Teaching towards knowledge integration in learning force and motion

Article *in* International Journal of Science Education · October 2019

DOI: 10.1080/09500693.2019.1672905

7 authors, including:

Ying Nie
South China Normal University
2 PUBLICATIONS **14** CITATIONS

SEE PROFILE

Yang Xiao
South China Normal University
12 PUBLICATIONS **69** CITATIONS

SEE PROFILE

Qiaoyi Liu
The Ohio State University
11 PUBLICATIONS **39** CITATIONS

SEE PROFILE

Jianwen Xiong
South China Normal University
54 PUBLICATIONS **306** CITATIONS

SEE PROFILE

International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: https://www.tandfonline.com/loi/tsed20

Teaching towards knowledge integration in learning force and motion

Ying Nie, Yang Xiao, Joseph C. Fritchman, Qiaoyi Liu, Jing Han, Jianwen Xiong & Lei Bao

To cite this article: Ying Nie, Yang Xiao, Joseph C. Fritchman, Qiaoyi Liu, Jing Han, Jianwen Xiong & Lei Bao (2019): Teaching towards knowledge integration in learning force and motion, International Journal of Science Education, DOI: 10.1080/09500693.2019.1672905

To link to this article: https://doi.org/10.1080/09500693.2019.1672905

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION
https://doi.org/10.1080/09500693.2019.1672905

Teaching towards knowledge integration in learning force and motion

Ying Nie[a,b,c], Yang Xiao [a,b], Joseph C. Fritchman [b], Qiaoyi Liu [b], Jing Han[b], Jianwen Xiong[a] and Lei Bao [b]

[a]School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou, People's Republic of China; [b]Department of Physics, The Ohio State University, Columbus, OH, USA; [c]Teaching and Research Office, Education Bureau of Zhangjiagang, Zhangjiagang, People's Republic of China

ABSTRACT
Knowledge integration is essential to achieve deep conceptual understanding, which requires students to develop well-connected knowledge structures through the central idea of a concept. To effectively represent and analyze knowledge integration, a conceptual framework model on force and motion is developed to map learners' knowledge structures in terms of how conceptual ideas and contextual conditions are connected. Two studies have been conducted. First, the misconceptions on force and motion held by Chinese middle school students are examined. Although the Chinese students had experienced extensive problem-drilling in instruction, which is notably different from that of the populations documented in the literature, their misconceptions are similar to those previously reported. This suggests that traditional problem-drilling does not substantively improve conceptual development. In addition, detailed analysis of students' qualitative and quantitative responses also suggests that students' misconceptions can be viewed through the conceptual framework model as local connections among subsets of contextual features, which indicate fragmented knowledge structures. The second study evaluates the effectiveness of a modified instruction that targets knowledge integration by explicitly emphasizing the learning of the central idea and making the needed connections with it. The results indicate that the modified instruction outperforms the traditional method in promoting knowledge integration.

ARTICLE HISTORY
Received 12 January 2019
Accepted 22 September 2019

KEYWORDS
Knowledge integration; physics education; conceptual development

Introduction

STEM learning requires students to acquire deep understanding of science concepts (National Research Council, 2011, 2012b). However, the results of traditional education often promote memorising rules and algorithms that may enable students to perform well on standardised tests but fail to develop deep understanding (Alonso, 1992). The

CONTACT Lei Bao bao.15@osu.edu Department of Physics, The Ohio State University, Columbus, OH 43210, USA; Jianwen Xiong jwxiong@scnu.edu.cn Department of Physics, School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou, 510006, People's Republic of China
Supplemental data for this article can be accessed at https://doi.org/10.1080/09500693.2019.1672905.

focus on traditional problem-drilling has been shown to have limitations such that it may actually promote memorisation of context-specific solutions with minimal generalisation rather than transitioning students from novices to experts (Chiu, Guo, & Treagust, 2007; Kim & Pak, 2002; Nakhleh, 1993; Stamovlasis, Tsaparlis, Kamilatos, Papaoikonomou, & Zarotiadou, 2005).

Recognising the failures of traditional teaching, research-informed teaching methods have evolved to include inquiry-based interactive-engagement elements in lecture, recitations, and labs, which aim to develop deep understanding by carefully targeting perceived deficits in student knowledge and actively encouraging students to explore and discuss, rather than rote memorisation (Alexopoulou & Driver, 1996; Crouch & Mazur, 2001; Keller et al., 2007; McDermott, 1996). These approaches help promote generalisation and deeper conceptual understanding by implicitly building connections between knowledge segments.

This study explores an alternative instructional method, which takes the knowledge integration perspective (Linn, 2005; Shen, Liu, & Chang, 2017) of learning and focuses the instruction on helping students develop and refine their knowledge structure toward an increasingly more coherently organised and extensively connected networks of ideas (Nordine, Fortus, Lehavi, Neumann, & Krajcik, 2018; Nordine, Krajcik, & Fortus, 2011). In addition, to effectively represent and analyze structural features of knowledge integration, a conceptual framework representation is developed to map the learners' knowledge structures in terms of how fragments and pieces of conceptual ideas and context conditions are connected.

The content topic of this study is about force and motion, which is a foundational concept in physics. A conceptual framework on force motion will be established based on existing literature on learning theories and students' misconceptions on the concept of force and motion. The conceptual framework is then used to analyze students' misconceptions with the knowledge integration perspective and to guide the development of a new instruction method to promote knowledge integration. Two empirical studies will be reported: (1) assessment of the misconceptions about force and motion held by Chinese middle school students who had experienced predominantly problem-drilling type of instruction, which is very different from the populations documented in the existing literature; and (2) the effectiveness of the modified instruction that targets knowledge integration.

Conceptual framework and misconceptions

The development of the conceptual framework representation has been informed by a large number of established learning theories in science education, in particular, the theories on conceptual change and knowledge integration. The theory on conceptual change focuses on the process and outcome of how students' pre-instructional cognitive structures are fundamentally restructured in developing science concepts (Duit & Treagust, 2003). In such cognitive restructuring, knowledge integration emphasises the development of increasingly more integrated networks of ideas, rather than a collection of disconnected ones (Linn, 2005).

In the conceptual change perspective, students' pre-instructional cognitive structures, which differ from expert conceptions, have commonly been referred as misconceptions

(Dykstra, Boyle, & Monarch, 1992; Halloun & Hestenes, 1985b). Students' misconceptions are strongly situated within contexts (context dependent) and exist as locally connected knowledge fragments that are unable to establish similarities and contrasts between contexts. Knowledge fragmentation has been demonstrated through many studies on expert-novice differences. A learner's knowledge organisation is shown to be a key factor separating experts from novices (Bagno, Eylon, & Ganiel, 2000; Chi, Feltovich, & Glaser, 1981; Jong & Ferguson-Hessler, 1986; Larkin, McDermott, Simon, & Simon, 1980; Smith, 1992). Experts' knowledge is organised around central principles, which guide problem solving and facilitate transfer between different knowledge domains and contexts (Brown, 1989; Perkins & Salomon, 1989; Salomon & Perkins, 1989). In contrast, novices lack a well-organised knowledge structure and solve problems relying on surface characteristics that are directly mapped to certain problem-solving outcomes through memorisation (Hardiman, Dufresne, & Mestre, 1989).

In the knowledge integration perspective, as a student progresses from lower to higher cognitive stages, the student's knowledge structure becomes more integrated and easier to transfer across contexts (less context specific) (Lee, Liu, & Linn, 2011; Linn, 2005; Shen et al., 2017). In assessment of knowledge integration, the ability to consistently use a *central idea* across a range of phenomena or contexts has been used as a key factor to evaluate students' achievement (Kubsch, Nordine, Neumann, Fortus, & Krajcik, 2018; Nordine et al., 2011). A central idea provides an anchor point to link other ideas, and thus acts as the central node for establishing well connected knowledge networks. It has also been shown that instruction emphasising central ideas are productive to promote knowledge integration and helps students develop deeper conceptual understanding (Nordine et al., 2011).

Synthesising the existing theories on conceptual change and knowledge integration, a number of key features of learning can be summarised to guide the development of the conceptual framework model. First, students often hold deeply rooted conceptions that are not in agreement with the science views (Duit & Treagust, 2003). Meanwhile, students' conceptions are usually context dependent and exist as disconnected knowledge fragments strongly situated within specific contexts (Minstrell, 1992; Bao & Redish, 2001, 2006). Following the knowledge integration perspective, productive instruction should help students develop a coherently integrated knowledge structure with well-connected networks of ideas rather than a fragmented knowledge structure consisting of collections of isolated knowledge pieces.

