P17: A cyclist is cycling at constant velocity on a straight road. The relationship between the position x (in meters) and the time t (in seconds) is given by $x = x_0 + vt$ with $x_0 < 0$ and $v > 0$. One of the graphs describes the same relationship between x and t. Which one?

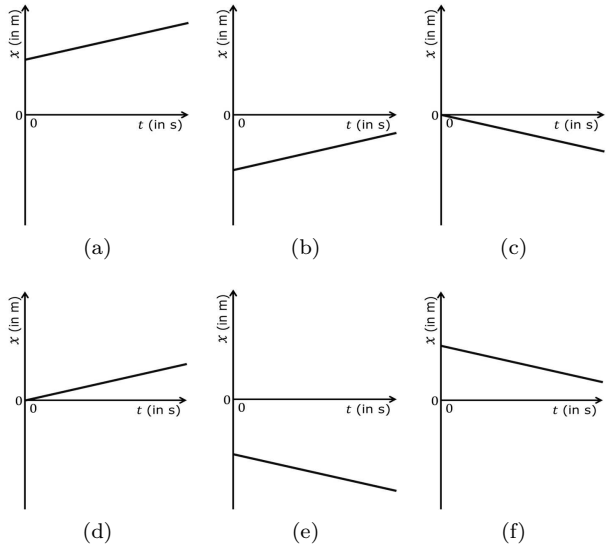


P18: A cyclist is cycling at constant velocity on a straight road. The relationship between the position x (in meters) and the time t (in seconds)

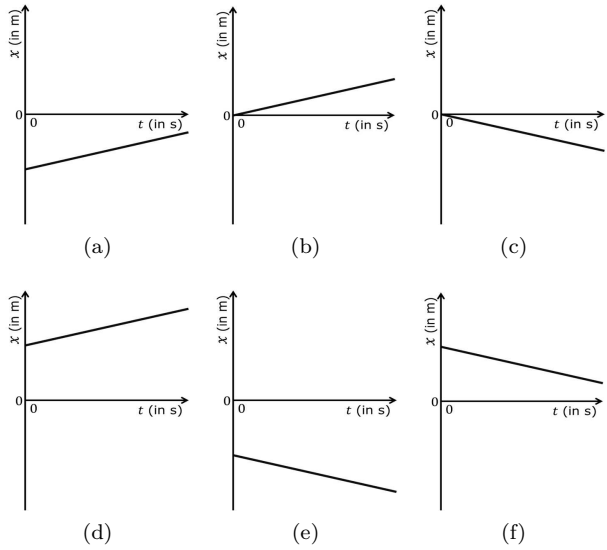
is given by $x = x_0 + vt$ with $x_0 = 0$ and $v > 0$. One of the graphs describes the same relationship between x and t. Which one?



P19: A cyclist is cycling at constant velocity on a straight road. The relationship between the position x (in meters) and the time t (in seconds) is given by $x = x_0 + vt$ with $x_0 > 0$ and $v < 0$. One of the graphs describes the same relationship between x and t. Which one?

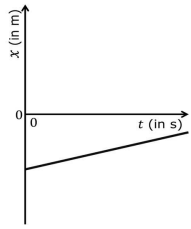


P20: A cyclist is cycling at constant velocity on a straight road. The relationship between the position x (in meters) and the time t (in seconds) is given by $x = x_0 + vt$ with $x_0 > 0$ and $v > 0$. One of the graphs describes the same relationship between x and t. Which one?



P21: A cyclist is cycling at constant velocity on a straight road. The graph

describes the relationship between the position x (in meters) and the time t (in seconds):



One of the equations describes the same relationship between x and t. Which one?

a) $x = x_0 + vt$ with $x_0 < 0$ and $v > 0$.

b) $x = x_0 + vt$ with $x_0 = 0$ and $v < 0$.

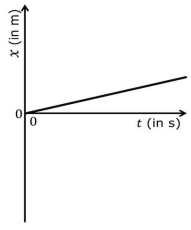
c) $x = x_0 + vt$ with $x_0 < 0$ and $v < 0$.

d) $x = x_0 + vt$ with $x_0 = 0$ and $v > 0$.

e) $x = x_0 + vt$ with $x_0 > 0$ and $v > 0$.

f) $x = x_0 + vt$ with $x_0 > 0$ and $v < 0$.

P22: A cyclist is cycling at constant velocity on a straight road. The graph describes the relationship between the position x (in meters) and the time t (in seconds):



One of the equations describes the same relationship between x and t. Which one?

a) $x = x_0 + vt$ with $x_0 = 0$ and $v > 0$.

b) $x = x_0 + vt$ with $x_0 < 0$ and $v < 0$.

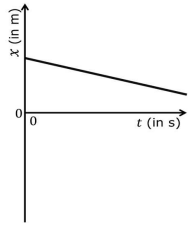
c) $x = x_0 + vt$ with $x_0 > 0$ and $v < 0$.

d) $x = x_0 + vt$ with $x_0 < 0$ and $v > 0$.

e) $x = x_0 + vt$ with $x_0 = 0$ and $v < 0$.

f) $x = x_0 + vt$ with $x_0 > 0$ and $v > 0$.

P23: A cyclist is cycling at constant velocity on a straight road. The graph describes the relationship between the position x (in meters) and the time t (in seconds):



One of the equations describes the same relationship between x and t. Which one?

a) $x = x_0 + vt$ with $x_0 > 0$ and $v > 0$.

b) $x = x_0 + vt$ with $x_0 = 0$ and $v > 0$.

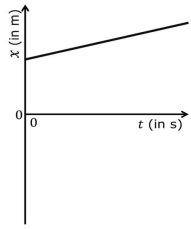
c) $x = x_0 + vt$ with $x_0 < 0$ and $v < 0$.

d) $x = x_0 + vt$ with $x_0 < 0$ and $v > 0$.

e) $x = x_0 + vt$ with $x_0 = 0$ and $v < 0$.

f) $x = x_0 + vt$ with $x_0 > 0$ and $v < 0$.

P24: A cyclist is cycling at constant velocity on a straight road. The graph describes the relationship between the position x (in meters) and the time t (in seconds):



One of the equations describes the same relationship between x and t. Which one?

a) $x = x_0 + vt$ with $x_0 > 0$ and $v < 0$.

b) $x = x_0 + vt$ with $x_0 < 0$ and $v > 0$.

c) $x = x_0 + vt$ with $x_0 = 0$ and $v < 0$.

d) $x = x_0 + vt$ with $x_0 = 0$ and $v > 0$.

e) $x = x_0 + vt$ with $x_0 > 0$ and $v > 0$.

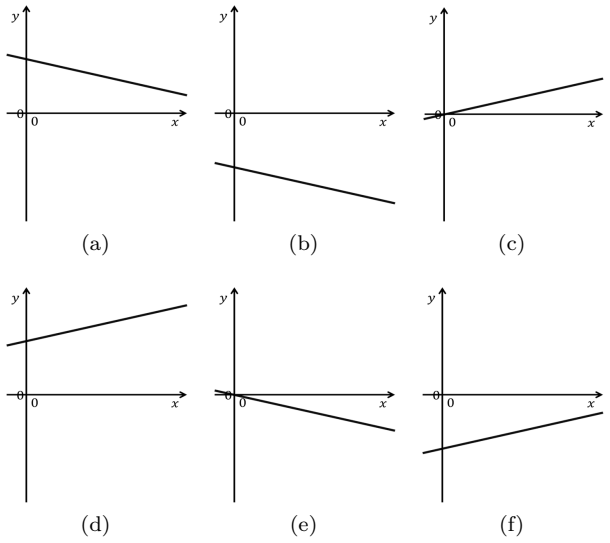
f) $x = x_0 + vt$ with $x_0 < 0$ and $v < 0$.

A.4.2 Mathematics

M1: The variable y is dependent of the variable x according to a first degree equation. The table describes this relationship between y and x:

x	y
0	-36
1	-30
2	-24
3	-18
4	-12

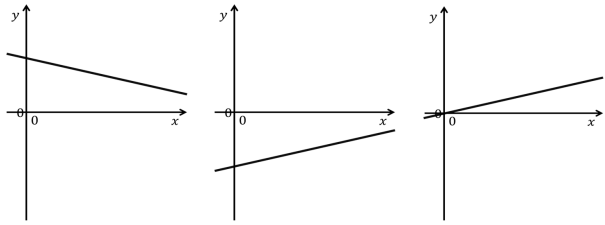
One of the graphs describes the same relationship between y and x. Which one?



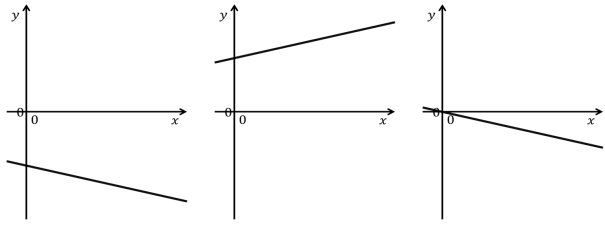
M2: The variable y is dependent of the variable x according to a first degree equation. The table describes this relationship between y and x:

One of the graphs describes the same relationship between y and x. Which one?

x	y
0	0
1	7
2	14
3	21
4	28



(a) (b) (c)

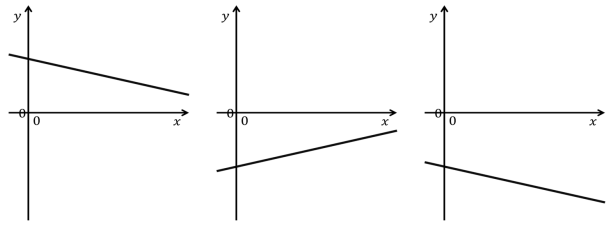


(d) (e) (f)

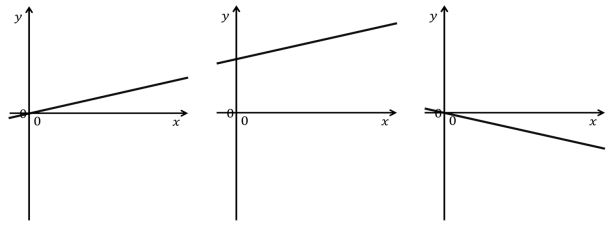
M3: The variable y is dependent of the variable x according to a first degree equation. The table describes this relationship between y and x:

x	y
0	24
1	20
2	16
3	12
4	8

One of the graphs describes the same relationship between y and x. Which one?



(a) (b) (c)

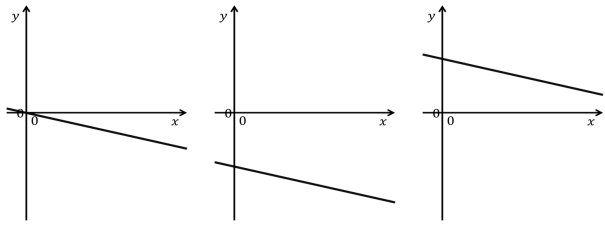


(d) (e) (f)

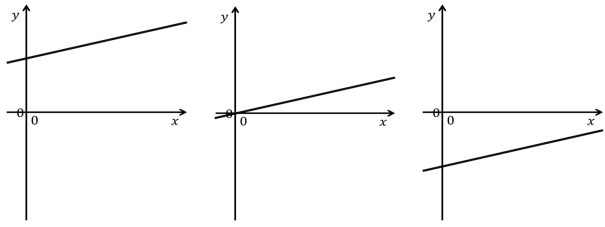
M4: The variable y is dependent of the variable x according to a first degree equation. The table describes this relationship between y and x:

x	y
0	18
1	21
2	24
3	27
4	30

One of the graphs describes the same relationship between y and x. Which one?

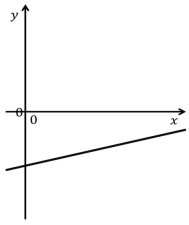


(a) (b) (c)



(d) (e) (f)

M5: The variable y is dependent of the variable x according to a first degree equation. The graph describes this relationship between y and x:



One of the tables describes the same relationship between y and x. Which one?

t	x
0	0
1	3
2	6
3	9
4	12

(a)

t	x
0	0
1	-3
2	-6
3	-9
4	-12

(b)

t	x
0	18
1	15
2	12
3	9
4	6

(c)

t	x
0	18
1	21
2	24
3	27
4	30

(d)

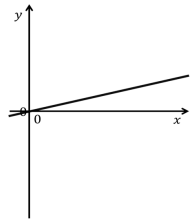
t	x
0	-18
1	-15
2	-12
3	-9
4	-6

(e)

t	x
0	-18
1	-21
2	-24
3	-27
4	-30

(f)

M6: The variable y is dependent of the variable x according to a first degree equation. The graph describes this relationship between y and x:



One of the tables describes the same relationship between y and x. Which one?

t	x
0	-24
1	-28
2	-32
3	-36
4	-40

(a)

t	x
0	0
1	4
2	8
3	12
4	16

(b)

t	x
0	24
1	28
2	32
3	36
4	40

(c)

t	x
0	-24
1	-20
2	-16
3	-12
4	-8

(d)

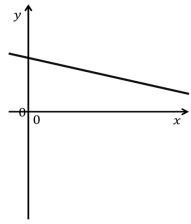
t	x
0	24
1	20
2	16
3	12
4	8

(e)

t	x
0	0
1	-4
2	-8
3	-12
4	-16

(f)

M7: The variable y is dependent of the variable x according to a first degree equation. The graph describes this relationship between y and x:



One of the tables describes the same relationship between y and x. Which one?

t	x
0	0
1	6
2	12
3	18
4	24

(a)

t	x
0	-36
1	-30
2	-24
3	-18
4	-12

(b)

t	x
0	0
1	-6
2	-12
3	-18
4	-24

(c)

t	x
0	36
1	30
2	24
3	18
4	12

(d)

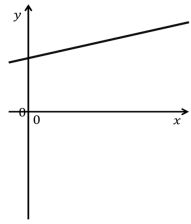
t	x
0	-36
1	-42
2	-48
3	-54
4	-60

(e)

t	x
0	36
1	42
2	48
3	54
4	60

(f)

M8: The variable y is dependent of the variable x according to a first degree equation. The graph describes this relationship between y and x:



One of the tables describes the same relationship between y and x. Which one?

t	x
0	0
1	-7
2	-14
3	-21
4	-28

(a)

t	x
0	0
1	7
2	14
3	21
4	28

(b)

t	x
0	-42
1	-35
2	-28
3	-21
4	-14

(c)

t	x
0	42
1	35
2	28
3	21
4	14

(d)

t	x
0	-42
1	-49
2	-56
3	-63
4	-70

(e)

t	x
0	42
1	49
2	56
3	63
4	70

(f)

