Rethinking Representational Competence: cognitive mechanisms, empirical studies, and the design of a new media intervention

A Thesis

Submitted to the

Tata Institute of Fundamental Research, Mumbai

for the degree of

Doctor of Philosophy (Ph. D.) in Science Education

by

Prajakt Pande

Homi Bhabha Centre for Science Education,

Tata Institute of Fundamental Research, Mumbai.

February, 2018

(Final Version Submitted in May, 2018)

Declaration

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions. The work was done under the guidance of Dr. Sanjay Chandrasekharan at the Tata Institute of Fundamental Research, Mumbai.

Prajakt Pande

In my capacity as supervisor of the candidate's thesis, I certify that the above statements are true to the best of my knowledge.

Sanjay Chandrasekharan

Dedicated to my parents and my wife

This journey would not have been possible

without your patience, support, encouragement and belief in me.

Acknowledgements

Recognition is due to the following persons and institutions for making this research possible:

My Ph. D. advisor, mentor and friend; Sanjay Chandrasekharan, I can't thank him enough for his dedication to and guidance in this research. Sanjay's patience, open-door policy and countless hours spent on academic discussion and reviews are gratefully and humbly appreciated. It has been a pleasure and privilege to have worked with an academic of this calibre, who still manages to harbour great modesty within. Sanjay, you are one of the coolest people I have ever known!

My thesis advisory committee members; Jayashree Ramadas and Savita Ladage, for their succinct and constructive inputs to the work. I remember how Jayashree and Savita took notes and regularly followed up on them during our formal as well as informal meetings about my progress, difficulties, future plans, and their suggestions and feedback on them. Those meetings were where I first learnt to present and defend my work.

The Director of HBCSE, K. Subramaniam and the Dean, Sugra Chunawala; for their support in both academic and non-academic matters. My family and I are particularly thankful for the genuine concern they expressed and the help they offered during the most difficult time in my health-life last year.

Late Prof. Chitra Natarajan for helping me through the initial stages of my proposal defense. Chitra, you will always be remembered for your jolly and lively nature!

Other faculty, students and academic staff at HBCSE for their feedback on my work during the ARMs, epiSTEMEs and other formal as well as non-formal events.

The entire administrative staff at HBCSE for encouragement and concern for the completion of my degree. They have been extremely helpful in a number of non-academic matters, making my life here in HBCSE easier, convenient and comfortable. Most people from the staff have been close friends during my stay; I will particularly miss participating in the cultural and sports events happening annually during the Dasara period at HBCSE.

My parents; who always encouraged me to pursue what I enjoyed most, and supported me in every possible way, whenever I found myself in trouble. This book is for you!

My wife and confidant; Chaitra, her belief in me and her complete unselfishness (well, not entirely!), unconditional love and compassion have been astonishing. Her companionship is valueless! Thank you for being with me, especially during the tough times, which is almost always during a Ph. D.

My sister Preeti; for boosting my confidence every now and then, especially during the early stages of graduate school when the mind is in great dilemmas related to life-choices! Her two adorable daughters (Anushka and Mugdha), who seek inspiration in me, always reminded me to not give up!

Members of the Learning Sciences Research Group (LSR) and friends; for the intense and fruitful academic (and non-academic) analysis, interactions and discussions, and for all the encouragement. Keep the faith!

My friends and extended family; who have been by my side throughout the many years of study. I am a doctor (well, the one with a Ph. D.) finally!

The Staff and students of Nutan Vidya Mandir, Jai Hind College, IIT Bombay, and our own CUBE programme; for facilitation and willingness to participate in my experiments.

Anonymous reviewers and editors of the different journals and conferences to which I submitted my work, for clear and concise comments.

Part of this work was possible because of funding support received from the Cognitive Science Initiative, Department of Science and Technology, Government of India.

List of publications (related to the dissertation work)

1. Pande, P., & Chandrasekharan, S. (under review). Expertise in chemistry is marked by sensorimotor tuning that helps pickup information quickly from external representations. **(Chapter 4)**

2. Pande, P., & Chandrasekharan, S. (2017). Representational competence: towards a distributed and embodied cognition account, *Studies in Science Education*, *53*(1), 1-43. UK: Routledge. **(Chapters 2 & 3)**

3. Pande, P., Shah, P. & Chandrasekharan, S. (2015). How do experts and novices navigate chemistry representations – an eye tracking investigation, In S. Chandrasekharan, S. Murthy, G. Banarjee, & A. Muralidhar (Eds.), *Proceedings of EPISTEME-6,* 102-109, HBCSE-TIFR, Mumbai, India. **(Chapter 4)**

4. Pande, P., & Chandrasekharan, S. (2014). Categorization Of Multiple External Representations By Chemistry Undergrads: An Eye-Tracking Study, In Y.-J. Lee, N. T-L. Lim, K. S. Tan, H. E. Chu, P. Y. Lim, Y. H. Lim, & I. Tan (Eds.), Proceedings of the International Science Education Conference 2014, 1393-1415. NIE-Singapore. **(Chapter 4)**

5. Pande, P., & Chandrasekharan, S. (2014). Integration of multiple external representations in chemistry: a requirements gathering study, In C.-C. Liu, Y. T. Wu, T. Supnithi, T. Kojiri, H. Ogata, S. C. Kong & A. Kashihara (Eds.), Proceedings of the 22nd International Conference on Computers in Education, 732-737. Japan: Asia-Pacific Society for Computers in Education. **(Chapter 4)**

6. Kothiyal, A., Majumdar, R., Pande, P., Agarwal, H., Ranka, A., & Chandrasekharan, S. (2014). How Does Representational Competence Develop? Explorations Using a Fully Controllable Interface and Eye-tracking, In C.-C. Liu, Y. T. Wu, T. Supnithi, T. Kojiri, H. Ogata, S. C. Kong & A. Kashihara (Eds.), Proceedings of the 22nd International Conference on Computers in Education, 738-743. Japan: Asia-Pacific Society for Computers in Education. **(Chapter 5)**

7. Pande, P., & Chandrasekharan, S. (2014). Eye-tracking in STEM education research: limitations, experiences and possible extensions, In Kinshuk & Murthy, S. (Eds.), Proceedings of The 6th IEEE International Conference on Technology for Education, 116-119. Kerala: IEEE. **(Chapters 4 & 5)**

8. Majumdar, R., Kothiyal, A., Pande, P., Agarwal, H., Ranka, A., Murthy, S., & Chandrasekharan, S. (2014). The enactive equation: exploring how multiple external representations are integrated, using a fully controllable interface and eye-tracking, In Kinshuk & Murthy, S. (Eds.), Proceedings of The 6th IEEE International Conference on Technology for Education, 233-240. Kerala: IEEE. **(Chapter 5)**

Contents

List of tables and charts

List of figures

List of boxes

List of abbreviations

ER	External Representation
MER	Multiple External Representations
IR	Internal Representation
RC	Representational Competence
STEM	Science, Technology, Engineering and Mathematics
TUF model	Transform, Unfreeze and Freeze model
DBR	Design-Based Research
DC	Distributed Cognition
EC	Embodied Cognition
RQ	Research Question
AOI	Area of Interest
CRM	Concepts, Reasoning and Modes of representations
RCA	Representational Construction Affordances

Abstract

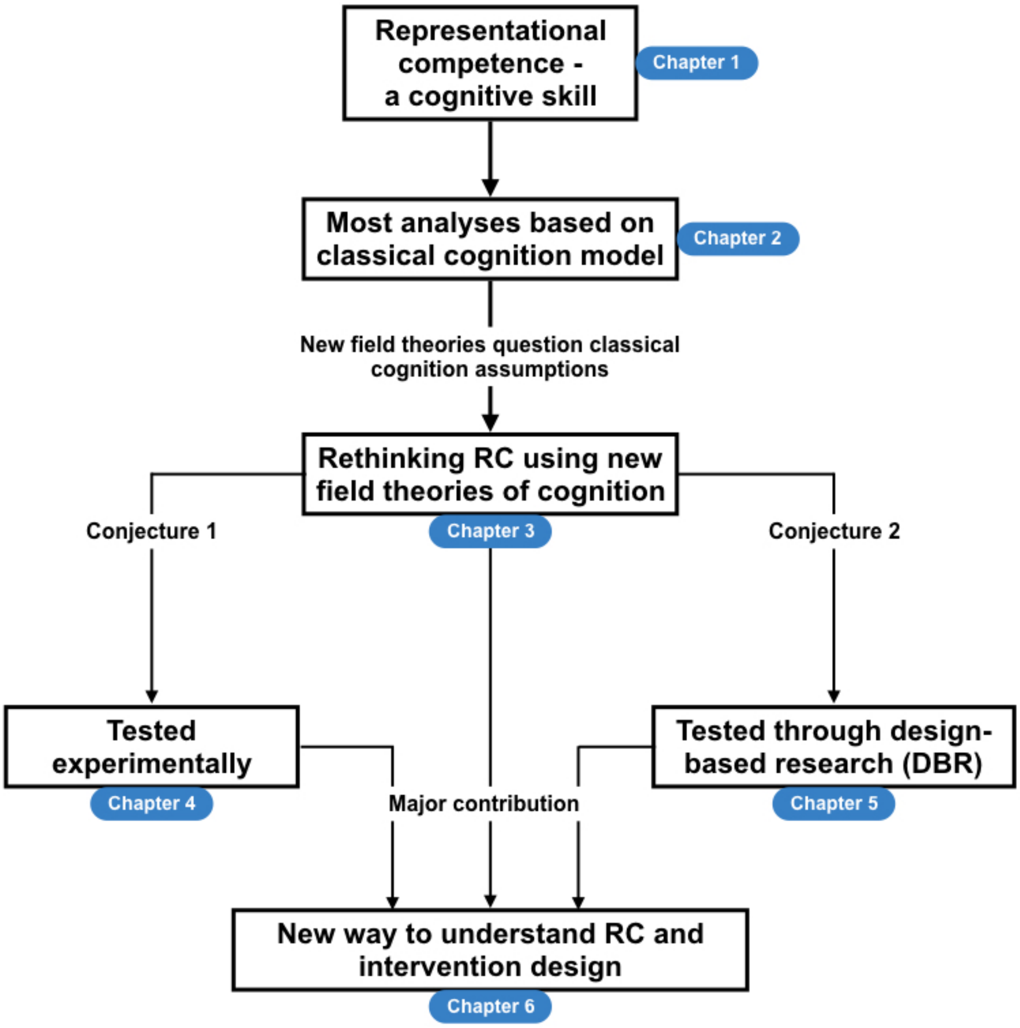
External representations (ERs), such as diagrams, equations, graphs, etc. are central to the practice and learning of science, mathematics, and engineering, as the phenomena and entities studied in these domains are often not available for direct perception and action. The ability to generate and use ERs in a domain in an integrated fashion, as well as perform transformations on the ERs, is termed representational competence (RC). Many learning difficulties are attributed to difficulties in achieving RC, particularly integration of ERs. RC thus presents a fundamental cognitive difficulty that cuts across different disciplinary domains, making it critical to develop teaching-learning strategies that help learners develop RC.

Most accounts of RC are grounded in the classical information processing model of cognition. In this model, a learner experiences high cognitive load during ER integration, as she tries to 'extract' information from ERs, internalize this information in the mind, and translate or process it to establish connections between the ERs. This characterization reduces the content of ERs to information, and treats ERs as 'vehicles' of information. This approach therefore does not seek to provide detailed accounts of the role played by ERs in cognition, and does not examine the cognitive mechanisms supporting integration of different ERs. Models based on this framework thus focus on processing cognitive load, and do not provide specific instructional design principles for effective development of RC.

Recent theories of cognition have moved away from this type of information processing models, to develop 'field' theories such as distributed and embodied cognition. These accounts suggest that ERs, and a learner's interaction with them, play a *constitutive* role in her learning of concepts. I extend this approach in this dissertation, to develop a theoretical model of the cognitive mechanisms underlying ER integration. This model focuses on how the cognitive system interacts with external representations, and the way integration abilities develop through this interaction. This mechanism model predicts that (i) the development of the ER integration ability would result in a reorganization of the sensorimotor system, and (ii) sensorimotor interaction would support ER integration and its development. To test these predictions, I developed two empirical studies, one based

on ER categorization tasks and eye tracking, and the other based on the design, development, and testing of an enactive new media intervention. The results from these studies broadly support the theoretical model. Based on these results, I outline some of the broader implications of the model and possible learning interventions.

Graphical abstract



Chapter 1: Introduction

Science often deals with entities and phenomena that cannot be directly observed and/or perceived, because they are too small (atoms, DNA, cells etc.), too big (galaxies, stars, tectonic plates etc.) happen in time-scales that are difficult to perceive (milliseconds, centuries, light years), and complex (feedback loops exist between levels and time scales). Understanding and analyzing these complex and imperceptible entities and phenomena require imagining them in detail, and developing indirect measures and novel representations (symbolic elements that stand in for the actual entities/phenomena) that help in this imagination. These representations are arrived at through practice and consensus in the scientific community. Even the perceivable entities and phenomena are not dealt with directly, as they need to be represented for various purposes, such as measuring, recording, observing, simultaneously dealing with multiple variables/factors/components, data handling, etc. External representations (ERs) are thus embedded in science practice, and they are critical for developing models, drawing inferences, making predictions, supporting claims and developing consensus. Ideas and content are distributed across ERs, and the learning and practicing of science are impossible without gaining expertise in interacting with ERs, thinking and imagining with them, and learning to generate them (Johnstone, 1991; Lesh, Post, & Behr, 1987; Tsui & Treagust, 2013). Imagined mental models, and ERs of these models, are developed over several iterations and revisions within science practice, where the internal and the external interact and help change each other (Nersessian, 2010). The final external representations and related internal models, which students are expected to learn in an integrated fashion, are often dense and opaque end-products, hiding the historical contexts and the problems through which they evolved.

The ability to generate and use ERs in an integrated fashion, as well as perform transformations on the ERs, is termed representational competence (abbreviated as RC, Kozma & Russell, 1997 & 2005). RC presents a fundamental cognitive difficulty that cuts across different domains such as science, mathematics and engineering (Pande & Chandrasekharan, 2017), making it critical to develop teaching-learning strategies that help learners in developing RC.

RC comprises of the following non-exclusive interrelated set of skills:

(a) Integrating internal and external representations, as well as different external representations

(b) generating ERs appropriate to the situation or problem

(c) communication using ERs

(d) reasoning using ERs

(e) choosing appropriate ERs based on the need of the situation/problem

(f) understanding and describing the different roles of an external representation in relation to other ERs

(g) critiquing ERs in terms of their strengths and shortcomings, etc. (Kozma & Russell, 1997; Kozma & Russell, 2005; Madden et al., 2011).

Figure 1.1 below situates this dissertation in relation to these facets of RC.

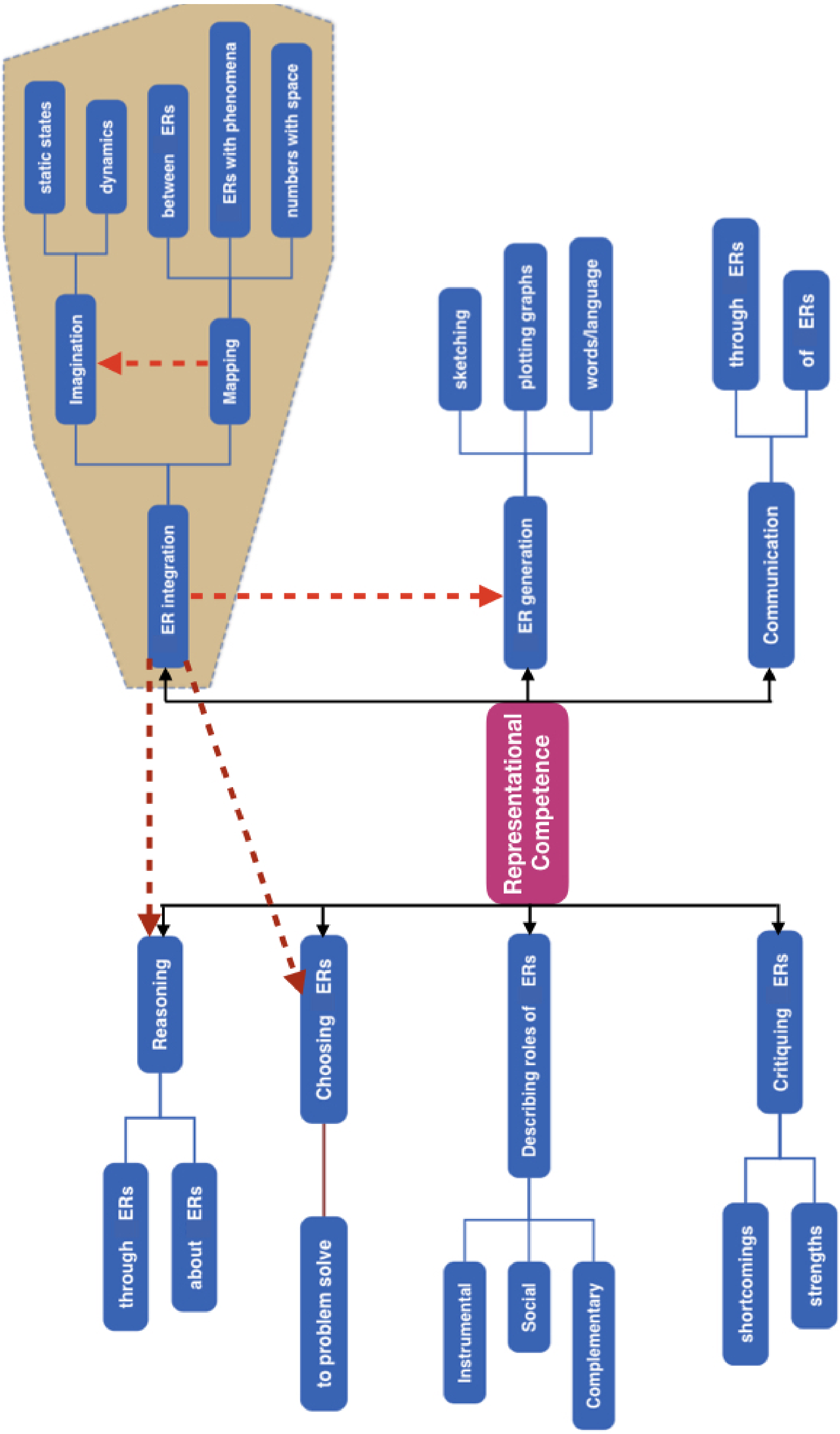


Figure 1.1 The abilities that comprise RC. The scope of this thesis is limited to ER integration, presented in the shaded area. Due to the interconnections between the abilities/concepts (not

indicated in the diagram to avoid complexity), the work developed in this dissertation extends to or includes concepts such as reasoning around ERs, choice of ERs or the relationships between them, and ER generation. These implicit relations are highlighted with dotted arrows.

In this dissertation, I focus on the ER integration sub-skill of RC (see box 1 for definition).

Box 1: Important concepts used in the document

ER integration is defined as the process of integrating ERs in a domain with the learner's internal (mental) schema, as she uses, understands and transforms between ERs in a domain (Kozma & Russell, 1997; Pande & Chandrasekharan, 2017; Pape & Tchoshanov, 2001).

The **sensorimotor system** comprises of the sensory (related to sensing the different states of the external environment as well as the body and internal organs), motor (related to action or production and regulation of body movement) involved in bodily movements. This system facilitates interactions of the learner's cognitive system with different components of the external environments, particularly ERs.

The sensorimotor system-based interaction (henceforth simply '**sensorimotor interaction**') is made possible due to "the capability of the central nervous system to integrate different sources of stimuli, and in parallel, to transform such inputs in motor actions", called '**sensorimotor integration**' (Machado et al., 2010). For instance, to perform an action as simple as picking up an object (say a box), the sensorimotor system integrates: current state or posture of the body, spatial relations between the body and the object, previous experiences about the object, regulation of body movement, etc. through information coming from the skin, muscles and joints, vestibular system (a system in the ear that tracks body's balance), the motor plan of the action, anticipatory adjustments of posture in relation to the motor plan and the position of the object, etc. (Machado etl al., 2010).

There is consensus in the education literature that many learning difficulties students face in these disciplines are attributable to problems in achieving RC, particularly ER integration (Chi, Feltovich & Glaser, 1981; Johnstone, 1991 & 2000; Johri, Roth & Olds, 2013, Larkin et al., 1980). Expert-novice studies of RC show significant differences between the two groups, in terms of the ability to understand individual representations, integrate ERs, and use and generate ERs for conceptual understanding, discovery and problem solving (Chi, Feltovich & Glaser, 1981; Larkin et al., 1980; Kohl & Finkelstein, 2008; Kozma & Russell, 1997). While students understand, and are able to use as well as generate, representations independently (diSessa, Hammer, Sherin & Kolpakowski, 1991;

diSessa & Sherin, 2001), they have great difficulty integrating ERs of a phenomenon (Knuth, 2000; Kozma & Russell, 1997; Wu & Shah, 2004).

1.1 The information processing model of RC

Performing tasks such as a simultaneous consideration of ERs, seeing the relationships between those ERs, interpreting them, reasoning about them in relation to the represented phenomena, etc. generate tremendous cognitive load on students' working memory (Johnstone 1982 & 1991), and one strand of literature considers this load to be at the root of the ER integration problem (Hinton & Nakhleh, 1999; Kohl & Finkelstein, 2008; Larkin et al., 1980).

Such cognitive load-based accounts of ER integration difficulties, which are currently dominant in the education literature, are rooted in the classical information processing model of cognition. This model is based on an analogy between computers and the human brain, and assumes that the learner’s mind, on encounter with an external representation (input), engages in information extraction (figure 1.2).

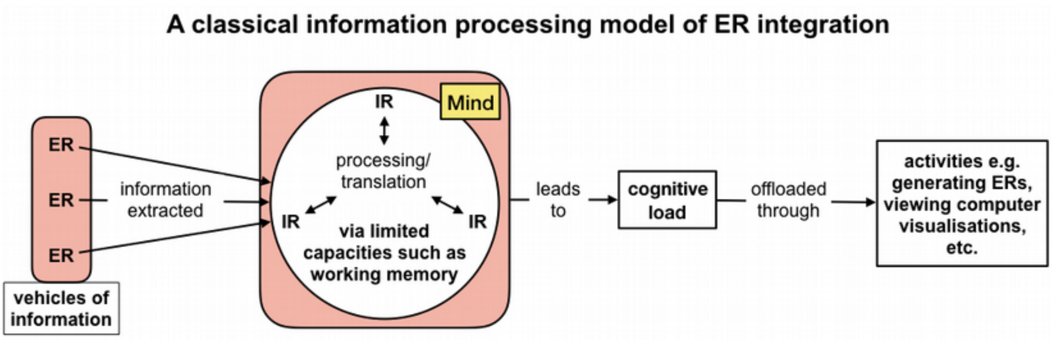


Figure 1.2 A classical information processing model of ER integration. In this model, meaning is ‘extracted’ through amodal, syntactic, processing of the information contained in ERs.

Correspondences between ERs are established through a translation process based on this extracted information. Such translation processes are considered to establish correspondences between ERs and the phenomena they represent, and also between the learner’s mental models (or internal representations, IRs) and the external representation. In this view, ERs act as 'vehicles', tools or transmission media, that carry the information, which is considered the key element the cognitive system works with. This translation process generates significant cognitive load, and learning difficulties are considered to arise because of this processing load, and the limitations of working memory in handling this load. Following from this view, the sole purpose of generating and using ERs during problem-solving is ‘offloading’ cognitive load. In this model, the extraction and translation of information are mediated mostly through mental capacities such as imagery and modality-independent (amodal) symbolic processing, as well as working memory (e.g. Gooding, 2006; Johnstone, 1982; Lesh et al., 1987; Tsui & Treagust, 2013; etc.). The limited nature of these processing resources are considered to be the root of problems in achieving ER integration. A central problem with this computer-inspired model is that it advocates that the mind (passively) receives information inputs from the external world, which it processes ‘inside’ (the skull) in coordination with capacities such as the working and long term memory, and produces an output (usually in the form of an) action.

These assumptions, particularly limited working memory capacity as the central processing bottleneck, have influenced many intervention designs. For instance, visualization software, interactive computer simulations, and virtual laboratories, are all designed to address working memory limitations. Ironically, the software interventions do not seek to augment the student's working memory and processing abilities, but only help offload some of the memory and processing load to the computer screen. Possibly because of this, such interventions have not been very successful in promoting RC (De Jong & van Joolingen, 1998; Rutten, van Joolingen & van der Veen, 2012). Further, by focusing on the "processor capacity" as well as the inaccessible nature of information extraction and translation processes, these models and interventions make the ER integration process, and the cognitive mechanisms underlying it, appear mysterious. Further, these models do not focus on the cognitive as well as practice elements that could lead to ER integration and its development.

1.2 The emerging model

The central assumptions of the information processing approach to cognition – that all cognitive processing is (or is best) done just by neural processes (inside the skull), and that external representations only help 'offload' information – have been seriously questioned by recent empirical and theoretical work in cognitive science, particularly by 'field' theories such as distributed cognition (DC) and embodied cognition (EC).

In the DC view, for instance, Kirsh (2010) outlines seven ways in which the external aspect of ERs, and our interactions with external representations, contribute to cognition:

(1) ERs change the cost structure of the inferential landscape.

(2) ERs provide a structure that can serve as a shareable object of thought.

(3) ERs create persistent referents.

(4) ERs facilitate re-representation.

(5) ERs are often a more natural representation of structure than mental representations.

(6) ERs facilitate the computation of more explicit encoding of information.

(7) ERs enable the construction of arbitrarily complex structure; and they lower the cost of controlling thought – they help coordinate thought.

"Jointly, these functions allow people to think more powerfully with ERs than without. They allow us to think the previously unthinkable" Kirsh (2010).

This approach mostly focuses on the distributed nature of cognitive processing and its advantages. However, understanding representational competence requires moving beyond just the recognition of the cognitive power of external representations: it needs a model of how new kinds of imagination is made possible by the coupling of ERs with the cognitive system (Chandrasekharan & Nersessian, 2015). This coupling is closely related to integration of ERs. Since different ERs capture different aspects of a phenomenon (Ainsworth, 1999 & 2008), they need to be integrated by the learner to understand the nature of that phenomenon. Any account of how ERs are used in learning, thus, needs to account for this integration process, particularly the role played by interactions with ERs and the cognitive processes involved in this integration.

In a related direction, recent work in embodied cognition by Landy, Allen, and Zednik (2014) articulates a distinction between syntactic/semantic approaches and *constitutive* approaches towards symbolic reasoning. In the first approach, symbols in ERs are considered to be internalized by the cognitive system, and then processed fully inside, i.e. just using neural processes (essentially the classical information processing model). In the constitutive account, the external symbols are part of cognition. Also, the external operations on them, as well as the sensorimotor system-based interaction processes (such as perception, physical manipulation, etc.) involved in these operations, are part of the cognition process. This constitutive view is supported by the fact that most scientific phenomena deal with entities not available to perception and action, and therefore the understanding of these entities is tightly intertwined with the external structures that stand in for these entities. The ERs thus play a twofold constitutive role in congaing these phenomena (stand-ins for imperceptible entities, structures that help constitute concepts), as understanding these imperceptible entities would be impossible without them. And since ERs are external structures, operations done on them are a critical component of understanding the entities and processes they stand in for. Figure 1.3 presents a graphic illustration of these ideas from the DC and EC theories.

The new 'field theories' of cognition emphasize interaction with external structures as the central process driving meaning and understanding. Extending this view to RC, interaction with external representations, particularly based on the sensorimotor system, would be key to ER integration. ER integration and the generation of concepts are also built on this sensorimotor integration, as interaction with ERs are based on the sensorimotor system, and such interactions exploit cognitive/ brain mechanisms similar to those involved in sensorimotor integration (Pande & Chandrasekharan, 2017).

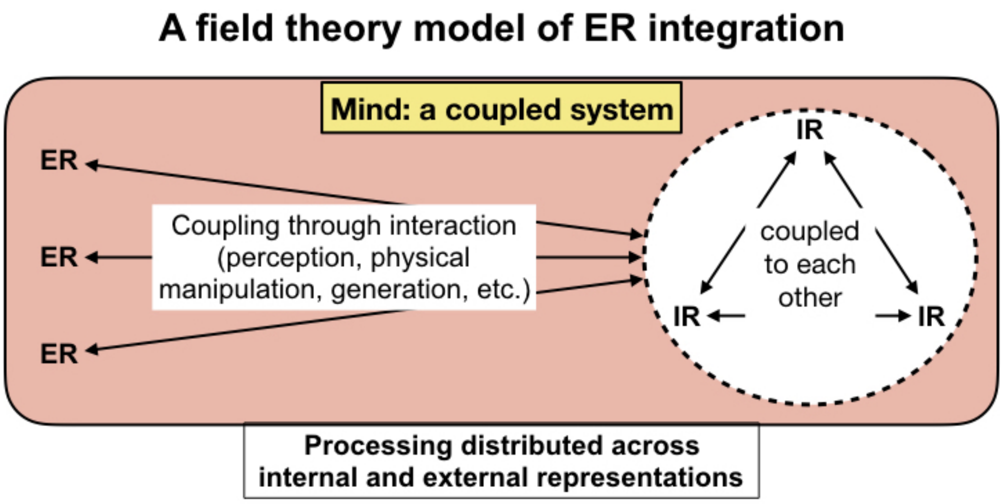


Figure 1.3 A general field theory model of cognition and ER integration. In this model, meaning is constituted through (bodily or sensorimotor system-based) interaction with the ERs. The mind is considered to be 'coupled' with ERs, and internal representations of ERs are considered to encode the sensorimotor aspects of the interactions.

A good example to illustrate the constitutivity position is provided by Landy and Goldstone (2007) who demonstrated how visual cues, such as spacing the elements in an arithmetic equation differently, or adding lines and circles around equations, influences problem solvers' symbolic reasoning abilities, such as following (or not following) the operator-precedence rule in arithmetic problems. This influence is a result of perceptual grouping, cued by the structural elements added to the equation, suggesting that external structures, and the perceptual as well as sensorimotor mechanisms involved in a problem solver's experiences with those external structures, constitute the processing and overall understanding (internal representations) of the symbols (Kirshner & Awtry, 2004; Landy et al., 2014).