Building on the existing learning theories, the conceptual framework is developed to clearly illustrate the knowledge structure of different states of learners' conceptions, from novices' to experts'. In a conceptual framework, a learner's ideas are activated by and depend on context features. An expert would then link ideas and contexts to form conceptual pathways around the central idea as a core node to establish a fully integrated knowledge structure. In contrast, novices often bypass the central idea and develop direct links using memorised algorithms or relations among surface features of a problem context (Chi et al., 1981; Bao & Redish, 2001, 2006). Although the novices' approach may produce correct results in limited situations, as the number of contexts and variables increases, mastery by memorisation will quickly become forbiddingly unfeasible in complex multivariate situations. On the other hand, relating the wide ranging situations to the central idea will aid in students' forming a well-integrated knowledge structure to support the construction of a deeper expert-like conceptual understanding.

It is worth noting that in existing literature, a popular approach to analyze conceptual connections is the use of concept maps (Novak & Cañas, 2006), which are graphical representations of relationships and links among key elements (definitions and terms) related to a concept, and also the connections between different concepts. At a first look, a concept map may appear similar to a conceptual framework, however, there is a fundamental difference. A concept map lacks the presentation of the involvement of contextual configurations and conditions and their connections to conceptual ideas, which together form the wide variety of conceptual pathways that a learner can develop. In contrast, these contextually manifested conceptual pathways are emphasised in a conceptual framework and become the essential features allowing more comprehensive representation and modelling of students' progression stages in knowledge integration.

The conceptual framework of force and motion

The first step to establish a conceptual framework is to identify the central idea of a concept. The governing principles of force and motion are largely summarised by Newton's first and second laws, which describes the relations between applied forces and kinetic changes. The fundamental force motion relation is that the acceleration of an object is proportional to the net force acting on the object, $\Sigma\vec{F} = \vec{F}_{net} = m\vec{a}$, which is identified as the central idea of force and motion concept in this study. It is important to note that the formula is a mathematical representation of the central idea, and not the central idea itself. The central idea involves two key components: the net force and the acceleration. When experts encounter any kinematic situation involving forces, they would connect everything through the central idea, which extends to form a problem-solving approach.

Among novice learners, a major difficulty regarding this central idea is the failure to distinguish the difference between the net force and an applied physical force (Thornton & Sokoloff, 1998). The net force is the sum of all applied physical forces, but is not actually an applied physical force itself. When students fail to distinguish the difference, they tend to broadly apply $F = ma$ to any given applied forces and may ignore the effects of some less obvious forces such as friction. Without a correctly-established central idea on net force, novice learners often fail to develop an expert-like integrated knowledge structure and rely on contextual details as resources to construct problem-solving pathways, leading to fragmented local connections among surface features (Hake, 1998; Bao, Hogg, & Zollman, 2002, Bao & Redish 2006).

To show the differences between expert and novice knowledge structures, a conceptual framework on force motion is illustrated in Figure 1. The top half of Figure 1 shows the experts' knowledge structure, in which elements (e.g. variables and relations) related to force and motion are fully connected in pathways through the central idea ($\Sigma F = ma$). Therefore, activating any elements within a pathway will activate the entire network, allowing an expert to have a deep and global conceptual stand point to analyze different contextual configurations and to transfer between known and novel situations. The expert's conceptual pathways demonstrate an integrated knowledge structure formed with a coherent network of connections between the net force and the key

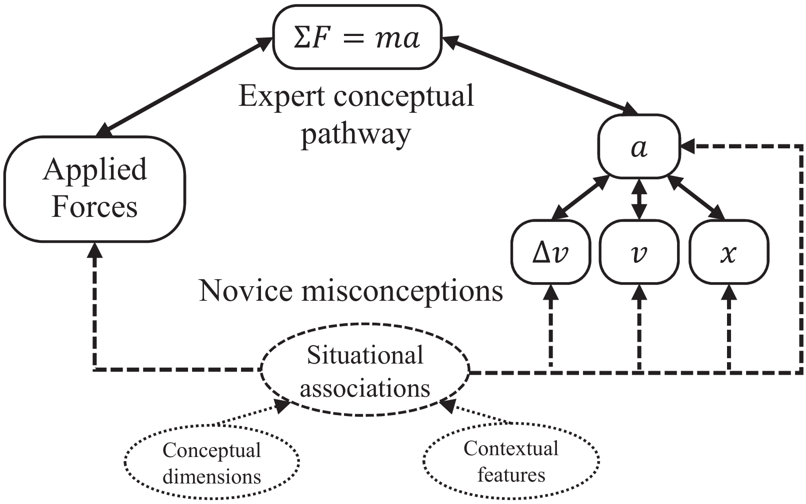

Figure 1. Conceptual framework of force and motion showing connections among variables, relations, and conceptual ideas. The two-way arrows indicate possible pathways of connections within a learner's knowledge structure. The solid lines represent experts' conceptual pathway, while the dashed lines represent novices' possible pathways.

variables describing the motion. Such a knowledge structure allows the expert to analyze unique contextual situations accurately by processing the contextual variables and relations through the central idea rather than local connections among contextual details.

The lower half of Figure 1 shows possible novices' knowledge structures, which consist of a wide variety of direct local connections among ideas and context features. For example, a novice often directly link an applied force with motion (v) in the same direction and vice-versa, leading to the most popular force motion misconception stating that there is always a force in the direction of motion (Gilbert & Watts, 1983; Trumper & Gorsky, 1996). These local connections form a fragmented knowledge structure, which express in contexts as the wide-ranging misconceptions. A non-exhaustive summary of typical variations of contextual features is given in Table 1. Obviously, the total combinations of the contextual configurations are immensely broad and difficult to be fully memorised. Therefore, the locally connected ideas developed by the novices can be functional to certain subsets of the context situations but can hardly provide a complete understanding for all the involved conditions. The fragmentation and strong context dependence also make it difficult for novices to transfer and generalise ideas across different contexts. For example, novice may view a person sitting in a chair and a book sitting on a table as two different situations with different applications of forces, although these are considered isomorphic from a Newtonian expert's point of view.

Conceptual framework analysis of misconceptions

Misconceptions on the force motion concept have been well studied and documented (Champagne, Klopfer, & Anderson, 1980; Gilbert & Watts, 1983; Halloun & Hestenes, 1985a; Harris, George, Hirsh-Pasek, & Newcombe, 2018; Trumper & Gorsky, 1996;

Table 1. A non-exhaustive compilation of conceptual ideas (in columns) and context variations (within each column) typically involved in force and motion scenarios.

Applied Forces	ΣF	a	Δv	v	x
• Acting in direction of motion • Acting in opposite direction of motion • Acting on object not in motion • Friction • No force	• $\Sigma F > 0$ • $\Sigma F = 0$ • $\Sigma F < 0$Z	• $a > 0$ • $a = 0$ • $a < 0$ • Accelerate • Constant motion • At rest • Decelerate	• Positive • Zero • Negative • Constant	• Increasing • Constant • Decreasing • At rest • In the direction of force • Opposite to the direction of force	• Increasing • Constant • Decreasing • Zero

Viennot, 1979). A general trend among the misconceptions is that novice learners intuitively enforce direct connections between applied forces and features of motion conforming to pre-Newtonian views of kinematics. To demonstrate the connection structures of misconceptions in the conceptual framework perspective, four examples of the common misconceptions are analysed below. Each of these can be described with the force-motion concept framework as pathways or links that directly connect specific contextual features, bypassing the central idea. The misconceptions are thus interpreted as specific situational pathways of the force-motion conceptual framework.

- Motion ($v \neq 0$) implies active force ($F \neq 0$): This misconception involves the belief that a moving object must carry a force to keep it moving (Viennot, 1979). Students with this misconception form direct links between the movement of the object (v or Δx) and the force acting on the object. Such links often have wide ranging variations depending on contextual configurations, and are activated based on specific details of a situation.
- No motion ($v = 0$) implies no force ($F = 0$): This misconception focuses on the belief that an object not in motion must not have any force acting on it (Gilbert & Watts, 1983; Halloun & Hestenes, 1985a; Viennot, 1979). This misconception forms a direct link between the stationary state of motion and the condition of no force, which is treated as being completely separate from the 'Motion' ($v \neq 0$) context. As a result, such knowledge is deeply situated in the specific contexts and forms an isolated fragment with few connections to other parts of the conceptual framework.
- Velocity is proportional to applied force: This misconception corresponds to the belief that motion and force are directly proportional (Champagne et al., 1980). Students with this type of misconception again form direct links between force and velocity, which are also extended with the proportional relations that add additional connections to the change in force and change in velocity or acceleration. Depending on the student's ability to generalise, variety of pathways may be needed for each variation of situation (e.g. increasing, constant, or decreasing velocity).
- Net force determines motion: This misconception combines aspects of other misconceptions in that if a net force acts on an object then there is motion, but if no net force acts on an object, then the object is either slowing down or stopped (Champagne et al., 1980; Halloun & Hestenes, 1985a). With this misconception, students actually develop some understanding of the central idea regarding the net force; however, the net force is incorrectly linked to specific states of motion in the same ways expressed in the above examples as if the net force is an applied force. Similarly, to handle the

variations of contextual configurations, a range of different pathways are needed for the variety of possible situations of net force and motion.