M9: The variable y is dependent of the variable x according to a first degree equation $y = ax + b$ with $a > 0$ and $b < 0$. One of the tables describes the same relationship between y and x. Which one?

t	x
0	-42
1	-49
2	-56
3	-63
4	-70

(a)

t	x
0	0
1	7
2	14
3	21
4	28

(b)

t	x
0	42
1	35
2	28
3	21
4	14

(c)

t	x
0	42
1	49
2	56
3	63
4	70

(d)

t	x
0	-42
1	-35
2	-28
3	-21
4	-14

(e)

t	x
0	0
1	-7
2	-14
3	-21
4	-28

(f)

M10: The variable y is dependent of the variable x according to a first degree equation $y = ax + b$ with $a > 0$ and $b = 0$. One of the tables describes the same relationship between y and x. Which one?

t	x
0	-18
1	-15
2	-12
3	-9
4	-6

(a)

t	x
0	0
1	-3
2	-6
3	-9
4	-12

(b)

t	x
0	18
1	15
2	12
3	9
4	6

(c)

t	x
0	0
1	3
2	6
3	9
4	12

(d)

t	x
0	-18
1	-21
2	-24
3	-27
4	-30

(e)

t	x
0	18
1	21
2	24
3	27
4	30

(f)

M11: The variable y is dependent of the variable x according to a first degree equation $y = ax + b$ with $a < 0$ and $b > 0$. One of the tables describes the same relationship between y and x. Which one?

t	x
0	0
1	4
2	8
3	12
4	16

(a)

t	x
0	-24
1	-20
2	-16
3	-12
4	-8

(b)

t	x
0	-24
1	-28
2	-32
3	-36
4	-40

(c)

t	x
0	0
1	-4
2	-8
3	-12
4	-16

(d)

t	x
0	24
1	28
2	32
3	36
4	40

(e)

t	x
0	24
1	20
2	16
3	12
4	8

(f)

M12: The variable y is dependent of the variable x according to a first degree equation $y = ax + b$ with $a > 0$ and $b > 0$. One of the tables describes the same relationship between y and x. Which one?

t	x
0	-36
1	-42
2	-48
3	-54
4	-60

(a)

t	x
0	0
1	6
2	12
3	18
4	24

(b)

t	x
0	-36
1	-30
2	-24
3	-18
4	-12

(c)

t	x
0	36
1	30
2	24
3	18
4	12

(d)

t	x
0	36
1	42
2	48
3	54
4	60

(e)

t	x
0	0
1	-6
2	-12
3	-18
4	-24

(f)

M13: The variable y is dependent of the variable x according to a first degree equation. The table describes this relationship between y and x:

x	y
0	-18
1	-15
2	-12
3	-9
4	-6

One of the equations describes the same relationship between y and x. Which one?

a) $y = ax + b$ with $a < 0$ and $b = 0$.

b) $y = ax + b$ with $a > 0$ and $b > 0$.

c) $y = ax + b$ with $a > 0$ and $b < 0$.

d) $y = ax + b$ with $a < 0$ and $b < 0$.

e) $y = ax + b$ with $a < 0$ and $b > 0$.

f) $y = ax + b$ with $a > 0$ and $b = 0$.

M14: The variable y is dependent of the variable x according to a first degree equation. The table describes this relationship between y and x:

x	y
0	0
1	6
2	12
3	18
4	24

One of the equations describes the same relationship between y and x. Which one?

a) $y = ax + b$ with $a > 0$ and $b < 0$.

b) $y = ax + b$ with $a > 0$ and $b = 0$.

c) $y = ax + b$ with $a < 0$ and $b > 0$.

d) $y = ax + b$ with $a < 0$ and $b < 0$.

e) $y = ax + b$ with $a > 0$ and $b > 0$.

f) $y = ax + b$ with $a < 0$ and $b = 0$.

M15: The variable y is dependent of the variable x according to a first degree equation. The table describes this relationship between y and x:

x	y
0	42
1	35
2	28
3	21
4	14

One of the equations describes the same relationship between y and x. Which one?

a) $y = ax + b$ with $a < 0$ and $b > 0$.

b) $y = ax + b$ with $a < 0$ and $b < 0$.

c) $y = ax + b$ with $a > 0$ and $b > 0$.

d) $y = ax + b$ with $a > 0$ and $b < 0$.

e) $y = ax + b$ with $a > 0$ and $b = 0$.

f) $y = ax + b$ with $a < 0$ and $b = 0$.

M16: The variable y is dependent of the variable x according to a first degree equation. The table describes this relationship between y and x:

x	y
0	24
1	28
2	32
3	36
4	40

One of the equations describes the same relationship between y and x. Which one?

a) $y = ax + b$ with $a > 0$ and $b = 0$.

b) $y = ax + b$ with $a > 0$ and $b > 0$.

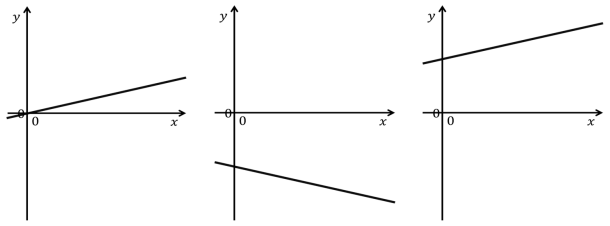
c) $y = ax + b$ with $a < 0$ and $b = 0$.

d) $y = ax + b$ with $a > 0$ and $b < 0$.

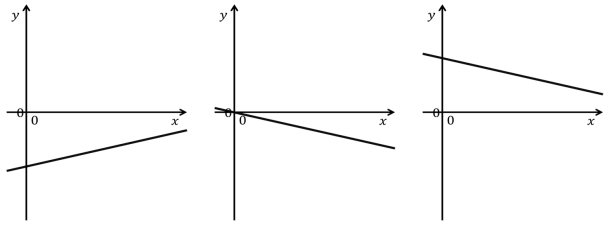
e) $y = ax + b$ with $a < 0$ and $b > 0$.

f) $y = ax + b$ with $a < 0$ and $b < 0$.

M17: The variable y is dependent of the variable x according to a first degree equation $y = ax + b$ with $a > 0$ and $b < 0$. One of the graphs describes the same relationship between y and x. Which one?

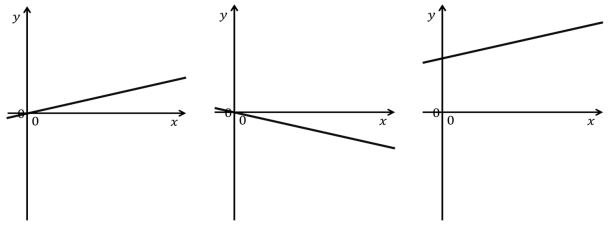


(a) (b) (c)

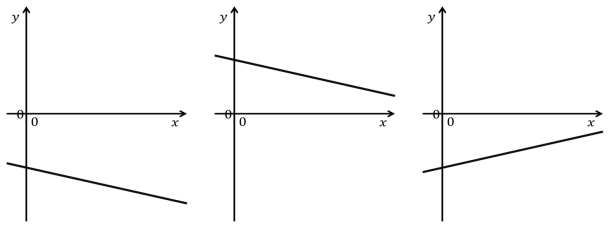


(d) (e) (f)

M18: The variable y is dependent of the variable x according to a first degree equation $y = ax + b$ with $a > 0$ and $b = 0$. One of the graphs describes the same relationship between y and x. Which one?

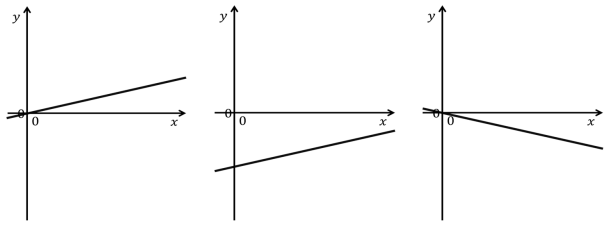


(a) (b) (c)

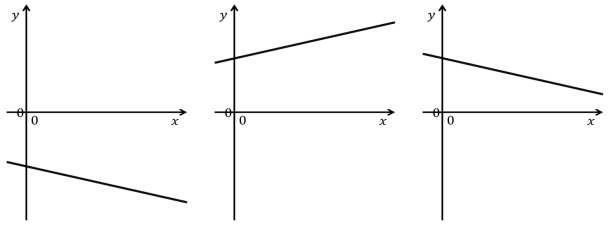


(d) (e) (f)

M19: The variable y is dependent of the variable x according to a first degree equation $y = ax + b$ with $a < 0$ and $b > 0$. One of the graphs describes the same relationship between y and x. Which one?

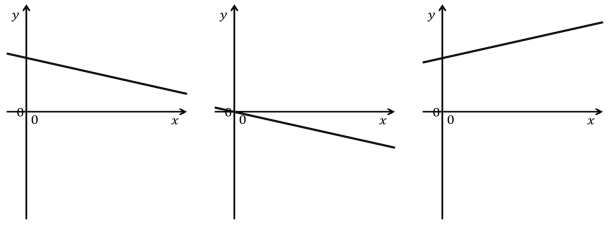


(a) (b) (c)

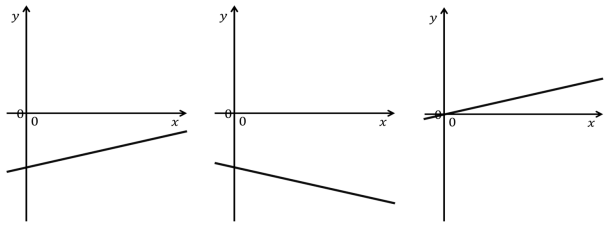


(d) (e) (f)

M20: The variable y is dependent of the variable x according to a first degree equation $y = ax + b$ with $a > 0$ and $b > 0$. One of the graphs describes the same relationship between y and x. Which one?

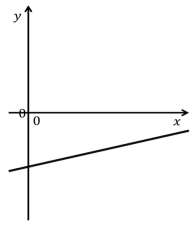


(a) (b) (c)



(d) (e) (f)

M21: The variable y is dependent of the variable x according to a first degree equation. The graph describes this relationship between y and x:



One of the equations describes the same relationship between y and x. Which one?

a) $y = ax + b$ with $a > 0$ and $b > 0$.

b) $y = ax + b$ with $a < 0$ and $b = 0$.

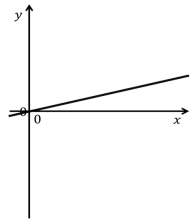
c) $y = ax + b$ with $a > 0$ and $b < 0$.

d) $y = ax + b$ with $a > 0$ and $b = 0$.

e) $y = ax + b$ with $a < 0$ and $b < 0$.

f) $y = ax + b$ with $a < 0$ and $b > 0$.

M22: The variable y is dependent of the variable x according to a first degree equation. The graph describes this relationship between y and x:



One of the equations describes the same relationship between y and x. Which one?

a) $y = ax + b$ with $a < 0$ and $b = 0$.

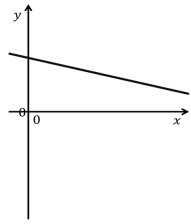
b) $y = ax + b$ with $a < 0$ and $b < 0$.

c) $y = ax + b$ with $a > 0$ and $b > 0$.

d) $y = ax + b$ with $a < 0$ and $b > 0$.

e) $y = ax + b$ with $a > 0$ and $b = 0$.

f) $y = ax + b$ with $a > 0$ and $b < 0$.



M23: The variable y is dependent of the variable x according to a first degree equation. The graph describes this relationship between y and x:

One of the equations describes the same relationship between y and x. Which one?

a) $y = ax + b$ with $a < 0$ and $b > 0$.

b) $y = ax + b$ with $a > 0$ and $b < 0$.

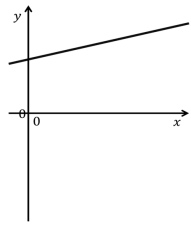
c) $y = ax + b$ with $a < 0$ and $b = 0$.

d) $y = ax + b$ with $a > 0$ and $b > 0$.

e) $y = ax + b$ with $a > 0$ and $b = 0$.

f) $y = ax + b$ with $a < 0$ and $b < 0$.

M24: The variable y is dependent of the variable x according to a first degree equation. The graph describes this relationship between y and x:



One of the equations describes the same relationship between y and x. Which one?

a) $y = ax + b$ with $a > 0$ and $b < 0$.

b) $y = ax + b$ with $a > 0$ and $b > 0$.

c) $y = ax + b$ with $a > 0$ and $b = 0$.

d) $y = ax + b$ with $a < 0$ and $b < 0$.

e) $y = ax + b$ with $a < 0$ and $b = 0$.

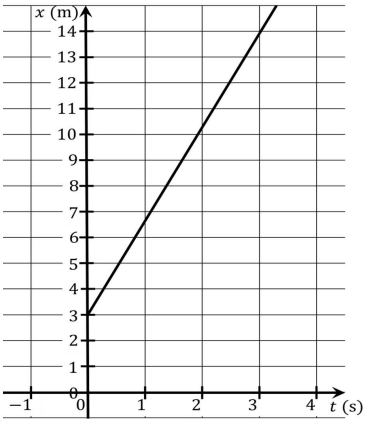
f) $y = ax + b$ with $a < 0$ and $b > 0$.

A.5 Conceptual understanding of linear functions

In this Appendix we provide the full test which was discussed and used in Chapter 5 to study the conceptual understanding of linear functions through the concepts y intercept and slope in linear functions in various representations. We first provide the 12-item kinematics part and then the as-isomorphic-as-possible 12-item mathematics part of the test.

A.5.1 Kinematics

K6: A train is moving along a straight track. The graph shows the position x of the train as function of the time t. The position is in meter, the time in seconds.



Determine the velocity of the train. Explain your answer.

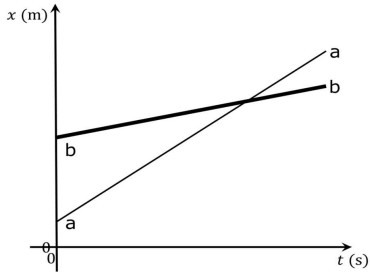
K11: A car is moving along a straight road. The position x of the car as function of the time t is given by

$$x = 5 - 10t.$$

The position is in meter, the time in seconds. Determine the initial position of the car. Explain your answer.

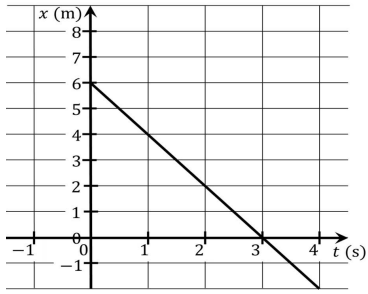
K2: Two cars a and b are moving along a straight road. The graph shows their position x as function of the time t. The position is in meter, the time in seconds.