Further evidence in support of the position comes from neurological studies investigating the use of mental abacus. Expert abacus users develop the ability to use an imagined internal abacus, on which they do visual and motor operations while solving complex arithmetic tasks. In contrast, students who are not familiar with the abacus imagine the standard written arithmetic algorithms (learned through paper and pencil operations) while solving the same arithmetic tasks. The interesting finding, however, is that these two operations in imagination (mental abacus, paper/pencil algorithms), which are *constituted* through interactions with different external structures, 'run' in different areas of the brain. f-MRI studies reveal that, in the case of mental abacus, predominantly visuo-motor areas of the brain are activated, whereas imagination of the paper/pencil-based algorithms mostly activates frontal areas of the brain (Chen et al., 2006; Hanakawa et al., 2003).

How can one explain this fMRI result using the classical cognition model? According to the classical information processing model, information in both the abacus as well as paper/pencil-based problem solving cases would be extracted in a symbolic form, and processed inside the brain amodally. As there is no visual or motor activity involved in processing the amodal symbolic operations, there should be no activation in the visuo-motor areas of the brain in either of the cases. In contrast, the field theory model, along with the theoretical position I propose here regarding the relation between sensorimotor integration and ER integration, suggest that as the mental abacus operations are learned with, and thus rely heavily on, visuo-motor operations, imagination based on stored abacus-based operations would activate visuo-motor areas of the brain significantly. Similarly, the imagined written algorithm operations are based on generating and manipulating text-based images in working memory, so these operations would activate the frontal areas more. This view accounts well for the fMRI data, and suggests that internal representations are generated through interaction, and they thus encode these interactions. These actions are activated during imagination based on the stored internal representations, such as the mental abacus. This analysis suggests that learning based on different external representations lead to different kinds of stored processes and imagined operations in the brain (figure 1.4).

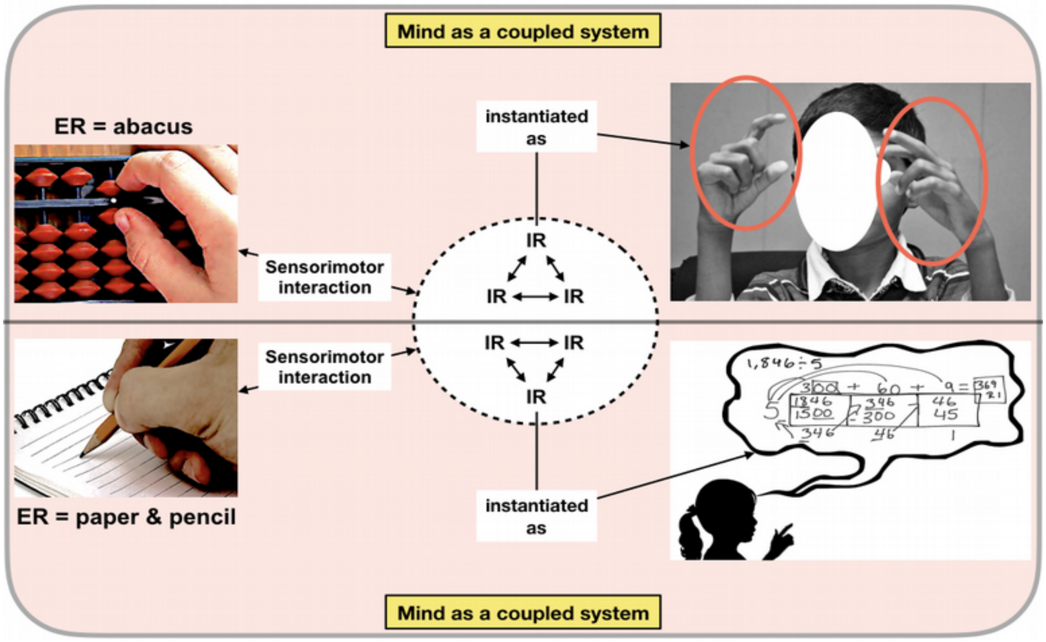


Figure 1.4 Development of different internal representations, based on sensorimotor interaction with different ERs. Expert abacus users develop an internal abacus learned through sensorimotor interaction with the physical abacus. This mental abacus is used to solve arithmetic problems mentally (in imagination), by 'running' the same sensorimotor interactions internally. Some of this covert sensorimotor processes 'leaks' into overt action, leading to gestures similar to the actions on the abacus (top panel). Problem-solvers not familiar with the abacus imagine written arithmetic algorithms, learned through paper/pencil-based interactions with the symbols and operations.

Extending this view, different operations in imagination would be made possible by different ERs, and integrated ERs. The integration process would also be driven by sensorimotor operations, as in the case of the physical abacus.

The constitutive view does not deny symbols or symbolic relations. In the above example, bead positions in the abacus are symbols that stand in for numbers. However, focusing on this symbolic nature directs the analysis away from the way the mental abacus (a thinking process) is generated from the physical abacus (a doing process), as the symbolic view would consider both as based on symbols. The constitutive view helps focus on the *processes* involved in this doing-to-thinking shift, as well as the cognitive and neural *mechanisms* involved (Rahaman et al., 2017), which leads to a richer understanding, and consequently, more detailed design directions. A symbol-based analysis would only provide a surface-level view of this change in cognition, and thus design directions based on cognitive load.

Importantly, the constitutive view offers the possibility of providing critical direction to the design of new computational media for learning, where embodied controllers such as multi-touch devices, Leap Motion, Kinect, Real Sense and Virtual Reality are used to develop new learning experiences through embodiment (Lindgren & Johnson-Glenberg, 2013), i.e. constitute new ways of integrating ERs (Abrahamson & Sánchez-García, 2016; Borar et al., 2017; Dickes et al., 2016; Karnam et al., 2016; Ottmar et al., 2015, Sinclair & De Freitas, 2014). The interconnections between ERs are considered to be created by actions, and not just by symbolic relations. ER integration is considered driven by the doing aspect, and not by the relations between symbols, even though the relations between symbols contribute to, or even make possible, the doing.

The three aspects of external representations discussed thus far (viz. power of external representations, constitution of concepts, and integration of external representations), are explored significantly in cognitive science and studies of scientific practice (Chandrasekharan & Nersessian, 2015; 2018), but are not addressed by current work in ER integration and RC, except in some isolated cases.

1.3 Sensorimotor markers of expertise

It is well known that expertise is marked by specific changes in the nature of cognition and perception, particularly related to problem-solving (e.g. response times, visual attention, etc.; NRC, 2000). These changes have been documented across multiple domains (e.g. chess, science, mathematics, social science, medicine, etc.; NRC, 2000). de Groot (1978), for instance, was among the first to demonstrate how expert chess players could almost instantaneously *see* problems, as well as possible moves to address those problems, when presented with different configurations of pieces on a chess board. Not only were experts quick to respond, they also suggested 'high quality' moves, in contrast to less experienced players. de Groot concluded that training in chess gradually reduced the time and efforts required to abstract patterns, and that the patterns were readily and directly *perceived* by expert chess players, thus marking the replacement of abstraction by perception (de Groot, 1978).

In education, a number of studies have explored how experts differ from novices in the way they pick up information during a problem situation, based on their (sensorimotor) experiences with the symbolic structures involved in that problem (Brathwaite et al., 2016; De Wolf et al., 2017; Kellman et al., 2010; Landy & Goldstone, 2007; Rivera & Garrigan, 2016). Closely related is a considerable amount of research on perceptual learning – a phenomenon characterized by changes in the process of information extraction, and changes in the perceptual-cognitive system (as well as mental models) of a learner as a result of visuo-spatial routines (perceptual manipulations theory; Landy et al., 2014), training and experience (e. g. Goldstone, 1998; Kellman & Garrigan, 2009). Kellman and colleagues (2010) for instance, show how transforming the structure of an algebraic equation affects the difficulty level as well as response times to solve that equation. They argue that people with different experiences with the different equation forms find some forms of equation more relevant than others, and that this relevance is established almost instantly after perceiving the problem, as indicated by response times.

There also exists a relatively smaller, yet significant chunk of studies in chemistry, biology and physics education which investigates differences in the perceptual processes between experts and novices (Pande & Chandrasekharan, 2017). Several studies specific to ER integration, which establish strong links between perceptual and cognitive processes (collectively referred to as perceptual-cognitive processes), as well as studies of mental models of abstract entities and phenomena (Landy, Allen & Zednik, 2014; Lowe & Schnotz, 2014; Rau, 2015) imply that identifying the markers of integration may help in understanding the nature of mental models. Eye movements and fixations are popularly used in studies investigating this link, as eye-behavior during a task is mostly implicit, (i.e. driven by task demands and not completely in the control of the agent), and thus is anchored closely to the perceptual-cognitive processes related to the task (Henderson & Ferreira, 2013; Irwin, 2004). The results from these studies suggest that training, and restructuring of prior knowledge based on training, reorganizes experts' perceptual-cognitive schemas (Cook et al., 2006; Kohl & Finkelstein, 2008). Process-based approaches, such as those focusing on the way participants navigate ERs, show some evidence in this direction. Using eye-tracking, Stieff, Hegarty and Deslongchamps (2011) show that students attend more to familiar visuo-spatial ERs during problem solving than the less known symbolic ERs. Cook et al. (2008) captured the number of transitions students made between molecular-to-molecular, macro-to-molecular, molecular-to-macro and macro-to-macro representations, in order to understand how students' prior knowledge interacted with the way they interpreted macro and molecular graphics of diffusion phenomena. On average, students with low prior knowledge made more transitions than students with high prior knowledge. Similar patterns of transitions between ERs are reported by Kohl and Finkelstein (2008) in a study examining the generation of ERs by three groups of participants – experts, weak novices and strong novices – while solving different sets of problems on electrostatics. The authors of these studies argue that, as the students with low-prior knowledge are less aware of the 'subtleties of representations and the conventions for interpreting them', they made frequent transitions between the ERs in order to better perceive and map features from one representation to the other (Kohl & Finkelstein, 2008).

However, not only are the results of these studies insufficient to make claims about the sensorimotor markers of the perceptual-cognitive processes involved in ER integration, but the theoretical considerations underlying these studies also are different from those investigating these markers, as many of these existing instances have been explained using the outdated 'top-down', 'bottom-up' information processing accounts (Gegenfurtner, Lehtinen, & Säljö, 2011; Goldstone, 1998; Kundel, Nodine, Conant, & Weinstein, 2007; Lowe, 2015). Moreover, though it is well known within scientific communities that experts 'act' on scientific concepts, procedures and representations in a significantly different manner than novices, very few intersubjective evidences for such constitutivity exist, apart from the expert-novice differences in patterns of attention over visual stimuli.

Such markers of sensorimotor changes based on science training, according to the view developed in this dissertation, are markers of changes in cognitive mechanisms associated with ER integration. The work outlined here thus brings together perspectives on ER integration, perceptual learning and constitutivity, and proposes that concepts are constituted through sensorimotor interaction, and this constitutivity process leads to perceptual learning, along with other sensorimotor changes.

1.4 Previous related work at HBCSE, TIFR.

Previous studies at HBCSE have examined how students integrate the dynamic structure-function relationships in scientific phenomena and entities through visualization, as well as the use of gestures and analogies. Subramaniam and Padalkar (2009), for instance, explored how and which ERs are used by adults to reason about astronomical phenomena such as an eclipse, involving the sun, earth and moon. They found that adults heavily rely on gestures to visualize and explain phenomena dynamics through static diagrams. Importantly, we benefit not only from the constant dynamic feedback available through gestures, but also from automatic associations gestures build between ourselves and the phenomena. These results indicate the close link between embodiment and RC. Mathai and Ramadas (2009) report similar findings in the context of structure-function relationships in middle-school biology. Padalkar and Ramadas (2009) designed and tested specific manipulative actions and pedagogic gestures to help middle-school students develop an integrated understanding of the dynamics of astronomical phenomena and their static models and diagrams. Building on this work, Srivastava and Ramadas (2013) demonstrated how use of gestures, analogies and perspective taking can together help integration of different external representations (such as 2-dimensional models and diagrams) by inducing mental simulation of 3-dimensional DNA structures.

The work presented in this dissertation integrates previous work across multiple studies done at the centre, by 1) focusing on ER integration as a general learning difficulty cutting across disciplines, and 2) developing a theoretical model of the cognitive mechanisms underlying ER integration based on new field theories in cognition. This work thus tightly connects the ER integration problem, and studies exploring ER integration, with recent cognitive science research.

1.5. Overview of the thesis

Building on the emerging models of cognition, and the three aspects of ERs (cognitive augmentation, constitutivity, integration) as well as the new understanding of the markers of expertise based thereupon, this dissertation develops:

(i) a new theoretical model of the cognitive mechanisms underlying ER integration and its development,

(ii) empirical studies to test this new model, and

(iii) a design that incorporates the model.

I begin with reviewing relevant literature (chapter 2), particularly the theoretical frameworks of ER integration and RC, and the empirical studies that investigate the nature of RC and its development in science, mathematics and engineering. This review brings together all the work done in RC, in many disparate areas, and identifies commonalities and differences in the research across several themes. The review finds that most research in this area, including intervention development, is either explicitly or implicitly based on the classical information processing model of cognition.

In chapter 3, I develop a distributed and embodied cognition account of ER integration, in contrast to the information processing accounts, for the following reasons:

1. One, current models of cognition reject the classical information processing approach; mental processes are now understood as distributed and embodied. Models of ER integration are models of cognition, and thus need to incorporate this theoretical shift, particularly because ERs are external (thus distributed), and working with ERs require sensorimotor interaction (embodied interaction).

2. Second, there is a parallel shift in the design of new computational media, where embodied controllers such as Leap Motion, Kinect, Real Sense and Virtual Reality are used to develop new learning experiences (Abrahamson & Sánchez-García, 2016; Dickes, Sengupta, Farris, & Basu, 2016), particularly to integrate ERs. These controllers have also been incorporated successfully into collaborative learning environments (Danish, Enyedy, Saleh, Lee & Andrade, 2015; Enyedy, Danish, Delacruz & Kumar, 2012; Enyedy, Danish & DeLiema, 2015). This design approach requires understanding the role of embodiment in ER integration and RC development.

3. Finally, the practice of science itself is now understood as distributed and embodied (Chandrasekharan, 2013; Chandrasekharan & Nersessian, 2015; Nersessian, 2010), and any future model of ER integration and RC development need to reflect this shift in our understanding of science practice.

The account developed in this chapter illustrates how ERs are understood by learners through an 'incorporation' process, where they become part of, and thus extend, the cognitive system, while also forming and extending the internal model of the scientific domain. This incorporation process is driven by sensorimotor actions/manipulations performed on the external representations, as well as through the exploration of many states of the external representations. Further, sensorimotor interactions with these ERs (overt as well as covert activation of the motor system) facilitate ‘capturing’ and ‘unfolding’ the different states of ERs, and these operations play a central role in ER integration.

Two interconnected conjectures, with empirical implications, emerge from this account:

- In this model, the development of the ER integration ability (expertise) would result in a reorganization of the cognitive system, particularly the sensorimotor system. This suggests that the process by which learners perceptually access ERs would change after significant training in a domain.
- Interaction, particularly based on the sensorimotor system, would support ER integration and its development.

To test these predictions, and thus also the theoretical model, I conceptualized two empirical projects.

The first project (chapter 4) sought to identify behavioral markers that could track sensorimotor changes as a learner interacted with scientific ERs, leading up to the development of constitutivity and ER integration. In this project, I first established the ER integration abilities of participants, who had different levels of education in chemistry. This was done by tracking how they related chemical phenomena and their dynamics, when presented with different static and dynamic ERs during a categorization task. I then looked for patterns in their eye gaze behavior, and correlated these patterns with participants’ ER integration abilities, to identify sensorimotor markers of ER integration. This project contributes to the existing work on the nature of expertise.

The second project (chapter 5) focused on the design, development and testing of a computer interface with fully manipulable ERs of a physical system. Besides being an intervention, the interface was also used as a ‘probe into the cognitive processes’, to explore how interactivity aids in ER integration.

Chapter 6 summa-rises the dissertation projects, and presents possible implications and contributions of this work, particularly in relation to enactive new-media designs supporting ER integration and RC development, as well as conceptual learning in science, mathematics and engineering.

The work reported here is among the first to:

(1) Weave together extensive and highly diverse theoretical as well as experimental work on ER integration from different disciplines.

(2) Objectively characterize the sensorimotor changes related to ER integration and RC facilitated by training in a domain.

(3) Design and test an enactive new-media[1] intervention based on DC and EC perspectives, exclusively targeting ER integration and RC development.

(4) Analyze in detail the relationship between interactivity, ER integration and learning.

(5) Conjecture thatusability and learnability design principles are not enough for the learning of complex representations and conceptual content based on computer and new media technologies.

Figure 1.5 below, outlines the dissertation and its logic, and presents the relationships between its chapters.

---

1 The term 'new media' broadly refers to different types of interactive (i.e. at least two-way) communication platforms (in contrast to mass media which are generally regarded as one-way platforms) based on computers (Manovich, 2001). Some mutually non-exclusive examples of new media are websites on the internet, virtual reality environments, multimedia environments, animations, computer games, human computer interfaces and the more advanced brain computer interfaces. In this dissertation, however, I am using the term new media to specifically refer to *modelling interfaces with virtual elements/representations* of target concepts which a user (or a learner) can *interact with or enact through multiple modalities* (auditory, tactile, visual, etc.) and/or *controllers* (e.g. mouse, gestures-based controllers such as LeapMotion, multi-touch devices, virtual reality gears, etc.)

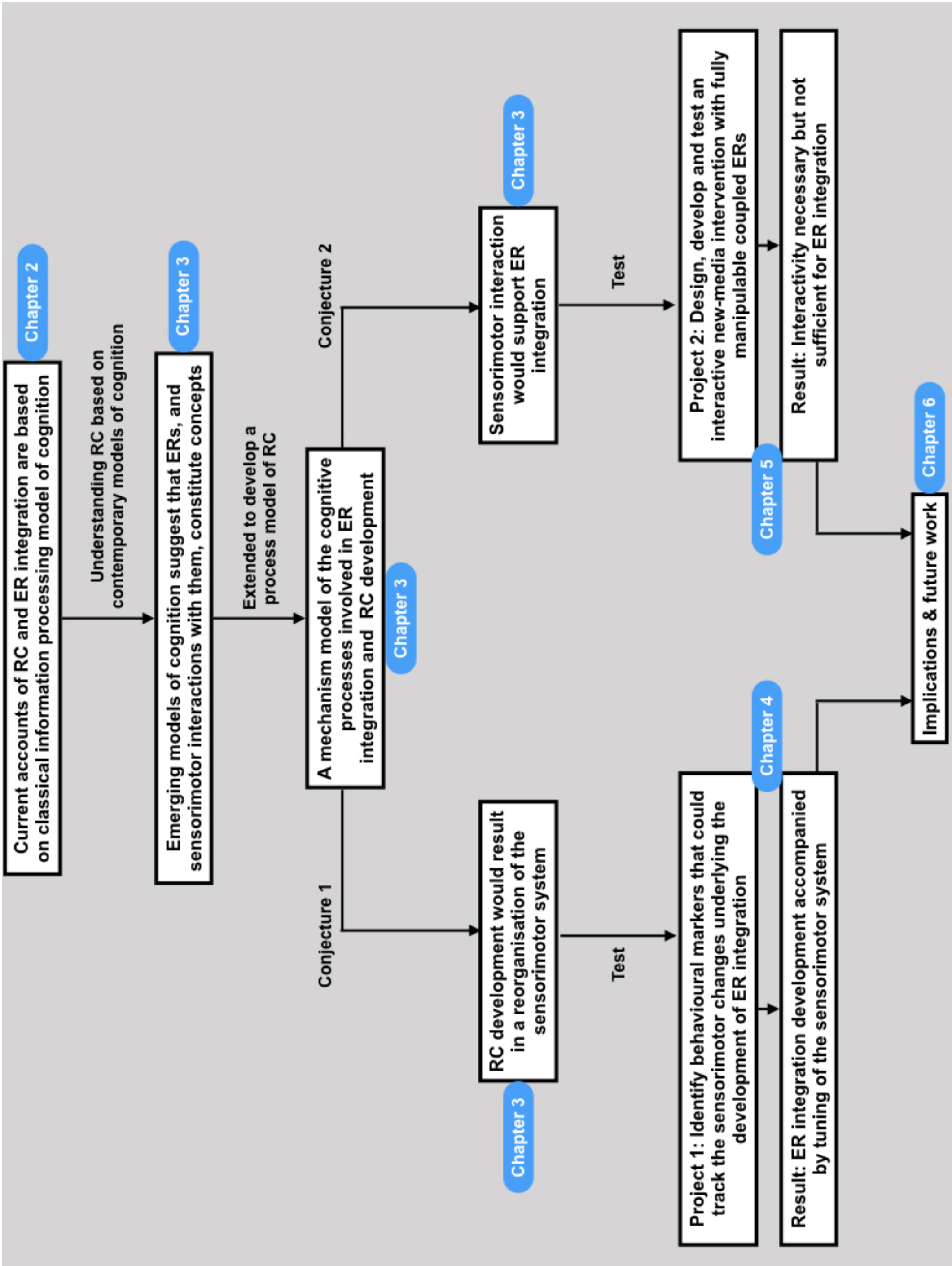


Figure 1.5 Outline of the thesis (graphical abstract).

Chapter 2: Bringing together research on ER integration in science, mathematics and engineering; identifying gaps

This chapter[i] presents a comprehensive review of existing work in RC, particularly on ER integration and its development, in science (chemistry, biology and physics), mathematics and engineering., This review brings together all the major (interrelated as well as disparate) theoretical and experimental work in RC, distributed across more than 170 papers, published in more than 35 different venues (in education research, cognitive science, learning sciences, educational technology, etc.) on topics related to chemistry, biology, physics, mathematics, and engineering.

In section 2.2, I examine nearly 30 different influential theoretical models of RC, ER integration, and their development. Section 2.3 more than 70 research papers reporting empirical investigations, in order to understand the different approaches to ER integration. A preliminary analysis revealed that a significant set of models and studies explicitly appeal to the classical information processing paradigm, while some other frameworks and studies implicitly assume classical information processing perspectives, but do not endorse this view explicitly. A third set of models and studies are neutral on the nature of ERs and RC. There is also a group of models and studies that subscribe to recent field theories of cognition, such as distributed and/or embodied cognition. This categorization, based on subscription to theoretical models and empirical studies and wider models of cognition, is captured in a chart presented towards the end of the chapter.

A different analysis was done to compare these theoretical models and empirical studies across the different disciplines mentioned above, in terms of: problems related to RC, nature of ERs, nature of learning difficulties, research methods employed, and the underlying theoretical assumptions. A discipline-based review captures this analysis for quick review, through boxes for each discipline, and a set of tables that brings together the diverse literature in a discipline-based categorization. Finally, I discus major findings from the review (Section 2.4).

2.1 Review methodology

Three different modes were employed to collect articles for this review: (a) keyword search on the ERIC database, (b) keyword search on Google Scholar, and (c) articles found relevant through cross-referencing. The following is a list of keywords used for methods (a) and (b): *scientific representations*, *learning – multiple representations*, *RC*, *RC in biology*, *RC in science* (then the word 'science' replaced with chemistry, biology, physics, mathematics, and engineering), *multiple external representations in science*. (the word 'science' then replaced with chemistry, biology, physics, mathematics, and engineering), *multiple representations in science* (the word 'science' then replaced with chemistry, biology, physics, mathematics, and engineering). Articles found using these keywords were filtered based on their date of publication, relevance and major discipline. Only articles published after 1990 were read and analyzed (with a few exceptional articles from before 1990 included due to their evident influence as well as frequent citations – e.g. Johnstone, 1982; Lesh et al., 1987, etc.). Further, only those studies/articles related to cognition research on multiple representations were included in the review. Articles exploring other dimensions, such as testing of a multi-representational user interface, use of multi-modal representations, and social significance of multiple representations were not included. Also, investigations of a single representational system (for instance only drawing) were not included, unless the studies had wider implication/ significance for the understanding of multiple representations.

2.2 Theoretical accounts of ERs and RC

A wide range of conceptual frameworks have been proposed to capture learning and cognition using ERs. I examine two kind of such models: (1) Models based on the relationship between the nature of a domain, ERs in that domain, and cognition, and (2) Developmental models. Models under the former section are further categorized into three interrelated but different subsets: one set focuses on the nature of knowledge in a domain, particularly pertaining to the space-time scales/levels, a second category focuses on reasoning through ERs, and a third set concerns mechanisms of ER cognition. Developmental models, on the other hand, focus more on the process of learning using ERs through stages of development, and are either based on (1) or are independent in a broad theoretical sense.

2.2.1 Models of ERs and cognition

2.2.1.1 Relationship between the nature of domain and ERs in that domain. Different scientific domains (biology, chemistry, physics, etc.) differ from each other in certain fundamental aspects, such as the nature and scale of problems, investigation methods, data, etc. These differences reflect in the nature of ERs used across these disciplines, and models of RC.

One of the first models of the relationships between the nature of a scientific domain, the ERs that constitute it, and a learner's interaction with those ERs, was proposed by Johnstone (1982). The model examines the visual-perceptual nature of representations used in science, particularly in chemistry. Chemistry ERs include the periodic table, chemical equations, graphs, molecular formulas, diagrams of experimental setups, diagrams depicting molecules, etc. Each of these conveys different information on chemical entities and phenomena. Johnstone's model, known as the model of 'three thinking levels', describes the way the discipline of chemistry is conceptually organized around these ERs.

According to Johnstone, knowledge in chemistry is distributed along the following three levels (figure 2.1):

a. *Descriptive/functional/macro level*, which deals with handling of materials, descriptions of phenomena and their properties, such as color, flammability, density, etc.

b. *Representational/symbolic level*, which deals with representations of chemical substances and phenomena using symbols, formulas, equations and conventions.

c. *Molecular and explanatory/micro/submicro level*, which captures the structure of chemical substances and phenomena, mechanisms of reactions, and the molecular/atomic interactions and changes that underlie chemical phenomena.

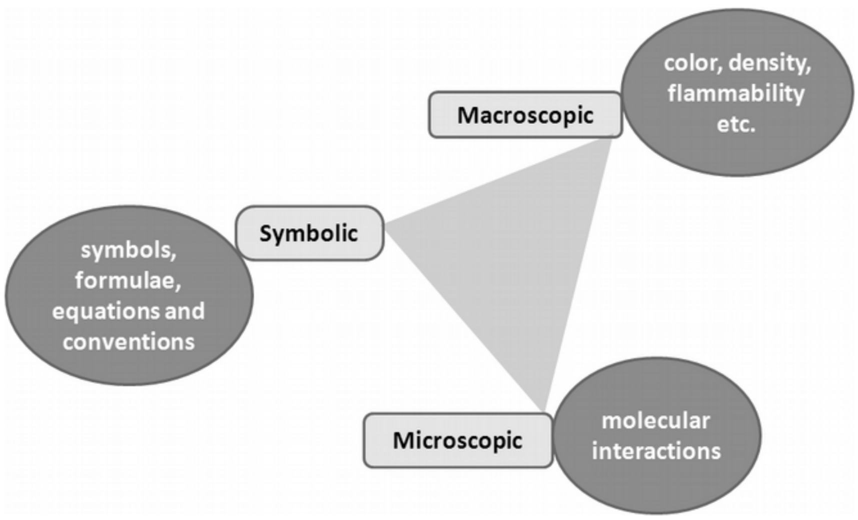


Figure 2.1 Johnstone's model of three thinking levels

The model considers ERs in chemistry as distributed across the three levels of thinking. Learning as well as doing chemistry requires, in this view, simultaneously processing the information gathered from ERs at all the three levels. This ability is characteristic of expertise in chemistry.

Supplementing this model of 'three thinking levels' with Baddeley's model of working memory, Johnstone (1991 & 2000) attributes students' difficulties in learning chemistry to the way this schema, the conceptual organization of chemistry, interacts with the limited capacity of the human working memory. According to Johnstone, the three-level schema puts significant load on a student's working memory as she attempts to understand a chemical reaction in terms of its equation and/or a graph (symbolic level) as well as the molecular mechanism (molecular level) of the reaction. As a result of the load, and the limited working memory capacity, students often ends up ignoring important features of the phenomenon, concentrating only on parts of it.

Several other models in chemistry education research attempt to conceptually organize chemical knowledge. Jensen (1998), for instance, replaces ‘macro’ level with ‘molar’ (referring to the perceivable stoichiometric ratios of chemical substances handled and used in carrying out reactions), retains the molecular level, and defines a third level called the electrical level, at which chemical phenomena are explained using subatomic particles (such as electrons) and their dynamics. Ben-Zvi, Eylon & Silberstein (1988) propose that single-particle modeling is sufficient to describe chemical properties of substances (but not physical properties). They then suggest that the sub-micro level be split into single-particle and multi-particle sub-micro levels of understanding chemical processes. A distinction between symbols used to denote chemical substances, and the numbers in stoichiometry, kinetics and mechanisms has also been proposed (Garforth, Johnstone & Lazonby, 1976; Nakhleh & Krajcik, 1994; Savoy, 1988).

Box 2 provides a quick summary of the discussion on the nature of chemistry ERs, as well as a review of ERs and RC in chemistry education literature.

Box 2: Nature of chemistry ERs

Nature of chemistry: Chemistry concerns the study of entities and phenomena which often cannot be observed directly at the level at which they occur (e. g. dissociation, neutralization, equilibrium dynamics, etc.). Teaching/learning and doing chemistry thus involves frequent use of different kinds of indirect representations, such as chemical equations, graphs, molecular formulas, diagrams of experimental setups, diagrams depicting molecules, etc., which help connect us indirectly with the actual chemical phenomena. Our thinking/understanding about chemical phenomena is essentially guided by these representations, and the nature of these representations influences teaching/learning/doing chemistry.