Using the conceptual framework shown in Figure 1 and Table 1, a student's misconceptions can be represented as numerous separate pathways and local connections formed as fragmented parts of the student's knowledge structure. Meanwhile, an expert's pathways would consist primarily a single coherent flow that links 'Applied Forces' through the central idea of net force and acceleration, and then extends to other motion variables (acceleration, velocity, and displacement) and vice versa. Motivated by the expert's knowledge structure, it can be implied that a possible way to improve students' learning may be to help them establish a good understanding of the central idea first, and then (more importantly) to aid them in building connections that link all contextual situations through the central idea. Similar approaches have also shown promising outcomes on student learning of energy concept (Nordine et al., 2011).

Research questions

The concept of force and motion is a fundamental topic in physics, which form the conceptual foundation of nearly all later physics topics. Therefore, the persistence of misconceptions after studying this topic is a major hurdle in the way of learning future physics concepts. In the US, instruction in introductory physics commonly emphasises developing conceptual understanding through inquiry based teaching methods such as Physics by Inquiry (McDermott, 1996), Workshop Physics (Laws, 2004), SCALE-UP (Beichner et al., 1999), and Discrepant Event (Anggoro, Widodo, Suhandi, & Treagust, 2019). While in China, although inquiry based teaching has been repeatedly proposed in the curriculum standards and guidelines since the year 2000, actual teaching practices in schools have not substantially changed due to the heavy influence from the National College Entrance Examination (NCEE; Wu, 2017; Zhu & Fan, 2012). As a result, in order to meet the NCEE requirement, Chinese educators have to emphasise problem-drilling in instruction, which doesn't typically transfer to foster conceptual development. However, neither country's instructional methods explicitly target knowledge integration. Nevertheless, their education goals expect that students would spontaneously develop a well-connected knowledge structure through a series of conceptual explorations or problem-drilling. It is then informative to examine how emphasis on traditional problem-drilling may impact conceptual understanding.

Based on the literature, knowledge integration is central to obtaining deep conceptual understanding, but is difficult to achieve in most traditional education settings. In the process of aiding students to develop an integrated knowledge structure, knowing which connections are essential within a conceptual framework can be a key factor to developing effective instruction. To facilitate knowledge integration, it is important to help students develop connections among fragmented knowledge pieces through a central idea. Explicit emphasis on learning the central idea and making connections through it can be a promising instructional approach, especially when such connections are difficult to be self-discovered during guided explorations.

Therefore, it is hypothesised that by helping students develop the central idea and the needed connections, students may develop a more integrated knowledge structure and achieve better learning performances. This leads to two main research questions:

1. Comparing to the literature, did the Chinese students, who were primarily educated through extensive problem-drilling, develop similar or different types of conceptual understanding?
2. How may a modified instruction method, which is designed to explicitly emphasise the central idea and its connections, impact student learning?

Method

Participants

The subjects of this study are 8th grade Chinese students taking middle school version of physics, equivalent to algebra-based physics courses in U.S. high schools. Two studies have been conducted with these students to answer the two research questions outlined in the previous section. Study 1 examines whether known common misconceptions among American students also prevail among the Chinese students, despite significantly different education backgrounds. Study 2 looks into the effectiveness of a new instructional approach on force and motion, which explicitly emphasises developing connections with the central idea of force and motion. The subjects in study 1 comprised 208 students (102 girls), whose mean age was 14.65 years (SD = 0.50). The subjects in study 2 comprised 171 students (77 girls), whose mean age was 14.32 years (SD = 0.54).

Experimental procedure

Study 1 investigates Chinese 8th grade students' potential misconceptions on force and motion. A total of 208 students from school S were selected to participate in this experiment after learning force and motion in their physics class. At that time, the students were expected to know how to measure and represent distance and velocity, develop ideas about force, and understand the relations between force and aspects of motion. All students were asked to complete a multiple-choice assessment, which will be introduced next. In addition, 14 out of the 208 students were randomly selected to participate in think-out-loud interviews. These interviews provide an opportunity to probe further the thought processes behind students' problem-solving approaches and reasoning. In particular, the results are inspected to check whether the students solve problems through memorised pattern matching or through concept-based reasoning.

Study 2 tests an education intervention with a different group of Chinese 8th grade students from school G, which is in the same geo-economic region of school S. A total of 171 students in four classes participated in this experiment. All students have completed two identical lessons on force and motion as part of their regular physics classes. All lessons are 45 min long. Then two classes with a total of 82 students were randomly selected as the treatment group, who received the new instruction in terms of two 45-minute lessons that emphasise using the central idea in problem-solving. The remaining two classes with 89 students became the control group, who continued their regular classes to receive two lessons of traditional problem-solving practices. In the two new lessons, the central idea of acceleration and net force was explicitly introduced along with practices for students to develop connections between the central idea and other related physical variables. The details of the lesson plan are outlined in the following section. Pre- and post- tests were administered before and after the two lessons for both treatment and control groups using the same assessment questions from Study 1. In addition, 18 students were selected from the 171 students to participate in think-out-loud interviews to further validate the quantitative outcomes.

Assessment instrument

In this study, students' understanding of force and motion was assessed with an instrument containing 15 questions in four different contextual categories including Force, No Force, Motion, and No Motion as shown in Table 2:

- Force: students are asked to identify the resulting motion of an object having a none-zero net force;
- No Force: students are asked to identify the resulting motion of an object with net force being zero;
- Motion: students are asked to identify applied force(s) acting on a moving object; and
- No Motion: students are asked to identify applied force(s) acting on a stationary object.

The four categories represent unique context configurations that can significantly impact students' performances (Alonzo & Steedle, 2009), especially for those who have fragmented knowledge structures. In contrast, students whom have an expert-like integrated knowledge structure are expected to perform consistently across the different contexts. Therefore, these questions provide a mean to assess students' knowledge structures by comparing the difference and consistency of students' performances on questions across the different contexts.

Within these 15 items, 8 questions (2, 3, 5, 6, 8, 10, 11 and 14) were adapted from the ordered multiple-choice (OMC) items developed by Alonzo and Steedle (2009) for their study on the Force and Motion learning progression. The other 7 items were adapted and modified from the tests used in the past Chinese Senior high school entrance examinations (e.g. Suzhou Education Examination Authority, 2017), which match closely the common problem-solving practices of the regular physics courses for the Chinese 8th grade students. The instrument is used both as the diagnostic post-test in Study 1, and as the pre- and post- tests in Study 2. In study 1, student's full responses to the diagnostic test were used to identify the specific misconceptions held by these students. In study 2, student's responses to the diagnostic test were dichotomously scored for quantitative analysis. The complete questions are included in the supplemental online materials. The face validity of the instrument was first evaluated by the group of researchers who designed this study and then by the extended group of physics teachers from the same

Table 2. Assessment questions used in this research.

Context	Questions
Force	1, 2, 7, 11
No Force	3, 9
Motion	4, 5, 8, 12, 13, 15
No Motion	6, 10, 14

school in which the study was conducted. The teachers whose students were participants of the studies were not exposed to the diagnostic tests before the studies were completed. Through the evaluation, all teachers and researchers agreed that the instrument was suitable for assessment of this student population. To support the reliability of the instrument, the Cronbach's α indexes of the pre-test (0.680) and post-test (0.706) in study 2 were computed, both of which achieved the minimally acceptable level (>0.65) (DeVellis, 2012).

Interview procedure

Interviews were conducted with a subset of the populations for both studies. Each interview lasted about half an hour. During the interviews, students were also asked to give their responses to each question and to write or sketch their work on their answer sheets. The interviews were audio taped and transcribed and analysed by a group of researchers including the authors and several similar level physics education specialists to identify students' misconceptions and reasoning processes. The practice of identifying students' misconceptions have been well-established with decades of research in the physics education community. In this study, the researchers all have received extensive training in this area and worked as a group to review students' responses and problem-solving procedures. Possible misconceptions were thoroughly discussed among the group and the ones commonly agreed by over 80% of the group were documented as the identified misconceptions.

Intervention lesson plan

In the schools involved in this study, the force and motion concepts were taught through four lessons that each lasts 45 min and consists of traditional lecturing, problem-solving practices, and a few lecture demonstrations. The new instruction intervention makes emphasis on developing the central idea and its applications in problem-solving. The instructional approach is designed to explicitly introduce the central idea and show its application in connecting and analysing contextual variables in problem-solving with lecture demonstrations and example problems. However, the instructional style is still similar to the traditional method, with the main change being the content focus.