Which car has the largest velocity? Mark the correct answer and explain.



Car a Car b Cars a and b have the same velocity

K7: A cyclist is riding along a straight bicycle path. The graph shows the position x of the cyclist as function of the time t. The position is in meter, the time in seconds.



Determine the initial position of the cyclist. Explain your answer.

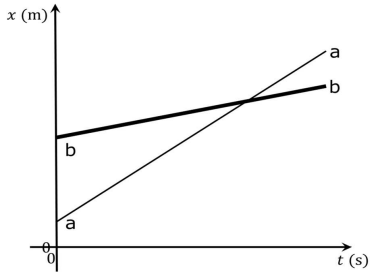
K12: A car is moving along a straight road. The position x of the car as function of the time t is given by

$$x = 3 - 12t.$$

The position is in meter, the time in seconds. Determine the velocity of the car. Explain your answer.

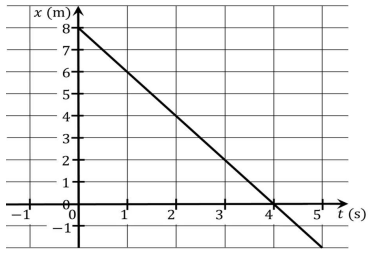
K1: Two cyclists a and b are riding along a straight bicycle path. The graph shows their position x as function of the time t. The position is in meter, the time in seconds.

Which cyclist has the largest initial position? Mark the correct answer and explain.



Cyclist a Cyclist b Cyclists a and b have the same initial position

K8: A car is moving along a straight road. The graph shows the position x of the car as function of the time t. The position is in meter, the time in seconds.



Determine the velocity of the car. Explain your answer.

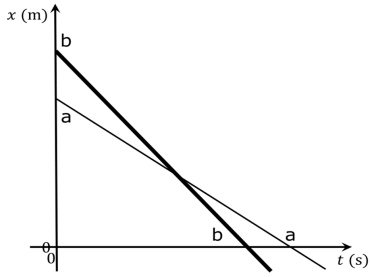
K9: A cyclist is riding along a straight bicycle path. The position x of the cyclist as function of the time t is given by

$$x = 2 + 6t.$$

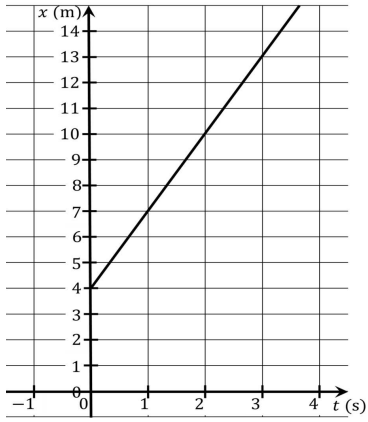
The position is in meter, the time in seconds. Determine the initial position of the cyclist.

K4: Two cyclists a and b are riding along a straight bicycle path. The graph shows their position x as function of the time t. The position is in meter, the time in seconds. Which cyclist has the largest velocity? Mark the correct answer and explain.

Cyclist a Cyclist b Cyclists a and b have the same velocity



K5: A car is moving along a straight road. The graph shows the position x of the car as function of the time t. The position is in meter, the time in seconds.



Determine the initial position of the car.

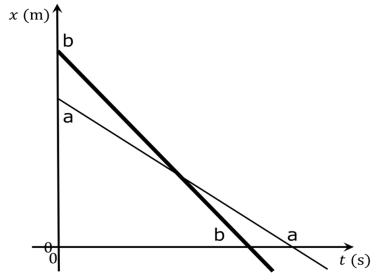
K10:A train is moving along a straight track. The position x of the train as function of the time t is given

$$x = 4 + 8t.$$

The position is in meter, the time in seconds. Determine the velocity of the train. Explain your answer.

K3: Two cars a and b are moving along a straight road. The graph shows their position x as function of the time t. The position is in meter, the time in seconds.

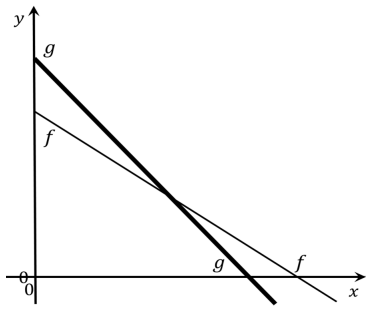
Which cyclist has the largest initial position? Mark the correct answer and explain.



Cyclist a Cyclist b Cyclists a and b have the same initial position

A.5.2 Mathematics

M3: The figure shows the graphs of two functions f and g.



Which of these functions has the largest function value of 0? Mark the correct answer and explain.

f g f and g have the same function value of 0

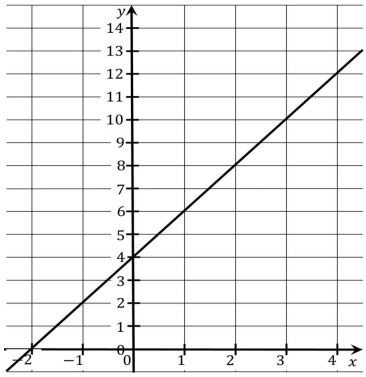
M10: A function f is given by

$$f(x) = 6x + 2.$$

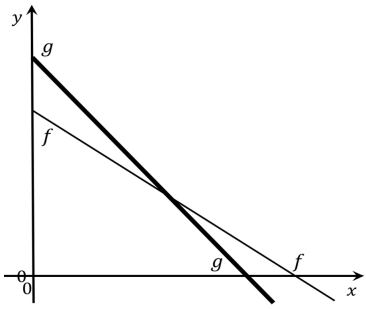
Determine the slope. Explain your answer.

M5: The figure shows the graph of a function f.

Determine the function value of 0.



M4: The figure shows the graphs of two functions f and g.



Which of these functions has the largest slope? Mark the correct answer and explain.

f g f and g have the same slope

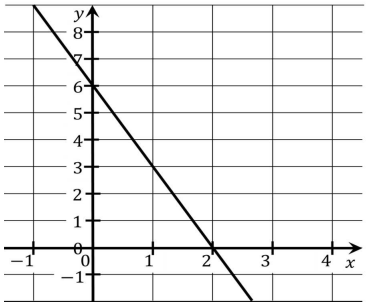
M9: A function f is given by

$$f(x) = 12x + 4.$$

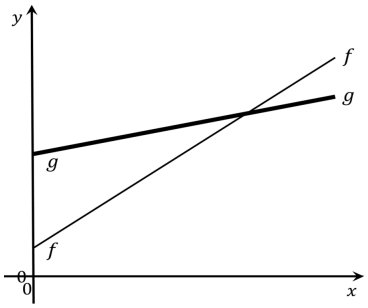
Determine the function value of 0.

M8: The figure shows the graph of a function f.

Determine the slope. Explain your answer.



M1: The figure shows the graphs of two functions f and g.



Which of these functions has the largest function value of 0? Mark the correct answer and explain.

f g f and g have the same function value of 0

M12: A function f is given by

$$f(x) = -6x + 12.$$

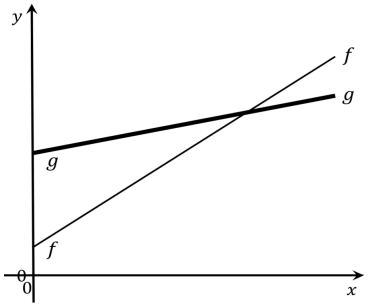
Determine the slope. Explain your answer.

M7: The figure shows the graph of a function f.

Determine the function value of 0. Explain your answer.



M2: The figure shows the graphs of two functions f and g.



Which of these functions has the largest slope? Mark the correct answer and explain.

f g f and g have the same slope

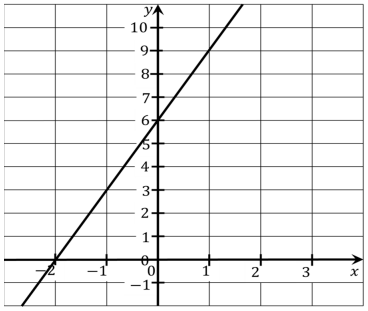
M11: A function f is given by

$$f(x) = -8x + 2.$$

Determine the function value of 0. Explain your answer.

M6: The figure shows the graph of a function f.

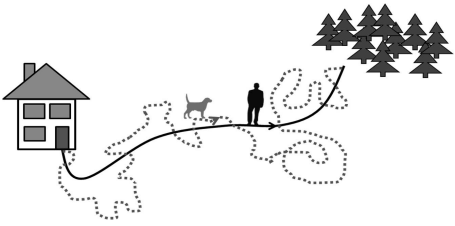
Determine the slope. Explain your answer.



A.6 Modified Kinematics Concept Test

In this Appendix we provide the modified Kinematics Concept Test as was discussed and used in Chapter 6 to study students' conceptual understanding of various kinematics concepts relevant to our students' curriculum.

Item 01: You and your dog Bobby start at home and walk to the park. On the way there, Bobby chases some rabbits, some squirrels and sniffs around. Who made the largest displacement with respect to the starting point at home, when you arrive at the park?



a) Bobby and I made the same displacement.
b) Bobby's displacement is larger than my displacement.
c) My displacement is larger than Bobby's displacement.

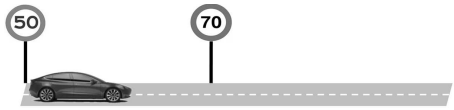
Item 02: You are going on vacation to Los Angeles. Your airplane takes off in Brussels and you fly directly to Los Angeles following the shortest route. John's airplane takes of the next day. He flies from Brussels to Los Angeles via a diverted route (a layover in Chicago). Who travelled the longest distance by airplane?

a) I and John travelled an equally long distance.



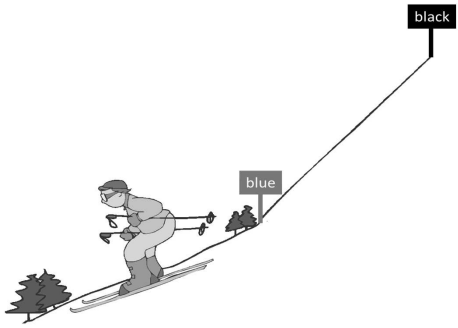
b) John travelled a longer distance.
c) I travelled a longer distance.

Item 001: Lisa is driving with her car to a friend. In doing so, she drives through two speed zones: first she drives 10 minutes at 50 km/h followed by 10 minutes at 70 km/h. What is her average speed along this route?



a) Somewhere between 50 km/h and 60 km/h.
b) 60 km/h.
c) Somewhere between 60 km/h and 70 km/h.
d) None of the above.

Item 002: Walter is skiing in a straight line and as fast as possible down the hill. He first follows a black slope and then a blue slope. In doing so, he first skies 2 km at 70 km/h and afterwards 2 km at 30 km/h. What is his average speed along this route?



a) Somewhere between 30 km/h and 50 km/h.
b) 50 km/h.
c) Somewhere between 50 km/h and 70 km/h.
d) None of the above.

Item 1: Two locomotives move in the same direction along a straight horizontal track. The figure shows the positions of the locomotives at instants 1 through 7. The time interval between the instants is always the same.



Is there an instant at which the two locomotives have the same speed?

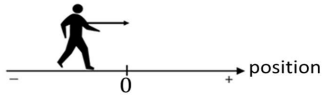
a) Yes, at instant 3.
b) Yes, at instant 5.
c) Yes, at instant 6.
d) Yes, at instants 3 and 6.
e) Yes, sometime between instants 4 and 5.
f) No.

Item 3: Two bodies move along the same straight path. The table shows the positions of the bodies at various instants.

Time in s	0	2	4	6	8	10	12	14
Position in m (body 1)	2	4	7	11	16	22	29	37
Position in m (body 2)	0	4	8	12	16	20	24	28

Is there an instant at which both bodies have the same speed?

a) Yes, at instant 2 s.
b) Yes, at instant 8 s.
c) Yes, at instants 2 s and 8 s.
d) Yes, somewhere between instants 4 s and 6 s.
e) No.

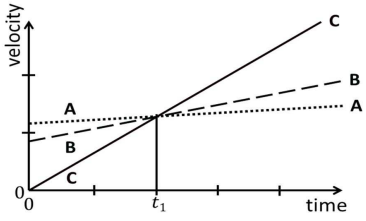


Item 5: A pedestrian walks with a constant speed along the position axis as illustrated in the figure.

Which statement about the pedestrian's velocity with respect to the position axis is correct?

a) It is zero.
b) It is always negative.
c) It is first negative, then positive.
d) It is always positive.
e) Nothing can be said about the sign of the velocity here.

Item 6: The velocity-time graph ($v(t)$-graph) illustrates the motion of three objects A, B and C along the same straight line.



Which statement about the accelerations of objects A, B and C at instant t_1 is correct?

a) Object A has the highest acceleration.
b) Object B has the highest acceleration.
c) Object C has the highest acceleration.
d) All objects have the same non zero acceleration.
e) All objects move with a constant velocity, therefore their acceleration is zero.

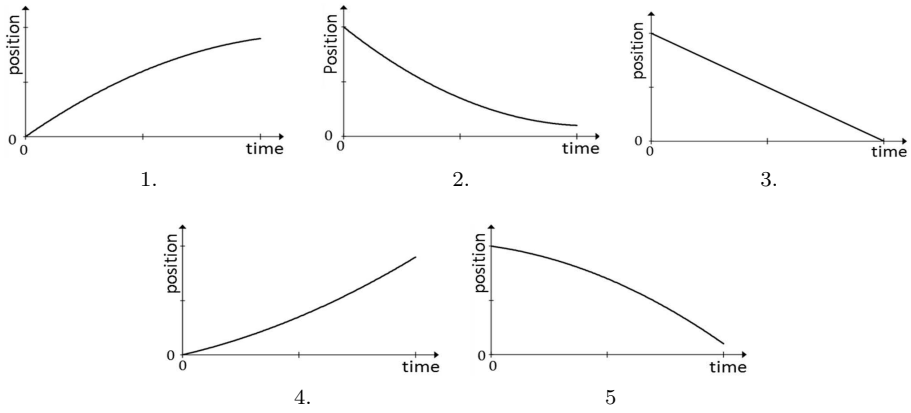
Item 9: A car is moving to the left and is slowing down. The direction of motion is illustrated in the figure.

Which statement about the acceleration of the car is correct?



a) The acceleration points to the left.
b) The acceleration points to the right.
c) Nothing can be said about the direction of the acceleration.
d) The acceleration has no direction.
e) The acceleration is zero.