Central external representations identified in previous research: Periodic table, chemical equations, concentration/energy graphs, molecular diagrams, static and dynamic visualizations, graphical projection for organic molecules (e.g. Fischer's projection formulas), informal and formal gestures for representing spatial configurations and conformations of molecules and stereo-isomers, animations and simulations.

Representational Competence definitions identified in the previous research: Requires integrating these MERs to understand chemical phenomena. Involves inter-relating and transforming between MERs in a dynamic fashion, generation of appropriate MERs.

Existing experimental approaches to study RC: Analyzing students' problem solving process, microgenetic studies, ethnographic observations of professional chemists, expert-novice comparison, prior knowledge & RC correlation, computer interface testing, eye-tracking

Major results from previous research studies: (a) Various learning difficulties in chemistry can be attributed to difficulties in understanding, handling and using MERs, (b) prior knowledge plays a key role in developing RC, but it does not guarantee it, and (c) While inter-relating MERs, novices rely upon surface features, whereas experts use chemical principles.

Main theoretical frameworks proposed in the previous research: Johnstone's model of three thinking levels (macro level, sub-micro level and symbolic level), based on the classical information processing model of cognition, particularly Baddeley's working memory model. Some studies combine Johnstone's model with distributed and situated cognition approaches.

Main existing intervention approaches: (a) development of visualization software that seek to lower working memory load, by simultaneously presenting on screen chemical equations, concentration and/or energy graphs, molecular-level animations and laboratory experiment video (e.g. SMV Chem, visChem, 4M:Chem, EduChem HS, etc.), (b) use of physical models, (c) laboratory-centered pedagogy, (d) MER centred problem-based curricula, and (e) integration of MERs through manipulable interfaces that demonstrate/show the interrelation between different representations

In biology, Kaptejin (1990) proposed a framework of biology ERs in relation to (a) the levels of biological organization, as well as (b) observability of the ERs, (i. e. one's ability to see entities and phenomena). Keptejin's model, similar to the Johnstone's model, has three distinct levels of ERs, viz. macro (organismic), micro (cellular) and molecular (biochemical). According to this model, one's ability to visualize entities and phenomena at all the three levels limits the understanding of biological phenomena. Box 3 provides a quick access to the nature of biology ERs, as well as a review of ERs and RC in biology education literature.

Box 3: Nature of biology ERs

Nature of biology: Knowledge in biology is distributed across several levels of organizations that are used to characterize biological systems (Dreyfus & Jungwirth, 1990; Kaptejin, 1990; Tsui & Treagust, 2013). The MERs used in biology parallel this multi-level structure. RC problem is more complex in this domain than in chemistry. Research on RC in biology is relatively recent in comparison to chemistry, physics, and mathematics.

Central external representations identified in previous research: Biochemical pathways, structural (configuration & conformation) representations of biomolecules, DNA/polypeptide helices (right/left handed), replication, transcription and translation mechanisms, planar depictions of animal body parts (ventral, dorsal), anatomical planes/sections (lateral, cross, sagittal, temporal, transverse, etc.), developmental and evolution time-scales, phylogenetic trees, punnett squares, Hardy-Weinburg equation, computational models of predator-prey dynamics and similar complex systems. These representations cut across multiple organizational levels, from molecules to cells, tissues, organisms, communities and ecosystems.

Representational Competence definitions identified in the previous research: Developing an integrated understanding of MERs that are distributed across different levels of organization.

Existing experimental approaches to study RC: (a) Student misconceptions in relation to multiple levels of organization, (b) how students interrelate MERs distributed across different levels of organization, (c) how experts and novices differ in the way they move between MERs and interlink the represented information, (d) eye-tracking, (e) relation between prior knowledge and MER transformation/interlinking.

Major results from previous research studies: (a) Generating MERs and relating them to previously encountered concepts helps in problem solving, (b) previous knowledge is strongly related with frequency of making transitions between MERs, (c) the static nature of MERs, which embed dynamic phenomena, could critically limit interrelating MERs.

Main theoretical frameworks proposed in the previous research: CRM model, Kaptejin's model of three levels, and the cube model supplemented by Baddeley's working memory model.

Main existing intervention approaches: Visualization software, mainly seeking to connect different levels of organization, guided by the information processing model of cognition. There also exist manipulable Netlogo models of complex, self organizing biological systems. Pedagogical interventions include problem-based learning (especially in genetics, biochemistry and medicine) and inquiry-based learning.

Tsui and Treagust (2013) recently proposed a more comprehensive framework of the conceptual organization of biology, termed the cube model (figure 2.2), which implies a three-dimensional knowledge structure.

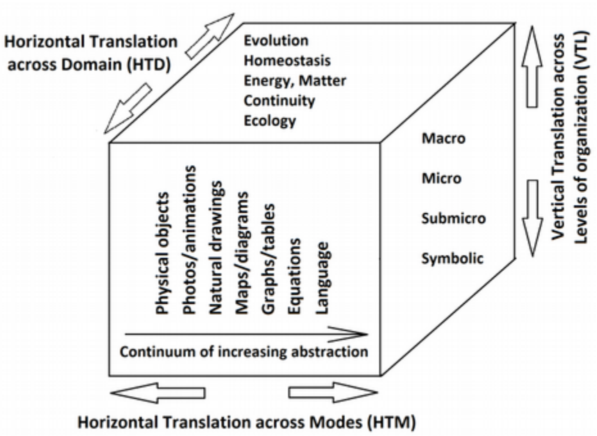


Figure 2.2 The cube model

In this model, knowledge in biology is spread across three different, but interdependent, dimensions, and learning in biology is marked by one's progress along these dimensions:

a) HTM: Horizontal Translation across Modes of representations, 'along a continuum of representations with increasing abstraction from real-life objects and actions to human language'

b) VTL: Vertical Translation across Levels of representations 'from the symbolic level (explanatory mechanisms), the submicro level (molecules), the micro level (organelles and cells), and the macro level (tissues, organs, systems, organisms, populations, and so on)'

c) HTD: Horizontal Translation across the Domain knowledge of biology, i. e. across evolution – homeostasis – energy – matter and organization – reproduction and genetics, etc.

Items from VTL can have one-to-many relationships with items from HTM, but not necessarily the other way around. For example, I often associate the term 'macroscopic' (VTL) with 'observable' (corresponding to worldly objects/actions, maybe even photographs/animations, essentially items along HTM). Also, graphs, tables and equations can all be counted under symbolic level or representations. However, equations would be strictly symbolic, and cannot be under the microscopic level. The VTLs are categories of representations similar to the levels of thinking, whereas HTMs are various modes through which information across those categories is obtained, presented and communicated. Although it is simple, comprehensive and unified, the cube model has limitations in capturing phenomena occurring over large temporal scales, such as evolution (Tsui & Treagust, 2013).

In mathematics, one of the most discussed and widely used conceptual frameworks is the Lesh Translation Model (figure 2.3), which is a network model developed to investigate student-generated representations and (information) translations between multiple representations. The model proposes that knowledge in mathematics is structured across five different, but interrelated and interconnected modes of representations, viz. (1) concrete/manipulable objects/situations (e. g. physical manipulatives such as tangram), (2) pictorial representations such as 2D/3D diagrams, (3) real-life contexts (e. g. acts of addition, sharing, etc.), (4) language (e. g. usage of mathematical terms such as 'addition' and 'subtraction'), and (5) written symbols (symbols denoting mathematical operations). Mathematical understanding is reflected in the ability to represent mathematical ideas in multiple ways across these five representational modes, and also in making connections and translations among them (Lesh et al., 1987).

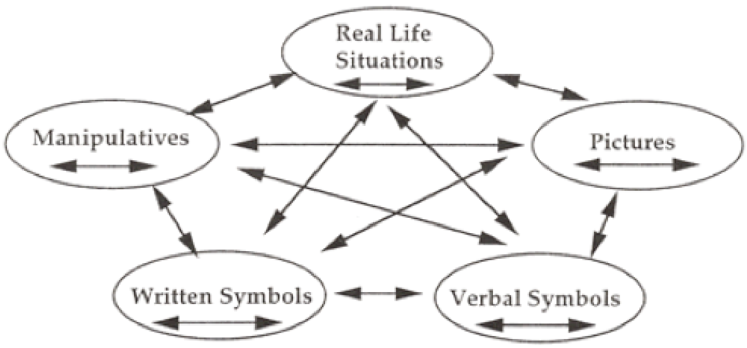


Figure 2.3 Lesh translation model

From a pedagogical perspective, the term 'translation' emphasizes interrelating information extracted from one representation with information from another. An expert would be fluent in translating between these proposed representational modes. The Lesh model has driven the conceptualization and development of a set of specific activities (called model eliciting activities) in mathematics and engineering pedagogy (Moore et al., 2013).

Due to the intertwined nature of physics (see box 4 for quick capture of commentary on the nature of physics and its ERs, and brief review of models), mathematics (box 5) and engineering (box 6), the Lesh translation model is equally applicable to ERs in physics and engineering. Manipulable models and prototypes of physical and engineering objects, free body diagrams, acts of navigation and motion, use of terms such as 'speed' and 'distance' in language, and symbols denoting physical properties of objects and phenomena such as 'force' and 'energy', are some examples of physics and engineering ERs belonging to the five modes of representations respectively.

As opposed to Lesh et al.'s network model, Roth and Tobin (1997) suggest a linear cascade model to explain the relationship between physics learning and practice, and the nature of physics ERs. This model emerged from an investigation aimed to understand how teachers use and translate between ERs while teaching in a physics class, and how this relates to students' difficulties in understanding the topic being taught – 'motion of a rolling ball on an inclined plane'. The authors propose a continuum of abstract and concrete representations, generalized from findings from the nature of ERs used in the classroom, to explain how the types of ERs in science, mathematics and engineering relate to student difficulties in interrelating information embedded in them. The continuum has more concrete representations (such as photographs and pictures of real-world objects/phenomena) on one end, more abstract representations (such as equations representing relationships between those worldly objects and phenomena) on the other end, and other representations placed in between, based on their abstract/concreteness. All these representations are separated by ontological gaps, and the distance between any two representations on the continuum is proportional to the ontological gap between them, which is in turn proportional to the difficulty to translate between them. In this view, students have conceptual difficulties because they lack an understanding of the translation process across items on the cascade (e.g. figure 2.4).

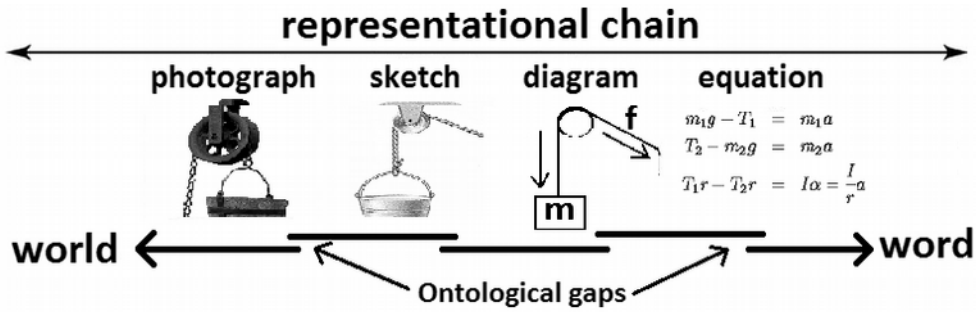


Figure 2.4 The representational chain model (adapted from Johri et al., 2013)

Johri, Roth and Olds (2013) refine this cascade model (figure 2.4), in the context of engineering design, focusing on the relationship between the world and language (and/or thought), where design moves through a series of representational transformations, which bring the world and the word closer.

The model situates words and abstract symbols on one extreme, while world (as experienced by a cognitive agent) on the other extreme of a continuum of representations. Starting with the notion of ontological gaps (Latour, 1993 cited in Roth & Tobin, 1997) between the worldly phenomena and their representations, Johri et al. (2013) argue that the representational translations are crucial in bringing the world and the word closer. In the natural sciences, this movement of the cognitive agent through a continuum of representations happens from the world to the word; whereas in engineering practice, it is the other way round (Johri et al., 2013; McCracken & Newstetter, 2001). In the context of education, Johri et al. state that difficulty in transformations between the kinds of representations results in difficulties in learning.

Box 4: Nature of physics ERs

Nature of physics: Physics aims to understand universal phenomena in terms of matter and its motion. It involves explaining the nature of physical worlds as small as atoms and as large as the universe itself, using abstract concepts such as force and energy, and time and space. Most frequently used MERs by physicists include: motion/kinematics graphs, mathematical equations, equations of laws (Ohm's law, Newton's laws, etc.), ray diagrams, force-body diagrams, circuit diagrams, Feynman diagrams, astronomy models, gestures (right/left hand thumb/grip rule, gestures in kinematics and astronomy), computational models, etc.

Central external representations identified in previous research: Motion/kinematics graphs, mathematical equations, equations of laws (Ohm's law, Newton's laws, etc.), ray diagrams, force-body diagrams, circuit diagrams, Feynman diagrams, astronomy models, gestures (right/left hand thumb/grip rule, gestures in kinematics and astronomy), computational models.

Representational Competence definitions identified in the previous research: While integration of MERs is necessary for physics learning, RC in physics is typically characterized by the ability to boil down all representations to abstract mathematical equations and working with them. Understanding and working at the level of abstract equations implies that one has an integrated understanding of MERs in physics.

Existing experimental approaches to study RC: Expert-novice comparisons of problem representations/comprehension, representational transformation between mathematical and real-world physical phenomena/entities, ability to generate and manage MERs while problem solving; meta-representational competence.

Major results from previous research studies: Experts are good at situating problems in relation to physics principles/laws, whereas novices rely on literal meanings of the problem statements. Experts are better able to generate MERs and fluently (yet less frequently) shift between them while problem solving.

Main theoretical frameworks proposed in the previous research: Meta-representational competence, native competence, expert-novice comparison model

Main existing intervention approaches: Computer simulations (PhET, Netlogo), problem-context-based curricula, computer visualization, and virtual laboratory.

Box 5: Nature of mathematics ERs

Nature of mathematics: The term 'multiple representations' has slightly different (and perhaps additional) connotations in mathematics than 'multiple representations' in the sciences. Multiple representations in mathematics do not necessarily mean representations belonging to multiple media/modes (see figure 4a). The use of multiple representations in mathematics starts very early, right from reciting of numbers and counting (connecting real world objects and numbers) to learning mathematical operations such as addition, subtraction and their real world meanings (through terms like 'lending' money or 'borrowing' it) and gaining expertise in the use of symbols and operators. MERs help mediate different facets or meanings of a mathematical concept, such as the concept of a number. An integrated understanding of these facets, mediated by MERs, is central to learning and doing mathematics.

Central external representations identified in previous research: Numbers, the real-number line, base-ten numerals, algebraic notations/equations, trigonometric models, Cartesian coordinate system, graphs, matrices, sets, geometries, metric spaces, programs, algorithms, logic, embodied and standardized measurement scales and tools (such as the ruler), real-world activities (borrowing, lending, interest, selling areas of land, etc.)

Representational Competence definitions identified in the previous research: The ability to shift between different MERs, particularly spatial and numerical representations, and transform between them.

Existing experimental approaches to study RC: (a) Abilities to inter-relate abstract MERs and concrete situations represented as diagrams, (b) difficulties in relating number/quantity and space in measurement tasks, and (c) reasoning during problem solving based on generation and interpretation of MERs, (d) ethnography.

Major results from previous research studies: Students lack the ability to develop/find equations relating to concrete situations. They also have difficulty integrating space and numerical value. However, students are good at inventing and using criteria to generate, use and interpret MERs to solve concrete problem situations.

Main theoretical frameworks proposed in the previous research: Level of abstraction in MERs, distinction between abstract and concrete sets. Expert-novice comparisons, Lesh translation model and Duval's stage/level models of mathematical competence.

Main existing intervention approaches: Model eliciting activities and problem-based approaches. Forcing of frequent transformations between abstract, realistic-concrete and language based MERs, computer simulations (GeoGebra, Netlogo), virtual/physical manipulatives (killmath)

Box 6: Nature of engineering ERs

Nature of engineering: Engineering involves designing, building, and maintaining different kinds of materials, systems, and processes, as well as a diverse set of structures such as machines, devices and buildings. Engineering activities heavily rely on the iterative generation and use of multiple representations (e.g. materials, inscriptions, sketches and drawings, functions, equations and graphs, prototypes), mathematical and physical models, and simulations. Inscriptions and MERs are crucial particularly to material activities (Roth & McGinn, 1998). These MERs mediate designing constrained real-world things, by making use of abstract ideas and operations (design mode), as well as connecting abstract operations with the behavior of real-world entities (testing mode), in a streamlined fashion. The ability to use MERs to achieve engineering ends is referred to as 'representational fluency' (equivalent to RC). Given its interdisciplinary nature, MERs in engineering are vast and diverse, and therefore the integration problem in engineering is more complex, compared to science. RC studies in engineering education are limited and relatively recent, and are mostly ethnographic investigations coupled with approaches from cognitive science. The next section discusses at some recent empirical work. The later sections discuss theoretical frameworks to characterize RC and student difficulties with MERs in engineering.

Central external representations identified in previous research: User requirements, sketches, drawings, functions, equations, materials, designs, models, prototypes, inscriptions, computer simulations, algorithms, programs.

Representational Competence definitions identified in the previous research: Integration of various abstract, concrete and spatio-temporally distributed MERs, often in the context of real-world problem-solving.

Existing experimental approaches to study RC: (a) Socio-cognitive aspects of engineering, design and technology, (b) ethnographic descriptions of real-world engineering problem solving, (c) how science, mathematics and engineering disciplines could be integrated by exploiting distributed MERs and concepts across these disciplines.

Major results from previous research studies: Integration of MERs is central to engineering learning and practice; MERs help offload information, and also helps add details to the design. Concepts are distributed across MERs, and so is engineering practice.

Main theoretical frameworks proposed in the previous research: (a) Representational-chain model, (b) Lesh translation model, (c) Distributed and situated cognition

Main existing intervention approaches: Model-eliciting activities (a problem-based pedagogy), computer mediated MER integration, visualizations and simulations, STEM integration activities.

2.2.1.2 Reasoning and ERs. The frameworks discussed under this section model students' interpretation of ERs and their reasoning in relation to scientific concepts.

Different external representations present different aspects of the world, and thus, serve different functions in cognition, communication and other activities. Sharon Ainsworth (1999 & 2008) presents three different functions of multiple representations – (a) they are complimentary to each other (as different representations provide different perspectives about the same phenomenon and/or entity), (b) one external representation may constrain the process of interpreting another (as a result of familiarity with it), and (c) multiple representations together help conceptual understanding (through representational integration). This proposal can be interpreted as suggesting possible ways in which a learner may use multiple representations to understand scientific concepts and reason about them. The model is employed by education researchers across many scientific disciplines (e. g. Won, Yoon & Treagust, 2014), for designing interventions as well as understanding the cognitive underpinnings of processing and integrating multiple representations.

Students' ability to interpret ERs in biochemistry is the focus of the model of Schönborn and Anderson (2009), which has three main intertwined components: concepts, reasoning and modes of representations (figure 2.5). According to the authors, this description of different abilities provides a framework for classification of expert ways of reasoning (i. e characterization of RC) and analyzing students' reasoning difficulties. For instance, experts are good at integrating any components the model describes, because they have the necessary conceptual knowledge, reasoning abilities and understanding of ERs to convey their conceptual knowledge, or reason about phenomena and/or entities.

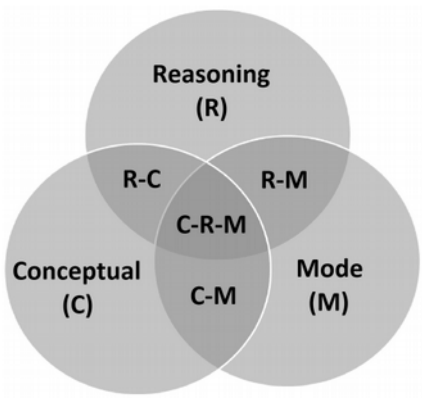


Figure 2.5 CRM model

Learning can be understood as development of connections between these components of the model, and a learner may exhibit the use or combination of any two or all the three components of the model (reasoning based on concepts, representation mediated reasoning, relationship between representations and concepts embedded in them). The authors emphasize reasoning using ERs (modes of representations) in learning, given the central role of representations in science cognition. The model suggests various abilities that characterize competence in science. However, from an assessment perspective, it generates a large number of possible abilities, to assess each of which would be is difficult and time consuming.

An alternative view is presented by Pape and Tchoshanov (2001), who explicate the distinction between internal and external representations, and recommend that thinking and reasoning through representations, (in the context of mathematics), is a result of the interaction of (a) internalization of external representations and (b) externalization of internal/mental images. In learning, the mental images of primary mathematical concepts (such as addition, say, using base-ten blocks) are gradually associated with external representations for these concepts (such as '+'). Also, a key aspect of RC in mathematics is the ability to associate abstract mathematical content with physical representations and vice versa. However, evidence of trade-offs during learning, between grounded (non-abstract, real world representations, such as a word problem) and abstract mathematical representations (such as algebraic expressions drawn from word problems), have been reported in the literature (Koedinger, Alibali & Nathan, 2008). The trade-off exists because there are cognitive costs to using the two types of representations (grounded and abstract). For example, as mathematical content to be learned gets more complicated, thinking in abstract representations becomes necessary, even though this is difficult.

An important contribution of this research stream is the discussion of the relationship between external and/or grounded and internal and/or abstract representations, which is overlooked by previously discussed models, and education research in general, despite being critical in understanding representational transformations, translations, coordination and reasoning (processes most of the models examine, but not from the perspective of the external-internal interaction).

The next model, Representational Construction Affordances (RCA) model (Prain & Tytler, 2012), implicitly assumes this internal-external representation distinction, and focuses on the relationship between the act of generating representations/artifacts of different kinds in scientific reasoning and conceptual understanding. RCA model (figure 2.6) – a Venn diagram of layered ovals of different sizes, with smaller oval(s) nested into larger oval(s) – concerns the relationships between broad and specific meaning-making practices in science around representational construction. The largest oval/layer signifies all the general material (instruments and artifacts) and symbolic tools (language, mathematics, gestures) offered by a culture. These general tools embed relatively specific representations (second oval) concerning epistemic and pedagogical practices around different knowledge systems (thought to be built on top of the general tools). Nested within the first two ovals (representational levels) are even more specific representational tools and practices concerning practice and pedagogy of science.

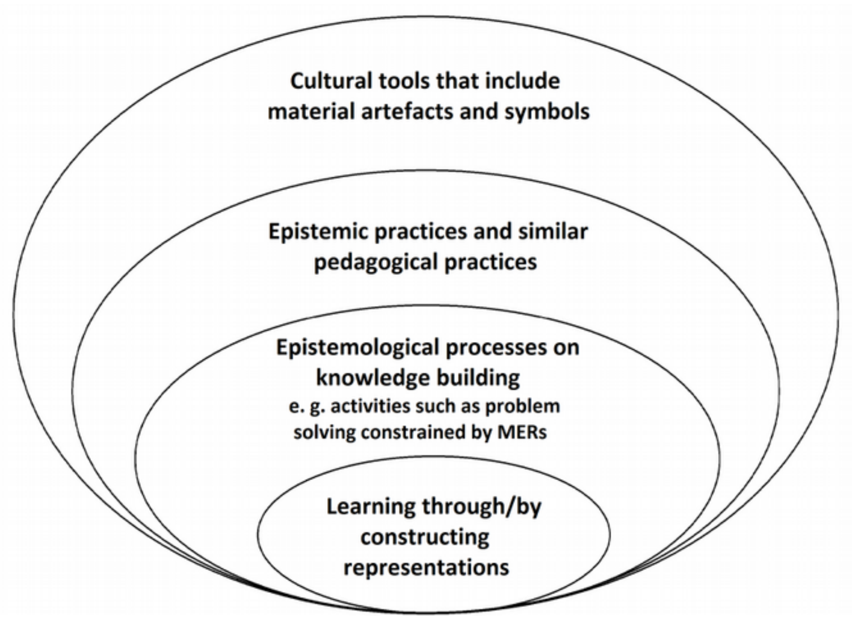


Figure 2.6 RCA model

This is a pan-domain model, and presents how representations 'productively constrain meaning-making practices in science and in science education, taking into account the interplay of diverse cultural and cognitive resources students use to achieve this meaning-making.' Representational fluency or flexibility can be understood as the ability to fluidly move between the general and specific representational systems as required, to facilitate meaning making. The authors stress the meaning-making point, and argue that a large fraction of the reasoning processes around ERs is informal in nature, i. e. not based on formal logic or other language-based systems (see also Tytler & Prain, 2010). Tytler, Prain, Hubber and Haslam (2013) support this argument further, by presenting case studies of students challenged to construct representations in order to solve problems on structure-function relationships in biology. They show that, during the problem solving process, students use visual and other non-formal modes of reasoning, along with linguistic forms of reasoning. Such informal modes of reasoning may be at the heart of ER integration, and thus, as Tytler et al indicate, may have significant teaching-learning implications.

In the distributed cognition view (e. g. Kirsh, 2010), ERs are integral to the cognitive processes of an agent, and there is a continuous, dynamic interaction between the agent's internal and external representations. The distributed cognition approach revolves around two core principles; first that "people establish and coordinate different types of structure in their environment" and "people offload their cognitive effort to the environment whenever practical" (Aurigemma et al., 2013). Aurigemma et al. (2013) extend these two core principles to propose a model of the engineering design process, where the transformation process (between and among multiple representations) rely not only on the dynamic interactions between the internal and external representations, but also on the representation building and other actions that the agent engages in. Building of external representations, in this view, not only offloads cognitive effort, but adds detail and constraints to the mental model and the reasoning of the agent, which would otherwise (as advocated by the classical information processing theories), run only in the head, and lack these details.

2.2.1.3 Mechanism of ER cognition. The models discussed in the above two sections focused more on the classification of external representations, and little on the mechanism of how the different kinds of ERs interact with a learner's mind. The frameworks reviewed below focus on this aspect.

Wu, Krajcik and Soloway (2001) propose a model of RC, examining the possible cognitive connections a learner could make between different available information sources, particularly external representations in chemistry. Informed by the general dual coding theory in cognition by Paivio (1991; elaborations and other versions by Mayer, 2005; Schnotz, 2002; Sweller & Chandler, 1991), the model implicitly assumes internal/mental representations and the external representations as distinct entities, and suggests that a cognitive system can be roughly represented into a 2X2 matrix (figure 2.7), made up of four different sub-systems: a conceptual system which is represented either (a) externally or (b) internally; and similarly, a visual system represented either (c) externally or (d) internally. The authors empirically verify that three specific kinds of cognitive connections are possible for a learner between her conceptual system and representations. The external and internal conceptual systems are connected (connection 1), as are the external and internal visual systems (connection 2). Moreover, the active learner also makes a connection between the internal conceptual and internal visual systems (connection 3). For instance, when a learner encounters an external conceptual stimulus, she actively interprets it (internally represents, connection 1). Similarly, the external visual stimulus is also interpreted (internally represented, connection 2). Often critical is the third connection, the connection between the internally represented conceptual and visual systems (Wu et al., 2001). Difficulties or errors in any of the three connections lead to difficulties in teaching-learning chemistry. This model is best understood as a model of the interaction between ERs and cognition than a model of levels.

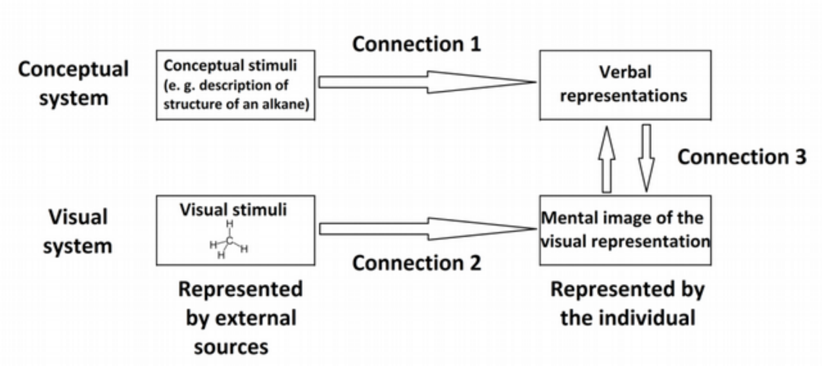


Figure 2.7 Wu et al.'s (2001) mechanism model of learning through ERs

Along similar lines, Schnotz (2002) describes a linear process of how ERs relate to cognition. According to his model, a learner initially perceives external representations (graphics or text) and creates a surface feature-based visual representation in the mind. This surface-feature based mental representation of an external representation or model is then mapped on to common features from other mental representations of external representations/models, which consolidates into a mental model of the subject matter (Schnotz, 2002). Such a mental model is more abstract than the surface-feature based visual representations, and is incomplete, erroneous or absent in novices, as their internal representation remains at the visual level due to lack of prior knowledge.

Based on an empirical study, Briggs and Bodner (2005) propose a model of problem solvers' ability to visualize molecules in a mental rotation task performed on an organic molecule. The results are interpreted in the form of a framework, which suggests that different components of a mental model are at work while handling multiple representations in organic chemistry. Four of the mental model components are: static representations viz. referents (physical objects), relations (spatial relation between referents; Gilbert, 2005), rules/syntax (order of referents guided by conceptual knowledge), and results (outcome/product of visualization). Another component is dynamic, and is rather an operation (e. g. visualization, rotation) performed on the static representations (Briggs & Bodner, 2005). Expert-novice differences can be explained on the basis of differences in static components, rules/syntax (conceptual understanding) and working memory. Unlike previous models, this model assumes the internal representations to be dynamic. However, the relationship between static and dynamic components is not clear. It is also not clear how the model would accommodate referents, relations and results that are dynamic in nature. The notion of conceptual knowledge is vague in the model. Moreover, the nature of conceptual knowledge could itself be dynamic than static.

2.2.2 Developmental models of RC

The focus of frameworks presented in this section is the process of RC development. These developmental models of RC may be informed by one or multiple theoretical assumptions discussed in the previous sections, and hence can be complementary to those models.