The students are split into a control group and a treatment group. The control group attended four lessons as part of their normal school curriculum. The treatment group attended the same first two lessons followed by two intervention lessons. The 'pre-test' was given to both groups after the first two lessons and the results demonstrated that the two groups were homogeneous at that point (results shown below). The post-test was given after both groups had completed four total lessons on force and motion.

For students in the treatment group, during their first intervention lesson the concept of applied forces was defined rigorously. By considering the combined effect of all applied forces, the idea of the net force was introduced. Since the Chinese 8th grade students had not yet learned vector calculations, only 1 dimensional cases were used. Examples of these cases included an object with one external force, an object with two forces acting in the same direction, and an object with two forces acting in opposite directions. Afterwards, the concept of acceleration was introduced through a demonstration conducted in front of the whole class, which involved a glider moving between two fixed photogates on an air track while under a constant external force. The experiment was conducted for several trials. In each trial, the constant external force was set at a different value and the time for the glider to travel between photogates as well as the instantaneous velocities at the photogates were recorded. Students observe and discuss the relation between the rate of change of velocity with time and the net force applied on the glider and were guided to develop a qualitative and operational understanding of acceleration and how it varies with the net force.

The second intervention lesson builds on the relation between the net force and the acceleration developed from the first lesson and helps students to establish the expert-like conceptual pathway on force and motion. That is, all applied forces are summed to obtain the net force, which leads to the acceleration, which in turn changes the velocity. Through problem-solving examples and practices, students were introduced to different combinations of acceleration and direction of initial velocity and were guided to relate such conditions to the corresponding motion of an object, which varies among accelerating, decelerating, constant velocity, or stationary. More details of the intervention lessons are provided in the online supporting materials.

Quantitative analysis

In Study 2, students' pre–post testing performances across item contexts and intervention conditions were compared to identify evidences for the effectiveness of the new instruction approach. To do this, three-way analysis of variance (ANOVA) and independent or dependent sample T-test were conducted. The ANOVA is a 2*2*4 mixed design, in which the instruction condition is the between subjects variable with two levels (traditional and new instruction). Test condition and question context are within subjects variables for two test times (pre- and post-test) and four context categories (force, no force, motion and no motion). For all T-tests, the Bonferroni correction was applied to correct p values for multiple comparisons among different item contexts (Simes, 1986).

Next, students were classified into different profiles through latent profile analysis (LPA) based on their dichotomous scores to the questions in the diagnostic test. LPA is part of the family of latent variable models, which assume that a latent variable is responsible for the relations between the observed variables (Collins & Lanza, 2010). These analysis methods are often used to characterise student' patterns of mental model in science education (Flaig et al., 2018; e.g. Schneider & Hardy, 2013; Straatemeier, van der Maas, & Jansen, 2008; van Schijndel, van Es, Franse, van Bers, & Raijmakers, 2018). For example, Schneider and Hardy (2013) classified children's understanding of floating and sinking into five profiles, i.e. misconception profile, fragmented profile, indecisive profile, prescientific profile, and scientific profile.

Using similar approaches, students' conceptual profiles of force and motion were identified based on their performance on different context categories. In this analysis, a total of four latent profile models (1-, 2-, 3- and 4-profile modes) were fitted with the data and were evaluated using Bayesian information criteria (BIC, Schwarz, 1978) and the entropy statistic (Pastor, Barron, Miller, & Davis, 2007) in order to identify the optimal number of profiles. Meanwhile, the Bootstrapped Likelihood Ratio Test (BLRT, Arminger, Stein, & Wittenberg, 1999) was also used to determine whether there were statistically significant differences between two models fitted to the data. The best fitting model should have small BIC, large entropy (>0.9), and is statistically different from other models in BLRT ($p < 0.05$) (Rosenberg, Schmidt, Beymer, & Steingut, 2018).

The identified best-fitting model is then used to produce the final solution for the profiles of student conception on force and motion. To determine the statistical significance among the different profile groups, two-way ANOVA (questions context * profile groups) was used to evaluate the differences among students' performances across the four different contexts as a function of profile groups. Follow-up T tests were also used to compare student performances between specific profile groups. The LPA and statistical analysis will provide evidence to answer the second research question. For example, if students in a profile group perform best on all the four context categories, these students may be considered as having developed a well-connected expert-like knowledge structure. On the other hand, if students in a profile group only perform better in some context categories, it is suggested that these students' knowledge structures are still fragmented with varying levels of partially connected local fragments.

As part of the evaluation on the effectiveness of the instruction intervention, transitions of the distributions of student profiles on pre- and post-test for treatment and control groups are compared. To compare pre- post-test differences, the distribution of pre-test is set as the baseline, whereas the distribution of post-test is set as the observed variation. To compare the differences between the two instruction conditions, the distribution of students going through traditional instruction is set as the baseline while the distribution of students going through the new instruction is set as the observed variation. All these distribution patterns from different instruction and pre- post-test conditions are compared using *Chi*-square test to determine the significance. This analysis again provides evidences for comparing the effectiveness of the different instruction approaches. It is expected that a more effective instruction method will have more students transitioning to the scientific profiles from pre-test to post-test.

Finally, as a technical note, all the ANOVA, *T*-test and *Chi*-square test and are conducted in *SPSS* 19.0. The LPA is conducted using the R package *tidyLPA* (Rosenberg et al., 2018).

Results of Study 1

To identify the misconceptions held by Chinese students, 208 students were given the force and motion assessment as a post-test immediately after they had completed the regular instruction on Newton's first and second laws. Among these students, 14 were randomly selected to participate in think-out-loud interviews. Their answers and reasoning were documented and analysed, from which four specific misconceptions were identified and described below (also see Table 3).

Motion implies applied force. (A thrown object carries the force of the throw)

Responses to question 5 on the diagnostic test reveal students' believing that a thrown object continues to experience the force of the throw throughout its motion. This question focuses on a stone thrown straight in the air and asks what forces act on the stone while it is still moving upwards, 65.38% correctly chose that only gravity acts on the stone, while 31.7% believe that both gravity and the force from the throw are acting on the stone. Out

Table 3. Students' responses on selected test items that assess common misconceptions on force motion.

Misconception Types	Test Item	Correct Choice (%)	Misconception Choice (%)
I. Motion implies applied force	Q5	A (65.38%)	C (31.70%)
II. No motion implies no applied force	Q6	D (56.73%)	B (37.02%)
III. Constant force implies constant velocity	Q2	A (14.42%)	C (57.69%)
IV. No force implies no motion.	Q3	C (54.81%)	A (28.37%)

of the 14 students interviewed, 7 believed that the force exerted by the person continues while the object is in motion.

No motion implies no net force.

Responses to question 6 on the diagnostic test reveal students' thinking that if an object has no motion, then it must experience no net force. This question asks why a stone thrown upward does not continue to move past its highest point, 56.73% correctly chose that gravity is responsible for slowing the stone to a stop, while 37.02% chose that gravity and the force from the throw balance at this point. Six of the 14 interviewed students also stated that because the object is stationary at the highest point during free fall, it must experience no force at that point.

Constant force implies constant velocity.

Responses to question 2 on the diagnostic test reveal prevalence of students' belief that application of a constant force causes an object to move at a constant velocity. In this question, a person is pushing a car along a horizontal surface with no friction between the car and the surface. Students are asked what happens to the car if the person applies a constant force. Only 14.42% of students correctly chose that the car will move with a continuously increasing speed, while 57.69% chose that the car would move with a constant speed.

Out of the 14 students interviewed, 10 held this misconception. These students directly relate an increase in an applied force to a corresponding increase in velocity due to certain real-life experiences, in which the involved frictional force was not clearly understood. They then conclude that objects experiencing a constant force must have a constant velocity.

No force implies no motion.

Responses to question 3 on the diagnostic test exemplify the misconception that directly links force and motion. This question again focuses on a person pushing a car on a frictionless surface just as in question 2 above. This time, the person suddenly stops applying the force and students are asked what happens to the car, 54.81% correctly responded that the car would move with a constant speed, while 34.52% responded that the car would either move with decreasing speed or immediately stop because there was no longer a force pushing the car.

Out of the 14 students interviewed, 8 held this common misconception. The reasoning by these 8 students were split. Four believed that if the external force decreases or disappears, the object will slow down or stop moving. The other four believed that stationary objects do not experience any force, so if an object doesn't experience an external force then it should not move.

Based on students' explanations, these misconceptions reflect a general belief that motion is directly related or even proportional to an applied force or net force, which was developed through experience of motion in a world with friction but without a clear understanding of how friction influences motion. The different misconceptions can be viewed as situational manifestations of this belief expressed within different contextual configurations such as those listed in Table 2. In responding to the questions, students were not consistently expressing this belief across different contexts. The result suggests that the misconceptions can be viewed, in the conceptual framework outlined in Figure 1, as different local connections linking certain subsets of the contextual conditions, forming a fragmented knowledge structure. Different students may develop different subsets of local connections for different contexts, causing the inconsistency in their reasoning.