Item 10: Which of these position-time graphs ($x(t)$-graphs) describes a motion with decreasing speed?

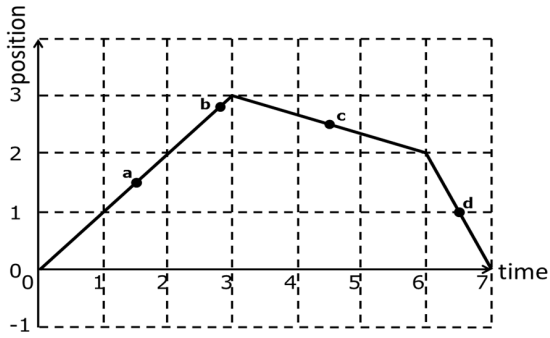


a) 1.
b) 1 and 2.
c) 2 and 5.
d) 4 and 5.
e) 2, 3 and 5.

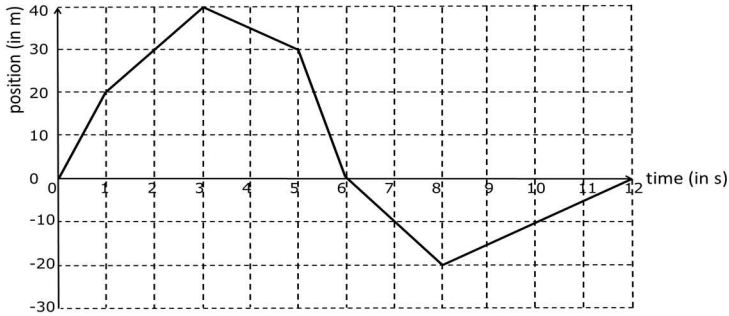
Item 11: The following position-time graph ($x(t)$-graph) illustrates the motion of an object. Four points are marked.

At which point(s) does the object have the highest speed?

a) At point a.
b) At point b.
c) At point c.
d) At point d.
e) At point a and at point b.



Item 16: The following position-time graph ($x(t)$-graph) illustrates the motion of an object.



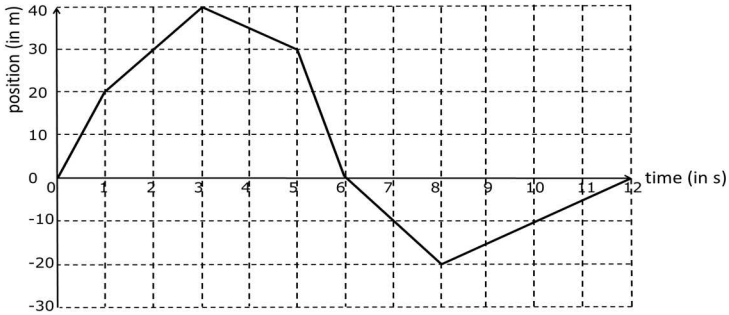
What is the speed at instant $t = 2$ s?

a) 0.1 m/s. c) 15 m/s. e) 30 m/s.
b) 10 m/s. d) 20 m/s. f) 60 m/s.

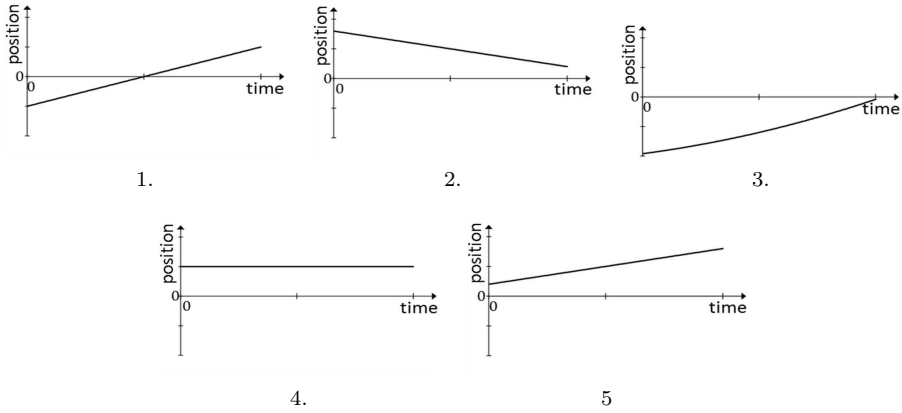
Item 20: The following position-time graph ($x(t)$-graph) illustrates the motion of an object.

What is the speed at instant $t = 10$ s?

a) 0.2 m/s. c) 1 m/s. e) 10 m/s.
b) 0.5 m/s. d) 5 m/s. f) 100 m/s.

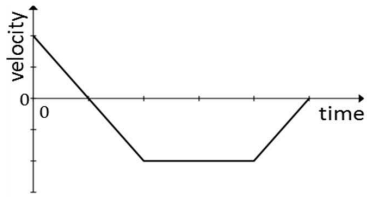


Item 21: Which of the following position-time graphs ($x(t)$-graphs) illustrates the motion of an object of which the speed is always positive?

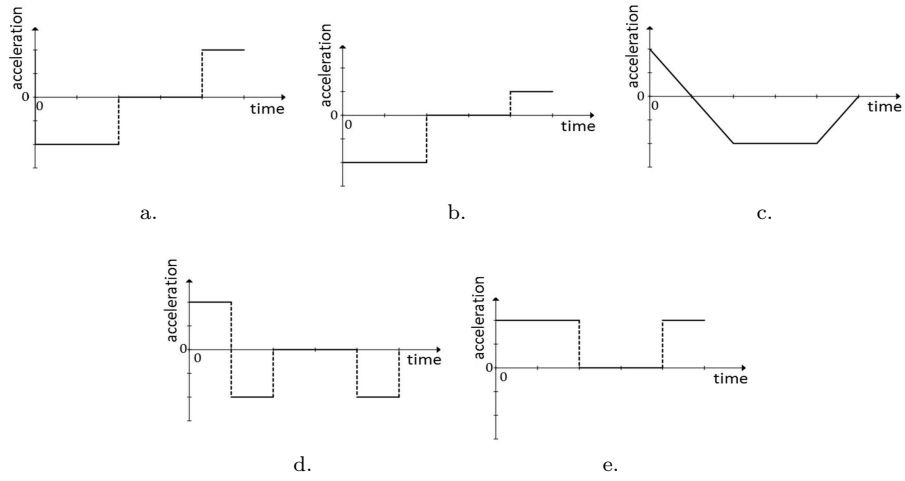


a) 5.
b) 1 and 5.
c) 2 and 5.
d) 4 and 5.
e) 1, 3 and 5.
f) 2, 4, 5.

Item 23: The following velocity-time graph ($v(t)$-graph) illustrates the motion of an object.

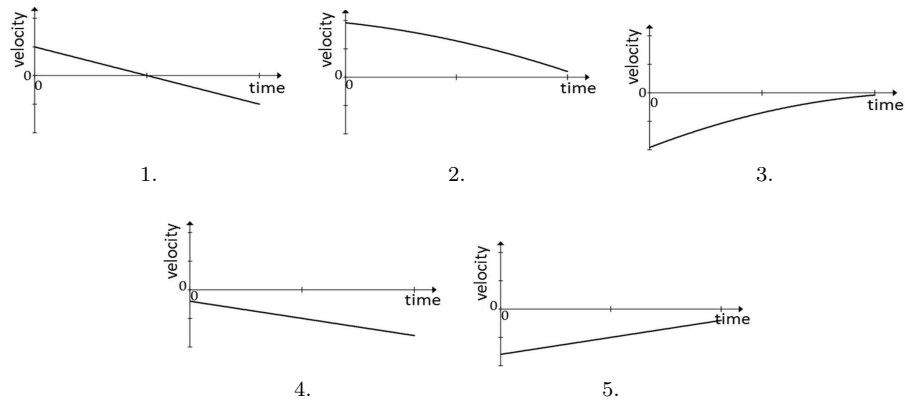


Which acceleration-time graph ($a(t)$-graph) illustrates the same motion?



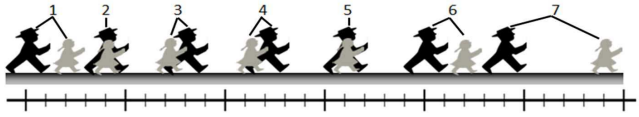
a) a.
b) b.
c) c.
d) d.
e) e.

Item 24: Which of the following velocity-time graphs ($v(t)$-graphs) illustrates the motion of an object of which the acceleration is always negative?



a) 4.
b) 1 and 4.
c) 4 and 5.
d) 1, 2 and 4.
e) 3, 4 and 5.

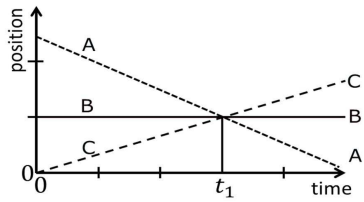
Item 25: Two pedestrians are walking along the same straight road. The figure shows the positions of the pedestrians at instants 1 through 7. The time interval between the instants is always the same.



Is there an instant at which the two pedestrians have the same speed?

a) Yes, at instant 2.
b) Yes, at instant 5.
c) Yes, at instant 4.
d) Yes, at instants 2 and 5.
e) Yes, sometime between instants 3 and 4.
f) No.

Item 26: The following figure shows the position-time graph ($x(t)$-graph) for three cars A, B and C moving along the same road.



For which car does the speedometer show the highest value at instant t_1?

a) Car A.
b) Car B.
c) Car C.
d) Car A and car C.
e) All cars move at the same speed.

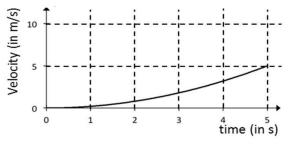
Item 27: Two objects are moving along the same straight road. The table shows the instant at which each object passes certain positions.

Is there a position at which the two bodies have the same speed?

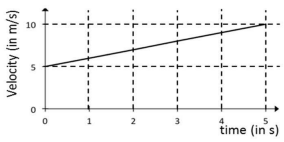
Position in m	10	20	30	40	50	60	70	80
Time in s (object 1)	3	4	6	9	13	18	24	31
Time in s (object 2)	0	4	8	12	16	20	24	28

a) Yes, at position 70 m.
b) Yes, somewhere between positions 20 m an 30 m.
c) Yes, at positions 20 m and 70 m.
d) Yes, somewhere between positions 40 m and 50 m.
e) No.

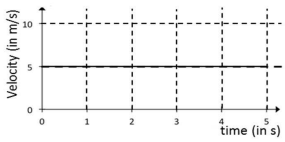
Item 30: The following velocity-time graphs ($v(t)$-graphs) illustrate the motion of five objects A, B, C, D and E. Which of these objects covers the shortest distance in the time interval shown?



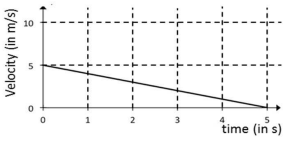
A.



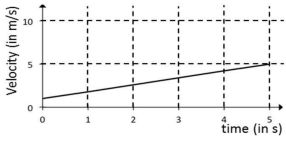
B.



C.



D.



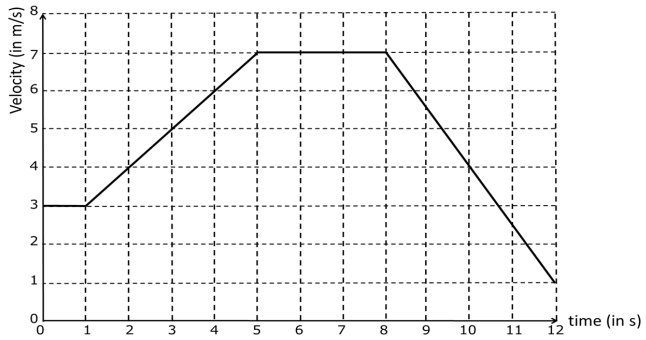
E.

a) A.
b) B.
c) C.
d) D.
e) E.

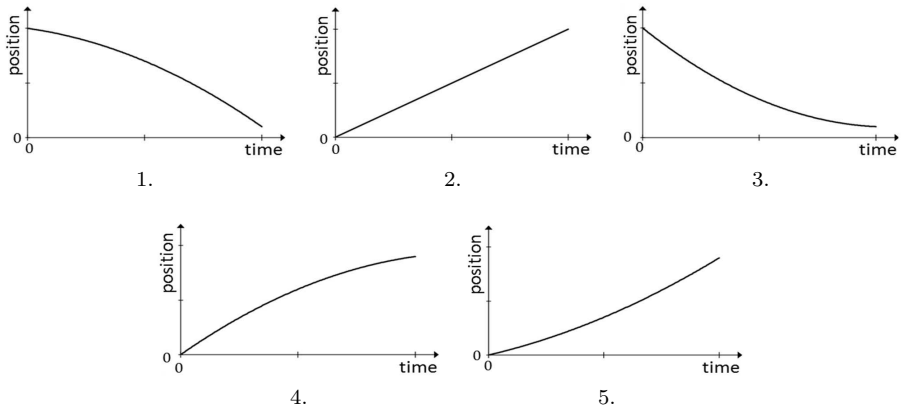
Item 32: The following velocity-time graph ($v(t)$-graph) illustrates the motion of an object along a straight line.

What is the magnitude of the acceleration of the object at instant $t = 3$ s?

a) 0.5 m/s^2
b) 1 m/s^2
c) 1.6 m/s^2 (= 5/3 m/s^2)
d) 4 m/s^2
e) 5 m/s^2
f) 15 m/s^2



Item 33: Which of the following position-time graphs ($x(t)$-graphs) illustrates the motion of an object of which the speed increases.

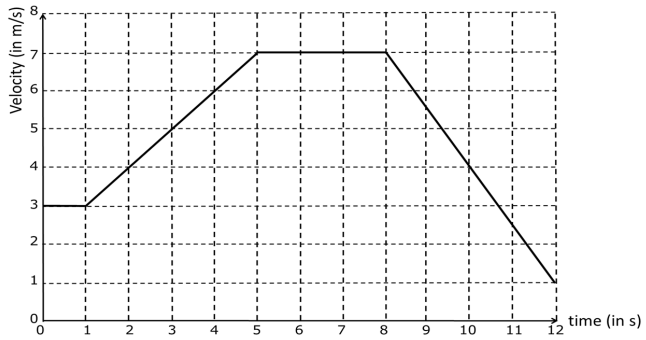


a) 5.
b) 1 and 5.
c) 3 and 4.
d) 4 and 5.
e) 2, 4 and 5.

Item 35: The following velocity-time graph ($v(t)$-graph) illustrates the motion of an object along a straight line.