Dreyfus (1991) provides a linear stage model based on the number and complexity of representations used by a learner simultaneously. The model proposes that ERs mediate the process of learning, which passes through four sequential stages: 1) using single representation, 2) using more than one representation in parallel, 3) making links between the representations used in parallel; and 4) integrating representations as well as flexibly moving between them. The author grants that the processes of representation (act of representing something) and abstraction as complementary processes moving in opposite directions. In other words, the act of representation is parallel to externalization, while abstraction connotes internalization.

A more sophisticated account of learning with multiple representations in mathematics is provided by Duval (2006). This model maintains that coordination between at least two representational forms, termed as registers, is necessary for comprehension of mathematical concepts. There are four such representational forms/registers: natural language, figures and diagrams, notation systems (symbols) and graphs. Learning with multiple representations involves students gaining more control over these registers. A learner initially stays within one register (e. g., carrying out calculations in only one notation system), then moves to conversions, where she changes the register (e. g. using notations/symbols like '+' to represent 'addition', a mathematical relationship originally described in language/words), and then finally achieves coordination among multiple registers.

Goldin and Kaput (1996) also provide a three-stage process of development of RC in mathematics. The stages are: (a) Inventive-semiotic stage, where a learner is introduced to new characters of a representational system (for instance numbers/counting) that symbolize aspects of familiar systems such as a real-life situation; (b) the use of this system as a template to learn a more sophisticated system of rules for the new symbol-configurations (for instance, the concept of a number) and diversities; and, (c) the new system, once learned and practiced with, becomes independent and detached from the earlier system of representations (for instance, doing arithmetic/algebraic exercises). The third stage indicates abstraction, and is particularly critical in characterizing RC in mathematics, since mathematicians often operate in the world of abstract entities.

An influential model of the different abilities of experts and learners working with ERs is given by Kozma and co-workers, who coined the term ‘representational competence’, to describe “a set of skills and practices that allow a person to reflectively use a variety of representations or visualizations, singly and together, to think about, communicate, and act on chemical phenomena in terms of underlying, imperceptible physical entities and processes” (Kozma, 2003; Kozma, Chin, Russell, & Marx, 2000; Kozma & Russell, 1997; Kozma & Russell, 2005; Madden, Jones & Rahm, 2011).

The authors characterize RC in terms of following specific skills in the context of chemistry: (a) using representations to describe chemical phenomena; (b) generating and/or selecting and explaining appropriate representations for a specific purpose; (c) identifying and analyzing different features of representations; (d) comparing and contrasting different representations and their information content; (e) making connections across different representations, mapping features of one type of representation onto those of another, and explaining the relationships between them; (f) understanding that the representations correspond to phenomena but are distinct from them; and (g) using representations in social discourse to support claims, draw inferences, and make predictions (Kozma & Russell, 2005).

Alternatively, several researchers opine that students' RC is often underestimated, despite reports suggesting difficulties in generation, selection, coordination and general handling of ERs among students (e. g. Izsák, 2011; Kieran, 1981; Leinhardt, Zaslavsky & Stein, 1990). Students exhibit better competence than previously thought (diSessa, Hammer, Sherin & Kolpakowski, 1991; diSessa, 2004; diSessa & Sherin, 2000). Preliminary research investigating the nature of untutored native competence among students in terms of content knowledge (whether inarticulate intuitions or articulable/potential principles), sources of such knowledge, and the possibilities of refining this knowledge indicate that students' capabilities with representations were often underestimated by prior studies (diSessa & Sherin, 2000). Students are capable of having deep and rich, although intuitive, ideas about dealing with and making sense of external representations in their own ways. This competence is referred to as 'native competence, or meta-representational competence' by diSessa & Sherin (2000), and constitutes the following abilities: (a) invent or design new representations, (b) critique and compare ERs, for their appropriateness and adequacy, (c) understand various functions of representations in context, and how representations serve such functions in that context, (d) explaining representations and (e) learning new representations quickly with minimal instructions (diSessa, 2004). The notion of meta-representational competence is different from RC in the following way: it is concerned with whatever students know about the act of representation and its products (meta-representation); it does not focus on representations used for instruction in a domain, or the standard school modes of reproduction and interpretation (diSessa & Sherin, 2000).

2.2.3 Summary

In this section I reviewed important theoretical frameworks for RC, categorizing them under two major emerging themes: models of RC concerning ERs and their cognition, and the developmental models of RC (figure 2.8).

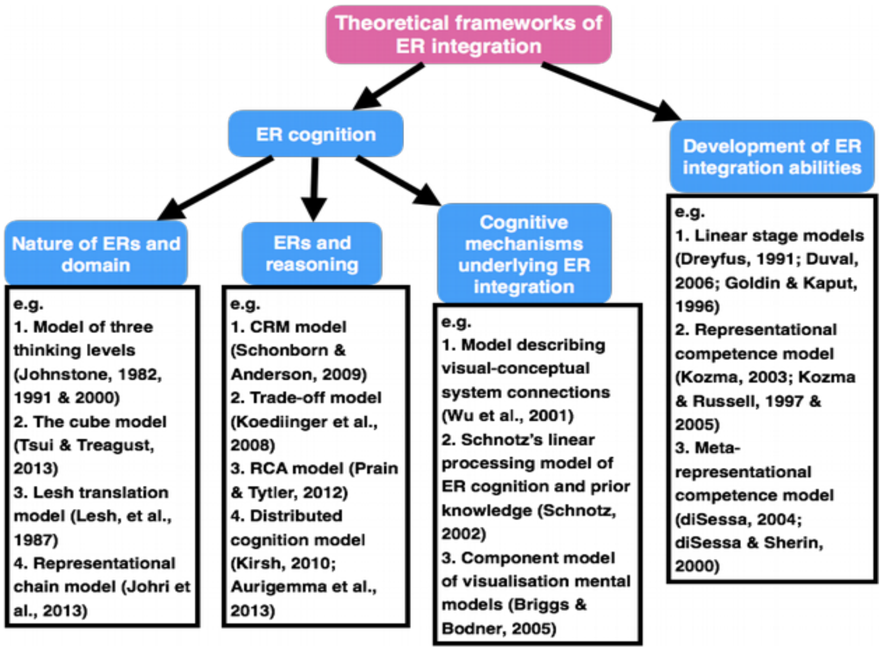


Figure 2.8 Broad categories of theoretical frameworks and models that emerged during the review.

Within the former category of models, I saw three different sets of models: one captures the relationship between nature of a domain, ERs in that domain, and cognition (Model of three thinking levels and its versions, cube model, Lesh translation model, model of ontological gaps, and representational chain model); the other captures how students reason about ERs (Ainsworth's model of function of representations, CRM model, internal-external/abstract-concrete ER trade-off model, RCA model, and distributed cognition framework); whereas a third set of theoretical frameworks models the cognitive mechanisms involved in the processing of ERs (Dual coding theory and models informed by this view, and the model of four cognitive components). Among the RC developmental frameworks, I reviewed stage models proposed by Dreyfus (1991), Duval (2006) and Goldin and Kaput (1998), Kozma and colleagues' RC characterization, and the model of native and meta-representational competence by diSessa and co-workers.

The section below discusses empirical studies investigating different aspects of RC, as well as the process of learning through ERs.

2.3 Empirical investigations of learning with ERs and RC

There is a vast literature reporting empirical investigations of RC, examining the learning and use of scientific and mathematical ERs, the use of ERs in science practice, and the nature of skills involved in RC. These studies are widely dispersed, and often published in discipline-specific venues. Only a few studies explicitly subscribe to one of the specific models discussed in previous sections (e. g. Aurigemma et al., 2013; Hinton & Nakhleh, 1999; Madden, Jones & Rahm, 2011; Moore et al., 2013, etc.). Most studies only broadly relate to the major theoretical frameworks of RC. As a consequence, there is no well articulated theoretical framework that helps integrate these disparate studies.

In this section, I try to bring together these disparate studies along two major themes based on the RC abilities they focus on: 1) linking ERs and translating between them, and 2) expert or student generation of ERs and their ER preferences.

Some studies discussed under each theme explicitly appeal to the classical information processing paradigm in order to explain student learning difficulties and/or expert-novice differences in relation to ERs, while some other studies implicitly assume classical information processing perspectives, but do not endorse this view explicitly. These studies focus on one or more of the following processes, usually identified with the classical information processing paradigm: working/short-term memory, long term memory, information storage (assumes a storage module), memory or information extraction (assumes searching), translation of information (assumes that information from one code is translated into other code(s)), and fully internal representations. A third set of studies are neutral on the nature of ERs and RC. Finally, some subscribe to recent cognition theories such as distributed and/or embodied cognition.

2.3.1 Linking ERs, translating and/or transforming between them

Students find it difficult to understand the interrelations between different symbolic representations, which capture different features or aspects of worldly phenomena. A wide range of studies have examined this difficulty. For instance, information about a chemical reaction is embedded in the symbols and numbers in the chemical equation representing that reaction. Being able to relate symbols and numbers with the dynamic reaction, by cross-linking the 'three thinking levels' (Hinton & Nakhleh, 1999), is one way to make sense of a chemical equation. Studies show that students lack a clear understanding of basic concepts such as oxidation numbers, ionic charge, atoms and atomic structure, formal rules for writing molecular formulas, as well as meaning of subscript letters, numbers and coefficients (Garforth, Johnstone & Lazonby, 1976; Savoy, 1988). Because of this, students face difficulties while dealing with chemical equations. In addition, students fail to associate the 'symbols and numerical answers with real objects and phenomena' when asked to explain different chemical equations (Herron and Greenbowe, 1986) using particulate drawing (Sanger, 2005). Studies examining how students balance chemical equations asking them to explain their balancing protocol reveal that many students balance chemical equations algorithmically i. e. without actually understanding the meaning of symbols and numbers (Hinton & Nakhleh, 1999; Nurrenbern & Pickering, 1987; Yarroch, 1985). Such an algorithmic approach to equations could be linked to the failure in understanding that the coefficient and subscript numerals are not just some numbers, but represent and quantify the particulate nature of matter. In other words, it is a failure in establishing correspondence between macro-level visible reality and the periodic table, symbols and numbers, chemical formulas and reaction mechanisms.

In mathematics, many studies examining the understanding of length and area measurement show that children often struggle to see the relationship between numbers, units and space; particularly how a numerical value is related to a spatial area (Battista, 2003; Battista & Clements, 1996; Kamii & Kysh, 2006; Pande & Ramadas, 2013), even though all these different representational systems model the same concept. Santos' (1996) examination of students' responses to contextualized hypothetical questions (such as '*how many tennis balls would it take to fill a classroom?*') reveal that students' use of numbers and algebraic as well as arithmetic operations is largely algorithmic.

A related set of studies analyzed the connections students and teachers make between ERs (particularly graphs, tables and pictorial representations) of mathematical functions (Çelik & Sağlam-Arslan, 2012; Hitt, 1998; Knuth, 2000) by documenting which ERs were preferred by the participants over others. Knuth (2000) presented high school students several function problems using algebraic and graphical representations, and asked them to solve each problem using either a graph or an equation, and then also produce an alternative solution method using the other representation. The author found that graphical representation provided during the study was often considered irrelevant by the students. Most students prefer algebraic/symbolic representations (Acevedo Nistal, van Dooren, Clarebout, Elen & Varschaffel, 2010; Acevedo Nistal, van Dooren & Varschaffel, 2012; also shown in probability tasks by Anastasiadou & Chadjipantelis, 2008). Students did not agree on which representation would be appropriate for a problem, and found it difficult to explicitly reason using chosen representations (Acevedo Nistal, van Dooren & Varschaffel, 2012). Learners also have a general difficulty establishing links between problem situations (word problem statements), graphs and functions and other symbolic representations (Billings and Klanderman, 2000).

Elia, Panaoura, Eracleous and Gagatsis (2007) used different tasks that required students to explicitly talk about their definitions and understanding of the concept of functions, identify correct algebraic functions in relation to diagrams of certain situations, and translate between multiple representations of algebraic functions. The authors report remarkable inconsistencies among students in relation to: (1) approaches to different representations of functions across tasks, and (2) definitions of functions and their ability to recognize the concept of function in different forms or problem solving tasks. They concluded that students tend to highly compartmentalize the concepts taught to them, based on differences in the situations and the representations encountered around those concepts. Students were also found to perform badly in relating situation diagrams and algebraic functions.

In chemistry, Kozma and Russell (1997) report an expert versus novice study, where they posed two tasks, a categorization task and a transformation task, to experts (practicing chemists) and novices (undergraduate students), individually. The authors wanted to know if the participants "*saw connections between different chemical visualizations corresponding to the same phenomena or if they understood something different for each type of visualization*". The first task required participants to group a set of 14 cards, with dynamic and still images (corresponding to several chemical reactions), into meaningful groups. The representations (dynamic and static images) included videos of the experiments, animations of the molecular events, graphs, and chemical equations. Observations revealed that novices formed their meaningful groups from a small number of cards, often from the same media type (e. g. all graphs as a category, all equations as another category and so on) while experts made larger groups, composed of multiple media forms. Experts gave largely conceptual reasons for forming particular groups, while novices' reasons were often based on surface features. In the second i. e. transformation task, participants were shown chemical equations, videos of experiments, dynamic graphs and animations of the molecular events of an experiment, one at a time. Participants were asked to transform the given representation to another form (such as drawing a graph corresponding to the given equation, selecting an animation that best corresponds to an equation, etc.).

The authors found that "*experts were much better than novices at providing verbal descriptions, due to their deeper understanding of chemical principles and concepts*". Also, experts were better than novices when transformations required a constructed response, such as drawing a graph or writing a chemical equation.

Similar findings are reported by Madden, Jones and Rahm (2011) in their examination of RC differences between first semester and advanced level chemistry students, in the context of ideal gas problems. The problems, aimed to investigate RC level among students, required the students to provide verbal descriptions of behavior of an ideal gas (from particulate level sketches, diagram and graphs), calculate (i. e provide a mathematical representation of) pressure exerted by a gas, and transform between these generated calculations, mathematical and verbal descriptions, graphs and particulate level sketches, etc. The authors used a modified version of Kozma's (2005) RC framework to analyze student performance, and found that students with less prior experience largely exhibited algorithmic use of the ideal gas law. Their use of equation, variables and values seemed to be disconnected from other representations, unlike students with more exposure.

Ben-Zvi, Eylon, & Silberstein (1987 & 1988) found that students' thinking relies primarily on perceptual/sensory information, and since the pedagogical practices while teaching symbols, equations, and operations do not seek to provide perceptual/sensory assistance, these aspects of science and mathematics are not understood by students in terms of their macro and micro-level instantiations. As a result, learners tend to concentrate more on the familiar representation(s) or algorithms in order to counter the cognitive load, and end up ignoring the relationships between concrete and abstract ERs (Johnstone, 1991; van Someren, Reimann, Boshuizen & de Jong, 1998).

Ozogul, Johnson, Moreno and Reisslein (2012) also focus on the load on working memory students experience while learning ERs. They examined the effects of various modes of integrating equations in circuit diagrams in the engineering domain, and found that undergrad students often fail in integrating the two kinds of representations, because of the increase in cognitive load during instruction. The ability to establish relevance, given the information depicted through ERs, is related to the amount of information working memory can handle (Chi et al., 1981). Higher prior knowledge facilitates identification of the relevant/necessary features in representations, and extraction as well as interlinking of information through these features (Chi et al., 1981; Cook, Wiebe & Carter 2008; Kozma & Russell, 1997; Larkin & Simon, 1987). This lowers cognitive load, as participants with higher prior knowledge can rely on their existing knowledge stored in long-term memory for information chunking.

The prior knowledge effect was demonstrated using eye-tracking by Cook et al. (2008) while examining the way students' prior knowledge interacted with how they interpreted macro and molecular graphics of diffusion phenomena. The authors captured and observed the number of transitions students made between molecular-to-molecular, macro-to-molecular, molecular-to-macro and macro-to-macro representations, using eye-tracking. On average, students with low prior knowledge made more transitions than students with high prior knowledge. Students with low prior knowledge focused more on surface features of representations (Kozma & Russell, 1997). Low-prior-knowledge students needed to make frequent transitions in order to map features from one representation to the other, trying to link them together. Similar patterns of transitions between ERs are reported by Kohl and Finkelstein (2008) in a study aimed to understand patterns of ER use during problem solving in electrostatics. Three groups of participants – experts, weak novices and strong novices – individually solved two different sets of problems. In one set, ERs were given to the participants. In the second set, word problems on electrostatics were given, and participants had to generate representations based on the textual description. Performance results showed multiple levels of competences across all the three groups, but experts (as originally designated) generally tended to successfully solve problems, making less number of to-and-fro transitions (measured as density of transitions between representations per minute) than the novices. Surprisingly, strong novices exhibited intermediate performance. Since participants with low prior knowledge are less aware of the 'subtleties of representations and the conventions for interpreting them', they may have needed more transitions to interpret the represented information (Cook et al., 2008) and relate it to information in other representations. Interestingly, researchers have found that although the domain knowledge and RC are interconnected, RC can be predicted from, but cannot guarantee, domain knowledge (Nitz, Nerdel & Prechtl, 2012; Nitz & Tippett, 2012).

A parallel set of studies argues that visuo-spatial thinking ability is fundamental to RC, although working memory capacity is the ultimate limiting factor. Bodner and Domin (2000) examined transforming of 2D representations into 3D and the reverse, and documented the difficulties students encounter with such transformations, especially in the context of organic chemistry. There is a deep relationship between students' mental rotation ability and their ability to transform 2D representations into mental 3D representations (Shubbar, 1990; Wu & Shah, 2004). Shubbar (1990) attributes students' difficulty in 2D-3D transformation to problems in either comprehending depth cues in 2D diagrams, or tracking the depth cues in molecular diagrams that depict chemical change. These multiple simultaneous activities put tremendous cognitive load on student, and are critical to learning difficulties (Wu & Shah, 2004). In chemistry, a learner needs to perform multiple operations at multiple spatial scales: atoms and molecules, their collective behavior, and properties and reaction mechanisms need to be imagined simultaneously in a consistent manner. Similarly, understanding biological phenomena such as evolution, for instance, requires traversing different levels of spatial scales, right from DNA mutation to changes in an organism across generations.

The empirical work discussed so far largely advocates vocabulary and ideas usually identified with the classical information processing theories in cognition.

Building from such studies, there have been a number of attempts since the early 1990s to develop chemical visualization/virtual manipulation software to help students develop RC. These interventions are based on the classical information processing approach to cognition, particularly Baddeley's working memory model (e. g. SMV Chem, visChem, 4M:Chem, EduChem HS, eChem, etc.). These interventions seek mostly to *display* multiple representations simultaneously on screen, to lower the load on students' memory.

There is also a significant number of studies in RC either disregarding the concepts such as working memory and/or cognitive load, or employing alternative perspectives to cognition. For instance, Kozma et al., (2000) anchor their work in the situated cognition perspective, which proposes that knowledge of a practitioner (say a chemist) is inseparable from the natural context of that practitioner (chemistry laboratory), and is therefore best investigated within the context of that practice. The researchers observed chemists and academicians practicing in laboratories, and reported that 'materializing' representations that could be perceived and manipulated, (i. e. creating and/or using ERs) helped participants operate on the otherwise non-perceptible entities and processes. ERs also helped chemists discuss problems; they used visualizations and structural diagrams to describe the composition and geometry of the compounds considered for synthesis, and used diagrams and equations to think through the possible reaction mechanisms. Kozma (2003) extended this earlier study with novices (undergraduate students), and reported lack of such RC in this group.

Shifting the focus from students, Stewart (1982 & 1983) argues that the origin of student difficulty in interlinking multiple representations lies in the teaching sequence of concepts. Taking examples from biology, the authors argue that teaching Mendelian genetics before cell division could be one reason why students fail to understand the connection between meiosis (micro level explanation) and Mendelian genetics (macro level). Longden (1982), on the other hand, situates the root of the problem in the static nature of diagrams used in science classrooms.

Schnepp and Nemirovsky (2001) emphasize the role of dynamics in understanding ERs in the context of physics, and argue that the recognition of motion in distance-time, velocity-time and other equations and graphs of motion requires merging perception with imagination. They found, through observing a calculus course for 12th graders, that students face difficulties in imagining motion depicted in mathematical representations of physical phenomena. Sometimes these depictions refer to physically impossible events. For instance, a distance-time graph depicting a plane after a slope refers to an object instantaneously stopping its motion. Imagination is key to recognizing the relevance of such representations to physical phenomena (Schnepp & Nemirovsky, 2001), and thus understanding the conceptual content of ERs. Thompson and Sfard (1994) argue that mere perception of a function, for instance, in its multiple representations such as a table and a graph, may not be sufficient for a student to realize the equivalence between those representations (also suggested by Kaput, 1995). It is extremely difficult to gauge if a student understands the continuity of that function, distributed across those multiple representations.

Similarly, White and Pea (2011), during observation of students collaboratively solving a set of decryption problems using a dynamically linked multiple representation environment (Code Breaker), discovered that although students may exhibit competence relative to a specific task during problem solving episodes, understanding that the concepts and mathematical operations are distributed across multiple representations may take numerous episodes of using multiple representations (Giere & Moffatt, 2003) as well as maneuvering different representational tools (Hutchins, 1995a; White & Pea, 2011) collaboratively in different problem situations.

Close to the view I advocate later in next chapter, a group of researchers argue that understanding scientific phenomena not only requires seeing the different connections between ERs, but also using those ERs and the connections between them, to build dynamic internal (mental) models that simulate the behavior of many individual components of real world events (Davidowitz, Chittleborough & Murray, 2010; Grove, Cooper & Cox, 2012; Levy & Wilensky, 2009) and effects of various parameters on such events. Difficulty in building consistent internal models of phenomena using ERs is a major problem identified among students. For instance, students are reported to have difficulties in mentally animating as well as simulating physical systems (such as flush-tank, gears; Hegarty, 2004; Schwartz & Black, 1996a & 1996b). This leads to problems in understanding and predicting system behavior and/or answering problems.

Unlike the previously discussed computer interventions (based on memory-based approach focusing on simultaneous display), recent work informed by these alternative perspectives focuses on interlinking representations through dynamic *manipulable* simulations, animations, and physical models. For instance, the manipulability feature in the Connected Chemistry Curriculum, based on the Netlogo 2D interface, may help students transform better between static and dynamic representations (such as equations, graphs and molecular simulations). The developers of this curriculum, through control-treatment group experiments where students' were asked to draw sub-microscopic pictures for certain chemical systems/reactions, report that the curriculum improves handling and understanding of multiple representations in chemistry, when compared to conventional text or lecture based curricula (Stieff & McCombs, 2006; Stieff & Wilensky, 2003).

Kothiyal et al. (2014) report in detail the development and testing of a *fully manipulable* simple pendulum simulation designed to help high school students integrate ERs around the concept of oscillation. The design principles behind this simulation are inspired by distributed and embodied cognition perspectives (e. g. external representations allow processing not possible/difficult to do in the mind, Kirsh, 2013; action patterns can activate concepts, hence actions and manipulations of the representations should be related to existing concepts, O'Malley & Soyer, 2012). Unlike the Netlogo and PhET simulations, this simulation focuses on the *enactivity*/manipulability of abstract ERs, particularly ERs such as equations/graphs, in order to give the learner maximum control over the behavior of the system through multiple modes. The authors claim that the *enactivity* of equations and abstract ERs is critical for understanding (implicitly as well as explicitly) the dynamic relationship between those ER, and thus imagine the represented entity/phenomenon.

Such manipulable interfaces have often been coupled with other scaffolds (such as exercises, quizzes, activities and teacher guides; Kukkonen, Kärkkäinen, Dillon & Keinonen, 2013; Varma & Linn, 2011) and these have been effective in improving students' representations and understanding. In organic chemistry, the activity of matching physical models to diagrams has been shown to provide (implicit) feedback to participants, leading to their improved performance during representational tasks (Padalkar & Hegarty, 2015). Computer interfaces have been explored from an assessment viewpoint in order to better characterize RC and multiple representational transformations among learners. Stieff, Hegarty and Deslongchamps (2011) examined students' use of a multi-representational molecular mechanics animation using eye-tracking, and observed that students mainly used graphical and model representations in animations, and often ignored the equation. Based on an eye-tracking exploration of participants' chemistry ER viewing as well categorization processes, Pande and Chandrasekharan (2014), and Pande, Shah and Chandrasekharan (2015) concluded that the richness of transitions, as well as the nature of transitions, between different parts of a representation (and/or different representations) could be considered a good marker of ER integration.

Interestingly, a group of researchers went beyond the 'traditional' enactive approaches to incorporate socio-cultural contexts and inter-student collaboration possibilities in the play-environments where a group of students interact with simulations (mixed-reality systems) with their full bodies by collectively and collaboratively enacting avatars representative of scientific concepts (Danish et al., 2015; Enyedy et al., 2012). In what the researchers called the Science through Technology Enhanced Play or STEP project (e.g. Danish et al., 2015), students move around in the classroom space, play-acting particulate matter (say water molecules) to arrive at an understanding of how particles in different states of matter (e.g. water, ice, vapour) would behave in a range of everyday situations such as a freezing cold day, heating or boiling water, etc.. The students see their ‘enaction’ projected into a computer simulation where an avatar in the form of a particle is displayed. The authors claim that this type of enaction is useful to direct students’ attention towards key concepts and help them make their own choices and decisions in arriving at a proper understanding of these concepts. Most intriguingly, their results show that mere embodiment and enaction may not be sufficient to bring clarity to nuanced concepts such as energy and its relationships with states of matter, as some (groups of) students tended to confuse between concepts such as energy and matter. However, this confusion was often clarified when the groups had non-enacting student observers, teachers and/or facilitators (Danish et al., 2015) helping the students reflect on their own actions and observed effects of those actions on avatar behaviour.

2.3.2 Generating ERs and representational preferences

Representations generated by students, and their choices of representations (e. g. which representation would help better in a given problem situation), are considered good indicators of misconceptions (as these reflect internal representations, Chi et al., 1981). They also suggest how different representations aid student thinking (as they support reasoning during the problem solving process, Izsák, 2011). Literature in this area documents different representational preferences among students, and suggests that students find it extremely difficult to generate ERs and use the generated ERs to reason about phenomena in systematic ways (Diezmann & English, 2001; Kamii et al., 2001; for detailed review, see Diezmann & English, 2001).

Many of the classic studies in science, mathematics and engineering education focus on the nature of representations participants generate during problem solving. These generated ERs are considered markers of participants' problem representations (or internal representations/mental constructs of problem situations). Extensive work, particularly in the 1980s, investigated the way experts and novices approach physics problems, and found certain key qualitative differences between the two groups, particularly in their problem representations. Chi et al.'s (1981) influential study, for instance, found that experts and novices categorize given physics problems into different groups. The categories and explanations generated by experts had few features in common with those provided by novices. Experts sorted problems on the basis of principles, such as Law of Conservation of Energy, which could be used to solve the problems. Novices, on the other hand, exhibited limited capabilities in going beyond surface features of the problem statements/diagrams (such as literal meanings of words) while categorizing problems. For instance, they put 'merry go round' and 'rotating disk' problems in the same category, as both involved rotating things. To explain these differences, Chi et al. (1981) postulated that differences in prior knowledge of the experts and novices make their problem schemata different from each other. The problem features engaged more tacit knowledge in the case of experts (Chi et al., 1981).

Interestingly, the authors found that both experts and novices used the same set of features in problem statement (and/or diagram), but the differences lay in the cues and interactions those features had with their prior knowledge and subsequent problem schemata. Participants' prior knowledge and their ability to identify patterns of meaningful information (in and using ERs) were closely related. Experts (generally assumed to possess denser domain knowledge) are more likely to extract task-relevant knowledge from a given representation or generate one to aid problem solving (Chi et al., 1981; Larkin, McDermott, Simon & Simon, 1980).

In Hmelo-Silver and Pfeffer's (2004) investigation of pictorial representations (and verbal responses) of models of aquatic systems generated by experts and novices, the experts were found to integrate structural-functional-behavioral information by dynamically imagining the mechanistic relationships between them; novices relied on the static structural features of the system components. Several other studies examining structure-function relationships report similar findings (e. g. Jacobson, 2001; Mathai & Ramadas, 2007; Subramaniam & Padalkar, 2009).

A related strand of research looks at how experts and novices differ in the way they use analogies to understand and explain biological phenomena (Dreyfus & Jungwirth, 1990). Student participants were asked to explain the meaning of various statements, such as – 'the nucleus controls the functioning of the cell'. Most participants used fallacious analogies, such as – 'just as brain controls the body', in their responses. Experts often used analogies from systems they understood better, but they also searched for potential mismatches in the analogies. The novices were satisfied with the criterion of familiarity with the system while choosing an analogy, and never checked the analogies for mismatches. The authors suggest that mismatches in analogies may result in inconsistencies among internal representations, difficulties in understanding ERs, and ultimately difficulties in understanding biological phenomena. Analogies are thus powerful yet risky tools in interlinking information at multiple levels of organization (Dreyfus & Jungwirth, 1990).

Santos (1996) examined students' responses to spatial problems that required generation and use of multiple representational approaches. When students were asked to estimate the number of tennis balls needed to fill a classroom, they often attempted to translate the word problem into calculations without completely understanding the problem. They tended to use arithmetic and/or algebraic approaches to solve such problems, and had difficulties in moving from the arithmetic representation to visual estimations. Billings and Klanderman (2000), in the context of problems on motion and related topics in physics, found that students (pre-service teachers), when given graphical representations showing relation between speed and other variables (time-distance graphs), excelled at generating symbolic representations and operating on them (e. g. calculating average speed). However, the same students struggled in generating reasonable graphical representations and interpreting them while designing question sets for school exams. Further analysis of the assignments and question sets submitted by the students revealed that students found it difficult to distinguish average speed from instantaneous speed, and even distance and speed. The slope of the line was often an area of misinterpretation and confusion.

On the other hand, students are reported to exhibit sophisticated reasoning around their choice of representations. For instance, fourteen-year olds, when posed with three kinds of tasks based on the design and working of a physical device (called winch, figure 2.9) in a study, generated many different equations, developed criteria to evaluate these equations, and finally selected some equations based on these criteria (Izsák, 2011). Interestingly, the participants developed as well as articulated their own criteria, such as '*single equation is better over multiple ones'* and that '*the expression must generate positive values for distance*', during the selection and evaluation of generated representations. Although students lacked coordination between their criteria while evaluating ERs generated by them, the fairly reasonable articulation of criteria hint at some competence among students in evaluating and integrating ERs.