Results of Study 2

In Study 2, students in the intervention and control groups were pre–post tested to evaluate the effectiveness of the new instruction intervention. The pre–post data along with interview results were further analysed to gain insights of fine grained conceptual states among students in the different instruction conditions. The outcomes of the analysis are given below.

Effectiveness of the instruction intervention

The performances of students enrolled in both instruction conditions are compared and summarised in Figures 2 and 3. To compare the changes in different question contexts between pre- and post-tests for the two instruction conditions, a three-way ANOVA was conducted. The results suggest significant interaction among instruction condition, test condition (pre- or post-test), and question contexts ($F(3, 167) = 5.136, p < 0.001$, $\eta_p^2 = 0.084$).

To evaluate the effectiveness of the instruction intervention, pre- and post-test performances of the two instruction groups were compared. As shown from Figures 2 and 3, the pre-test scores of the two groups are statistically identical. Interactions between instruction conditions and question contexts on the pre-test were analysed with a two-way ANOVA, which reveals insignificant interactions on the pre-test ($F(3, 167) = 0.683$, $p = 0.564$, $\eta_p^2 = 0.012$). Therefore, these two groups of students can be considered as coming from a single homogeneous population.

On the post-test, the overall performances of the two groups depart significantly (see Figure 2): the students having the new instruction made significant pre–post improvement ($t(81) = 7.83$, $p < 0.001$, $d = 0.86$), while the students in the traditional instruction didn't change much ($t(88) = 0.79$, $p = 0.431$, $d = 0.08$). In addition, the post-test ANOVA reveals significant interactions ($F(3, 167) = 4.488$, $p = 0.005$, $\eta_p^2 = 0.075$),

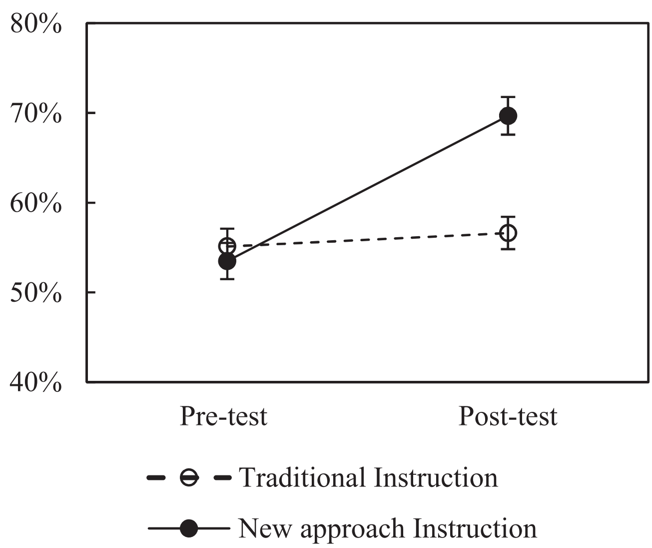

Figure 2. The overall pre-post testing performance of students in the control and treatment groups. The error bars represent the standard errors of the means.

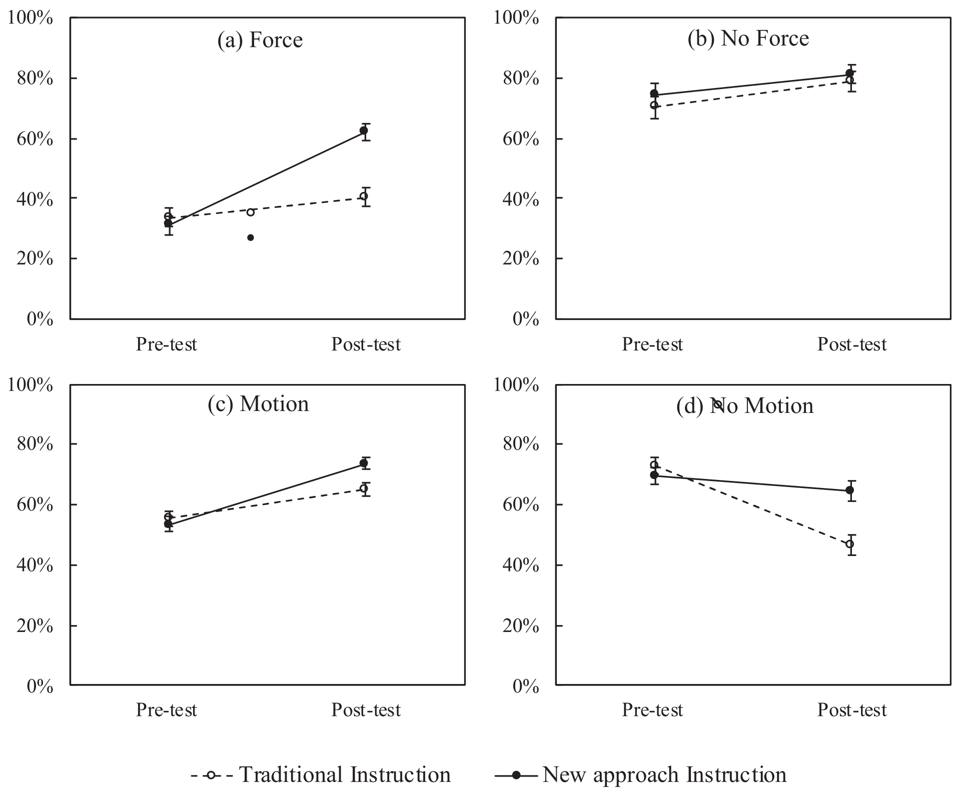

Figure 3. Pre-post testing performance of students in the control and treatment groups in four context categories. The error bars represent the standard errors of the means.

with main effects for both item contexts and instruction conditions ($F(3, 167) = 39.846$, $p < 0.001$, $\eta_p^2 = 0.417$; $F(1, 169) = 18.617$, $p < 0.001$, $\eta_p^2 = 0.099$ respectively). Based on the ANOVA and T-test results, the new instruction method shows a positive effect on students' learning.

Interactions between test conditions and question contexts are also significant from the results of two-way ANOVA for each instruction condition ($F_{trad}(3, 86) = 22.800$, $p < 0.001$, $\eta_p^2 = 0.443$; $F_{new}(3, 243) = 20.606$, $p < 0.001$, $\eta_p^2 = 0.203$). This demonstrates that students answer questions in contexts differently on the pre- and post- tests, regardless of the instruction condition. However, the variation of ways in which students answer these questions is affected by the instruction condition.

When inspecting across the different context categories, students' performances vary substantially (see Figure 3). For the context of Force, students in the new instruction had much larger improvement ($t(81) = 8.41$, $p < 0.001$, $d = 0.93$) over the traditional instruction group ($t(88) = 2.55$, $p = 0.012$, $d = 0.27$). For the context of No Force, students in both instruction groups perform similarly with a small pre–post improvement ($t_{trad}(88) = 2.470$, $p = 0.015$, $d = 0.26$; $t_{new}(81) = 1.620$, $p = 0.109$, $d = 0.18$). For the context of Motion, the group having new instruction again made more improvement ($t(81) = 8.00$, $p < 0.001$, $d = 0.88$) than the students with traditional instruction ($t(88) = 4.38$, $p < 0.001$, $d = 0.46$). Finally, for the context of No Motion, both groups had negative changes but the students having the new instruction have a much smaller decrease ($t(81) = -1.24$, $p = 0.218$, $d = -0.14$) than the students with traditional instruction ($t(88) = -7.43$, $p < 0.001$, $d = -0.79$). When directly comparing pre–post score gains of the traditional and new instruction groups across the four context categories, the differences are statistically significant on three context categories, including Force ($t(169) = 5.33$, $p < 0.001$, $d = 0.83$), Motion($t(169) = 3.10$, $p < 0.001$, $d = 0.48$), and No Motion ($(169) = 4.09$, $p < 0.001$, $d = 0.63$) while the difference on No Force is insignificant ($t(169) = -0.323$, $p = 0.747$, $d = 0.05$).

The results discussed above provide encouraging evidence for the effectiveness of the new instruction approach, which seem to have improved students' overall performance and performances on the Force and Motion context categories. The No Force context reveals no significant improvement, which may be related to the emphasis of the new instruction method on the central idea of force and its relation to acceleration and then motion. This method does not explicitly emphasise on what occurs when there is no net force.

For the No Motion context, students in both instruction groups had negative pre–post changes, however, the students in the new instruction had a significantly smaller negative shift compared to the students with traditional instruction. However, understanding the possible causes for the negative change will need further exploration, which is beyond the scope of this study. The results on the contexts of No Motion and No Force warrant further research and refinement on instruction method.