What is the distance covered by this object between instants $t = 1$ s and $t = 5$ s?

a) 4 m.
b) 5 m.
c) 8 m.
d) 16 m.
e) 20 m.
f) 32 m.



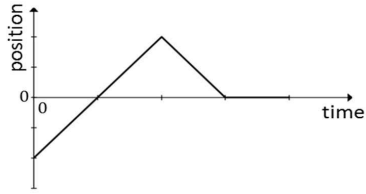
Item 36: A helicopter is landing. It moves vertically downward and decreases its speed. The direction of motion is indicated in the figure.



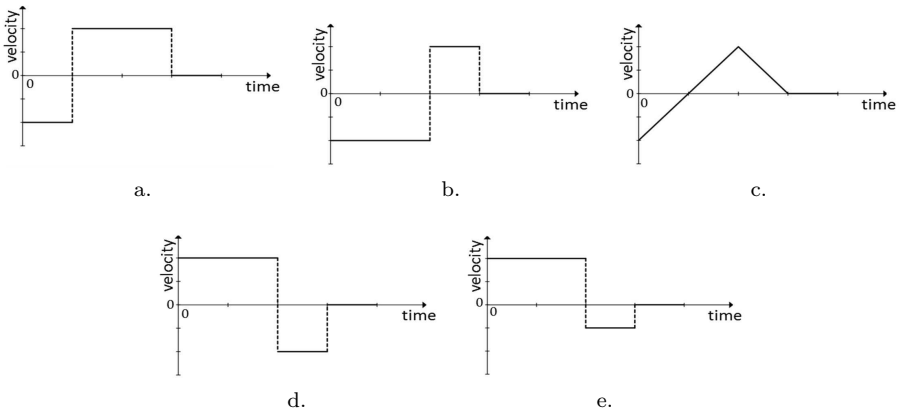
Which statement about the acceleration of the helicopter is correct?

a) The acceleration is zero.
b) The acceleration is directed downward.
c) The acceleration is directed upwards.
d) Nothing can be said about the direction of the acceleration.
e) The acceleration does not have a direction.

Item 37: The following position-time graph ($x(t)$-graph) illustrates the motion of an object.

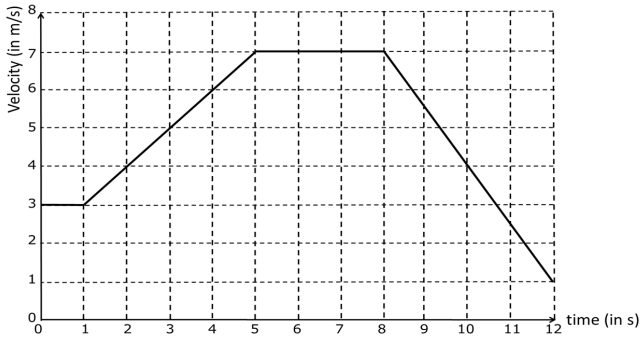


Which velocity-time graph ($v(t)$-graph) illustrates the same motion?



a) a. c) c. e) e.
b) b. d) d.

Item 38: The following velocity-time graph ($v(t)$-graph) illustrates the motion of an object along a straight line.

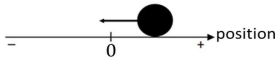


What is the distance covered by this object between instants $t = 8$ s and $t = 12$ s?

a) 1.5 m. c) 6 m. e) 16 m.
b) 5 m. d) 12 m. f) 24 m.

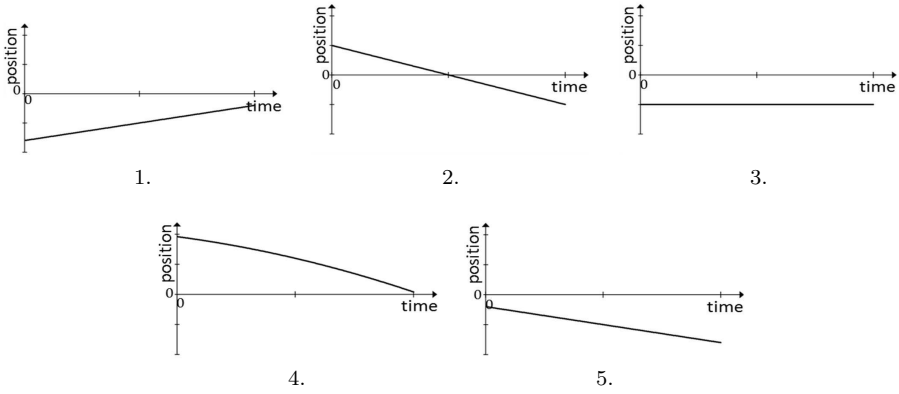
Item 41: A bal rols with a constant velocity along the position axis as shown in the figure.

Which statement about the velocity of the bal with respect to the position axis is correct?



a) It is zero.
b) It is always positive.
c) It is always negative.
d) It is first positive, then negative.
e) Nothing can be said about the sign of the velocity here.

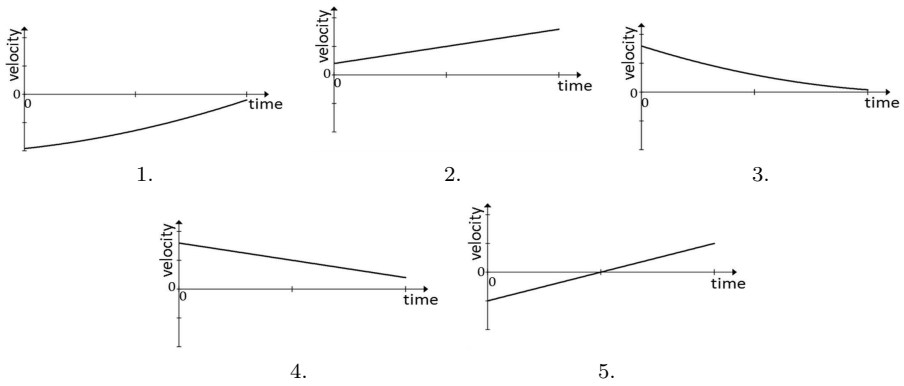
Item 43: Which of the following position-time graphs ($x(t)$-graphs) illustrates the motion of an object of which the velocity is always negative?



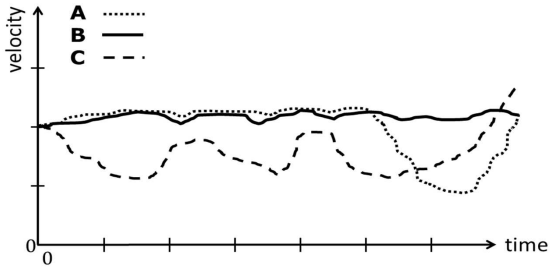
a) 5.
b) 1 and 5.
c) 2 and 5.
d) 1, 3 and 5.
e) 2, 4 and 5.
f) 3 and 5.

Item 48: Which of the following velocity-time graphs ($v(t)$-graphs) illustrates the motion of an object of which the acceleration is always positive?

a) 2.
b) 2 and 4.
c) 2 and 5.
d) 1, 2 and 5.
e) 2, 3 and 4.



Item 49: The following velocity-time graph ($v(t)$-graph) illustrates the velocities of three runners A, B and C during the same time interval.



Which runner covers the largest distance during this time interval?

a) Runner A.
b) Runner B.
c) Runner B and runner C.
d) All runners cover the same distance.
e) The graph does not provide sufficient information to determine this.

Bibliography

[1] A. L. Porter and I. Rafols, "Is science becoming more interdisciplinary? measuring and mapping six research fields over time," *Scientometrics*, vol. 81, p. 719, Apr 2009.

[2] M. Sanders, "STEM, STEM education, STEMmania," *The Technology Teacher*, vol. 68, no. 4, pp. 20–26, 2008.

[3] H. Knipprath, L. Thibaut, M. Buyse, S. Ceuppens, H. D. Loof, J. D. Meester, L. Goovaerts, A. Struyf, J. B. Pauw, F. Depaepe, J. Deprez, M. D. Cock, L. Hellinckx, G. Langie, K. Struyven, D. V. D. Velde, P. V. Petegem, and W. Dehaene, "Stem education in flanders: How stem@school aims to foster stem literacy and a positive attitude towards stem," *IEEE Instrumentation Measurement Magazine*, vol. 21, pp. 36–40, June 2018.

[4] W. E. Forum, "The future of jobs: Employment, skills and workforce strategy for the fourth industrial revolution," in *World Economic Forum*, 2016.

[5] M. V. Bøe, E. K. Henriksen, T. Lyons, and C. Schreiner, "Participation in science and technology: young people's achievement-related choices in late-modern societies," *Studies in Science Education*, vol. 47, pp. 37–72, mar 2011.

[6] R. W. Bybee, J. A. Taylor, A. Gardner, P. Van Scotter, J. C. Powell, A. Westbrook, and N. Landes, "The bscs 5e instructional model: Origins and effectiveness," *Colorado Springs, Co: BSCS*, vol. 5, pp. 88–98, 2006.

[7] M. Honey, G. Pearson, and H. Schweingruber, *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*. Washington, D.C., USA: National Academies Press, feb 2014.

[8] D. onderwijs en vorming, "Stem-kader voor het vlaamse onderwijs - principes en doelstellingen," tech. rep., Vlaamse overheid, Brussels, Belgium, 2015.

[9] D.-G. for Research and I. E. Commission), "Science education for responsible citizenship," tech. rep., European Union, Brussels, Belgium, 2015.

[10] N. R. Council, *Successful K-12 STEM Education: Identifying Effective Approaches in Science, Technology, Engineering, and Mathematics*. Washington, DC: The National Academies Press, 2011.

[11] S. Marginson, R. Tytler, B. Freeman, and K. Roberts, "STEM: Country comparisons," Tech. Rep. June, Australian Council of Learned Academies, Melbourne, Australia, 2013.

[12] M. Caprile, R. Palmén, P. Sanz, and G. Dente, "Encouraging STEM Studies for the Labour Market," 2015.

[13] A. P. Carnevale, N. Smith, and J. Strohl, "Help wanted: Projections of job and education requirements through 2018," tech. rep., Georgetown University Center on Education and the Workforce, Washington D.C., USA, 2010.

[14] M. Hossain and M. G. Robinson, "How to Motivate US Students to Pursue STEM (Science, Technology, Engineering and Mathematics) Careers," *US-China Education Review*, vol. 4, pp. 442–451, 2012.

[15] J. Rothwell, "Still searching: Job vacancies and STEM skills," tech. rep., Metropolitan Policy Program: Brookings Institution, Washington D.C., USA, 2014.

[16] N. A. of Sciences, N. A. of Engineering, and I. of Medicine, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. Washington, DC: The National Academies Press, 2007.

[17] I. V. Mullis, M. O. Martin, and T. Loveless, "years of timss: International trends in mathematics and science achievement, curriculum, and instruction," *TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA)*, 20.

[18] C. Schreiner and S. Sjøberg, "The ROSE project An overview and key findings," tech. rep., University of Oslo, Oslo, Norway, 2010.

[19] B. Thomas and J. J. Watters, "Perspectives on Australian, Indian and Malaysian approaches to STEM education," *International Journal of Educational Development*, vol. 45, pp. 42–53, nov 2015.

[20] OECD, *PISA 2015 Results (Volume I)*. Paris, France: The Organisation for Economic Co-operation and Development (OECD), 2016.

[21] D. Sands, "Concepts and conceptual understanding: What are we talking about?," *New Directions in the Teaching of Physical Sciences*, pp. 7–11, 06 2014.

[22] L. C. McDermott, "Oersted medal lecture 2001: "physics education research—the key to student learning"," *American Journal of Physics*, vol. 69, no. 11, pp. 1127–1137, 2001.

[23] J. Clement, "Students' preconceptions in introductory mechanics," *American Journal of Physics*, vol. 50, no. 1, pp. 66–71, 1982.

[24] J. M. Aguirre, "Student preconceptions about vector kinematics," *The Physics Teacher*, vol. 26, no. 4, pp. 212–216, 1988.

[25] J. M. Aguirre and G. Rankin, "College students' conceptions about vector kinematics," *Physics Education*, vol. 24, no. 5, p. 290, 1989.

[26] D. Hestenes, M. Wells, and G. Swackhamer, "Force concept inventory," *The Physics Teacher*, vol. 30, no. 3, pp. 141–158, 1992.

[27] J. Woolnough, "How do students learn to apply their mathematical knowledge to interpret graphs in physics?," *Research in Science Education*, vol. 30, pp. 259–267, Sep 2000.

[28] D. Gil-Perez and A. M. Pessoa de Carvalho, "Physics teacher training: Analysis and proposals," in *Connecting research in physics education with teacher education* (A. Tiberghien, L. Jossem E., and J. Barojas, eds.), pp. 42–47, International Commission on Physics Education, 1997.

[29] J. Smith, A. Disessa, and J. Roschelle, "Misconceptions reconceived: A constructivist analysis of knowledge in transition," *Journal of the Learning Sciences*, vol. 3, 11 1993.

[30] D. Hestenes and M. Wells, "A mechanics baseline test," *The Physics Teacher*, vol. 30, no. 3, pp. 159–166, 1992.

[31] R. J. Beichner, "Testing student interpretation of kinematics graphs," *American Journal of Physics*, vol. 62, no. 8, pp. 750–762, 1994.

[32] P. Nieminen, A. Savinainen, and J. Viiri, "Force concept inventory-based multiple-choice test for investigating students' representational consistency," *Phys. Rev. ST Phys. Educ. Res.*, vol. 6, p. 020109, Aug 2010.

[33] F. M. Goldberg and L. C. McDermott, "Student difficulties in understanding image formation by a plane mirror," *The Physics Teacher*, vol. 24, no. 8, pp. 472–481, 1986.

[34] F. M. Goldberg and L. C. McDermott, "An investigation of student understanding of the real image formed by a converging lens or concave mirror," *American Journal of Physics*, vol. 55, no. 2, pp. 108–119, 1987.

[35] A. Lichtenberger, C. Wagner, S. I. Hofer, E. Stern, and A. Vaterlaus, "Validation and structural analysis of the kinematics concept test," *Phys. Rev. Phys. Educ. Res.*, vol. 13, p. 010115, Apr 2017.

[36] E. S. Jennifer, "Mathematical beliefs and conceptual understanding of the limit of a function," *Journal for Research in Mathematics Education*, vol. 31, p. 258, 05 2000. Copyright - Copyright National Council of Teachers of Mathematics May 2000; Last updated - 2014-05-16; CODEN - JRMEDN.

[37] P. Barniol and G. Zavala, "Test of understanding of vectors: A reliable multiple-choice vector concept test," *Phys. Rev. ST Phys. Educ. Res.*, vol. 10, p. 010121, Jun 2014.