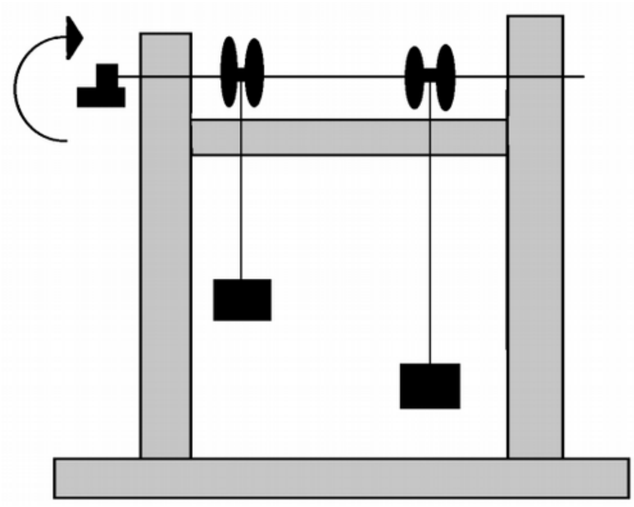


Figure 2.9 The winch (adapted from Izsák, 2011). The author posed three problems to the students based on the design and mechanism of working of this device viz. (a) *predict the distance between the weights after an arbitrary number of cranks*/handle rotations; (b) *determine whether and, if so, when one weight will ever be twice as high as the other and* (c) *determine whether and, if so, when the weights will meet at the same height*

Such seemingly pragmatic representational preference tendencies (reasoning) among students are highly context-dependent, and there may be significant individual variations (Çikla & Çakiroglu, 2006). For instance, equations are preferred during mathematical situations, whereas graphs are used with contextualized word/mathematics problems (Keller & Hirsch, 1998; Scanlon, 1998). Students feel comfortable in using only symbols in fraction problems, whereas they encounter difficulties in relating concrete models/visuals of fractions with number-line, verbal and symbolic representations (Biber, 2014; Brenner, Herman, Ho & Zimmer, 1999).

Novices have relatively unstable internal representations of problem situations than experts (Anzai, 1991; Anzai & Yokohama). This possibly arises from the limited capacity of their working memory. In a problem solving experiment, expert and novice participants were asked to predict the behavior of a constrained system (figure 2.10), a yo-yo made by connecting the centers of two circular disks with an axle. The system was kept on a table, in such a way that the yo-yo disks could roll, but not slide (Anzai & Yokohama, 1984; re-described in Anzai, 1991).

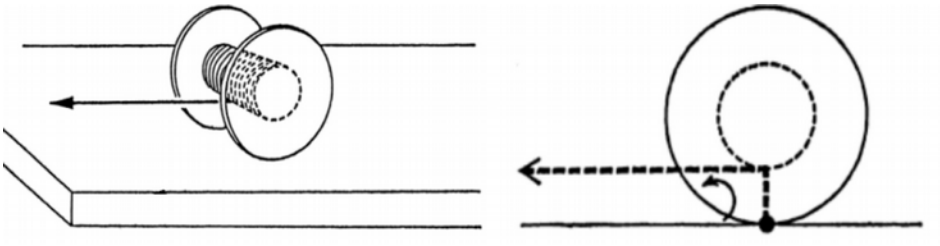


Figure 2.10 Yo-yo on a table problem (adapted from Anzai & Yokohama, 1984). The problem diagram is on the left, and on the right is its abstract diagram

Participants were asked whether the yo-yo would roll to the left or to the right (left being the correct answer). Experts applied the lever-fulcrum principle to answer this problem correctly, while novices related the problem to real-world situations, trying to erroneously animate the yo-yo and thus performing poorly. However, changing the problem representation to abstract diagrams helped some novices to answer the yo-yo problem correctly (Anzai, 1991), possibly because of the reduction in cognitive load provided by the more directed animation of the movement of yo-yo. This result supports the suggestion from Larkin (1982, cited in Anzai, 1991) and Anzai (1991) that experts and novices tend to use qualitatively different internal spatial representations to solve problems. Experts could be performing better in this generation task because producing and/or using diagrams allow computationally more efficient search for stored tacit information, and inference based on this information, compared to symbols and sentences (Kozma, 2003; Larkin & Simon, 1987).

A related strand of research examines ways to improve generation and integration of ERs among students, using different approaches. Cardella et al. (2006) investigated the role of sketching and ERs, through a verbal-protocol-case-study analysis of engineering students' representations and representational activities during a design problem solving process. Participants had a general tendency to use the given ERs, rather than generate new ones. However, those who could generate relevant ERs, such as problem statements, sketches and calculations, exhibited 'information gathering' during the process, to progress through the task. Based on this result, the authors suggest that generation of ERs compensates for the limitations of imagery. Successful problem solvers tend to construct more accurate, complete and abstract representations (numerical/symbolic/mathematical forms) of the problems, than the unsuccessful ones. Representations generated by successful problem solvers also evolved over time in terms of their abstractness (Domin and Bodner, 2012; Sevian et al., 2015), suggesting addition of information from participants' prior knowledge.

Reisslein, Moreno and Ozogul (2010) distinguish between abstract and contextualized engineering ERs, and emphasize the use of instructions that involve both these kinds. The instructions allow students to get more opportunities to interconnect abstract ERs, such as equations, with real-life situations. The study assessed learning outcomes and program rating, using a survey based on a post-test. Three groups of participants received three different types of ER-mediated instructions: abstract (e. g. equations), contextualized (e. g. only circuit diagrams) and combined (both equations and circuit diagrams). Participants who received the combined contextualized-abstract instruction scored higher on the post-test, produced better problem representations, and rated the program's diagrams and helpfulness higher than their counterparts.

Based on the Lesh translation model (discussed in 2.1), model eliciting activities – particularly specific and goal-directed activities that involve building a working model of phenomena using ERs – have been reported to be effective interventions in engineering education (e. g. Diefes-Dux, Moore, Zawojewski, Imbrie & Follman, 2004; Lesh & Doerr, 2003 cited in Moore et al., 2013). Moore et al. (2013) investigated how engineering students used representations and representational fluency in modeling heat exchange, and what role representations and representational fluency played in conceptual development during this activity. The students were expected to develop a model to 'predict the interface temperature and the sensation felt by human skin when touching a utensil made of a specified material at a given temperature'. Student-generated representations were grouped under the five categories provided by the Lesh translation model, viz., concrete, pictorial, symbolic, language and realistic. Model development was found to be a function of representational fluency, involving not only generation and use of ERs, but also translation across the five categories of representations, and among multiple representations belonging to the same category. Going through the process of model development also often improved representational fluency among students (Moore et al., 2013).

Kindfield (1994) demonstrated, through an empirical study, the role of diagram generation in tweaking the mechanism of working memory, by connecting external representations/models, internal representations, and conceptual knowledge. The study analyzed students' ability to generate diagrams during meiosis (cell division) problem solving, and the quality of the generated diagrams among participants with varying degree of formal training in genetics (meiosis – cell division). Two criteria were used to distinguish the participants: (a) number of different representations of chromosomes used to reason about meiosis, and (b) nature and timing of inclusion of different features of representations. These two criteria determine a 'knowledge-dependent representational variability' (Kindfield, 1994). Similar to RC, this concept captures the quantity and quality of variations in the use and generation of ERs. Expert problem solvers exhibited knowledge-dependent representational variability, fine-tuned their diagrams according to the nature of the task, and used them systematically during reasoning. A cyclical approach was observed in experts' problem solving process. They first drew diagrams that offloaded their mental model, and then paused over the drawn figure, where they offloaded the computation of chromosomal configurations (essential for correct reasoning), and then drew again to externalize and check solutions, while also keeping track of the previous steps. Expert problem solvers thus exhibit better working memory skills mediated by diagrams (as the cyclical approach indicates) than novices, and these memory skills and conceptual knowledge co-evolve (Kindfield, 1994).

Anzai's (1991) suggests repeated generation of ERs as an intervention to improve RC. In a repeated physics problem solving experiment, where a student solved the a set of problems many times, and drew diagrams for each problem every time she solved that problem, new inference strategies were learned during the many iterations over time. The structure and quality of diagrams, and the relevance of the different components and elements in relation to the solution, increased dramatically over time, indicating that the student learned to make better transformations of the problem statements into sketches, diagrams and finally abstract free body diagrams (Anzai, 1991). Izsák (2011) calls this process of repeated representation generation and problem solving *adaptive interpretation,* which involves cycles of ER generation and self evaluation. ERs are generated first to interpret problems, then to solve problems; then ERs are generated again, and they are evaluated; this process continues until one gets a grip on the problem.

The repeated generation of sketches and ERs is believed to augment thinking and generation of ideas relevant to the process (Purcell & Gero, 1998). During engineering design, ERs such as graphics and sketches actively bring together an agent's explicit conceptual knowledge, cognitive experiences (Herbert, 1988) as well as implicit understanding of system behavior. ERs can be easily and flexibly manipulated according to the needs of the design problem. Students who do sketching during problem solving are more likely to formulate the problem precisely, meet relatively more problem constraints, and also produce quality designs solutions, indicating that sketching helps in the overall design process. Also, different representations, such as problem statements, diagrams, equations and verbal descriptions, all serve different purposes to students, depending on the progress towards the solution (Cardella et al., 2006). The finding about sketching, particularly how it helps in meeting problem constraints, hints that sketching facilitates the dynamics of the design process, and also understanding the dynamics of the product, by providing an external memory while thinking and designing, and ultimately lowering the cognitive load (Cardella et al., 2006).

Purcell & Gero's (1998) coherent and detailed review of various empirical and theoretical accounts of sketches in engineering design argues that generation of representations during the design process facilitates reinterpretation of the design itself, and this eventually leads to the emergence of new ways of *seeing* into the design. For instance, using a design process analysis of participants from mechanical, instructional and architecture design, Goel (1995) showed that the structure of their sketches improved as the design process progressed. The designers gradually added precise details to an initial vague sketch. However, during the process, the designers often try out different design ideas in the sketches (lateral transformations), one (single/set) of which is then narrowed down and fixed as a theme to which details are added (vertical transformation). These processes of lateral and vertical transformations can be seen as a result of reinterpretation (Purcell & Gero, 1998; also designated as 'restructuring' by Cardella et al., 2006). Sketching facilitates reinterpretation by creating a perceptual space, and subsequently a conceptual space (Herbert, 1988), of many relevant ideas. The best ones are then chosen for refining.

Reinterpretation through the generation of ERs is also reported by Aurigemma et al. (2013), who observed a bioengineering researcher designing a 'Lab-on-a-chip' (LoC) device in an integrated systems biology lab. The authors, motivated by the distributed and situated cognition frameworks, report that the design activities were driven by generation of ERs, going back and forth between the ERs and the prototypes, and modifying both ERs and the prototype iteratively as a result of constant reinterpretation, in order to arrive at a well-functioning version of the prototype. The participant iteratively went through various drawings of the device, inscribed (numerals and calculations) on the drawings, generated her own drawings, imagined the structure-function relationships of the various components of the device, and tried to map them onto the requirements (often numerical in nature). She also used modeling and simulation software such as COMSOL and MATLAB to explore different design possibilities and constraints, and integrated the results with the prototype. The models output numerical data, which the student integrated with her drawings, her imagined functioning of the design, and the actual test results she got from the physical prototype. Much of the cognitive activity of the student, thus, involved inferring dynamic information (in this case, the flow of liquid through the LoC device) from multiple representations (such as test results, drawings, and numerical data) that were all static in nature.

2.3.3 Summary

This section reviewed empirical studies that investigated the ways in which students and experts established links between ERs and translated between them, and the patterns of ER generation as well as preferences of students and experts. Figure 2.11 below presents a summary.

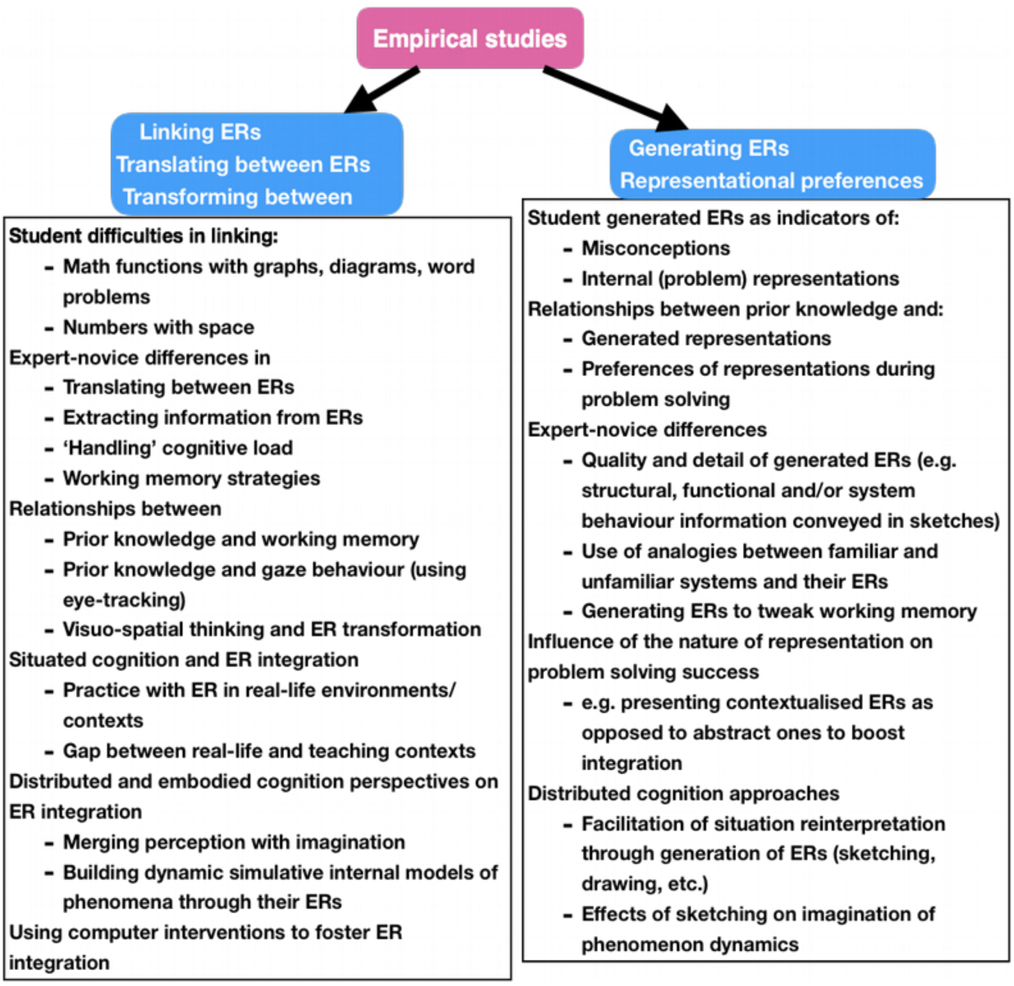


Figure 2.11 Overview of the empirical studies reviewed.

I discussed how studies (from science, mathematics and engineering) reviewed under each theme assumed the classical information processing model, by highlighting classical cognition ideas such as working memory load, information storage, information extraction and translation. I also discussed studies that take a relatively neutral stance on the nature of ERs and RC, as well as studies that subscribe to recent cognition theories such as distributed and/or embodied cognition.

In the following section, I bring together the major themes revealed by the extensive literature review presented thus far.

2.4 Findings from the review

Below are the four major findings from the review of theoretical models and empirical studies of RC.

2.4.1 Ambiguous use of the term 'representation'

The term 'representation' is used often in science education literature. The review revealed that it is used in an ambiguous way, referring to internal representations, external representations, or both. Some notable exceptions include: (a) problem solving studies in physics education research, where the term 'problem representations' refers to problem solver's internal representations of the problem presented (e. g. Chi et al., 1981), and (b) several studies explicitly using the term ERs or external representations (e. g. Mammino, 2008; Nakhleh & Postek, 2008), particularly studies employing distributed cognition frameworks (e. g. Pande & Chandrasekharan, 2014; Aurigemma et al., 2013). Since the distinction between internal and external representations is not usually made, the problem of how the external and internal representations interact is rarely examined, particularly how they interact to raise/lower cognitive load, and support the imagination.

One way to think about representations in science, mathematics and engineering is to consider equations, graphs, etc. as external manifestations of experts' internal models. These external representations augment cognition by offloading memory/ processing as well as providing novel ways of combining elements (Aurigemma et al., 2013; Kirsh, 2010). Another approach would be to consider these external media as providing starting points for the learner to develop rich internal representations and their manipulations. A third approach would be to consider the external and internal representations as being coupled, and constantly interacting with each other (Chandrasekharan, 2014). Since the literature does not make the distinction between internal and external representations, these possibilities are not examined.

2.4.2 Different nature of ERs, and RC across disciplines

Many ERs share structural commonalities across disciplines because of the intertwined nature of these disciplines. For instance, ERs in mathematics (such as equations) appear in physics, chemistry, engineering and even biology in various forms. However, there exist subtle discipline-dependent differences between these ERs and their affordances. Tables 1-6 below present a comparison of ERs in chemistry, biology, physics, mathematics and engineering across certain comparison criteria: examples of discipline-specific problems (table 1), nature of ERs[footnote1] (table 2), general learning difficulties and their nature (table 3), widely used research methods (table 4), important theoretical frameworks (table 5), and major interventions (table 6). Each table has disciplines in the first column and criterion and specific items arranged in the second column. The disciplines (first column) have been arranged sequentially across rows starting with chemistry, followed by biology, physics, mathematics and engineering on the basis of (observed) increasing complexity in the nature of ERs and the RC problem.

Table 2.1 Trends in the literature on examples of problems pertaining to ERs and RC.

Discipline	Examples of problems
Chemistry	1. Balancing chemical equations, 2. Plotting concentration graphs, 3. Imagining reaction mechanisms, 4. Imagining chemical equilibrium, 5. Representing chemical equilibrium

Biology	1. Understanding biological phenomena at multiple levels of organization (molecular, cellular, tissue, organ, organ system, organism, community, ecology and evolution), 2. Correspondence between levels (e.g. genotype/micro with phenotype/macro)
Physics	1. Producing problem-situation representations, 2. Producing mathematical models of physical phenomena/entities*
Mathematics	1. Relating concepts of number, mathematical operations, & fractions to their ERs (digits, '+', '-' signs, decimals, etc.), 2. Implicit understanding of reasoning underlying symbol systems & symbolic operations needed for working with the mathematical representation
Engineering	1. Problems in designing, building devices, 2. Developing processes and systems, 3. Creating scale models, endurance-performance tests, simulations, 4. Relating engineering practice to ERs
*Mistakes in mathematical representation are not considered, as most of the literature focuses on 'physics reasoning', and 'half-way' representations that help in the process of formalization. Such representations are needed before the final mathematical equations.	

Table 2.1 captures how the disciplines differ in the nature of problems they deal with in relation to ERs, although many problems in a discipline may be interdependent and/or tightly intertwined with those in the other. For instance, relating the concept of number, mathematical operations performed on numbers and fractions to their ERs (problem in mathematics) is fundamentally linked to balancing chemical equations (chemistry).

Table 2.2 Trends in the examples of ERs and their nature.

Discipline	ERs and their nature
Chemistry	1. Periodic table, 2. Chemical equations, 3. Concentration-energy graphs, 4. Molecular diagrams, 5. Observable properties, 6. Animations & simulations **(Well defined, convention/rule-based, constrained – i.e. not very flexible, little scope for generation)**
Biology	1. Biochemical pathways, 2. Structures of biomolecules, 3. Phylogenetic trees, 4. Computational models of complex systems **(Well defined, rule-based, inclusive of but more diverse than chemistry ERs)**

Physics	1. Problem statements, 2. Problem situation, 3. Sketches & diagrams, 4. Mathematical equations, 5. Simulations **(Less defined, more customizable & less constrained, provides space for free ER generation)**
Mathematics	1. Digits, 2. Mathematical operations/procedures, 3. Symbols, 4. Equations, 5. Functions, 6. Charts, 7. Diagrams **(Well defined but more complicated ER system; allows representing a concept entirely using different representations)**
Engineering	1. Text, 2. Materials, 3. Inscriptions, sketches & drawings, 4. Mathematical formulae, equations & functions, 5. Prototypes & physical models **(highly open ended, more complicated than the previous cases, use ERs from multiple disciplines)**

As can be noticed in table 2.2, ERs in chemistry are more defined and constrained in nature than those in other disciplines. For instance, there are certain conventions that guide the denoting of chemical elements, compounds and other substances, writing chemical equations, plotting graphs, and drawing atomic/molecular diagrams. The periodic table is a well defined, conventionalized and compact representation of chemical elements, and their properties, and is fundamental to chemistry. Given these conventions, one has very little scope to freely generate ERs and/or alter standard chemical representations while learning/ doing chemistry. There are thus limitations to the manner of using ERs in chemistry. ERs in biology are more diverse than those in chemistry. Biology inherits certain representational systems (ERs) from chemistry, for example, chemical/biochemical equations, graphs. Phylogenetic trees in biology, like chemical equations, are ERs that are strictly conventionalized. However, macro-level biological diagrams and descriptions are quite flexible for customized usage. ERs in mathematics are highly conventionalized and rule-based. But unlike chemistry, a single concept in mathematics (such as a ‘number’; figure 2.12) can be represented in multiple ways.

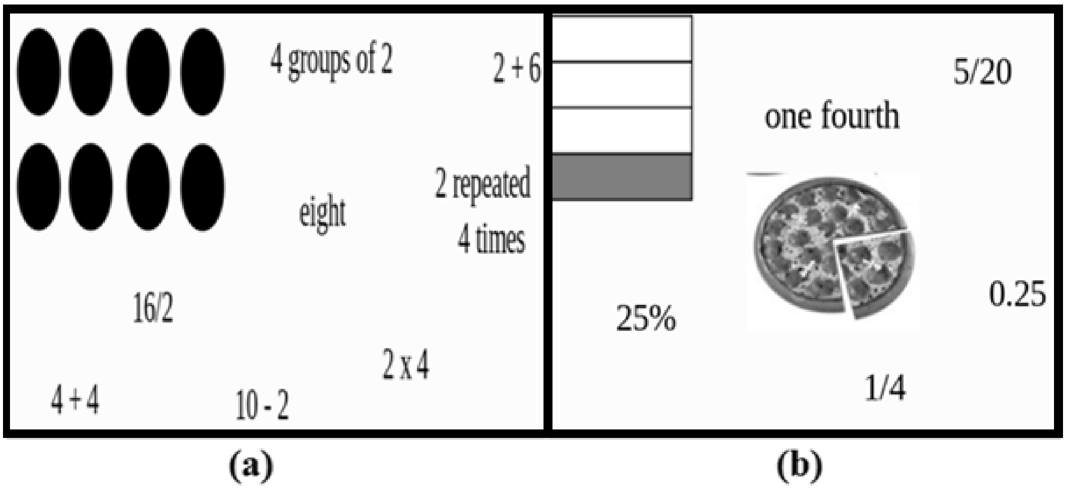


Figure 2.12 Multiple representations in mathematics. (a) Numeral '8' can be represented in various forms. For instance, '2 + 6' and '8' are both in symbolic form, but are different kinds of representations for the same concept. (b) Fraction ¼ can be represented in different ways. This makes RC in mathematics more complex to define and characterize. Lesh, Post and Behr (1987) suggest two categories of such representations; opaque and transparent. For instance, in (a), the representation 'eight' and 'four black ovals' convey the exact amount of information about the entity they represent (here a number). However, consider the representation '2x4' which leads to the number/quantity '8', but in addition conveys that the '8' is a multiple of 2 (or 4). This information is not conveyed by the earlier two representations. In this case, the earlier two representations are transparent, whereas the latter is an example of opaque representation (Zazkis and Gadowsky, 2001). Students often ignore/overlook the transparent or opaque nature of representations when they are asked questions based on these representations.

This makes usage of ERs in mathematics more flexible for the learner or doer. Physicists employ mathematical ERs and representational systems in solving physics problems. Use of diagrams in physics is conventionalized, but the learner has enough space to generate diagrams in her own way; she can scribble and represent situations under study in multiple ways and perspectives. Engineering borrows ERs from many of these disciplines, and from areas other than the core scientific domains, such as social sciences, humanities, economics, etc. There is ample space for engineers to freely generate and play with ERs, prototypes and models.

It is extremely difficult to use one kind of external representation to capture every detail (feature) of the entity or phenomenon it represents. Naturally this means multiple external representations (ERs) exist to meet this difficulty. In addition, each representation facilitates different perspectives towards entities and phenomena, as well as different affordances or action possibilities (both implicit and explicit). The exposure to multiple points of views and affordances enriches one's experiences around what is being represented, ultimately improving conceptual understanding. ERs are complementary to each other in terms of the information they convey (Ainsworth, 1999 & 2008; Kelly & Jones, 2008; Kozma & Russell, 1997; Mayer, 2005; Stieff & McCombs, 2006; Tsui & Treagust, 2003; Wilensky, 1999). On the other hand, the fact that concepts related to scientific phenomena and objects can be represented in multiple ways implies that these ideas are distributed across multiple representations. This means that the aspects of RC, particularly interconnecting information distributed across ERs, explaining the relationships between them, and mapping features of one type of representation onto those of another are different across disciplines.

2.4.3 Integration of ERs: A general cognitive difficulty

Despite the differences in the nature of ERs between all the above disciplines, it is evident from in table 2.3 that the RC problem is constituted by the following learning difficulties common to all the disciplines: visualization of/through ERs, generation of ERs to represent entities and phenomena, visualizing and understanding entities and phenomena from ERs, interrelating information from ERs, and transforming between ERs.

Table 2.3 Similarities and differences between the disciplines in the nature of general learning difficulties.

Discipline	Learning difficulties & their nature
Chemistry	1. Visualization, 2. Interconnection between ERs, 3. Representational transformation, 4. Transformation between static-dynamic ERs, 5. Conceptual integration across ERs, 6. Representation abstraction
Biology	1. Understanding levels of organization, 2. Visualization, 3. Transition between ERs, 4. Representational transformation, 5. Conceptual integration across ERs
Physics	1. Visualization, 2. Transformation between static-dynamic ERs, 3. Generation of ERs 4. Transformation between mathematical & real world physical ERs, 5. Representation abstraction

Mathematics	1. Transformation between spatial (e.g. area/volume) & numerical (e.g. units/numbers) ERs, 2. Generating ERs, equations, 3. Thinking in equations/functions, 4. Comprehending problem representation-situation, 6. Representation abstraction
Engineering	1. Visualization, 2. Transformation between static-dynamic ERs, 3. Generation of ERs, 4. Modeling

Thus, difficulty in these operations underlies mastering ERs in a given discipline, and this difficulty leads to many different learning problems in that discipline. Supporting this view, in a specific knowledge domain, processing and understanding of ERs, and the ability to fluidly generate and use ERs in an integrated fashion (for conceptualization, discovery and communication), are indicative of expertise in that domain. This suggests that integration of ERs (RC) is a general cognitive difficulty.

2.4.4 Focus on classical information processing model of cognition

Commonalities between the disciplines observed in tables 2.3 through 2.6, as well as sections 2.1 and 2.2 in the review show that most theoretical accounts of RC, as well as empirical studies and interventions across the domains, have been either explicitly or implicitly informed by classical information processing models of cognition (Ainsworth, 1999 & 2008; Johnstone, 1982; Wilensky, 1999).

Table 2.4 Trends in the nature of widely employed research methods across the disciplines.

Discipline	**Research methods**
Chemistry	1. Problem posing/solving, 2. Microgenetic, 3. Ethnography, 4. Expert-novice, 5. Prior knowledge & RC correlation, 6. Interface testing, 7. Eye-tracking
Biology	1. Prior knowledge & RC correlation, 2. Expert-novice, 3. Eye-tracking 4. Interface testing
Physics	1. Expert-novice, 2. ER generation & analysis, 3. Problem solving case- studies, 4. Microgenetic, 5. Design-based research, 6. Interface testing, 7. Eye-tracking
Mathematics	1. Prior knowledge & RC correlation 2. Expert-novice, 3. Problem posing/solving, 4. ER generation
Engineering	1. Ethnography, 2. Design and problem solving case-studies, 3. Design-based research, 4. Interface testing

Table 2.5 Important theoretical frameworks of RC and learning with ERs across the disciplines.

Discipline	Important theoretical frameworks
Chemistry	1. Johnstone's model of three thinking levels and working memory, 2. Wu et al.'s ER comprehension model, 3.Abstractness of representations, 4. Distributed and situated cognition
Biology	1. Multiple levels of organization, 2. Cube model, 3. CRM mode
Physics	1. Expert-novice qualitative differences (information processing model), 2. Meta-representational/native competence, 3. Abstractness of representations
Mathematics	1. Lesh Translation Model, 2. Duval's levels of RC, 3. Abstractness of representations
Engineering	1. Representational chain model, 2. Lesh translation model, 3. Situated & distributed cognition approaches

Table 2.6 Notable trends in the major interventions across the disciplines to address the problem of RC.

Discipline	Major interventions
Chemistry	1. Computer visualization tools (visChem, 4M:Chem), 2. Computer simulations, 3. Problem-based Curricula, 4. Conceptual change model, 5. Laboratory integration, 6. Sequential ER introduction
Biology	1. Computer visualization tools (evolution animations), 2. Computer simulations (Netlogo models), 3. Problem-based Curricula, 4. Laboratory integration
Physics	1. Computer simulations (PhET, Netlogo), 2. Problem-context-based learning, 3. Computer visualization, 4. Virtual laboratory
Mathematics	1. Computer simulations (GeoGebra, Netlogo), 2. Problem-context-based learning, 3. Virtual/physical manipulatives
Engineering	1. Computer visualization & simulations, 2. Model eliciting activities, 3. Design & technology activities, 4. Problem-based teaching-learning, 4. STEM integration

Chart 1 below presents categories of theoretical models and empirical studies based on their subscription to major theories of cognition.