The above analysis corroborates Alonzo and Steedle's (2009) finding that students' performances in the four contexts are not directly dependent on each other. This lends support to the idea of a fragmented knowledge structure: knowledge and learning are strongly situated within each context, in which the learning took place and are largely separated from other contexts in a novice learner's knowledge structure.

Students' conceptual profiles and development

Latent profile analysis (LPA) is conducted to classify students into different conceptual profiles based on students' dichotomous scores to the questions in the diagnostic test. Statistics of the different profiles and pre–post transitions are analysed to help clarify students' conceptual development pathways through different instructions.

In this analysis, a total of four latent profile models (1-, 2-, 3- and 4-profile modes) have been fitted with the data and the fitting evaluation results are summarised in Table 4. From Table 4, the 3-profile model appears to have the best overall fitting. It has the smallest BIC, although its entropy value is slightly lower than that of the 2-profile model. The results of BLRT more clearly support that the 3-profile model fits significantly better than the 2-profile model ($p < 0.001$).

Using the 3-profile model, students are classified into 3 profile groups. Performances of these three groups of students are detailed in Table 5 and illustrated in Figure 4, with sub-scores on the four context categories. The results include both pre- and post-test scores for students as independent measures such that the total sample size becomes $N = 342$.

Profile 1 includes the best performing students ($N = 72$), who had achieved relatively high scores in all contexts. In particular, these students scored perfectly on the No Force context and there were no significant differences among other three contexts. Therefore, these students can be considered as having developed sufficiently well-connected knowledge structures and obtained a 'proficient' level of understanding on the force motion concept.

Profile 2 consists of students with 'mixed' performances ($N = 132$). These students scored perfectly on the No Force context, but performed mediocrely on Motion and No Motion contexts, and worst on the Force context. The result suggests that students in profile 2 started to make significant improvement on one of the four contexts but still maintained poor to mediocre performances on the remaining three contexts. The differences in conceptual development among the different contexts are evidences of students' having moderately fragmented knowledge structures. Therefore, students in profile 2 can be considered as the 'mixed' group, with mixed levels of development among different contexts.

Finally, Profile 3 is considered as the 'novice' group with students having low performance across all contexts. For this group of students, there are also variations among the different contexts; their scores on Motion and No Motion contexts are slightly better than those on Force and No Force contexts. These variations reflect the outcomes of students' applying different contextually situated misconceptions held by these students.

Table 4. Statistics of latent profile analysis.

Number of Profiles	BLRT (p-value)	BIC	entropy
1	/	339.106	1.000
2	80.21 (0.001)	288.071	0.968
3	133.25 (0.001)	183.999	0.950
4	13.37 (0.073)	199.807	0.908

The p-value listed under BLRT reflects the statistical significance when comparing the current model with the previous model. The results show that the 2-profile model fits significantly better than the 1-profile model ($p < 0.001$), and the 3-profile model is also significantly better than the 2-profile model ($p < 0.001$). But the 4-profile model is not significantly better than the 3-profile model ($p = 0.073$). This supports that the 3-profile model is the best fitting model.

Table 5. Pre-post testing performances of students in different profile groups. The scores are in percentages and the standard errors are given in brackets.

	Number of students									
		Traditional		New						
Profile	Total	Pre	Post	Pre	Post	Force	No Force	Motion	No Motion	Total
1	72	10	17	9	36	81.25 (1.83)	100 (0.00)	81.71 (1.59)	85.19 (2.37)	84.72 (1.74)
2	132	35	40	39	18	25.57 (1.58)	100 (0.00)	61.74 (1.76)	60.1 (2.60)	56.87 (1.80)
3	138	44	32	34	28	36.23 (2.27)	40.58 (1.67)	51.57 (1.81)	54.83 (2.58)	46.67 (2.09)

The results in Table 5 are evaluated for statistical significance with ANOVA, which shows significant interactions between profiles and item contexts ($F(6, 1017) = 86.394, p < 0.001$, $\eta_p^2 = 0.338$). Post-hoc comparisons suggest that the proficient group performs significantly better than the mixed or novice group regarding the total score and scores on Force, Motion and No Motion contexts (see Table 6 for the pair-wise comparisons). Here, the acceptable level of p-value is set as 0.017 based on the Bonferroni correction procedure for multiple comparisons (i.e. 0.05/3 = 0.017 for 3 comparisons between any two of the three profiles). Although the proficient group is significantly better than the mixed group on the Force, Motion, and No Motion contexts by a large margin, the two are identical in the No Force context. Meanwhile, the proficient group is significantly better than the novice group on all contexts by a large margin. Comparing the mixed and novice groups, the largest and most significant difference is on the No Force context. The two perform similarly on the No Motion context and the mixed group is moderately ahead in the Motion context, while the novices are moderately ahead on the Force context.

Transitions of students among different profiles from pre- to post-test can provide useful information on the possible progressions of students' learning. In Figure 5, the distributions of profiles and their conditional probabilities for transitions from pre- and post-test are illustrated. Regarding the profile distributions under different conditions, chi-

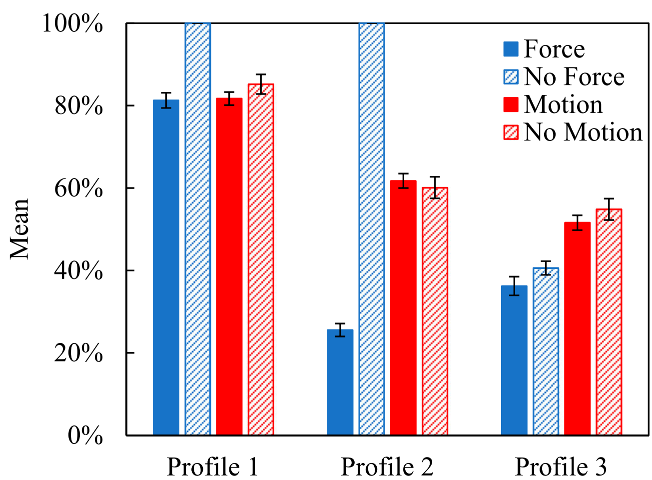

Figure 4. Performance means of students categorised in three latent profiles.

Table 6. Pair-wise t-test comparison between profiles. Here, *d* is the Cohen's *d*.

	Profile 1 vs. Profile 2				Profile 1 vs. Profile 3				Profile 2 vs. Profile 3			
Context	*df*	*t*	*p*	*d*	*df*	*t*	*p*	*d*	*df*	*t*	*p*	*d*
Force	202	22.0	<0.001	3.24	208	15.4	<0.001	1.93	268	−3.82	<0.001	0.47
Motion	202	8.42	<0.001	1.106	208	12.5	<0.001	1.60	268	4.03	<0.001	0.49
No Motion	202	7.14	<0.001	0.94	208	8.66	<0.001	1.14	268	1.44	0.152	0.18
No Force	202	0	1	0	208	35.6	<0.001	3.75	268	35.6	<0.001	4.25
Total	202	15.7	<0.001	1.11	208	19.8	<0.001	1.37	268	5.22	<0.001	0.32

squared tests show that there is no difference between profile distributions of students in the two instructional conditions on the pre-test ($\chi^2(2) = 1.266, p = 0.531$), which further confirms the equivalence of the two sample groups. From pre- to post-test, students in the traditional instruction group reveal no difference in their profile distributions ($\chi^2(2) = 4.043,\ p = 0.132$). For students receiving new instruction, their pre–post profile distributions are significantly different ($\chi^2(2) = 24.517, p < 0.001$ Cramer's $V = 0.387$).

When comparing the pre–post profile transitions, the transition patterns are also significantly different between the two instruction groups. For example, the favoured transitions include those from profile 2 or 3 to profile 1. The group receiving new instruction had significantly more students transitioning to profile 1 from profile 2 or 3 than the traditional instruction group did ($\chi_2^2(2) = 12.054,\ p < 0.001$, Cramer's $V = 0.404$, Adjusted standardized residuals$_{2\to1} = 2.9$, $\chi_3^2(2) = 9.797$, $p < 0.001$, Cramer's $V = 0.354$, Adjusted standardized residuals$_{3\to1} = 3.1$). It is also noted that students started in profile 1 on pre-test were in small numbers for both groups. Therefore, comparisons on students who had started and remained in profile 1 on pre- and post-test cannot be conducted due to statistical uncertainties. However, it would be interesting to investigate this subgroup of students in future with a larger sample size. To sum up, the analysis of profile transitions provides fine grained evidence

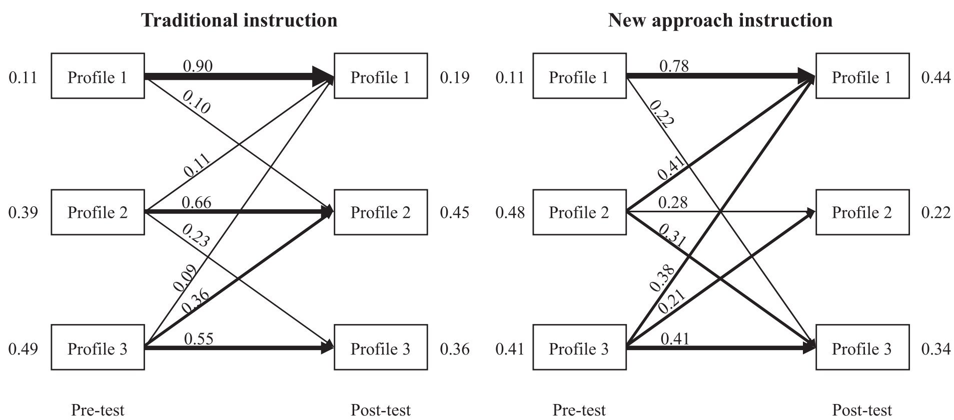

Figure 5. Profile distributions and pre-post transitions in different instruction conditions. The arrows represent the observed transitions. The fractions next to the profile labels give the probability distributions in pre- or post-test. The fractions above the arrows and the thickness of the arrows show the conditional probabilities of transitions among students categorised in different profiles.

to support that the new instruction approach can significantly improve students' learning towards the favoured conceptual profile.