[38] M. Yerushalmy, "Designing representations: Reasoning about functions of two variables," *Journal for Research in Mathematics Education*, vol. 28, p. 431, 07 1997. Copyright - Copyright National Council of Teachers of Mathematics Jul 1997; Last updated - 2014-05-23; CODEN - JRMEDN.

[39] R. Ely, "Nonstandard student conceptions about infinitesimals," *Journal for Research in Mathematics Education*, vol. 41, p. 117, 03 2010. Copyright - Copyright National Council of Teachers of Mathematics Mar 2010; Document feature - ; Tables; People - Leibniz, Gottfried Wilhelm (1646-1716; Last updated - 2014-05-16; CODEN - JRMEDN; SubjectsTermNotLitGenreText - Leibniz, Gottfried Wilhelm (1646-1716.

[40] S. M. Matteson, "Mathematical literacy and standardized mathematical assessments," *Reading Psychology*, vol. 27, pp. 205–233, 04 2006. Date revised - 2006-09-01; Last updated - 2016-09-27; CODEN - RRPSDW; SubjectsTermNotLitGenreText - Mathematics (51850); Reading Processes (71150); Comprehension (13950); Achievement Tests (00100); Reading Ability (70500); Teacher Education (87850.

[41] T. Dreyfus and T. Eisenberg, "Intuitive functional concepts: A baseline study on intuitions," *Journal for Research in Mathematics Education*, vol. 13, no. 5, pp. 360–380, 1982.

[42] T. H. Kjeldsen and J. Lützen, "Interactions between mathematics and physics: The history of the concept of function—teaching with and about nature of mathematics," *Science & Education*, vol. 24, pp. 543–559, Jul 2015.

[43] E. F. Redish and E. Kuo, "Language of physics, language of math: Disciplinary culture and dynamic epistemology," *Science & Education*, vol. 24, pp. 561–590, Jul 2015.

[44] E. B. Pollock, J. R. Thompson, and D. B. Mountcastle, "Student Understanding Of The Physics And Mathematics Of Process Variables In P-V Diagrams," in *Physics Education Research* (L. Hsu, C. Henderson, and L. McCullough, eds.), vol. 951 of *American Institute of Physics Conference Series*, pp. 168–171, Nov. 2007.

[45] W. M. Christensen and J. R. Thompson, "Investigating graphical representations of slope and derivative without a physics context," *Phys. Rev. ST Phys. Educ. Res.*, vol. 8, p. 023101, Jul 2012.

[46] M. Planinic, Z. Milin-Sipus, H. Katic, A. Susac, and L. Ivanjek, "Comparison of student understanding of line graph slope in physics and mathematics," *International Journal of Science and Mathematics Education*, vol. 10, pp. 1393–1414, Dec 2012.

[47] M. Planinic, L. Ivanjek, A. Susac, and Z. Milin-Sipus, "Comparison of university students' understanding of graphs in different contexts," *Phys. Rev. ST Phys. Educ. Res.*, vol. 9, p. 020103, Jul 2013.

[48] L. Ivanjek, A. Susac, M. Planinic, A. Andrasevic, and Z. Milin-Sipus, "Student reasoning about graphs in different contexts," *Phys. Rev. Phys. Educ. Res.*, vol. 12, p. 010106, Feb 2016.

[49] B. Keogh and S. Naylor, "Concept cartoons, teaching and learning in science: an evaluation," *International Journal of Science Education*, vol. 21, no. 4, pp. 431–446, 1999.

[50] E. Mazur, *Peer instruction: A User's Manual*. Upper Saddle River, NJ: Prentice Hall, 1997.

[51] K. Miller, J. Schell, A. Ho, B. Lukoff, and E. Mazur, "Response switching and self-efficacy in peer instruction classrooms," *Phys. Rev. ST Phys. Educ. Res.*, vol. 11, p. 010104, Feb 2015.

[52] M. J. Lage, G. J. Platt, and M. Treglia, "Inverting the classroom: A gateway to creating an inclusive learning environment," *The Journal of Economic Education*, vol. 31, no. 1, pp. 30–43, 2000.

[53] D. Mills, B. McKittrick, P. Mulhall, and S. Feteris, "CUP: cooperative learning that works," *Teaching Physics*, vol. 34, no. 1, pp. 11–16, 1999.

[54] T. R. Kelley and J. G. Knowles, "A conceptual framework for integrated STEM education," *International Journal of STEM Education*, vol. 3, p. 11, dec 2016.

[55] L. D. English, "Stem education k-12: perspectives on integration," *International Journal of STEM Education*, vol. 3, p. 3, Mar 2016.

[56] T. Hatch, "The differences in theory that matter in the practice of school improvement," *American Educational Research Journal*, vol. 35, no. 1, pp. 3–31, 1998.

[57] G. Venville, L. J. Rennie, and J. Wallace, "Disciplinary versus Integrated Curriculum: The challenge for school science," *The New Critic*, no. 10, pp. 1–9, 2009.

[58] D. S. Hartzler, *A meta-analysis of studies conducted on integrated curriculum programs and their effects on student achievement*. PhD thesis, School of Education, Indiana University, USA, 2000.

[59] K. Becker and K. Park, "Effects of integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students' learning: A preliminary meta-analysis," *Journal of STEM Education: Innovations and Research*, vol. 12, no. 5 & 6, pp. 23–37, 2011.

[60] R. H. Bruning, J. G. Schaw, M. M. Norby, and R. R. Ronning, *Cognitive psychology and instruction*. Columbus, Ohio, USA: Pearson, 4th ed., 2004.

[61] L. Thibaut, S. Ceuppens, H. De Loof, J. De Meester, L. Goovaerts, A. Struyf, J. Boeve-De Pauw, M. De Cock, W. Dehaene, J. Deprez, L. Hellinckx, H. Knipprath, G. Langie, K. Struyven, D. Van de Velde, P. Van Petegem, and F. Depaepe, "Integrated stem education: A systematic review of instructional practices in secondary education," *European Journal of STEM Education*, vol. 3, no. 1, p. 02, 2018.

[62] J. McKinley, "Critical argument and writer identity: Social constructivism as a theoretical framework for efl academic writing," *Critical Inquiry in Language Studies*, vol. 12, no. 3, pp. 184–207, 2015.

[63] L. C. McDermott, "Students' conceptions and problem solving in mechanics," in *Connecting research in physics education with teacher education* (A. Tiberghien, L. Jossem E., and J. Barojas, eds.), pp. 42–47, International Commission on Physics Education, 1997.

[64] D. Perkins, *Making learning whole: how seven principles of teaching can transform education*. San Francisco, California, USA: Jossey-Bass, 1st ed., 2009.

[65] G. Pearson, "National academies piece on integrated stem," *The Journal of Educational Research*, vol. 110, no. 3, pp. 224–226, 2017.

[66] STEM@school, "Stem@school module 31 - the green wave (autonomous driving car)," 2018.

[67] STEM@school, "Stem@school module 32 - a museum security system," 2018.

[68] STEM@school, "Stem@school module 33 - a rehabilitation device," 2018.

[69] STEM@school, "Stem@school module 41 - a device to throw/catch (tennis) balls," 2018.

[70] STEM@school, "Stem@school module 42 - a passive house," 2018.

[71] H. De Loof, "Educating engaged and competent students for stem: effects of integrated stem education," 2019.

[72] J. De Meester, "Designing istem learning materials for secondary education," 2019.

[73] L. Goovaerts, "Development and assessment of istem competencies," 2019.

[74] A. Struyf, "How to engage students for stem? analysing the ingredients for a motivational cocktail in the learning environment," 2019.

[75] L. Thibaut, "Implementing integrated stem: teachers' attitudes, instructional practices and students' learning outcomes," 2018.

[76] A. R. P. de Ataíde and I. M. Greca, "Epistemic views of the relationship between physics and mathematics: Its influence on the approach of undergraduate students to problem solving," *Science & Education*, vol. 22, pp. 14052–1421, 2013.

[77] R. A. Karam, "Matemática como estruturante e física como motivação: uma análise de concepções sobre as relações entre matemática e física," *Proceedings VI Encontro Nacional de Pesquisa em Educação em Ciências*, 2007.

[78] D. De Bock, W. Van Dooren, and L. Verschaffel, "Students' understanding of proportional, inverse proportional, and affine functions: two studies on the role of external representations," *International Journal of Science and Mathematics Education*, vol. 13, pp. 47–69, Mar 2015.

[79] C. M. Czerniak, W. B. Weber Jr, A. Sandmann, and J. Ahern, "A literature review of science and mathematics integration," *School Science and Mathematics*, vol. 99, no. 8, pp. 421–430, 1999.

[80] D. E. Meltzer, "Relation between students' problem-solving performance and representational format," *American Journal of Physics*, vol. 73, no. 5, pp. 463–478, 2005.

[81] P. B. Kohl and N. D. Finkelstein, "Student representational competence and self-assessment when solving physics problems," *Phys. Rev. ST Phys. Educ. Res.*, vol. 1, p. 010104, Oct 2005.

[82] V. overheid, "Secundair onderwijs - tweede graad aso - natuurwetenschappen - vakgebonden eindtermen," 2012.

[83] T. N. C. of Teachers of Mathematics, "Principles and standards for school mathematics," 2000.

[84] A. Acevedo Nistal, W. Van Dooren, G. Clarebout, J. Elen, and L. Verschaffel, "Conceptualising, investigating and stimulating representational flexibility in mathematical problem solving and learning: a critical review," *ZDM*, vol. 41, pp. 627–636, Oct 2009.

[85] M. De Cock, "Representation use and strategy choice in physics problem solving," *Phys. Rev. ST Phys. Educ. Res.*, vol. 8, p. 020117, Nov 2012.

[86] S. E. Ainsworth, P. A. Bibby, and D. J. Wood, "Analysing the costs and benefits of multi-representational learning environments," in *Learning with multiple representations*, pp. 120–134, Pergamon Amsterdam, 1998.

[87] R. Even, "Factors involved in linking representations of functions," *The Journal of Mathematical Behavior*, vol. 17, no. 1, pp. 105 – 121, 1998. Representations and the Psychology of Mathematics Education: Part I.

[88] K. N. Bieda and M. J. Nathan, "Representational disfluency in algebra: evidence from student gestures and speech," *ZDM*, vol. 41, pp. 637–650, Oct 2009.

[89] S. Ainsworth, P. Bibby, and D. Wood, "Examining the effects of different multiple representational systems in learning primary mathematics," *The Journal of the Learning Sciences*, vol. 11, no. 1, pp. 25–61, 2002.

[90] R. Duval, "A cognitive analysis of problems of comprehension in a learning of mathematics," *Educational Studies in Mathematics*, vol. 61, pp. 103–131, 2006.

[91] D. Kirsh, "Thinking with external representations," *AI & SOCIETY*, vol. 25, pp. 441–454, Nov 2010.

[92] P. B. Kohl and N. D. Finkelstein, "Effect of instructional environment on physics students' representational skills," *Phys. Rev. ST Phys. Educ. Res.*, vol. 2, p. 010102, Jan 2006.

[93] M. T. H. Chi, P. J. Feltovich, and R. Glaser, "Categorization and representation of physics problems by experts and novices," *Cognitive Science*, vol. 5, no. 2, pp. 121–152, 1981.

[94] P. B. Kohl and N. D. Finkelstein, "Patterns of multiple representation use by experts and novices during physics problem solving," *Phys. Rev. ST Phys. Educ. Res.*, vol. 4, p. 010111, Jun 2008.

[95] F. M. Goldberg and J. H. Anderson, "Student difficulties with graphical representations of negative values of velocity," *The Physics Teacher*, vol. 27, no. 4, pp. 254–260, 1989.

[96] A. Acevedo Nistal, W. Van Dooren, and L. Verschaffel, "What counts as a flexible representational choice? an evaluation of students' representational choices to solve linear function problems," *Instructional Science*, vol. 40, pp. 999–1019, Nov 2012.

[97] D. Niemi, "Assessing conceptual understanding in mathematics: Representations, problem solutions, justifications, and explanations," *The Journal of Educational Research*, vol. 89, no. 6, pp. 351–363, 1996.

[98] K. Adu-Gyamfi and M. J. Bossé, "Processes and reasoning in representations of linear functions," *International Journal of Science and Mathematics Education*, vol. 12, pp. 167–192, Feb 2014.

[99] P. W. Thompson, "Notations, conventions, and constraints: Contributions to effective uses of concrete materials in elementary mathematics," *Journal for research in mathematics education*, pp. 123–147, 1992.

[100] R. Cox and P. Brna, "Supporting the use of external representations in problem solving: The need for flexible learning environments.," 1995.

[101] S. Ainsworth, D. Wood, and C. O'Malley, "There is more than one way to solve a problem: Evaluating a learning environment that supports the development of children's multiplication skills," *Learning and Instruction*, vol. 8, no. 2, pp. 141–157, 1998.

[102] T. Seufert, "Supporting coherence formation in learning from multiple representations," *Learning and Instruction*, vol. 13, no. 2, pp. 227 – 237, 2003. External and Internal Representations in Multimedia Learning.

[103] D. Treagust, G. Chittleborough, and T. Mamiala, "The role of submicroscopic and symbolic representations in chemical explanations," *International Journal of Science Education*, vol. 25, no. 11, pp. 1353–1368, 2003.

[104] M. Yerushalmy, "Student perceptions of aspects of algebraic function using multiple representation software," *Journal of Computer Assisted Learning*, vol. 7, no. 1, pp. 42–57, 1991.

[105] H. Tabachneck, A. Leonardo, and H. Simon, "How does an expert use a graph? a model of visual and verbal inferencing in economics," in *Proceedings of the 16th annual conference of the Cognitive Science Society*, vol. 842, p. 847, Erlbaum Hillsdale, NJ, 1994.

[106] D. Corradi, J. Elen, and G. Clarebout, "Understanding and enhancing the use of multiple external representations in chemistry education," *Journal of Science Education and Technology*, vol. 21, pp. 780–795, Dec 2012.

[107] G. Leinhardt, O. Zaslavsky, and M. K. Stein, "Functions, graphs, and graphing: Tasks, learning, and teaching," *Review of Educational Research*, vol. 60, no. 1, pp. 1–64, 1990.

[108] T. Wemyss and P. van Kampen, "Categorization of first-year university students' interpretations of numerical linear distance-time graphs," *Phys. Rev. ST Phys. Educ. Res.*, vol. 9, p. 010107, Feb 2013.

[109] J. D. Davis, "Real-world contexts, multiple representations, student-invented terminology, and y-intercept," *Mathematical Thinking and Learning*, vol. 9, no. 4, pp. 387–418, 2007.