Three main assumptions can be isolated from the review of such models and theories. These are usually also identified with the classical information processing approaches.

(a) the mind extracts information from ERs, which acts as 'vehicles', or transmission media, for the information,

(b) ERs and the concepts they represent are linked through some form of information 'translation', and (c) the translation is mediated mostly through mental capacities such as imagery and/or amodal symbolic forms, as well as working memory (e. g. Johnstone, 1982; Gooding, 2006; Tsui & Treagust, 2013; etc.)

These assumptions, particularly limited working memory capacity as a central processing bottleneck, can be seen as influencing many intervention designs. For instance, ER visualization software used in chemistry, interactive computer simulations, and virtual laboratories, are designed to address working memory limitations. Ironically, the software interventions do not seek to augment the student's working memory and processing abilities, but only help offload some of the memory and processing load to the computer screen. Possibly because of this, such interventions have not been very successful in promoting RC (De Jong & van Joolingen, 1998; Rutten, van Joolingen & van der Veen, 2012). Further, by focusing on the "processor capacity", as well as the inaccessible nature of the information extraction and translation processes, these models and interventions make RC appear mysterious. They do not focus on the cognitive as well as practice elements that could lead to RC development (For instance, how and why are certain interventions effective in the development of RC? What role does practice play in the RC development process, apart from enhancing working memory load abilities? What is the nature of interaction between internal and external representations? What is the role played by interactivity in simulations and other software?).

Different from such load and translation accounts, a third major chunk of models and studies take a relatively neutral stance on the nature of ERs and RC, but these approaches do not seek to generalize, and provide detailed accounts of the cognitive processes involved in ER integration. A final set of models and studies subscribes to recent cognition theories such as distributed and/or embodied cognition.; However, they fail to provide a general framework for ER integration. Without such a general account, it is difficult to develop focused (educational) technology interventions that support RC, particularly interventions that take into account the differences in ERs across different disciplines. I propose such an account in the next chapter.

Chart 1 Categorization of theoretical models and empirical studies based on their subscription to general cognitive theories

Chapter 3: Towards a distributed and embodied cognition account of ER integration

The extensive literature review presented in the previous chapter reveals that current work on ER integration and RC across the different disciplines fails to provide a general framework of ER integration, without which, it is difficult to develop focused educational technology interventions that support the development of ER integration. Here[ii] I propose a model of the cognitive processes involved in a generic ER integration problem, and then outline a generic account of the cognitive mechanisms underlying these processes, using perspectives from recent cognitive theories, particularly distributed and embodied cognition.

The model and the account I propose make explicit their departure from classical cognitivist assumptions identified in the review in following ways:

- Firstly, I emphasize the distinction between internal and external representations, considering the two as dynamically coupled, through constant interactions between the learner and external representations. My focus is on how different external representations are integrated. But since this integration process is closely coupled with the formation of an internal model of the domain, the model also considers integration of ERs and internal models.
- I focus on the way the cognitive system *interacts* with ERs, as opposed to the view: that all ERs embed information; that this abstract information is isolated from the external structure and pulled inside by the cognitive system (somehow); and that cognition arises from the manipulation of this information inside the head. Our account is thus inspired by recent 'field' theories of cognition, particularly the idea of 'constitutivity', which treats external symbols as part of cognition. The external operations on ERs, and the sensorimotor processes involved in these operations, are part of cognizing the concepts instantiated by the ER (Landy et al., 2014).

Since the focus of the new account is to help individual learners integrate ERs, I will be taking an individual-focused approach to distributed cognition, which Hutchins has recently termed 'extended cognition' (Hutchins, 2014). He distinguishes it from traditional DC, which he considers a system-level theory. Similarly, since representation (internal and external) is the focus of the account, I do not consider radical embodied cognition accounts that reject internal representation (such as ecological psychology and dynamic systems theory), which consider sensorimotor interaction with external entities *both necessary and sufficient* for cognition. I accept the argument that sensorimotor interaction is necessary, and develop a framework that is based on a coupling between sensorimotor interaction and representation, using the common coding model, which is a representationalist position within embodied cognition (Chandrasekharan & Osbeck, 2010). The account I develop thus includes the key tenets of DC (cognition as a process distributed across people and artifacts, interaction between internal and external representations), and embodied cognition (enactive and modal internal models, participatory relationship with external entities). The following are the central theoretical assumptions of the account I propose:

1) Internal representations have a model-like structure (mental models), and they can run independent of external representations, to provide knowledge. This process provides capabilities different from the processing of external representations.

2) Internal models have an enactive/simulative nature (Nersessian, 2010; Chandrasekharan, 2009), and they are dynamic, with a neural network like structure.

3) Internal models interact with external ones, and they are built and extended through this interaction process. This interaction augments cognition, and it provides capabilities different from the offline processing above.

4) The enactive nature of internal models is the key feature that enables the processing of dynamics, which is the central explanatory process in science and engineering.

5) The integration of ERs is based on dynamics, and the mental simulation of dynamics. The motor system is the key player in the simulation of dynamics (Schubotz, 2007).

6) The motor system is the central mediator in ER integration, as it is the major integrating system in the body (actions require integrating perception, proprioception, muscles, balance etc. in complex ways), and this integration capability is reused in ER integration. This explains why the enaction/interaction features provided by new-media technologies help in understanding and learning science and engineering (and make discoveries possible using new interactive simulation systems such as *FoldIt*), and also why integration of ERs based on static media is harder. This view also explains why activity-based classroom interventions facilitate the integration process.

Before discussing how distributed and embodied approaches to cognition could be extended to develop an account of the ER integration problem, I characterize the cognitive processes involved in a generic integration problem by developing a model. This generic model – termed the TUF model -- can be used to examine the different cognitive frameworks, to see which provides a better understanding of this problem.

3.1 The TUF model: capturing the general cognitive processes involved in ER integration

The generic case of integration of ERs in science and engineering involves the observed (or described) actual dynamic behavior of a physical system (such as a falling object, a pendulum or a chemical process), an equation capturing the behavior, and graphs that display the equation's output for some sets of values. The transition to the equation is often mediated by geometric structures, such as free-body diagrams, and there may be other structural representations involved, such a molecular models. Broadly though, the learner needs to develop an integrated internal representation of the three modes – the phenomenon, its equation and the graphs. If structural representations are present, the integration process has to deal with one more level of complexity. An indicator of integration is the ability to transform smoothly between the three modes. This transformation is difficult, because it requires shifting between spatial and numerical modes (e. g. graph and equation), as well as dynamic and static modes (e. g. phenomenon and equation). Even the spatial to numerical transformation requires understanding dynamics, as the students need to understand how the values in the equation get translated into a graph, which requires thinking of various values of equations and 'movements' of the graph based on these values. Thus, to integrate the ERs, the student needs to "unfreeze" the static representations, by generating their dynamic behavior in imagination, and then connect these dynamics with the dynamic behavior of the phenomenon. In the other direction, students also need to be able to "freeze" the behavior of real-world systems into equations, so that limit cases and other variations can be explored and combined. I call this model of the cognitive processes involved in ER integration as the TUF model – transform-unfreeze-freeze model (figure 3.1).

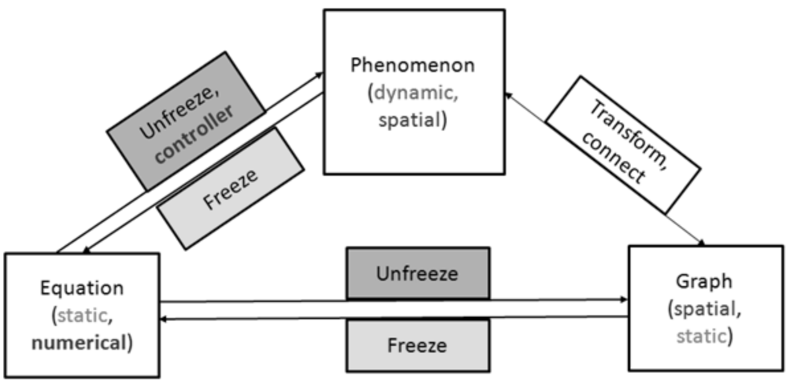


Figure 3.1: The TUF (transform-unfreeze-freeze) model depicting the processes involved in ER integration.

3.2 A DC and EC-based account of the cognitive mechanisms underlying the TUF model

The generic structure presented in figure 3.1 above suggests that a mechanism account of the cognitive processes involved in ER integration would need to address two important questions:

- *how external representations connect with imagination*
- *how dynamic behavior could be imagined from static external representations*

The first question is related to the new view that ERs are thinking and learning devices, that the process of interacting with ERs augments cognition, and that this interaction is a central component in forming internal models of imperceptible phenomena. An account of the interaction process, and its role in imagination (forming internal models), is needed to understand ER integration.

Next, the central component of models in science and engineering is dynamics, and the integration of ERs requires (and happens through) understanding of dynamics, particularly the way it is captured by ERs. The second question concerns how this dynamics is processed by the cognitive system, specifically how it is generated from ERs (which are mostly static), and how interactivity contributes to the understanding of dynamics.

Once we have an understanding of these processes, we would be able to design interventions, particularly new-media technology interventions that allow learners to quickly integrate ERs. Answering these two questions is not easy, as it requires bringing together complex literatures that cut across many areas of cognitive science. Answering the first question requires understanding how external representations are processed by the cognitive system. In our view, this question is best addressed within the distributed cognition (DC) framework (Hutchins, 1995a; Hutchins 1995b), which was developed to study cognitive processes in complex (usually technical and scientific) task environments, particularly environments where external representations and other cognitive artifacts are used by groups of people. The DC approach was first outlined by Cole and Engestrom (1993), Pea (1993), and Salomon (1993), and apart from the currently dominant model presented by Hutchins (1995a, 1995b), significant contributions to the initial framework were made by Cox (1999), Hollan, Hutchins and Kirsh (2000), and Kirsh (1996, 2001, 2010). Most work in DC is focused on understanding how internal and external representations work together to create and help coordinate complex socio-technical systems. The primary unit of analysis in DC is a distributed socio-technical system, consisting of people working together (or individually) to accomplish a task and the artifacts they use in the process. The people and artifacts are described, respectively, as agents and nodes. Behavior is considered to result from the interaction between external and internal representational structures.

The canonical example of external representational structures in DC is the use of speed bugs in a cockpit (Hutchins, 1995a). Speed bugs are physical tabs that can be moved over the airspeed indicator to mark critical settings for a particular flight. When landing an aircraft, pilots have to adjust the speed at which they lose altitude, based on the weight of the aircraft during landing for that particular flight. Before the origin of the bugs, this calculation was done by pilots while doing the landing operation, using a chart and calculations in memory. With the bugs, once these markers are set between two critical speed values (based on the weight of the aircraft for a particular flight), instead of doing a numerical comparison of the current airspeed and wing configuration with critical speeds stored in memory or a chart, pilots simply glance at the dial to see where the speed-indicating needle is in relation to the bug position (figure 3.2).

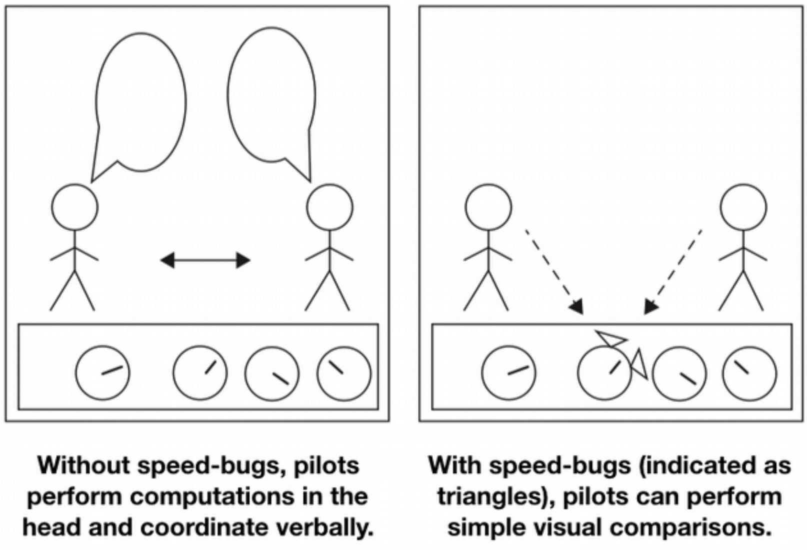


Figure 3.2 The cockpit environment as a DC system (adapted from Chandrasekharan & Tovey, 2012).

This external representation allows pilots to 'read off' the current speed in relation to permissible speeds using perception. They can then calibrate their actions in response to the perceived speed difference. The speed bugs (an external artifact) thus lower the pilot's cognitive load at a critical time period (landing), by cutting down on calculations and replacing these complex cognitive operations with a perceptual operation. The setting of the speed bugs also leads to a public structure, which is shared by everyone in the cockpit. This results in the coordination of expectations and actions between the pilots. These two roles of the speed bug (lowering cognitive load and promoting coordination between pilots) are difficult to understand without considering the human and the artifact as forming a distributed cognitive system.

This account focuses on memory offloading, but it has been extended in two ways: 1) to show how processing, particularly mental rotation, is lowered using external manipulations that serve as 'epistemic actions' (Kirsh, 2010; Kirsh & Maglio, 1994) and 2) how imagination is augmented by active manipulation, particularly in computational models (Chandrasekharan & Nersessian, 2015; Chandrasekharan, 2014; Marshal, 2007). These studies, and other similar ones showing how external representations are used to generate action patterns (Bodemer et al., 2004; Martin & Schwartz, 2005) suggest that the brain 'incorporates' the external representations (Chandrasekharan, 2014; Rahaman et al., 2018) as part of the imagination system. This incorporation process is considered to be driven by actions/manipulations done on the representations, and the exploration of many states of the representations. This incorporation view is different from the classical information processing view, where the information encoded in the representation is extracted by the cognitive system, and all cognitive operations are internal operations done on this extracted information. The new approach suggests that actions and manipulations on ERs lead to the ERs getting incorporated -- becoming part of the cognitive system. In this view, it would be possible to improve the process of integration (of the imagination and the external representation), by restructuring the latter to support actions and manipulations, say by using new-media technology approaches, or classroom interventions based on inquiry and activities (Lehrer & Schauble, 2006; Tytler, Prain, Hubber & Waldrip, 2013). Such an approach to developing RC would be quite different from the approach based on cognitive load, as the incorporation approach tries to support the integration process directly using manipulations and feedback, rather than through simultaneous presentation of ERs to lower cognitive load.

The above account provides a rudimentary 'incorporation' model of how external representations connect with imagination (see Chandrasekharan, 2014; 2009; Rahaman et al, 2018, for details), and brings us to the second question: *How is dynamics generated from static representations?* Embodied cognition research argues that the brain and all cognitive processes developed for action, and the body and the motor system are therefore closely involved in most cognitive operations. Supporting this theoretical view, there is evidence that the motor system is used while generating dynamic information from static images (such as system drawings, see Hegarty, 2004) and vice versa. Common instances of this generation include: judging the sense of speed of a vehicle from its tire-marks (or judging tire-marks given speed), judging the sense of force from impact marks (or judging impact marks, given force), sense of movement speed from photos of action (say soccer), sense of movement derived from drawings, cartoons, sculptures, etc. Experimental evidence for the use of the motor system in this process comes from the work on the Two-Thirds Power Law for end-point movements such as drawings and writings. The law relates the curvature of a drawing trajectory with the tangential velocity of the movement that created the drawing/writing. The human visual system deals more effectively with stimuli that follow this law than with stimuli that do not. When the curvature-velocity relationship does not comply with the power law, participants misjudge the geometric and kinematic properties of dynamic two-dimensional point-displays (Viviani & Stucchi 1989; 1992). Also, the accuracy of visuo-manual and oculomotor 2D tracking depends on the extent to which the target's movement complies with the power law. This relation allows humans to judge the speed in which something was drawn, using curvature information, and vice versa (judge curvature given speed). This capacity is presumably what we use when we judge speed from tire marks, and also evaluate drawings and paintings. Recent experimental evidence shows that observers simulate the drawing actions of a painter while observing paintings (Taylor, Witt & Grimaldi, 2012). There is also evidence that object-related hand actions are evoked while processing written text (Bub & Masson, 2012).

Such predictions can also work the other way, where given a dynamic trace, we can imagine and predict the static sample that comes next. In one experiment, dynamic traces of handwriting samples were shown to participants. They were then shown some samples of written letters (such as *l*, *h* etc.), and asked to judge which letter came next to the shown trace. Participants could identify the letter following the trace more accurately (Kandel, Orliaguet & Viviani, 2000) when the trace followed the Two-Thirds power law, i. e. the angular momentum of writing was related to curvature in a way laid out by the law. Accuracy went down significantly for traces that did not follow this relation. Based on this and other experiments, Viviani (2002) argues that "in formulating velocity judgements, humans have access to some implicit knowledge of the motor rule expressed by the Two-thirds

Power Law". Much of the experimental evidence in this domain is about the replication of biological movements from static images, but everyday experience (such as the tire mark case) suggests that non-biological movements can also be replicated, and it is highly likely that this process also is based on motor system activation (Chandrasekharan, 2014; Schubotz, 2007).

Box 7: TUF model summary

The TUF account proposes that to integrate the ERs, a student needs to

1. "**unfreeze**" the static representations, by generating their dynamic behaviour in imagination
2. **transform** between phenomena and cues of dynamic behaviour present in the static graphs
3. "**freeze**" the imagined (and perceived) behaviour of real-world systems into equations, so that limit cases and other variations can be explored and combined.

All these processes are mediated by the sensorimotor system.
The cognitive/neural mechanisms involved are: (a) 'incorporation' of the ER, and (b) mental simulations based on covert activation of the sensorimotor system.

This account suggests that the motor system needs to be activated to start the "unfreezing" of ERs, to generate dynamic content using the static representation. It is possible that this activation process is difficult to do for novices, and enactive computer interventions that allow manipulations on the ERs could help trigger this activation, thus setting the unfreezing process in motion. Note that this approach is different from the designs suggested by the cognitive load account, where manipulation of ER is not the central feature of the intervention. Also, this approach is in synergy with the 'incorporation' account provided by recent work in distributed cognition (Chandrasekharan, 2014; Chandrasekharan & Nersessian, 2015; Rahaman et al., 2018), as it suggests manipulation of the ERs as a way of promoting incorporation of the external representation with the imagination system. A related idea is that actions done on ERs with dynamic content would help improve integration, as the action system is involved in processing dynamics, and it is also the central integrating system in the body. This view provides an explanation for why interactivity provided by new-media technologies helps improve understanding and integration, and understanding and integration is limited with static media (Majumdar et al., 2014).

The above brief review of distributed cognition and embodied cognition approaches, and how they could together provide a general cognitive account of the ER integration problem, presents just an outline of the way these theoretical frameworks could contribute to our understanding of ER integration and RC. As of now, the two theoretical approaches only provide a way of understanding the "unfreezing" aspect of ER integration, and how external representations could be incorporated into imagination. The frameworks do not provide a clear way of understanding how dynamic processes are "frozen" into equations. Future work in these areas, particularly in close collaboration with science education research and (new media-based) educational technology development, may help provide a better understanding of this problem, and ER integration and RC in general. Design-based Research (Cobb, Confrey, diSessa, Lehrer & Schauble, 2003) provides an ideal way to bring together these disciplines to address the RC problem, as it offers a way of developing interventions that could test hypotheses about RC integration as well as cognitive processes. The work reported in this thesis is focused in this direction, focusing on new media[iii], because, unlike static media, they provide 1) the possibility of making dynamics embedded in formal notations explicit, and 2) action-based manipulation of this dynamics. The dynamics and the active manipulation of the dynamics is considered central to the integration process.

Given the involvement of sensorimotor mechanisms in the incorporation, imagination and integration processes, as well as the relationships between action-perception-imagination capacities, two interconnected conjectures, with empirical implications, emerge from this theoretical account:

1. In this model, the development of the ER integration ability would result in a reorganization of the cognitive system, particularly the sensorimotor system. This suggests the way learners perceptually access ERs would change after significant training in a domain.

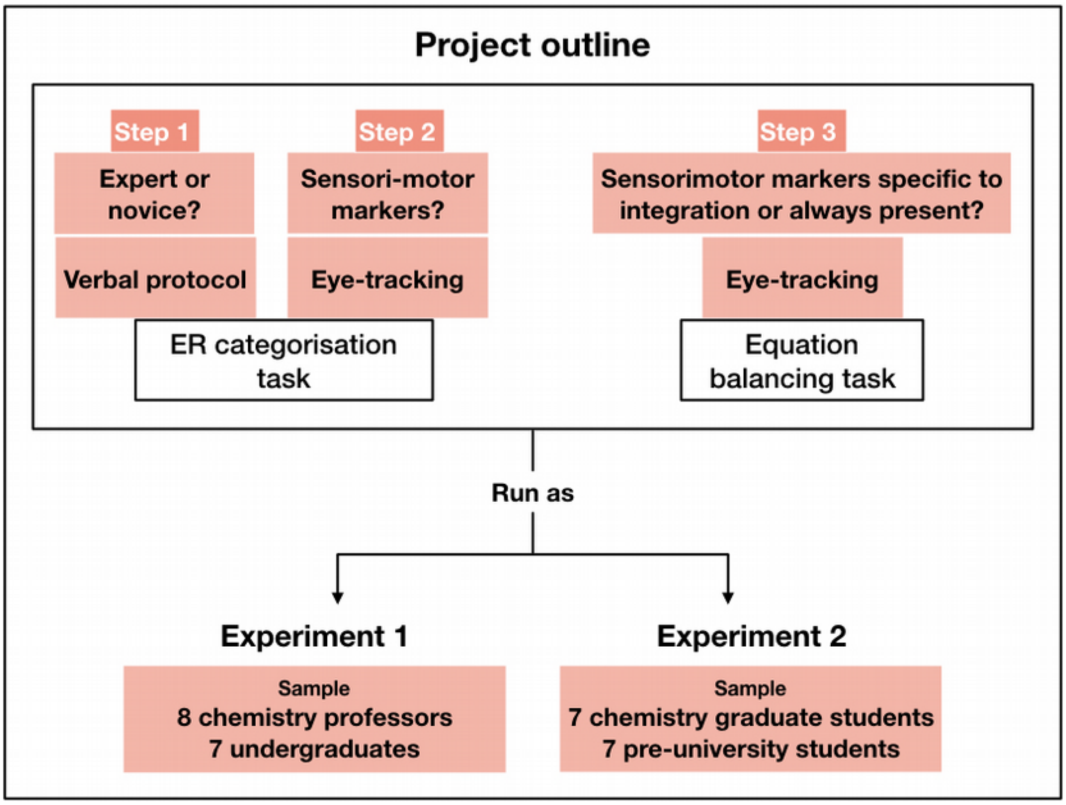
2. Interaction, particularly based on the sensorimotor system, would support ER integration and its development.

To test these predictions, and therefore our theoretical model, I conceptualized the following two empirical projects:

The first project is situated at the interface between chemistry education and cognitive science, and sought to identify possible behavioral markers that could track the sensorimotor changes accompanying the development of constitutivity and ER integration, through chemistry education. Different groups of participants with varying levels of education in chemistry performed tasks with general chemistry ERs while I captured their gaze behavior. I then searched for patterns of development in ER integration, by comparing the level of education of these participants with their ER integration abilities and patterns of behavior.

The second project sought to test the second conjecture, by answering the following question: how can the DC and EC-based theoretical account be utilized to design and build effective new media-based enactive learning environments that support ER integration through sensorimotor interaction? It focused on the design, development and testing of an interactive computer interface with fully manipulable ERs, developed as an intervention to support ER integration at middle-school level. This interface provides coupled ERs of a phenomenon to a learner, in their static as well as dynamic states. Testing the interface also contributed to further understanding of the predictions made by the first conjecture.

Chapter 4: Does ER integration based on chemistry training change the sensorimotor system?



This project sought to test the first conjecture – whether achieving ER integration (i.e. expertise) results in changes in the sensorimotor system of a learner. Here, I bring together perspectives on ER integration, perceptual learning and constitutivity, and propose that perceptual learning could emerge with constitutivity. Accordingly, markers of perceptual learning (predominantly eye-behavior), based on science training, are markers of changes in cognitive mechanisms associated with ER integration.

To test the first conjecture, I first identify expertise (ER integration) related to training. Then markers of sensorimotor behavior, particularly eye-movements, are identified in relation to the development ER integration, as a learner progressed in her training.

To understand any possible correlation between ER integration abilities and sensorimotor behavior, I executed the following two steps of analysis, using a categorization task (step 1), and eye-tracking (step 2):

- Step 1: How do participants with different levels of education in chemistry differ in ER integration abilities (or expertise)?
- Step 2: What are the sensorimotor markers associated with the development of ER integration in chemistry?

I then considered the alternate explanation that any sensorimotor changes seen are always present, and they are thus not necessarily markers of expertise. To test this possibility, I did a further step of analysis, using an equation balancing task that did not require ER integration, and tracking eye movements during this task.

- Step 3: Are these markers always present, or only triggered while solving tasks related to ER integration?

I conducted two experiments to address these questions. In order to obtain as distinct a result as possible in relation to ER integration abilities, the first experiment was conceptualized as an expert-novice investigation, involving chemistry professors (experts) and undergraduate students (novices). These participants (i) viewed and categorized a set of chemistry ERs, while their gaze behavior was recorded using an eye-tracker, and (ii) balanced a set of unbalanced chemical equations, while their gaze behavior was recorded using an eye-tracker. To understand how ER integration ability, constitutivity and the underlying cognitive mechanisms develop, I replicated the experiment with two more groups of participants, viz. pre-university students and doctoral students (experiment 2), thus making the study a cross-sectional investigation.

During both the experiments, I used a Tobii X2-60 eye-tracker to obtain fine-grained data on participant's eye-movement as they viewed the representations and also video recorded the experiment sessions to capture verbal and gesture data.

In the next sections, I discuss in detail the two experiments, and show how the differences between the participants could be accounted for by a fine-tuning of the sensorimotor system resulting from training in chemistry. While doing the above, I also demonstrate several novel methods of gaze and behavior data analysis, as well as visualization and interpretation, based on the TUF model and the theoretical account outlined in chapter 3. These methods provide important insights about expert behavior, thus taking a step towards behavioral characterization of the cognitive processes involved in ER-based constitution of concepts and integration, particularly in chemistry.

4.1 Experiment 1

The first experiment sought to characterize differences in the ER integration abilities of participants with different levels of expertise and education in chemistry (step 1 above). Sensorimotor markers were then identified and related to those abilities (step 2 above). To ensure that the ER integration abilities (and hence the sensorimotor markers) between the groups were as distinct as possible, chemistry experts and novices were studied. Chemistry professors (experts) and chemistry undergraduate students (novices) performed two tasks; an ER categorization task (adapted from Kozma & Russell, 1997) and a chemical equation balancing task.

4.1.1 Sample

8 chemistry professors (expert group, code-named FC; 4 female) and 7 chemistry undergraduate students (novice group, code-named UG; 4 female) from a leading university in mid-western India volunteered to participate in the study. Participants in the expert group were involved in research and teaching in chemistry for at least 5 years after earning their doctorates in their respective specializations. All the seven undergraduate students were in the fourth semester at the time of the study. Informed consent was obtained from all participants.

Each participant individually performed two tasks during the experiment.

4.1.2 Categorization task

This task is an adapted version of Kozma & Russell's (1997) ER categorization task, which is an ideal tool to establish differences in ER integration abilities between experts and novices (step 1). The task also provides researchers the opportunity to observe participants' interaction with ERs (step 2), as it involves participants viewing and categorizing a set of chemistry ERs into meaningful categories. The adapted version assessed the RC abilities of our experts and novices, and tested how well our sample compared with the existing expert-novice studies of RC. Once the expert/novice status was established, behavior and eye-movement data was analyzed to identify sensorimotor markers of ER integration, by correlating them with participants' performance.

Five general chemical reactions were chosen from undergraduate general chemistry textbooks, used widely in the geographical area of the study. These reactions were fairly representative of their respective reaction type, and were deemed familiar to undergraduate students in the area. The five reactions were: a simple strong acid-strong base neutralization reaction, a precipitation reaction, NO_2-N_2O_4 gas equilibrium reaction, and two complex-ion equilibrium reactions. The task involved the following four ERs corresponding to each reaction: a chemical equation, a graph (except for the precipitation reaction), a video of laboratory personnel performing the reaction in a laboratory, and a 3D molecular animation. Chemical equations and approximate graphs for each reaction were generated using an image processing software. The graphs included: a titration curve (representing the neutralization reaction), a concentration vs. solubility curve (related to one of the complex-ion equilibrium reaction), and two curves depicting relationships between temperature vs. concentration (of which one represented the NO_2-N_2O_4 equilibrium and the other was related to a complex-ion equilibrium reaction). No graph representation was used for the precipitation reaction. Free and open source demonstration videos of the five chemical reactions were procured from on-line resources. The videos were not annotated (i. e. they were devoid of any textual information regarding the substances, apparatus or procedure involved in the demonstration). For molecular animations, to control for design variation, a professional 3D animator was recruited to develop 3D molecular animations for the five chemical reactions. The animations were designed by the first author in consultation with a chemistry expert, and developed by the 3D animator. Each animation depicted approximate molecular dynamics of the corresponding reaction (e.g. displacement of molecules or ions, particulate collision, molecular aggregation, etc.), and did not have any other embedded representation, such as text, narrative, graphs or equations (see Appendix 1 for a link to a sample animation). The animations uniformly used space-filling models of atoms/molecules and followed a CPK coloring scheme consistent with Jmol (2014).

This resulted in 19 representations (Appendix 1) corresponding to the five different chemical reactions. To make these representations more convenient for handling, an image of each representation was color printed and pasted on a 3×4 inch cardboard. For animation and demonstration videos, a static picture of an important moment captured as a screen-shot was used for printing as a card.

4.1.3 Equation balancing task

This was a confirmation task to test if the sensorimotor markers are specific ER integration or not (step 3). In this task, each participant was presented with six unbalanced chemical equations (presented one after the other) of different general chemical reactions (e.g. Hinton & Nakhleh, 1999; Nurrenbern & Pickering, 1987), and was asked to perform a non integration – balancing those equations, while their eye-movements over the stimuli were captured.