Results of qualitative analysis of Study 2

Qualitative data from student interviews were collected and analysed in this study to gain deeper understanding of possible students' conceptual development pathways. After the post-test a total of 18 students from the sample pool of students were randomly selected to participate in the 'think-out-loud' interviews. Based on students' responses, three main themes of problem solving behaviours emerge.

Proficient students ($N = 9$) were able to give accurate analysis on most force and motion problems. Specifically, they recognised the net force acting on objects as the key conceptual underpinning and can correctly relate this to acceleration, which extends to descriptive variables of motion such as velocity and displacement. Their implied knowledge structure is very similar to that of an expert illustrated in the top half of Figure 1.

In contrast, novice students ($N = 6$) made a range of direct connections between force and motion, which were expressed in a variety of misconceptions developed through prior experiences. All students believed that force and motion are directly and mutually corresponding to each other: whenever there is a force, there must be motion; and vice versa. Furthermore, the net force concept didn't seem to hold a place among these students, instead, they would relate applied forces directly to an object's motion. In response to situations involving a cart moving on a surface with no friction, these students had responses such as, 'if the cart only experiences a constant external force, the cart will remain moving at a constant velocity,' and 'when the external force is removed, the cart will stop.' These have been previously documented in the literature as some of the most common misconceptions on force and motion.

For mixed students ($N = 3$), their overall knowledge structure appeared similar to that of the novices, but with a slight alteration. Uniquely, these students seemed to recognise that zero net force results in constant velocity and were able to quickly recall this in relevant contexts. This allowed them to correctly solve a number of mechanics problems related to the equilibrium state (i.e. zero net force). However, they did not have a fully developed concept of net force and were unable to correctly analyze situations with a non-zero net force. For example, these students would give responses such as, 'if the spaceship is moving at a constant velocity, then it must either experience a zero net force or no force at all,' yet they'd also reply, 'when only one constant external force is exerted on an object, the velocity of the object will increase to a maximum, and then continue to move at a constant velocity.' The latter response is again a typical variation of the misconception that velocity is proportional to an applied force.

Based on the students' responses in the interviews, it appears that mixed students solve problems through pattern matching based on their previous experiences on similar problems instead of deep understanding of the fundamental concept. This reflects a fragmented knowledge structure formed with memorised situations rather than a fully connected conceptual framework. Therefore, it appears that the transition from novice to mixed group can be achieved with substantive success in the traditional instruction (see Figure 5) through drilling on force and motion problems. However, transition to the proficient level was nearly nonexistence in the traditional instruction (see Figure 5 for the pre–

post transition probabilities), but seemed to have been achieved by the new instruction. The emphasis on building connections with the central idea seems effective in promoting knowledge integration.

Discussion and conclusion

The first study showed that the focus on problem-drilling in Chinese instruction did not minimise the existence of common misconceptions on force and motion. New instruction methods are needed to aid in conceptual development, which were explored in the second study. The results show that a new instruction approach designed to target forming connections between fragments of the force and motion concept through the central idea is effective in helping students develop a more integrated knowledge structure and achieve better performance in conceptual testing.

The results of this study support the use of the conceptual framework analysis outlined in Figure 1. Experts correctly focus on the central idea of the concept ($F_{net} = ma$), regardless of context. Novice learners primarily rely on memorizations of direct associations between force and motion, which are locally linked and partially functional for limited problem contexts. Furthermore, evidence for fragmentation of the student knowledge structure is demonstrated by inconsistent beliefs between contextually situated understandings and changes when transitioning from one profile to another. An expert is able to view a set of context as conceptually isomorphic, while the mixed and novice students often treat each context as a unique situation with a context specific understanding disconnected from other related situations.

The key feature of the conceptual framework analysis is the explicit emphasis on the central idea of a concept and its role as the central node connecting a network of ideas to form an integrated knowledge structure (National Research Council, 2012a, 2013). The central idea network is also the focus of the new instruction method. Without a good understanding of the central idea, students fall back upon other learning strategies, primarily basing their beliefs on prior experiences and memorising context-matched solutions. Therefore, research into identifying the central idea of a concept and helping students make connections through the central idea may provide useful insights on developing effective instruction for knowledge integration.

Limitations and suggestions for future studies

When interpreting the result of this study, it is important to note a possible weakness of this design, which engages a randomised approach with a single sample ($N_{total} = 171$). Due to this limitation in sampling, the results cannot be generally extended to students from other types of education systems. It would be beneficial to investigate if similar results can be reproduced with students taught in different types of instruction.

In addition, the statistical treatment may be excessive for this small-scale data. Although the conclusions from this study are supported by statistical analysis as being statistically significant, the confidence of such outcomes can be improved with a larger sample size, which could be further explored in future studies for replication and extension. Using a larger sample can also reveal more finer-grained conceptual profiles of students' understanding, which may help establish a progression of knowledge integration.

Furthermore, deeper insights of students' knowledge integration can be gained by additional quantitative and qualitative data collection methods such as concept-mapping and network analysis (Kubsch et al., 2018; Novak & Cañas, 2006).

Finally, the hypothesis of promoting knowledge integration of conceptual understanding through the central idea of a concept has been proposed and tested in this study with a single content topic. It would be beneficial to explore the general applicability of this method with different content domains, which are being pursued in follow-up studies.

Acknowledgements

Any opinions expressed in this work are those of the authors and do not necessarily represent those of the funding agencies.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Science Foundation Award DUE-1712238.

ORCID

Yang Xiao http://orcid.org/0000-0002-2571-7759
Joseph C. Fritchman http://orcid.org/0000-0002-5496-0379
Qiaoyi Liu http://orcid.org/0000-0002-1613-7891
Lei Bao http://orcid.org/0000-0003-3348-4198

References

Alexopoulou, E., & Driver, R. (1996). Small-group discussion in physics: Peer interaction modes in pairs and fours. *Journal of Research in Science Teaching*, *33*(10), 1099–1114. doi:10.1002/(SICI)1098-2736(199612)33:10<1099::AID-TEA4>3.0.CO;2-N

Alonso, M. (1992). Problem solving vs. conceptual understanding. *American Journal of Physics*, *60*(9), 777–778. doi:10.1119/1.17056

Alonzo, A. C., & Steedle, J. T. (2009). Developing and assessing a force and motion learning progression. *Science Education*, *93*(3), 389–421.

Anggoro, S., Widodo, A., Suhandi, A., & Treagust, D. F. (2019). Using a discrepant event to facilitate preservice elementary teachers' conceptual change about force and motion. *Eurasia Journal of Mathematics, Science and Technology Education*, *15*, 8. doi:10.29333/ejmste/105275

Arminger, G., Stein, P., & Wittenberg, J. (1999). Mixtures of conditional mean- and covariance-structure models. *Psychometrika*, *64*(4), 475–494.

Bagno, E., Eylon, B.-S., & Ganiel, U. (2000). From fragmented knowledge to a knowledge structure: Linking the domains of mechanics and electromagnetism. *American Journal of Physics*, *68*(S1), S16–S26.

Bao, L., Hogg, K., & Zollman, D. (2002). Model analysis of fine structures of student models: An example with Newton's third law. *American Journal of Physics*, *70*(7), 766–778. http://dx.doi.org/10.1119/1.1484152

Bao, L., & Redish, E., F. (2001). Concentration analysis: A quantitative assessment of student states. *American Journal of Physics*, *69*(S1), S45–S53. http://dx.doi.org/10.1119/1.1371253.