[110] D. E. Trowbridge and L. C. McDermott, "Investigation of student understanding of the concept of velocity in one dimension," *American Journal of Physics*, vol. 48, no. 12, pp. 1020–1028, 1980.

[111] R. J. Beichner, "Testing student interpretation of kinematics graphs," *American Journal of Physics*, vol. 62, no. 8, pp. 750–762, 1994.

[112] D. E. Trowbridge and L. C. McDermott, "Investigation of student understanding of the concept of acceleration in one dimension," *American Journal of Physics*, vol. 49, no. 3, pp. 242–253, 1981.

[113] N. S. of Professional Engineers, "NSPE Position Statement No. 02-1768," 2018.

[114] R. W. Bybee, "Advancing stem education: A 2020 vision," *Technology and Engineering Teacher*, vol. 70, no. 1, p. 30, 2010.

[115] J. M. Furner and D. D. Kumar, "The mathematics and science integration argument: A stand for teacher education," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 3, no. 3, pp. 185–189, 2007.

[116] J. A. Beane, "Curriculum integration and the disciplines of knowledge," *The Phi Delta Kappan*, vol. 76, no. 8, pp. 616–622, 1995.

[117] E. T. Hinde, "Revisiting curriculum integration: A fresh look at an old idea," *The social studies*, vol. 96, no. 3, pp. 105–111, 2005.

[118] N. Mustafa, Z. Ismail, Z. Tasir, and M. N. H. Mohamad Said, "A meta-analysis on effective strategies for integrated stem education," *Advanced Science Letters*, vol. 22, no. 12, pp. 4225–4228, 2016.

[119] J. L. Riskowski, C. D. Todd, B. Wee, M. Dark, and J. Harbor, "Exploring the effectiveness of an interdisciplinary water resources engineering module in an eighth grade science course," *International journal of engineering education*, vol. 25, no. 1, p. 181, 2009.

[120] H.-H. Wang, T. J. Moore, G. H. Roehrig, and M. S. Park, "Stem integration: Teacher perceptions and practice," *Journal of Pre-College Engineering Education Research (J-PEER)*, vol. 1, no. 2, p. 2, 2011.

[121] L. S. Nadelson and A. L. Seifert, "Integrated stem defined: Contexts, challenges, and the future," *The Journal of Educational Research*, vol. 110, no. 3, pp. 221–223, 2017.

[122] M. Stohlmann, T. J. Moore, and G. H. Roehrig, "Considerations for teaching integrated stem education," *Journal of Pre-College Engineering Education Research (J-PEER)*, vol. 2, no. 1, p. 4, 2012.

[123] L. Hardy, "High tech high," *American School Board Journal*, vol. 188, no. 7, pp. 12–15, 2001.

[124] E. W. Eckman, M. A. Williams, and M. B. Silver-Thorn, "An integrated model for stem teacher preparation: the value of a teaching cooperative educational experience," *Journal of STEM Teacher Education*, vol. 51, no. 1, p. 8, 2016.

[125] L. Shulman, "Knowledge and teaching: Foundations of the new reform," *Harvard Educational Review*, vol. 57, no. 1, pp. 1–23, 1987.

[126] H. El-Deghaidy and N. Mansour, "Science teachers' perceptions of stem education: Possibilities and challenges," *International Journal of Learning and Teaching*, vol. 1, no. 1, pp. 51–54, 2015.

[127] A. Asghar, R. Ellington, E. Rice, F. Johnson, and G. M. Prime, "Supporting stem education in secondary science contexts," *Interdisciplinary Journal of Problem-Based Learning*, vol. 6, no. 2, p. 4, 2012.

[128] B. S. Barrett, A. L. Moran, and J. E. Woods, "Meteorology meets engineering: an interdisciplinary stem module for middle and early secondary school students," *International Journal of STEM Education*, vol. 1, p. 6, Sep 2014.

[129] L. Gentile, L. Caudill, M. Fetea, A. L. Hill, K. Hoke, B. Lawson, O. Z. Lipan, M. Kerckhove, C. A. Parish, K. J. Stenger, *et al.*, "Challenging disciplinary boundaries in the first year: A new introductory integrated science course for stem majors," *Journal of College Science Teaching*, vol. 41, no. 5, p. 44, 2012.

[130] T. J. Moore, S. S. Guzey, and A. Brown, "Greenhouse design: An engineering unit," *Science Scope*, vol. 37, no. 7, pp. 51–57, 2014.

[131] R. Gresnigt, R. Taconis, H. van Keulen, K. Gravemeijer, and L. Baartman, "Promoting science and technology in primary education: a review of integrated curricula," *Studies in Science Education*, vol. 50, no. 1, pp. 47–84, 2014.

[132] S. Gopalakrishnan and P. Ganeshkumar, "Systematic reviews and meta-analysis: Understanding the best evidence in primary healthcare," *Journal of Family Medicine and Primary Care*, vol. 2, no. 1, pp. 9–14, 2013.

[133] L. S. Uman, "Systematic reviews and meta-analyses," *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, vol. 20, no. 1, p. 57, 2011.

[134] J. Doust, E. Pietrzak, S. Sanders, and P. Glasziou, "Identifying studies for systematic reviews of diagnostic tests was difficult due to the poor sensitivity and precision of methodologic filters and the lack of information in the abstract," *Journal of clinical epidemiology*, vol. 58, no. 5, pp. 444–449, 2005.

[135] M. B. Miles, A. M. Huberman, M. A. Huberman, and M. Huberman, *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage, 1994.

[136] P. A. Ertmer and T. J. Newby, "Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective," *Performance Improvement Quarterly*, vol. 26, no. 2, pp. 43–71, 2013.

[137] J. G. Greeno, A. M. Collins, and L. B. Resnick, "Cognition and learning," *Handbook of educational psychology*, vol. 77, pp. 15–46, 1996.

[138] M. N. Dilshad, "Learning theories: Behaviorism, cognitivism, constructivism," *International Education and Research Journal*, vol. 3, no. 9, 2017.

[139] T. Stavredes, "Learning theory," in *Effective online teaching: Foundations and strategies for student success*, pp. 33–42, San Francisco, CA: John Wiley & Sons, 2011.

[140] A. C. Clark and J. V. Ernst, "A model for the integration of science, technology, engineering, and mathematics," *Technology and Engineering Teacher*, vol. 66, no. 4, p. 24, 2006.

[141] H. El-Deghaidy, N. Mansour, M. Alzaghibi, and K. Alhammad, "Context of stem integration in schools: Views from in-service science teachers," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 6, pp. 2459–2484, 2017.

[142] S. S. Guzey, T. J. Moore, and M. Harwell, "Building up stem: An analysis of teacher-developed engineering design-based stem integration curricular materials," *Journal of Pre-College Engineering Education Research (J-PEER)*, vol. 6, no. 1, p. 2, 2016.

[143] R. E. Satchwell and F. L. Loepp, "Designing and implementing an integrated mathematics, science, and technology curriculum for the middle school.," *Journal of Industrial Teacher Education*, vol. 39, no. 3, pp. 41–66, 2002.

[144] E. H. Mohd Shahali, L. Halim, M. S. Rasul, K. Osman, and M. A. Zulkifeli, "Stem learning through engineering design: Impact on middle secondary students' interest towards stem," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 5, pp. 1189–1211, 2017.

[145] R. Driver, H. Asoko, J. Leach, P. Scott, and E. Mortimer, "Constructing scientific knowledge in the classroom," *Educational researcher*, vol. 23, no. 7, pp. 5–12, 1994.

[146] P. Eastwell, "Social constructivism.," *Science Education Review*, vol. 1, no. 3, pp. 82–86, 2002.

[147] R. S. Prawat and R. E. Floden, "Philosophical perspectives on constructivist views of learning," *Educational Psychologist*, vol. 29, no. 1, pp. 37–48, 1994.

[148] B. for Kids Partnership for 21st Century Learning (P21), "Framework for 21st century learning."

[149] L. Bryan, T. Moore, C. Johnson, and G. Roehrig, "Integrated STEM education," in *STEM roadmap: A framework for integration* (C. C. Johnson, E. E. Peters-Burton, and T. J. Moore, eds.), pp. 23–37, London, UK: Taylor & Francis, 1st ed., 2015.

[150] A. C. Burrows, J. M. Breiner, J. Keiner, and C. Behm, "Biodiesel and integrated stem: Vertical alignment of high school biology/biochemistry and chemistry," *Journal of Chemical Education*, vol. 91, no. 9, pp. 1379–1389, 2014.

[151] L. Debs and T. Kelley, "Gathering design references from nature," *Technology and Engineering Teacher*, vol. 75, no. 1, p. 10, 2015.

[152] S. S. Guzey, T. J. Moore, M. Harwell, and M. Moreno, "Stem integration in middle school life science: Student learning and attitudes," *Journal of Science Education and Technology*, vol. 25, pp. 550–560, Aug 2016.

[153] P. R. Hernandez, R. Bodin, J. W. Elliott, B. Ibrahim, K. E. Rambo-Hernandez, T. W. Chen, and M. A. de Miranda, "Connecting the stem dots: measuring the effect of an integrated engineering design intervention," *International Journal of Technology and Design Education*, vol. 24, pp. 107–120, Feb 2014.

[154] R. K. James, C. E. Lamb, D. L. Householder, and M. A. Bailey, "Integrating science, mathematics, and technology in middle school technology-rich environments: A study of implementation and change," *School Science and Mathematics*, vol. 100, no. 1, pp. 27–35, 2000.

[155] C. A. Mathis, E. A. Siverling, A. W. Glancy, and T. J. Moore, "Teachers' incorporation of argumentation to support engineering learning in stem integration curricula," *Journal of Pre-College Engineering Education Research (J-PEER)*, vol. 7, no. 1, p. 6, 2017.

[156] G. H. Roehrig, T. J. Moore, H.-H. Wang, and M. S. Park, "Is adding the e enough? investigating the impact of k-12 engineering standards on the implementation of stem integration," *School Science and Mathematics*, vol. 112, no. 1, pp. 31–44, 2012.

[157] M. Stohlmann, T. J. Moore, J. McClelland, and G. H. Roehrig, "Impressions of a middle grades stem integration program," *Middle School Journal*, vol. 43, no. 1, pp. 32–40, 2011.

[158] S. L. Stump, J. A. Bryan, and T. J. McConnell, "Making stem connections," *The Mathematics Teacher*, vol. 109, no. 8, pp. 576–583, 2016.

[159] C. G. Valtorta and L. K. Berland, "Math, science, and engineering integration in a high school engineering course: A qualitative study," *Journal of Pre-College Engineering Education Research (J-PEER)*, vol. 5, no. 1, p. 3, 2015.

[160] J. G. Wells, "Pirposal model of integrative stem education: Conceptual and pedagogical framework for classroom implementation," *Technology and Engineering Teacher*, vol. 75, no. 6, p. 12, 2016.

[161] M. D. Merrill, "First principles of instruction: a synthesis.," in *Trends and Issues in Instructional Design and Technology, 2nd Edition* (R. A. Reiser and J. V. Dempsey, eds.), pp. 62–71, Upper Saddle River, NJ: Merrill/Prentice Hall, 2007.

[162] J. J. van Merriënboer and P. Kirschner, *Ten steps to complex learning: A systematic approach to four-component instructional design*. London: Lawrence Erlbaum Associates, 2007.

[163] A. E. Lawson, "Relationships among level of intellectual development, cognitive style, and grades in a college biology course," *Science Education*, vol. 64, no. 1, pp. 95–102, 1980.

[164] R. K. Purser and J. W. Renner, "Results of two tenth-grade biology teaching procedures," *Science Education*, vol. 67, no. 1, pp. 85–98, 1983.

[165] R. E. Mayer, "Should There Be a Three-Strikes Rule Against Pure Discovery Learning?," *American Psychologist*, vol. 59, no. 1, pp. 14–19, 2004.

[166] L. B. Buck, S. L. Bretz, and M. H. Towns, "Characterizing the level of inquiry in the undergraduate laboratory," *Journal of College Science Teaching*, vol. 38, no. 1, pp. 52–58, 2008.

[167] R. S. Matthews, J. L. Cooper, N. Davidson, and P. Hawkes, "Building bridges between cooperative and collaborative learning," *Change: The Magazine of Higher Learning*, vol. 27, no. 4, pp. 35–40, 1995.

[168] D. W. Johnson and R. T. Johnson, "Making cooperative learning work," *Theory Into Practice*, vol. 38, no. 2, pp. 67–73, 1999.

[169] J. De Groof, V. Donche, and P. Van Petegem, *Onderzoekend leren stimuleren: effecten, maatregelen en principes*. Leuven: Acco, 2012.

[170] P. Black and D. Wiliam, "Assessment and classroom learning," *Assessment in Education: Principles, Policy & Practice*, vol. 5, no. 1, pp. 7–74, 1998.

[171] S. M. Gómez Puente, M. van Eijck, and W. Jochems, "A sampled literature review of design-based learning approaches: a search for key characteristics," *International Journal of Technology and Design Education*, vol. 23, pp. 717–732, Aug 2013.

[172] S. W. Lee and S. Min, "Riding the implementation curve: Teacher buy-in and student academic growth under comprehensive school reform programs," *The Elementary School Journal*, vol. 117, no. 3, pp. 371–395, 2017.

[173] E. Ketelaar, D. Beijaard, H. P. Boshuizen, and P. J. D. Brok, "Teachers' positioning towards an educational innovation in the light of ownership, sense-making and agency," *Teaching and Teacher Education*, vol. 28, no. 2, pp. 273 – 282, 2012.

[174] G. I. A. Inc., "Google sketchup."

[175] S. Ceuppens, J. Deprez, W. Dehaene, and M. De Cock, "Design and validation of a test for representational fluency of 9th grade students in physics and mathematics: The case of linear functions," *Phys. Rev. Phys. Educ. Res.*, vol. 14, p. 020105, Aug 2018.

[176] E. F. Redish, "Problem solving and the use of math in physics courses," in *World View on Physics Education in 2005*, Focusing on Change, (Delhi), pp. 1–10, August 21-22 2005.

[177] H. Hudson and W. McIntire, "Correlation between mathematical skills and success in physics," *American Journal of Physics*, vol. 45, no. 5, pp. 470–471, 1977.

[178] E. Deliyianni, A. Gagatsis, I. Elia, and A. Panaoura, "Representational flexibility and problem-solving ability in fraction and decimal number addition: A structural model," *International Journal of Science and Mathematics Education*, vol. 14, pp. 397–417, Jul 2016.