This task exploits an extremely well established and popular experimental paradigm in cognitive psychology – interference. If sensorimotor changes are task-general (i.e. not specific to ER integration), the participants, particularly experts, when presented with chemical equations (stimulus similar to one of the stimuli in the ER integration or categorization task), would exhibit eye-movements (markers of sensorimotor changes) similar to those seen on equations presented during the categorization task, which required ER integration.

Note that solving this task did not require imagining the dynamics of chemical phenomena, and hence, ER integration, as balancing equations is mostly based on algorithms. This task also intended to test whether such a presentation of a representation (in a problem that does not require ER integration for successful completion) automatically triggers among participants, particularly experts, an imagination of the represented chemical process.

Six chemical equations of different general chemical reactions were randomly chosen from an undergraduate textbook for this task. Unbalanced and/or partial versions of these equations were generated and typed in appropriate font and font-size in a text-presentation program capable of full-screen slideshow (see appendix 2 for screenshot of each equation). Each slide presented one equation; the sequence of the slides/equations was predetermined, and was maintained across participants.

Two chemistry experts and one cognitive science expert collectively discussed the usability and validity of the ERs used in both the tasks, for content, conceptual and representational appropriateness. Their comments and suggestions were incorporated in the ER designs before rendering the images, movies and cards for presentation during the tasks.

4.1.4 Research questions

The conjecture from the theoretical model was operationally captured by the following specific research questions (RQs):

1. Do groups of participants with different levels of training in chemistry differ in categorizing general chemistry ERs and explaining the relationships between them? (Categorization task; step 1)

2. Do these participants differ in reasoning about the mapping between dynamic and static ERs? (Categorization task; step 1)

3. What eye-movement patterns do participants exhibit while observing the ERs? What are the between-group similarities and differences in eye-movement? (Categorization task; step 2)

4. What eye-movements do the participants exhibit while observing static unbalanced chemical equations? How do the groups differ? (Balancing task; step 3)

5. What do these patterns suggest about ER integration and RC development? (both tasks)

4.1.5 Experiment protocol

Participants performed the experiments individually. Each participant completed the balancing task first, followed by the categorization task, to avoid any possible priming effects (based on the exposure to different ERs of chemical reactions as well as the act of performing ER categorization) while perceiving unbalanced chemical equations.

Each participant sat in front of a laptop screen at a distance of 50-70 cm. The laptop was attached with Tobii X2-60 portable eye-tracker (Tobii Technologies, Sweden, sampling rate of 60Hz). The researcher sat next to the participant, (but not very close, to ensure that there was no interference in gaze data collection) and controlled the stimulus presentation using mouse and keyboard. Once both the participant and the researcher were comfortable in their positions, eye-tracker calibration procedure (5-point calibration, Tobii Technology, 2014) was completed before proceeding to the tasks.

4.1.5.1 Balancing task

On successful calibration of the eye tracker, the participant was introduced to the equation balancing task and was given appropriate instructions (see Appendix 3).

The six unbalanced equations were presented as a slideshow to the participant in a predetermined randomized sequence. The participant was asked to balance the presented equation mentally, and could take as much time as s/he needed. After the participant was ready with an answer, s/he could tell the researcher to edit the unbalanced equation appropriately using mouse and keyboard (e.g. enter 2 as a coefficient to the first reactant, add a missing product, etc.) The participant could skip solving any equation; however, s/he was not allowed to return to the skipped problem, to avoid complications in the eye-behavior data recording and processing.

4.1.5.2 Categorization task

Next, the researcher introduced the participant to the categorization task. See Appendix 3 for the instructions each participant received from the researcher before beginning the task. This task took place in two phases viz., ER observation phase, and categorization phase.

In the ER observation phase, each participant viewed the 19 representations (images and movies) on the laptop screen, presented one at a time in a predetermined randomized sequence. This presentation sequence was maintained for every participant. While viewing, the participant was handed over a printed card of the corresponding representation last viewed. There was no time limit for viewing each representation. The animation and demonstration movies could be played as many times as the participant wished. However, going back to a previously shown representation and shuffling through cards collected by the participant was not allowed, to avoid complications in the eye-movement data.

Once the participant had viewed all the 19 representations as well as had collected all the corresponding cards with her, s/he was asked to group the cards into chemically meaningful categories. This marked the beginning of categorization phase. The participant was free to implement any grouping criteria, make any number of categories, and place any number of cards into a category. There was no time limit to this phase. After categorization, the participant showed the researcher the different categories s/he made, and justified her categorization scheme in terms of relationship between the representations. The participant was then asked to perform a second trial, i.e. one more round of categorization, with a different grouping scheme. The second trial was to facilitate the emergence of non-spontaneous, well-thought or alternative schemes, if any. It was observed during a pilot study preceding this experiment that sometimes even expert participants spontaneously employed conceptually superficial grouping criteria in the first round. Moreover, some participants had explicitly expressed their wish to perform a second trial during the pilot.

Each participant took 15 minutes on average to complete the balancing task and 40-60 minutes for the categorization task.

4.1.6 Data Sources

The entire experiment was video recorded using a Sony camcorder (DCR SR40) mounted on a tripod. Eye-tracking was used to obtain fine-grained data about participants' eye-movement and gaze behavior as they observed the representations.

The main sources of data for the experiment were: (a) researcher's documentation of the participant responses during the balancing task, the categories made by the participants, and their verbal justifications during the categorization task, (b) video recordings of the balancing and categorization processes, and (c) dynamic eye-movement and gaze-behavior data, as well as screen-activity-capture during the balancing task and the ER viewing phase of the categorization task.

4.1.7 Methods of verbal data analysis (step 1)

4.1.7.1 Categorization task

The video recordings were transcribed and annotated by the author with the help of the notes collected during the study. Annotated transcripts were then coded by the author, for analysis of the nature of categories generated during both rounds of categorization. Based on the transcripts (participant's category justification), each category of representations generated was assigned to one of the five different category types using the scheme shown in table 4.1.

Table 4.1 Category coding scheme (based on Kozma and Russell, 1997).

Nature of categories	Criteria	Example
Conceptual	Chemically meaningful combinations of cards, supplemented with correct conceptual description of grouping criteria	Associations of cards depicting equilibrium reactions, or precipitation reaction
Mixed	Categories with correct or plausible combinations of cards, where some associations and/or representations are explained using chemical concepts while others are explained using visual features	A category made with, say 4 cards depicting equilibrium reaction, of which two cards are explained using the concept of equilibrium, while the other two using feature-similarity such as Δ (heat symbol) and a burner
Feature-based	Associations of cards explained purely on the basis of visual features of the representations grouped together	Associating an animation showing molecules settling down with a laboratory demonstration showing precipitation; explained in words such as, 'both settling down'
Media-based	Combinations of cards based on the medium of representation	All molecular animations as a category, all graphs as another, etc.
Inappropriate or incorrect	Incorrect or meaningless combinations of cards, categories not belonging to any of the above category types	An association between equation of a precipitation reaction and a video showing effect of temperature on a chemical equilibrium

The following data were obtained and tabulated for each participant per categorization trial, post-coding: Total number of cards used, total number of categories generated, and number of categories of each type (e. g. number of conceptual categories, number of media-based categories, and so on). For each participant, only the categorization trial in which s/he exhibited the highest performance (as indicated by the mean number of conceptual and mixed categories, in contrast to media-based or feature-based categories) was considered for group level analysis – hereafter referred to as the 'best of two'. Means and standard deviations were calculated from these data for each group. A student t-test was performed to check for significance of between-group differences.

4.1.7.2 Balancing task

This task was conducted to test if the sensorimotor changes are task-general (i.e. not specific to ER integration), and if a context-independent presentation of a representation automatically triggers among participants, particularly experts, an imagination of the represented chemical process. Given this goal, the accuracy results of participants were not considered for analysis. This was a confirmation task.

4.1.8 Methods of eye-behavior data analysis (steps 2 & 3)

4.1.8.1 Categorization task (step 2)

The unprocessed raw data collected during the viewing phase of this task are in the form of dynamic screen-activity recordings (video), and time-stamped logs of eye and mouse activity. For a detailed analysis, these data needed to be filtered and refined using Tobii Studio (gaze-data analysis package from Tobii Technology). Tobii Studio version 3.2 (Tobii Technology, 2014) was used to analyze the eye-tracking data. In the first step, separate segments of viewing data per representation (see section 5.4.2 of Tobii Studio User's Manual Version 3.4.5 for procedure details) were generated for each participant. For a total of 19 representations, the segmentation yielded 19 segments per participant. The segments from all the participants, for each representation, were compiled to generate a scene. Each scene thus contained the viewing data for all the participants for the corresponding representation (e.g. Scene-A would have all the participants' segments for representation 'A', similarly, scene-B would have all the participants' viewing data for representation B, and so on). This yielded a total of 19 scenes for the 19 different representations. The following data were generated per scene (i.e. per representation), for each participant: Total viewing duration in seconds (i.e. time spent viewing that scene or representation), mean viewing duration in seconds, total number of saccades, and mean number of saccades. A saccade is defined as a rapid movement of the eyes between two consecutive fixations (Tobii Technology, 2014; see Appendix 4 for gaze parameter definitions).

To obtain specific gaze-data, different non-overlapping Areas of Interest (AOIs) were defined and generated for each of the 19 scenes. I report data related to a total of 9 scenes corresponding to the static representations (i. e. 4 graphs and 5 equations), as the analysis of gaze-data related to the remaining 10 dynamic scenes (animations and demonstration videos) were not done, as this is beyond the scope of this dissertation due to the extreme complexity involved. Four AOIs were created for scenes corresponding to graphs: origin, X-axis, Y-axis, and curve(s) (see figure 4.1a). The scenes corresponding to chemical equations had the following AOIs: arrow, reactant-1 (R1), reactant-2 (R2), and so on, and product-1 (P1), product-2 (P2) and so on (figure 4.1b).

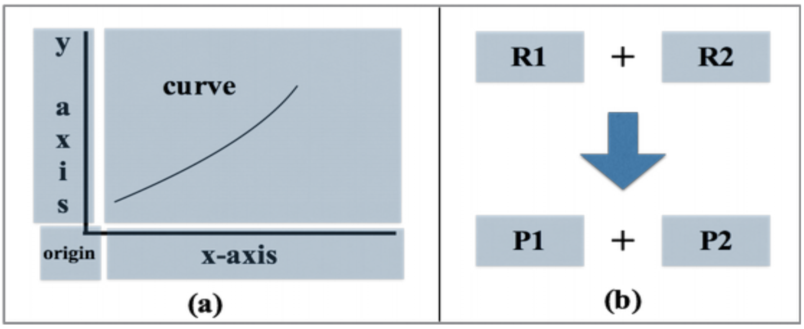


Figure 4.1 AOIs for (a) graph representations, (b) chemical equations. Each shaded box is a separate AOI. AOI shapes and sizes may differ from scene to scene.

The following data were generated per AOI per scene, for each participant: Total viewing duration in seconds, fixation index (count), fixation duration in seconds, saccade index, and AOI hit (shown as either 1 or 0; 1 denotes that activity was recorded in that AOI, whereas 0 indicates that no activity happened in that AOI). To compare the two groups of participants, the following metrics were calculated for each group using these data: mean viewing duration in seconds per scene for equations, graphs, animations and demonstration videos; mean number of saccades per scene for equations, graphs, animations and demonstration videos; mean and percent fixation count, mean and percent fixation duration.

Further, I examined the nature of gaze transitions over the static ERs, i.e. graphs and chemical equations. Gaze transitions may be identified in multiple ways, and are understood as systematic eye-movements between fixations. The nature of gaze transitions is an important marker of comparison and integration between two AOIs and/or representations, and the content they embed. For our analysis, a transition is an event between any two successive eye-fixations occurring in two different AOIs. Note that gaze transitions are not the same as saccades. As discussed earlier, saccades are the eye movements between two consecutive fixations, irrespective of AOIs. Transitions, however, are those saccades between consecutive fixations happening between two different AOIs. Consider two AOIs *x* and *y,* for instance*;* now suppose if the first few fixations happen in *x* and the next fixation(s) happen in *y*, our algorithm would register only one transition between *x* and *y*. However, the eye-tracker will register many saccades between the fixations irrespective of *x* and *y*. Hence, all transitions are saccades, but all saccades are not transitions. See Appendix 5 for details of the processing and analysis steps taken to generate transition data.

For the graph representations, I discuss transitions with respect to transition diagrams composed of different boxes and links between them. The boxes refer to the different AOIs, links between them refer to transitions and the direction of each link refers to the direction of that transition.

For chemical equations, I split the transitions into two types: long transitions and short transitions. Long transitions are defined as the gaze transitions occurring within two distantly related AOIs, whereas short transitions are the gaze transitions happening over two closely related AOIs. For instance, in figure 4.2 below, any direct transition between the two reactants (R1 and R2) or between the two products (P1 and P2) would be counted as a short transition or short jump, while transitions between a reactant (say R2) and a product (say P1) would be counted as long transitions or long jumps.

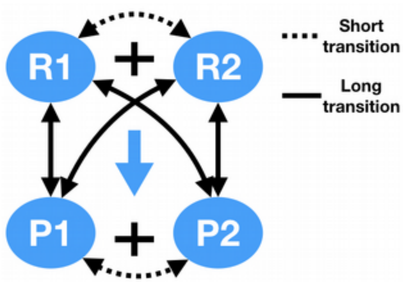


Figure 4.2 Long and short jumps between AOIs of a chemical equation.

Note that 'short' and 'long' refer to the conceptual relation between the AOIs and not the spatial relation between the AOIs in the representation. Therefore, transitions between R2 and P1 are treated as long transitions, irrespective of how closely they are situated in space. This is because they are on two different sides of a balanced chemical equation/process. In short, long transitions signify transitions between reactants and products, whereas short transitions signify transitions within either reactants or products.

Two unique overall indicators of the specific gaze activity, inertia and volatility, were defined and calculated using the transition data as follows:

- Inertia = the number of transitions made to the same AOI/total number of transitions
- Volatility = 1 – inertia

In other words, volatility indicates how flexible a participant is in moving between new AOIs and exploring novel relationships between AOIs. Inertia is a measure of how fixated a participant is to one or a limited set of AOIs.

Findings from some of the existing work (e.g. Cook, 2006) suggest that novices are likely to exhibit either very high or very low inertia values. This is because they may 'see' only limited relations between ERs or their elements or are desperately looking for relations all over the place and are clueless about where to look for such relations. Experts, on the other hand, are expected to show moderate inertia/volatility values as they are neither fixated at one AOI nor clueless about the relations.

4.1.8.2 Balancing task (step 3)

The data analysis protocol was similar to the one followed in analyzing gaze data obtained while participants viewed chemical equations during the categorization task. However, parameters relating to visual attention only, such as number of fixations and fixation duration, were ignored, as the attention patterns are likely to be different in the two tasks, even if the same equation is presented, due to the inherent differences in the nature of these tasks. While discussing eye-movement results in this task, emphasis is, thus, given ton more general parameters related to the perception process, such as the nature of transitions and volatility, than to parameters of attention.

AOIs similar to those depicted in figure 4.1b were used to extract total viewing duration in seconds, fixation index (count), fixation duration in seconds, saccade index, and AOI hit. The number of transitions between the AOIs were calculated using this data, which were subjected to further classification similar to that described in figure 4.2. Inertia and volatility measures were also obtained from the transition data.

4.1.9 Findings

4.1.9.1. Step 1 (RQs 1 and 2) -- Establishing ER integration differences

Unlike results from previous studies (e.g. Kozma & Russell, 1997), our experts and novices did not differ in the total number of categories generated (experts = 38, novices = 40), mean number of categories made (experts mean = 4.75, S.D. = 1.66; novices mean = 5.71, S.D. = 1.11), average number of cards used (experts = 17.75, S.D. = 0.89; novices = 18, S.D. = 2.44) and the average number of cards used in a category (experts mean = 3.62, S.D. = 0.83; novices = 3.73, S.D. = 0.98). However, experts made significantly more conceptual categories (mean = 1.5, S.D. = 1.19) than novices (mean = 0.28, S.D. = 0.49). They also tended to make significantly more mixed categories (mean = 1.83, S.D. = 0.65) than novices (mean = 0.71, S.D. = 0.76) on average (RQ 1). Table 4.2 shows the mean category distribution for the two participant groups across the above defined five types of categories (best of two).

Table 4.2 Nature of categories for experts and novices (best of two rounds/trials)

Group/Nature of category	Experts (Mean)	Novices (Mean)
Conceptual	1.5 (1.19)*	0.28 (0.49)*
Mixed	1.83 (0.65)**	0.71 (0.76)**
Surface feature-based	1.16 (0.83)**	3.29 (1.38)**

Media-based	0.67 (1.07)	0.57 (0.53)
Inappropriate	0 (0)*	0.85 (1.07)*

*significant at p<.05, **significant at p<.01 between the groups.

The two groups also differed significantly in relying on feature-based criteria for categorization (mean feature-based categories for experts, mean = 1.16, S.D. = 0.83; for novices, mean = 3.29, S.D. = 1.38). However, the two groups did not differ in the number of media-based categories (mean for experts = 0.67, S.D. = 1.07; mean for novices = 0.57, S.D. = 0.53). Every novice made nearly one incorrect or inappropriate ER combination per trial on an average (mean = 0.85, S.D. 1.07), unlike experts who did not provide any inappropriate or incorrect justification.

Participants from both the groups seemed to struggle in interpreting animations (RQ 2), although only one of them (an expert, FC1) reported this explicitly.

"To be very frank.. this molecule thing (pointing to animation cards).. I was a little (unsure – shakes hand). Because…No.. I'm really confused with these ones (pointing to the three ungrouped animation cards)." (FC1)

A few experts and one novice (UG4) interpreted molecular dynamics more accurately than their other group-mates. Though only one novice (UG6) had left an animation uncategorized, as opposed to three experts (FC1, FC2, and FC6 – who left 3, 2 and 2 animations uncategorized respectively), more animations ended up under the relatively low rated feature-based, media-based, and incorrect or inappropriate categories for novices than experts. This is also consistent with the nature of categories novices made. Participants from both the groups were similar in reporting details of molecular dynamics while explaining animations. However, experts were slightly better at linking the molecular dynamics depicted in animations with other ERs, in a chemically appropriate manner, while novices were often overconfident about their interpretations of animations, among other ERs.

Experts' reasoning for the mappings between animations and other ERs often went beyond the immediate 'chemical' inferences drawn from events observable in the animations (such as molecular aggregation, breaking or formation of molecules, etc.) Not only did experts describe the process dynamics of chemical reactions under a set of conditions (RQ 2), but they also explained how the reaction dynamics may change if a certain condition during the chemical process is altered. For instance, an expert – FC5 – provided the following justification on grouping ERs related to solubility and precipitation phenomena together:

"When you add ammonia.. you are getting silver ammoniate. Now this solubility.. in the beginning it <u>increases with the concentration</u>... whenever you add the precipitating reagent you are getting a precipitate. <u>When you are adding a different amount of the precipitating reagent, that amounts in the complete precipitate</u>. But <u>when you add excess</u>... that results in dissolution. That is what is (also) reflected in this.. suppose if you add.. PbI_2 you are getting a solid. But <u>(if) you add excess of K</u>.. ultimately you are getting a clear solution." (FC5)

It is clear from the underlined fragments of FC5's statement that she imagined counter-factual situations, or situations that were beyond the currently represented information during the task, as none of the ERs included in this task which represented precipitation reaction between KI and PbI_2 had any explicit information about adding excess of K or its effect on the precipitation dynamics.

Next, in relation to the identification of general behavior patterns (RQ 3), two experts began the categorization phase by first arranging the cards medium-wise, and then proceeded to relate the ERs (figure 4.3). These initial actions, termed 'epistemic actions' (Kirsh & Maglio, 1994), help organize the task space and are known to reduce cognitive load.

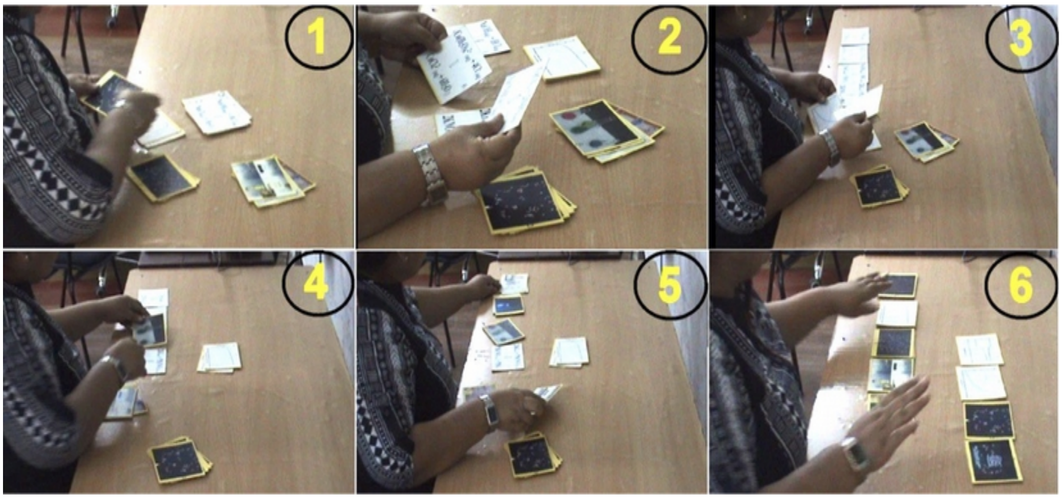


Figure 4.3 An episode of epistemic actions – in (1), she is seen sorting the cards media-wise and keeping them on the table as four different stacks. With such an arrangement, she is able to decrease the chaos and 'see' conceptual relationships between the cards. In (2) she picks up two cards from the stack of chemical equations and compares them, while the other three sets of cards (animations, video-snapshots and graphs) lie on the table. In (3), the participant has spread all the equation cards – another epistemic action performed to improve perceptual reach. She is also spotted comparing the graph cards (held one in each hand) either with each other or the equation cards. In (4 and 5), FC2 can be seen comparing different cards and placing them together. (6) depicts the completion of FC2's categorization task.

A few other expert participants found it useful to initially spread the cards on the table in no particular fashion – another instance of epistemic action, performed possibly to obtain an overview of all the cards together.

4.1.9.2. Step 2 (RQs 3 and 5; gaze-behavior data): -- Identification of sensorimotor markers

Novices spent more time viewing each representation than the experts on average (table 4.3). Novices also recorded significantly more saccades per scene than experts (p<.01, table 4.4). These viewing duration as well as saccade results are consistent with results from Cook et al., (2008) and Kohl and Finkelstein (2008).

Table 4.3 Mean viewing duration in seconds per scene

Group/Nature of category	Experts (Mean)	Novices (Mean)
Equations	13.65 (8.45)	26.94 (21.84)
Graphs	21.43 (12.44)	35.25 (24.06)
Animations	38.75 (26.34)	69.65 (31.76)

Demonstrations	64.98 (38.18)	125.28 (73.11)

Note: Differences not statistically significant.

Table 4.4 Mean saccades per scene across the different representation media

Group/Nature of category	Experts (Mean)	Novices (Mean)
Equations	66.37 (49.30)	131.68 (108.11)
Graphs	100.22 (73.39)	169.91 (100.59)
Animations	90.22 (56.69)	142.17 (72.64)
Demonstrations	171.9 (109.75)	354.34 (227.91)

Note: All the between group values are significant at $p<.01$.

Novices also spent more time, as well as fixated more number of times, on average, in each AOI across all the equations than experts (figures 4.4 a & b; difference not significant).

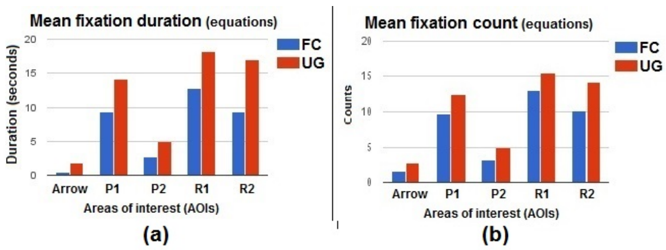


Figure 4.4. (a) Mean fixation duration per AOI, (b) Mean fixation count per AOI, across all the equations.

Interestingly, participants in both the groups seemed to find more relevant information about chemical equations from the reactant-related AOIs of the equations than the products, or the process symbol – arrow. Both the groups combined spent 63% of the total fixation time in the reactant AOIs across all equations, against just 34% in the product AOIs (Significant at $p<0.0001$). Similar trends were observed for fixation count presented in figure 4.5 below.

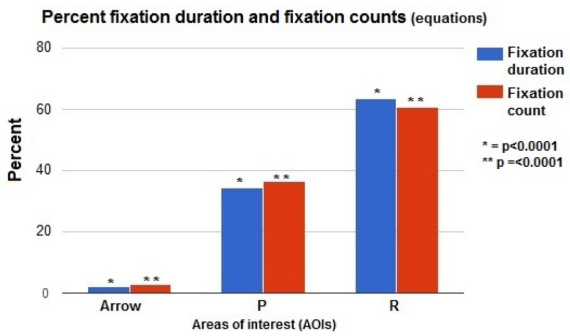


Figure 4.5. Percent fixation duration (blue) and percent fixation count (red) for rounded up AOIs across all the chemical equations.

Across the graphical representations, novices spent longer in all the AOIs (except Y-axis), and visited the AOIs more frequently on average than experts (figure 4.6, difference statistically not significant).

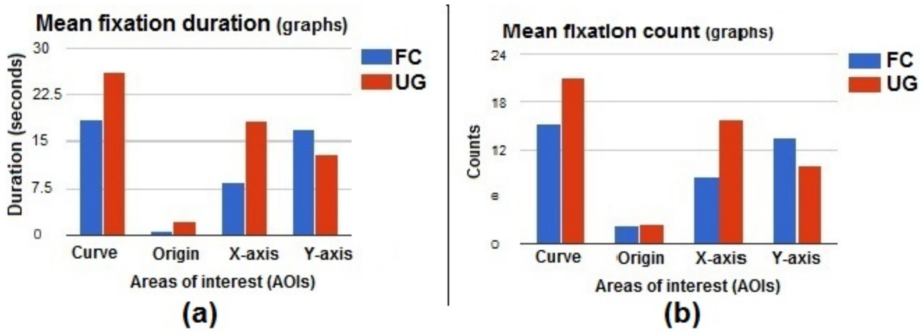


Figure 4.6 (a) Mean fixation duration per AOI, (b) Mean fixation count per AOI, across all the graphical representations.

From the gaze transition data, I found that experts and novices did not differ in the number of between-AOI transitions across graphical representations (experts: total = 662 transitions, mean = 20.69 per graph, S.D. = 3.64; novices: total = 541 transitions, mean = 19.32 per graph, S.D. = 7.91). Figure 4.7 shows a normalized distribution of transitions for experts and novices between the different AOIs across the graph representations.

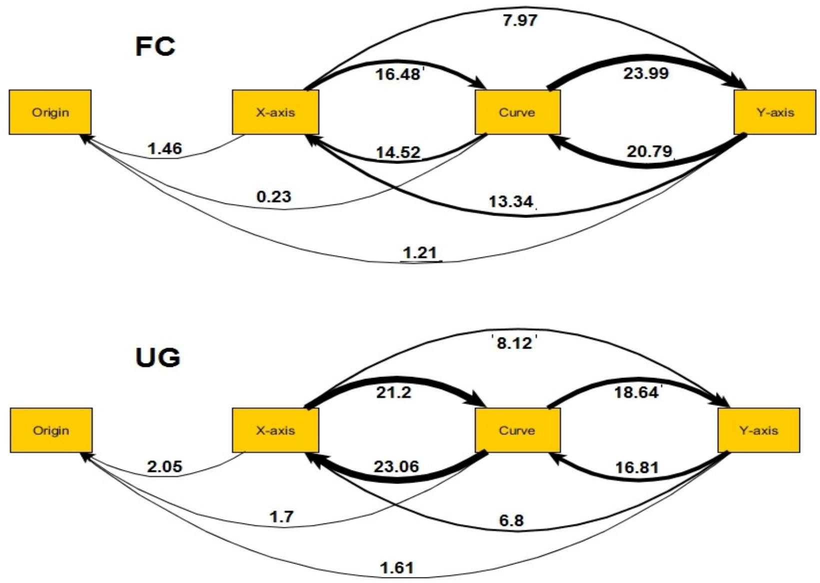


Figure 4.7 Percent transitions between AOIs averaged across all graphical representations. Each box represents an AOI. Direction of the arrow indicates direction of the transition. The thickness and the numbers on the arrows indicate the relative number of transitions between those two AOIs. FC = experts, UG = novices.

Experts transited more frequently between the curve and Y-axis AOIs by a considerable margin than between the curve and X-axis AOIs. Novices showed the exact opposite pattern; however, neither within-group nor between-group differences were statistically significant. Coincidentally, X-axis in each graph showed the independent variable, whereas Y-axis depicted the dependent variable. The dependent variable indicates properties of a reaction system. Experts appear to be interested in deriving meaning from how the dependent variable is responding to the independent variable (process dynamics – RQs 2 and 4), while novices may be trying to figure out what would the response be. However, this can only be speculated from the gaze patterns, as none of the experts' transcripts reveal any corroborative evidence. Both the groups exhibit frequent transitions between the curve of the graph and the axes (close to 80%). Transitions between the two axes are relatively less frequent, while transitions to and from the origin are negligible.

For equations, there was no difference in the number of transitions between the two groups (experts: sum = 323 transitions, mean = 13.46 per equation, and S.D.

= 5.66; novices: sum = 258 transitions, mean = 12.9 per equation, S.D. = 5.57). However, experts performed significantly more long transitions (mean = 6.58, S.D. = 1.39) than novices (mean = 3.95, S.D. = 1.06) at $p = 0.001$. Novices tended to perform significantly more short transitions (mean = 8.95, S.D. = 1.48) than the experts (mean = 6.75, S.D. = 1.2) at p=.01 (figure 4.8).