Bao, L., & Redish, E. F. (2006). Model Analysis: Representing and Assessing the Dynamics of Student Learning. *Physical Review Special Topics - Physics Education Research*, *2*(1), 010103. http://dx.doi.org/10.1103/physrevstper.2.010103

Beichner, R., Bernold, L., Burniston, E., Dail, P., Felder, R., Gastineau, J., … Risley, J. (1999). Case study of the physics component of an integrated curriculum. *American Journal of Physics*, *67*(S1), S16–S24.

Brown, A. (1989). Analogical learning and transfer: What develops? In S. Vosniadu, & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 369–412). New York: Cambridge U.P.

Champagne, A. B., Klopfer, L. E., & Anderson, J. H. (1980). Factors influencing the learning of classical mechanics. *American Journal of Physics*, *48*, 1074–1079.

Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152.

Chiu, M.-H., Guo, C.-J., & Treagust, D. F. (2007). Assessing students' conceptual understanding in science: An introduction about a national project in Taiwan. *International Journal of Science Education*, *29*(4), 379–390.

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis*. doi:10.1002/9780470567333

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

DeVellis, R. F. (2012). *Scale development: Theory and applications* (Vol. 26). Thousand Oaks, California, US: Sage publications.

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, *25*(6), 671–688.

Dykstra, D. I., Jr., Boyle, C. F., & Monarch, I. A. (1992). Studying conceptual change in learning physics. *Science Education*, *76*(6), 615–652.

Flaig, M., Simonsmeier, B. A., Mayer, A.-K., Rosman, T., Gorges, J., & Schneider, M. (2018). Conceptual change and knowledge integration as learning processes in higher education: A latent transition analysis. *Learning and Individual Differences*, *62*, 49–61. doi:10.1016/j.lindif.2017.12.008

Gilbert, J. K., & Watts, D. M. (1983). *Concepts, misconceptions and alternative conceptions: Changing perspectives in science education*.

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

Halloun, I. A., & Hestenes, D. (1985a). Common sense concepts about motion. *American Journal of Physics*, *53*(11), 1056–1065.

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

Hardiman, P. T., Dufresne, R., & Mestre, J. P. (1989). The relation between problem categorization and problem solving among experts and novices. *Memory & Cognition*, *17*(5), 627–638.

Harris, J., George, N. R., Hirsh-Pasek, K., & Newcombe, N. S. (2018). Where will it go? How children and adults reason about force and motion. *Cognitive Development*, *45*, 113–124. doi:10.1016/j.cogdev.2018.01.002

Jong, A. J. M. d., & Ferguson-Hessler, M. G. M. (1986). Cognitive structures of good and poor novice problem solvers in physics. *Journal of Educational Psychology*, *78*(4), 279–288. doi:10.1037/0022-0663.78.4.279

Keller, C., Finkelstein, N., Perkins, K., Pollock, S., Turpen, C., & Dubson, M. (2007). Research-based practices for effective clicker use. *AIP Conference Proceedings*, *951*, 128–131. AIP.

Kim, E., & Pak, S.-J. (2002). Students do not overcome conceptual difficulties after solving 1000 traditional problems. *American Journal of Physics*, *70*(7), 759–765.

Kubsch, M., Nordine, J., Neumann, K., Fortus, D., & Krajcik, J. (2018, June 23). *Measuring integrated knowledge – a network analytical approach*. 2. Retrieved from http://ccl.northwestern.edu/2018/ICLS2018Volume3_proceedings.pdf

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*(4450), 1335–1342.

Laws, P. W. (2004). Workshop Physics Activity Guide, Module 4: Electricity and Magnetism. In *Workshop Physics Activity Guide*. Wiley-VCH.

Lee, H.-S., Liu, O. L., & Linn, M. C. (2011). Validating measurement of knowledge integration in science using multiple-choice and explanation items. *Applied Measurement in Education*, *24* (2), 115–136. doi:10.1080/08957347.2011.554604

Linn, M. C. (2005). The knowledge integration perspective on learning and instruction. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 243–264). doi:10.1017/CBO9780511816833.016

McDermott, L. C. (1996). *Physics by inquiry: An introduction to physics and the physical sciences* (Vol. 1). New York, NY: John Wiley & Sons.

Minstrell, J. (1992). Facets of students' knowledge and relevant instruction. In R. Duit, F. Goldberg, & H. Niedderer (Eds.), *Proceedings of the international workshop: Research in physics learning-theoretical issues and empirical studies* (pp. 110–128). Kiel: The Institute for Science Education.

Nakhleh, M. B. (1993). Are our students conceptual thinkers or algorithmic problem solvers? Identifying conceptual students in general chemistry. *Journal of Chemical Education*, *70*(1), 52. doi:10.1021/ed070p52

National Research Council. (2011). *Assessing 21st century skills: Summary of a workshop*. doi:10.17226/13215

National Research Council. (2012a). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Presshttps://doi.org/10.17226/13165.

National Research Council. (2012b). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: The National Academies Press.

National Research Council. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Presshttps://doi.org/10.17226/18290.

Nordine, J., Fortus, D., Lehavi, Y., Neumann, K., & Krajcik, J. (2018). Modelling energy transfers between systems to support energy knowledge in use. *Studies in Science Education*, *54*(2), 177–206. doi:10.1080/03057267.2018.1598048

Nordine, J., Krajcik, J., & Fortus, D. (2011). Transforming energy instruction in middle school to support integrated understanding and future learning. *Science Education*, *95*(4), 670–699. doi:10.1002/sce.20423

Novak, J. D., & Cañas, A. J. (2006). The origins of the concept mapping tool and the continuing evolution of the tool. *Information Visualization*, *5*(3), 175–184.

Pastor, D. A., Barron, K. E., Miller, B. J., & Davis, S. L. (2007). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology*, *32*(1), 8–47. doi:10.1016/j.cedpsych.2006.10.003

Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher*, *18*(1), 16–25.

Rosenberg, J. M., Schmidt, J. A., Beymer, P. N., & Steingut, R. R. (2018). *Interface to mclust to easily carry out Latent Profile Analysis [Statistical software for R]*. Retrieved from https://github.com/jrosen48/tidyLPA

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational Psychologist*, *24*(2), 113–142.

Schneider, M., & Hardy, I. (2013). Profiles of inconsistent knowledge in children's pathways of conceptual change. *Developmental Psychology*, *49*(9), 1639–1649. doi:10.1037/a0030976

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

Shen, J., Liu, O. L., & Chang, H.-Y. (2017). Assessing students' deep conceptual understanding in physical sciences: An example on sinking and floating. *International Journal of Science and Mathematics Education*, *15*(1), 57–70. doi:10.1007/s10763-015-9680-z

Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika*, *73*(3), 751–754.

Smith, M. U. (1992). Expertise and the organization of knowledge: Unexpected differences among genetic counselors, faculty, and students on problem categorization tasks. *Journal of Research in Science Teaching*, *29*(2), 179–205. doi:10.1002/tea.3660290207
Stamovlasis, D., Tsaparlis, G., Kamilatos, C., Papaoikonomou, D., & Zarotiadou, E. (2005). Conceptual understanding versus algorithmic problem solving: Further evidence from a national chemistry examination. *Chemistry Education Research and Practice*, *6*(2), 104–118.
Straatemeier, M., van der Maas, H. L. J., & Jansen, B. R. J. (2008). Children's knowledge of the earth: A new methodological and statistical approach. *Journal of Experimental Child Psychology*, *100*(4), 276–296. doi:10.1016/j.jecp.2008.03.004
Suzhou Education Examination Authority. (2017). *The physicstest forJunior High School graduation and Senior High School entrance*. Suzhou.
Thornton, R. K., & Sokoloff, D. R. (1998). 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*, *66*(4), 338–352. https://doi.org/10.1119/1.18863.
Trumper, R., & Gorsky, P. (1996). A cross-college age study about physics students' conceptions of force in pre-service training for high school teachers. *Physics Education*, *31*(4), 227–236.
van Schijndel, T. J. P., van Es, S. E., Franse, R. K., van Bers, B. M. C. W., & Raijmakers, M. E. J. (2018). Children's mental models of prenatal development. *Frontiers in Psychology*, *9*, doi:10.3389/fpsyg.2018.01835
Viennot, L. (1979). Spontaneous reasoning in elementary dynamics. *European Journal of Science Education*, *1*(2), 205–221.
Wu, W. (2017). Relationship between science teachers' conceptions of assessment of students' academic performance and their instructional approaches. In L. L. Liang, X. Liu, & G. W. Fulmer (Eds.), *Chinese science education in the 21st Century: Policy, practice, and research: 21* 世纪中国科学教育政策、实践与研究 (pp. 259–293). doi:10.1007/978-94-017-9864-8_11
Zhu, Z., & Fan, X. (2012). Reflections on teachers' conceptions of teaching. *Chinese Education & Society*, *45*(4), 42–55. doi:10.2753/CED1061-1932450403