[179] M. Hill, M. D. Sharma, J. O'Byrne, and J. Airey, "Developing and evaluating a survey for representational fluency in science," *International Journal of Innovation in Science and Mathematics Education (formerly CAL-laborate International)*, vol. 22, no. 6, pp. 22–42, 2014.

[180] S. Ainsworth, "The functions of multiple representations," *Computers & education*, vol. 33, no. 2, pp. 131–152, 1999.

[181] S. Ainsworth, "Deft: A conceptual framework for considering learning with multiple representations," *Learning and Instruction*, vol. 16, no. 3, pp. 183 – 198, 2006.

[182] A. H. Schoenfeld, J. P. Smith, and A. Arcavi, "Learning: The microgenetic analysis of one student's evolving understanding of a complex subject matter domain," *Advances in instructional psychology*, vol. 4, pp. 55–175, 1993.

[183] L. C. McDermott, M. L. Rosenquist, and E. H. van Zee, "Student difficulties in connecting graphs and physics: Examples from kinematics," *American Journal of Physics*, vol. 55, no. 6, pp. 503–513, 1987.

[184] L. Bollen, M. De Cock, K. Zuza, J. Guisasola, and P. van Kampen, "Generalizing a categorization of students' interpretations of linear kinematics graphs," *Phys. Rev. Phys. Educ. Res.*, vol. 12, p. 010108, Feb 2016.

[185] A. A. Nistal, W. V. Dooren, and L. Verschaffel, "Improving students' representational flexibility in linear-function problems: an intervention," *Educational Psychology*, vol. 34, no. 6, pp. 763–786, 2014.

[186] P. Engelhardt, "An introduction to classical test theory as applied to conceptual multiple-choice tests," in *Getting Started in PER*, vol. 2, per central (online), April 2009.

[187] A. C. Ornstein, "Norm-referenced and criterion- referenced tests: An overview," *NASSP Bulletin*, vol. 77, no. 555, pp. 28–39, 1993.

[188] R. J. McCowan and S. C. McCowan, "Item analysis for criterion-referenced tests," *Online Submission*, 1999.

[189] K.-Y. Liang and S. L. Zeger, "Longitudinal data analysis using generalized linear models," *Biometrika*, vol. 73, no. 1, pp. 13–22, 1986.

[190] M. Stokes, C. Davis, and G. Koch, *Categorical Data Analysis Using the Sas®System, 2Nd Edition*. SAS Institute Inc., Cary, NC, USA, second ed., 2000.

[191] C.-f. Sheu, "Regression analysis of correlated binary outcomes," *Behavior Research Methods, Instruments, & Computers*, vol. 32, pp. 269–273, Jun 2000.

[192] J. Docktor and K. Heller, "Gender differences in both force concept inventory and introductory physics performance," *AIP Conference Proceedings*, vol. 1064, no. 1, pp. 15–18, 2008.

[193] S. Andersson and A. Johansson, "Gender gap or program gap? students' negotiations of study practice in a course in electromagnetism," *Phys. Rev. Phys. Educ. Res.*, vol. 12, p. 020112, Aug 2016.

[194] S. Bates, R. Donnelly, C. MacPhee, D. Sands, M. Birch, and N. R. Walet, "Gender differences in conceptual understanding of newtonian mechanics: a uk cross-institution comparison," *European Journal of Physics*, vol. 34, no. 2, p. 421, 2013.

[195] S. J. Pollock, "Comparing student learning with multiple research-based conceptual surveys: Csem and bema.," *AIP Conference Proceedings*, vol. 1064, no. 1, pp. 171–174, 2008.

[196] K. Wilson, D. Low, M. Verdon, and A. Verdon, "Differences in gender performance on competitive physics selection tests," *Phys. Rev. Phys. Educ. Res.*, vol. 12, p. 020111, Aug 2016.

[197] H. Dawkins, H. Hedgeland, and S. Jordan, "Impact of scaffolding and question structure on the gender gap," *Phys. Rev. Phys. Educ. Res.*, vol. 13, p. 020117, Sep 2017.

[198] H. Broekkamp and B. van Hout-Wolters, "The gap between educational research and practice: A literature review, symposium, and questionnaire," *Educational Research and Evaluation*, vol. 13, no. 3, pp. 203–220, 2007.

[199] J. Sweller, J. J. Van Merrienboer, and F. G. Paas, "Cognitive architecture and instructional design," *Educational psychology review*, vol. 10, no. 3, pp. 251–296, 1998.

[200] J. J. Kaput, "Technology and mathematics education," in *Handbook of research on mathematics teaching and learning* (D. A. Grouws, ed.), pp. 515–556, New York, NY: Macmillan Publishing Company, 1992.

[201] M. Scaife and Y. Rogers, "External cognition: how do graphical representations work?," *International journal of human-computer studies*, vol. 45, no. 2, pp. 185–213, 1996.

[202] D. L. McGee and R. Martinez-Planell, "A study of semiotic registers in the development of the definite integral of functions of two and three variables," *International Journal of Science and Mathematics Education*, vol. 12, no. 4, pp. 883–916, 2014.

[203] N. Presmeg, "Semiotics and the "connections" standard: Significance of semiotics for teachers of mathematics," *Educational Studies in Mathematics*, vol. 61, pp. 163–182, Feb 2006.

[204] A. F. Heckler and T. M. Scaife, "Adding and subtracting vectors: The problem with the arrow representation," *Physical Review Special Topics-Physics Education Research*, vol. 11, no. 1, p. 010101, 2015.

[205] S. Ceuppens, L. Bollen, J. Deprez, W. Dehaene, and M. De Cock, "9th grade students' understanding and strategies when solving $x(t)$ problems in 1d kinematics and $y(x)$ problems in mathematics," *Phys. Rev. Phys. Educ. Res.*, vol. 15, p. 010101, Jan 2019.

[206] B. Ibrahim and N. S. Rebello, "Representational task formats and problem solving strategies in kinematics and work," *Phys. Rev. ST Phys. Educ. Res.*, vol. 8, p. 010126, Jun 2012.

[207] A. A. Nistal, W. V. Dooren, and L. Verschaffel, "Students' reported justifications for their representational choices in linear function problems: an interview study," *Educational Studies*, vol. 39, no. 1, pp. 104–117, 2013.

[208] M. Hill and M. D. Sharma, "Students' representational fluency at university: A cross-sectional measure of how multiple representations are used by physics students using the representational fluency survey," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 11, no. 6, pp. 1633–1655, 2015.

[209] M. Bassok and K. J. Holyoak, "Interdomain transfer between isomorphic topics in algebra and physics.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 15, no. 1, p. 153, 1989.

[210] P. Barniol and G. Zavala, "Force, velocity, and work: The effects of different contexts on students' understanding of vector concepts using isomorphic problems," *Phys. Rev. ST Phys. Educ. Res.*, vol. 10, p. 020115, Aug 2014.

[211] G. Zavala and P. Barniol, "A detailed analysis of isomorphic problems: The case of vector problems," in *Physics Education Research Conference 2013*, PER Conference, (Portland, OR), pp. 377–380, July 17-18 2013.

[212] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, 1960.

[213] A. J. Viera, J. M. Garrett, *et al.*, "Understanding interobserver agreement: the kappa statistic," *Fam Med*, vol. 37, no. 5, pp. 360–363, 2005.

[214] R. L. Doran, *Basic Measurement and Evaluation of Science Instruction*. Washington, DC: National Science Teachers Association, 01 1980.

[215] M. Hill, M. D. Sharma, and H. Johnston, "How online learning modules can improve the representational fluency and conceptual understanding of university physics students," *European Journal of Physics*, vol. 36, no. 4, p. 045019, 2015.

[216] R. K. Thornton and D. R. Sokoloff, "Assessing student learning of newton's laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula," *American Journal of Physics*, vol. 66, no. 4, pp. 338–352, 1998.

[217] D. N. Perkins and G. Salomon, *Transfer of learning*, vol. 2, pp. 6452–6457. Oxford, England: Pergamon Press, 1992.

[218] J. Mestre, "Transfer of learning: Issues and research agenda," 2002.

[219] A. L. Roberts, M. D. Sharma, S. Britton, and P. B. New, "An index to measure the ability of first year science students to transfer mathematics," *International Journal of Mathematical Education in Science and Technology*, vol. 38, no. 4, pp. 429–448, 2007.

[220] S. Britton, P. B. New, M. D. Sharma, and D. Yardley, "A case study of the transfer of mathematics skills by university students," *International Journal of Mathematical Education in Science and Technology*, vol. 36, no. 1, pp. 1–13, 2005.

[221] S. Ceuppens, J. Deprez, W. Dehaene, and M. De Cock, "Students' conceptual understanding of kinematics in an integrated stem curriculum," *European Journal of STEM Education*.

[222] J. De Meester, H. Knipprath, M.-P. Buyse, S. Ceuppens, H. De Loof, L. Goovaerts, L. Hellinckx, A. Struyf, L. Thibaut, D. Van de Velde, J. Boeve-de Pauw, M. De Cock, F. Depaepe, J. Deprez, G. Langie, K. Struyven, P. Van Petegem, and W. Dehaene, "Leren ontwerpen van stem-integrerend leermateriaal," *Tijdschrift voor Lerarenopleiders*, vol. 38, no. 4, pp. 27–36, 2017.

[223] G. Zavala, S. Tejeda, P. Barniol, and R. J. Beichner, "Modifying the test of understanding graphs in kinematics," *Phys. Rev. Phys. Educ. Res.*, vol. 13, p. 020111, Aug 2017.

[224] D.-H. Nguyen and N. S. Rebello, "Students' understanding and application of the area under the curve concept in physics problems," *Phys. Rev. ST Phys. Educ. Res.*, vol. 7, p. 010112, Jun 2011.

[225] W. Feller, *An introduction to probability theory and its applications vol. 1*. New York, USA: Wiley, 1957.

[226] S. W. Brahmia, "Negative quantities in mechanics: a fine-grained math and physics conceptual blend?," in *Physics Education Research Conference 2017*, PER Conference, (Cincinnati, OH), pp. 64–67, July 26-27 2017.

[227] R. Rosenthal, R. L. Rosnow, and D. B. Rubin, *Contrasts and effect sizes in behavioural research: A correlational approach*. Cambridge, UK: Cambridge University Press, 2000.

[228] A. Field, *Discovering statistics using IBM SPSS statistics*. Sage Publications, Inc, 4th ed., 2005.

[229] D. Kaltakci-Gurel, A. Eryilmaz, and L. C. McDermott, "Development and application of a four-tier test to assess pre-service physics teachers' misconceptions about geometrical optics," *Research in Science & Technological Education*, vol. 35, no. 2, pp. 238–260, 2017.

[230] I. Caleon and R. Subramaniam, "Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves," *International Journal of Science Education*, vol. 32, no. 7, pp. 939–961, 2010.

[231] I. S. Caleon and R. Subramaniam, "Do students know what they know and what they don't know? using a four-tier diagnostic test to assess the nature of students' alternative conceptions," *Research in Science Education*, vol. 40, pp. 313–337, May 2010.

[232] K. C. D. Tan, N. K. Goh, L. S. Chia, and D. F. Treagust, "Development and application of a two-tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis," *Journal of Research in Science Teaching*, vol. 39, no. 4, pp. 283–301, 2002.

[233] S. Ceuppens, J. Deprez, W. Dehaene, and M. De Cock, "Tackling misconceptions in geometrical optics," *Physics Education*, vol. 53, no. 4, p. 045020, 2018.

[234] F. M. Goldberg and L. C. McDermott, "Student difficulties in understanding image formation by a plane mirror," *The Physics Teacher*, vol. 24, pp. 472–481, nov 1986.

[235] K. Cummings and E. Grillo, "Conceptual underpinnings of students' ability to understand reflections from a plane mirror," in *AIP Conference Proceedings*, pp. 185–188, AIP, 2005.

[236] I. Bianchi and U. Savardi, "What fits into a mirror: Naïve beliefs about the field of view.," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 38, no. 5, pp. 1144–1158, 2012.

[237] G. Tural, "Cross-Grade Comparison of Students' Conceptual Understanding with Lenses in Geometric Optics," *Science Education International*, vol. 26, no. 3, pp. 325–343, 2015.

[238] D. Kaltakci-Gurel, A. Eryilmaz, and L. C. McDermott, "Identifying pre-service physics teachers' misconceptions and conceptual difficulties about geometrical optics," *European Journal of Physics*, vol. 37, p. 045705, jul 2016.

[239] C. W. Anderson and E. L. Smith, "Children's conceptions of light and color: understanding the role of unseen rays," tech. rep., The Institute for Research on Teaching, Michigan: East Lansing, 1986.

[240] K. Wosilait, P. R. L. Heron, P. S. Shaffer, and L. C. McDermott, "Development and assessment of a research-based tutorial on light and shadow," *American Journal of Physics*, vol. 66, pp. 906–913, oct 1998.

[241] L. Viennot, F. Chauvet, P. Colin, and G. Rebmann, "Designing strategies and tools for teacher training: The role of critical details, Examples in optics," 2005.

[242] S. M. Pompea, E. F. Dokter, C. E. Walker, and R. T. Sparks, "Using Misconceptions Research in the Design of Optics Instructional Materials and Teacher Professional Development Programs," in *Education and Training in Optics and Photonics*, Optical Society of America, 2007.

[243] M. Allen, *Misconceptions in Primary Science*. Berkshire, UK: Open University Press, McGraw-Hill Education, 1st ed., 2010.

[244] L. C. McDermott, "Guest comment: How we teach and how students learn - a mismatch?," *American Journal of Physics*, vol. 61, no. 4, p. 295, 1993.

[245] P. Colin and L. Viennot, "Using two models in optics: Students' difficulties and suggestions for teaching," *American Journal of Physics*, vol. 69, pp. S36–S44, jul 2001.

[246] A. Santana, Y. Rodríguez, and E. A. Gómez, "Construction of ray diagrams in geometrical optics: a media-focused approach," *Physics Education*, vol. 47, no. 6, p. 715, 2012.

[247] F. Favale and M. Bondani, "Misconceptions about optics: An effect of misleading explanations?," in *Education and Training in Optics and Photonics* (M. F. P. C. Martins Costa and M. Zghal, eds.), Optical Society of America, jul 2013.

[248] B. Vuilleumier, ""optics: Law of reflection" from the wolfram demonstrations project." http://demonstrations.wolfram.com/OpticsLawOfReflection/ Accessed on 2018-02-12.

FACULTY OF SCIENCE
DEPARTMENT OF PHYSICS AND ASTRONOMY
PHYSICS EDUCATION RESEARCH GROUP
Celestijnenlaan 200c box 2406
B-3001 Leuven