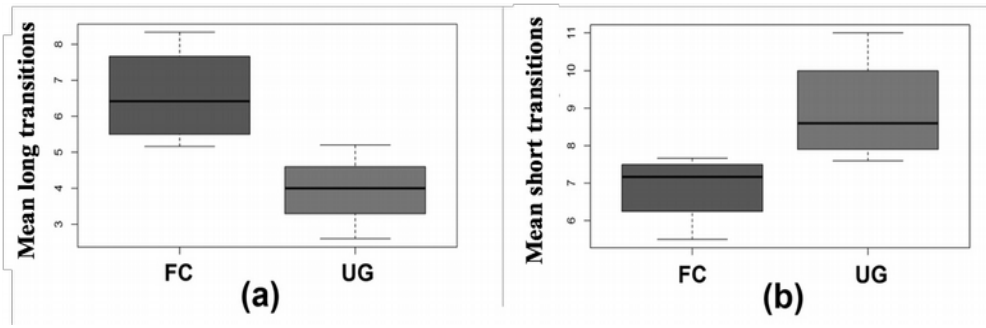


Figure 4.8 Box plots depicting (a) mean number of long transitions performed by experts and novices, differences significant at p=0.001; and (b) Short transitions for experts (FC) and novices (UG), differences significant at p=0.01.

Experts' gaze transitions have significantly larger proportion of long transitions (mean percent = 48.92, S.D. = 4.23) than novices (mean percent = 30.62, S.D. = 2.58). Conversely, they tend to perform significantly fewer short transitions than the novices (Figure 4.9).

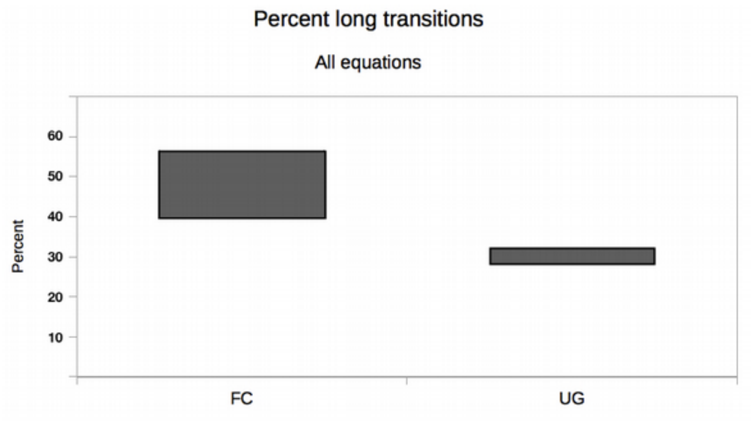


Figure 4.9 Box plots showing percent long transitions made by experts (FC) and novices (UG) across all the equations. An inverse would show percent short transitions.

In terms of volatility across all equations– a general measure of how flexible a participant is to explore different parts of a representations in relation to each other, novices show a significantly lower mean volatility index of 0.25 (S.D. =

0.09) in relation to experts' mean value of 0.33 (S.D. = 0.05) at p = 0.05 (figure 4.10). In other words, experts are almost 1.5 times more flexible in navigating or exploring the different AOIs while observing equations.

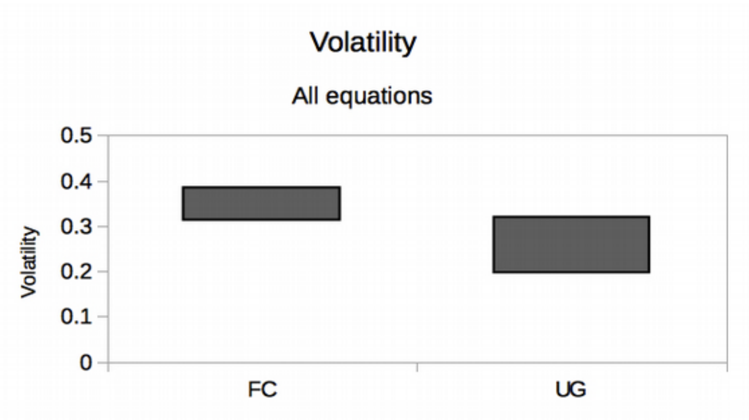


Figure 4.10 Volatility box-plots for all equations. Experts show a mean volatility index of 0.33 as opposed to novices whose index 0.25. The difference is significant at p=0.05.

Across the graphical representations, the between-group difference is not significant (figure 4.11). However, novices show considerably lower volatility values (experts mean = 0.38, S.D. = 0.05; novices mean = 0.33, S.D. = 0.06).

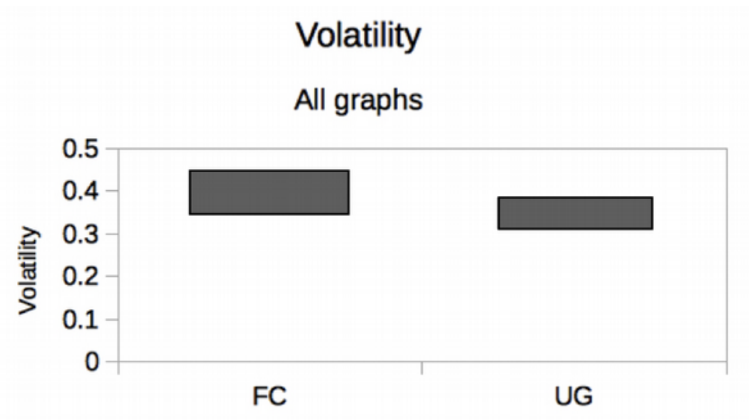


Figure 4.11 Volatility box-plots across all graphical representations. Values are statistically not significant.

4.1.9.3. Step 3 (RQ 4, equation-balancing task) -- Confirming the relation between ER integration and sensorimotor markers

There was no difference between experts and novices in terms of the proportion of long transitions while viewing the unbalanced chemical equations. On an average, 26.82 percent transitions of the total transitions experts made were long transitions (S.D. = 5.49), while the proportion of long transitions for novices was 26.4 percent (S.D. = 6.64). However, the proportion of long transitions for both the groups are consistently lower than those each group exhibited during the categorization task.

The two groups did not vary in terms of volatility values: For experts, the mean volatility index was 0.47 (S.D. = 0.09); whereas novices recorded a volatility index of 0.41 (S.D. = 0.05). However, unlike the long transition values, the volatility values are consistently higher for both the groups than those observed for the groups during the categorization task. This between-task difference could be a result of the inherent differences in the nature of the tasks.

4.1.10 Discussion

Step 1 results show that chemistry professors (experts, FC) were more competent at conceptually relating and grouping ERs than novices. The statistically significant differences in the nature of categories between the two groups match the relative competence levels reported in previous expert-novice studies. The categorization schemes also indicate that novices rely more on surface-features of ERs and find it difficult to infer and imagine chemical reaction dynamics.

Analysis of experts' verbal responses related to category justification, in terms of ER relationships, revealed that not only did some experts 'unfreeze' successfully the components of static ERs, into an understanding of their dynamic behavior in relation to other representations, but also imagined certain counterfactual situations, as well as how the behavior of those components would be affected under such situations. For instance, none of the ERs related to precipitation reaction between KI and PbI_2 included in the categorization task had any explicit information about adding excess of K or its effect on the precipitation dynamics. However, FC5 – an expert – imagined beyond the 'given' information in the ERs, and explained verbally how solubility of the elements and compounds involved in such reactions is affected in relation to their concentration. While provision of conceptual explanations of this kind is a commonly reported characteristic of domain expertise (e. g. Chi, Feltovich & Glaser, 1981), our interpretation focuses on the ability to imagine counterfactual dynamics related to chemical processes.

Interestingly, in step 2 analysis, some experts were found to use strategies such as spreading the cards on the table and/or devise preliminary criteria to arrange the ERs on the table, etc., before proceeding to form finer categories. All these actions are identified as ‘epistemic actions’, which experts in various domains are known to perform, in order to change certain structures in their environment to search for a solution or strategy while performing a task, and/or lower the cognitive load generated in a situation (Kirsh, 2010), as well as see newer relationships between the ERs (Aurigemma et al., 2013). Kirsh and Maglio (1994) contrast epistemic actions with pragmatic actions, in that performing the latter brings a person physically closer to the solution of the problem, while not serving any specific cognitive role (although they may end up playing such a role). The actions some of our experts performed, such as those illustrated in figure 5.3, cannot be considered as pragmatic actions given their role in helping the experts gain newer insights into the relationships between the ERs. It is clear that these relationships, on whose basis the experts refined their categories, ‘appeared’ to the experts after they spread the cards on the table or arranged them media-wise initially. Why would they otherwise refine their categories? Their epistemic actions helped them organize the task space and eventually related the ERs in newer ways that were either partially imagined in the beginning or not imagined at all.

Further support on the distinction between experts and novices from eye-tracking revealed that the gaze behavior of experts is considerably different from that of novices. The fixation duration, fixation count, mean viewing duration and frequency of saccades data, which indicate the time to pickup information from a location in a representation, share similarities with the eye-tracking results in Cook et al. (2008) and Kohl and Finkelstein (2008), and suggest that our novices took significantly more time as well as effort to relate the different parts of a representation in comparison to experts, who exhibited relatively stable attention patterns.

Novices found it difficult to bind their internal representations of static and dynamic ERs into a coherent integrated mental model. This is why we see novices making significantly fewer specific saccades (transitions), though their frequency of overall saccades is very high. This also indicates that a high average frequency of saccades may not necessarily mean that a participant's search and coordination between features of the ERs were systematic. Specific saccades are required to establish efficient mappings between the ERs. Overall, novices' attempts to integrate different information in a representation, as well as across the ERs, is haphazard.

Experts, on the other hand, were more balanced than the novices. The frequency of transition between two AOIs is a measure of the comparison made by a participant between the contents of the AOIs as well as ER integration (Holsanova, 2014). Novices' gaze transitions across equations were skewed more towards 'within the reactant and product species', pointing to fewer attempts at establishing coordination between species on the two sides of the arrow of a dynamic chemical process. For experts, however, the proportion of gaze transitions between reactants and products as well as within reactants and products is almost evenly distributed. This implies that the experts coordinated equally well between reactants and products, and between two reactants or two products. This is another evidence pointing to their relative stability in navigating representations.

For graphical representations, experts seemed to be coordinating more between the independent variable (X-axis) and the curve; the shape of the curve ultimately shows the behavior of dependent variable given changes in the independent variable. This indicates that the experts were interested in deriving the process dynamics in terms of how dependent variable responds to the independent variable in a graph. This is consistent with experts' chemically meaningful explanations regarding the dynamics of the phenomena inferred from ERs. In contrast, novices may have tried to deduce the end product of the reaction – a rather static understanding of the reaction (Talanquer, 2013), and faced difficulties in inferring effects of independent variable(s) on the behavior of dependent variable(s) through the shape of the curve.

Evidences from our pilot experiment in the past have indicated that this between-group difference in gaze behavior is a function of growing expertise, as undergraduate students who performed the categorization task in ways similar to experts (Pande & Chandrasekharan, 2014) tended to exhibit similar gaze patterns across graph representations and equations to those of experts (Figure 4.12).

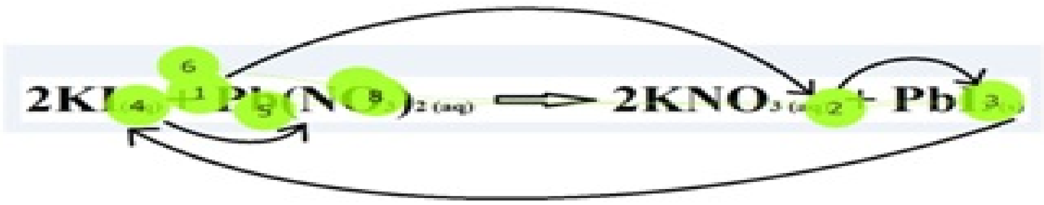


Figure 4.12 An instance of gaze behavior of an expert-like undergraduate student (Pande & Chandrasekharan, 2014). This student makes less fixations and equally frequent short and long transitions.

In contrast, the gaze patterns of a relatively novice candidate revealed a relatively linear sequential scanning of the equation with long pauses in between (figure 4.13).

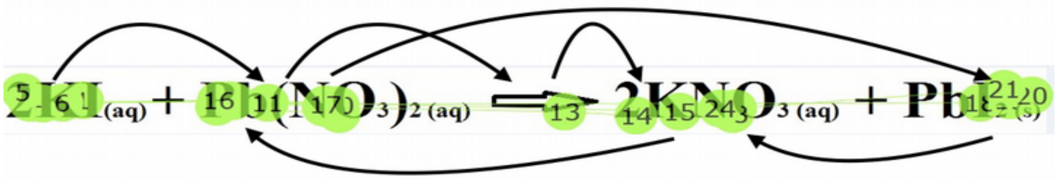


Figure 4.13 An instance of a novice student's eye-gaze sequence. This student largely follows the sequence of the equation with long pauses between transitions. He also makes twice as many short transitions as long transitions.

This novice student did not seem to follow a particular element or molecule for its changing states. For instance, her first 6 fixations occurred on 2KI, followed by another 6 fixations on the next reactant molecule. She returned to $Pb(NO_3)_2$ after one fixation on the arrow and a few fixations on the first product, but the fixation points (14, 15 and then 16, 17) do not seem to follow any specific element, coefficient or subscript. Subsequent fixations (points 18-21 on PbI_2 and 22-24 on NO_3) indicated a similar pattern.

In the balancing task, no differences between the two groups in terms of gaze behavior were observed. This indicates that ER integration is context specific and that the presentation of chemical equations in a non-ER integration-related context does not trigger simulation and/or imagination similar to that involved in ER integration tasks.

Experiment 2 reported in the following sections is a cross-sectional investigation, which aimed to explore sensorimotor markers of this ER integration development in chemistry, where two more groups of participants performed the categorization and balancing tasks.

4.2 Experiment 2 (Cross-sectional investigation of ER integration development)

The first experiment successfully identified the sensorimotor markers of the cognitive mechanisms underlying ER integration, primarily in terms of gaze-behavior. The second experiment sought to understand how ER integration and the underlying cognitive mechanisms develop during chemistry education, by studying cross-sections of the chemistry education process. In the first experiment, two important cross-sectional groups – chemistry professors (FC) and undergraduate students (UG) – were studied. The second experiment complements findings from the first experiment, by replicating the categorization and balancing tasks with two more groups of participants viz., pre-university students and chemistry graduate students.

4.2.1 Sample

7 pre-university (i.e. 11^{th} grade) students (code-named PU; 2 female) pursuing general science and mathematics courses from a junior college, and 7 chemistry graduate students (GS; all male) pursuing Ph. D. in chemistry from different

major universities across Western India volunteered to participate in the experiment. Informed consent was obtained from all participants.

Each participant individually performed the two tasks (categorization and balancing) during the experiment.

Data from experiment 1 for chemistry professors (FC1-FC8) and undergraduate students (UG1-UG7) were compared with the participants from this experiment to identify developmental trends based on education.

4.2.2 Experiment protocol and data analysis

Since this is a replication of experiment 1, the tasks, methodology and data analysis steps taken were exactly similar to experiment 1.

The categorization as well as behavior and gaze data analysis was similar to experiment 1, except that the focus in this experiment was solely on the interaction process and not on attention. Hence, parameters such as the number of fixation or fixation duration were ignored.

4.2.3 Findings

4.2.3.1 Step 1 (RQs 1 and 2): Establishing ER integration differences

Table 4.5 below presents the distribution of categories among the four groups of participants.

Table 4.5 Distribution of categories across the four participant groups.

Group/Nature of category	**FC** Mean (SD)	**GS** Mean (SD)	**UG** Mean (SD)	**PU** Mean (SD)
Conceptual	1.5 (1.19)	0.5 (0.84)	0.28 (0.49)	0.14 (0.38)
Mixed	1.83 (0.65)	1.5 (1.38)	0.71 (0.76)	1.86 (1.35)
Feature-based	1.16 (0.83)	3.0 (2.37)	3.29 (1.38)	1.29 (1.38)
Media-based	0.67 (1.07)	1.33 (2.06)	0.57 (0.53)	2.29 (2.63)
Inappropriate	0 (0)	0.83 (1.17)	0.85 (1.07)	0.86 (1.21)

As expected, chemistry professors have the highest mean number of conceptual categories among the four groups, while pre-university students the lowest. There are no differences between the professors, graduate students and pre-university students in the average mixed categories made, except for novices who made nearly half as mixed categories as any other group. Professors and pre-university students make significantly fewer feature-based categories than the other two groups, whereas graduate students and pre-university students have slightly higher average number of media-based categories. Finally, doctoral, undergraduate and pre-university students were similar in terms of making erroneous associations between the ERs.

Figure 4.14 below presents the overall trends in the normalized proportions of the different kinds of categories each group made in terms of stacked radar charts.

A weak developmental trend among the four cross-sectional groups is visible in the radar charts. Starting from participants with less experience in chemistry (PU and UG) to those with more experience (GS and FC), the radar plots show a clear shift from a largely media and feature-based categorization scheme to a more conceptual one. The doctoral (GS) and undergraduate (UG) students both show clear tendencies towards feature-based categorization schemes, whereas both the professors (FC) and pre-university (PU) students tend to be more diverse in their categorization schemes, although in considerably different ways.

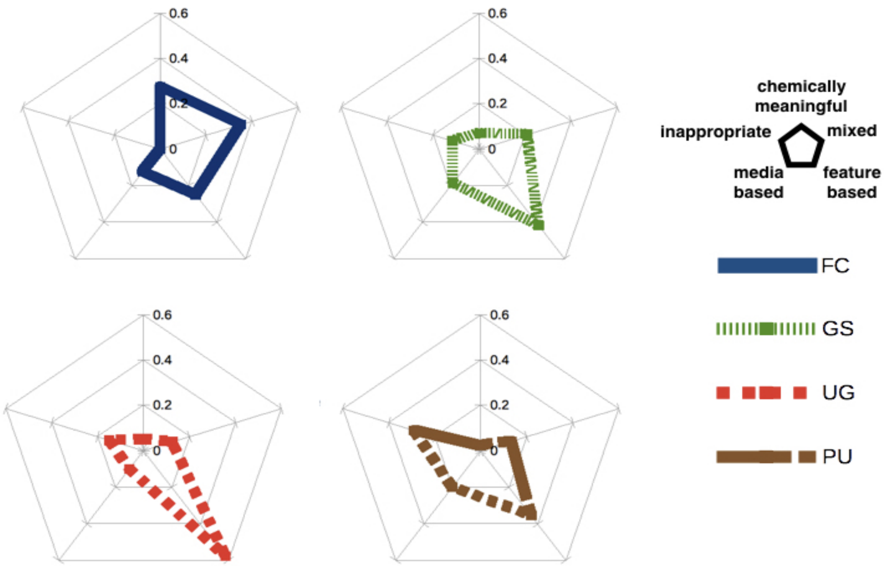


Figure 4.14 Radar charts showing cumulative proportion trends of category types. Each corner of a pentagon radar plot represents a kind of ER category as indicated in the top right corner.

4.2.3.2. Step 2 (RQs 3 and 5) -- Identification of sensorimotor markers

The first experiment demonstrated how educational experience in chemistry may be influencing the way in which participants navigate ERs. The gaze behavior patterns mark these differences in navigation and the underlying cognitive processes. To specifically focus on navigation, I ignore markers of attention (parameters such as fixation count, visit duration, etc.), and discuss only dynamic eye-movement related data from this experiment.

Figure 4.15 plots the average saccade frequency for participants from the four groups while viewing the different types of representations during the categorization task.

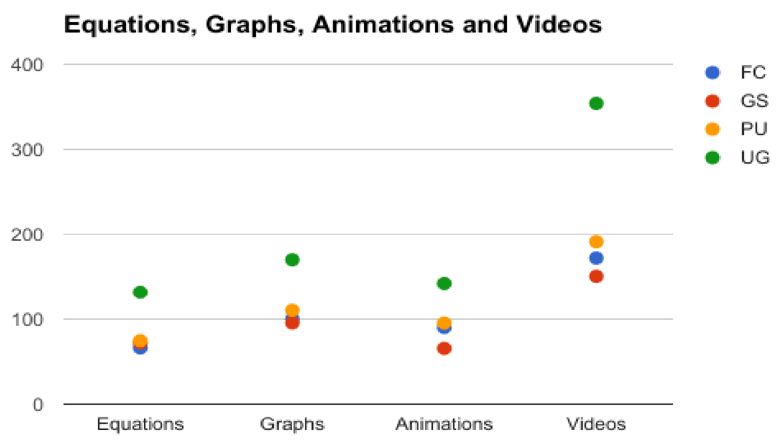


Figure 4.15 Average number of saccades per representation for the four participant groups.

Undergraduates recorded significantly more saccades per representation than the other three groups (p<.01).

In terms of specific saccades (transitions) across graphical representations, chemistry professors transited more frequently between curve and Y-axis by a considerable margin than between curve and X-axis (figure 4.16 below), whereas undergraduates showed the exact opposite pattern, as also noted previously in experiment 1 results.

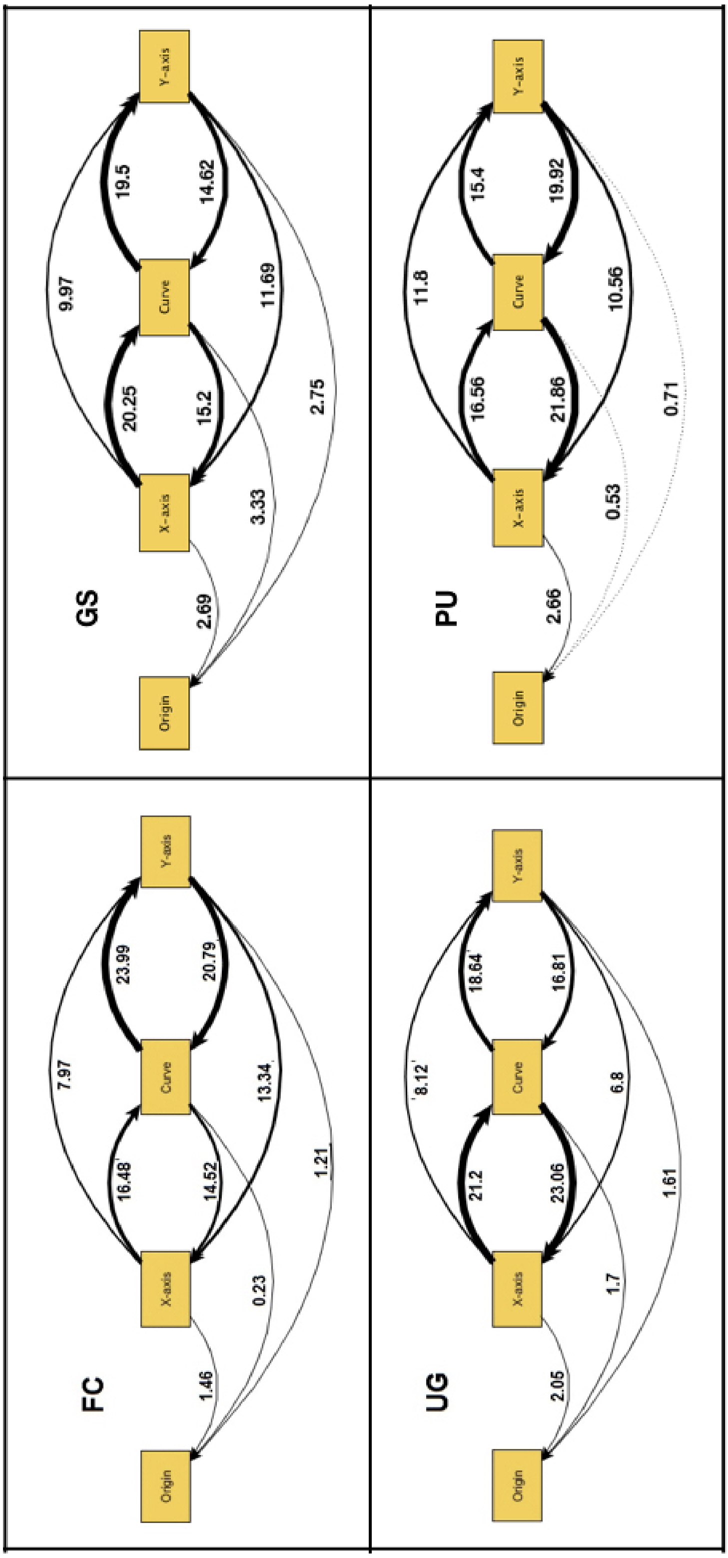


Figure 4.16 Normalized transitions between AOIs of graph representations for all the groups. Each box represents an AOI. Direction of the arrow indicates direction of the transition. The thickness and the numbers on the arrows indicate the relative number of transitions between those two AOIs.

Different from the professors and undergraduates, both the doctoral students and pre-university students transited equally often between X-axis and curve, and Y-axis and curve. The proportion of transitions between the two axes was indifferent for all the four groups. All the groups also exhibit frequent transitions between the curve of the graph and the axes (close to 80%) and, as can be observed from figure 4.16, transitions between the two axes are relatively less frequent, while transitions to and from the origin are almost negligible.

While viewing equations, chemistry professors (mean = 48.92, S.D. = 4.23), doctoral students (mean = 51.36, S.D. = 4.08) and pre-university students (mean = 53.08, S.D. = 3.83) performed significantly higher proportion of long transitions than undergraduates (mean = 30.62, S.D. = 1.58), as evident from figure 4.17 which depicts interquartile box plots for normalized proportion of long transitions among the four groups.

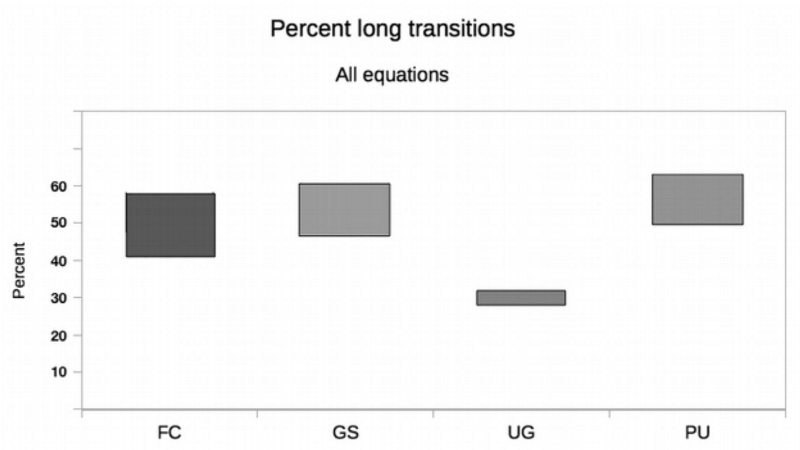


Figure 4.17 Percent long transitions between the different AOIs across all equations. The means and standard deviation values for each group are as follows: FC = 48.92 (4.23), GS = 51.36 (4.08), UG = 30.62 (1.58) and PU = 53.08 (3.83).

Figure 4.18 below presents box plots for volatility values across all static ERs (all equations + all graphs).

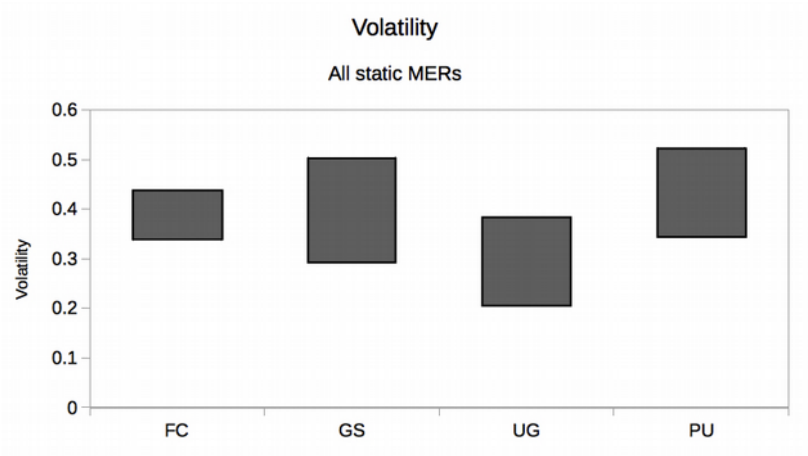


Figure 4.18 Box plots for volatility values across all static ERs combined (all equations + all graphs).

Among all the participants, undergraduates were found to be the least volatile (mean = 0.28, S.D. = 0.09). They were an extremely inert group in terms of gaze transitions, indicating that they hesitated in mapping between different parts of a representation. Although these differences are not statistically significant. Chemistry professors (mean = 0.36, S.D. = 0.05) and doctoral students (mean = 0.37, S.D. = 0.1) were moderately volatile, while the pre-university students reported the highest volatility (mean = 0.39, S.D. = 0.07), though the differences between the groups are only indicative and not significant.

4.2.3.3 Step 3 (RQ 4): Confirming the relation between ER integration and sensorimotor markers

While the percent long transitions have decreased during balancing as compared to the categorization task for all the participants on an average, there are no differences between the groups (figure 4.19). Mean and standard deviation values: FC = 0.27 (0.05), GS = 0.28 (0.03), UG = 0.26 (0.07), PU = 0.31 (0.06).

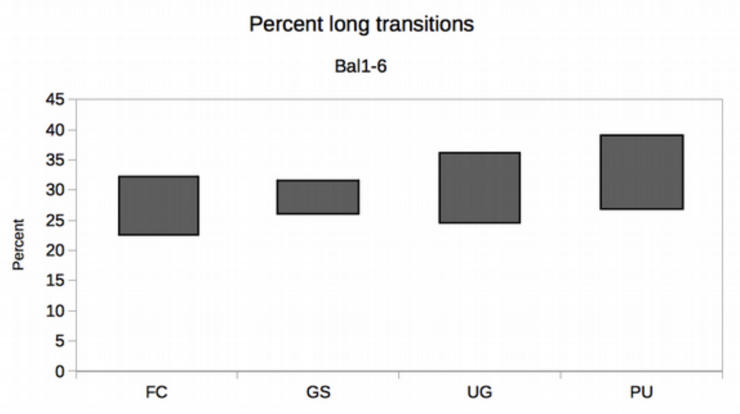


Figure 4.19 Mean percent long transitions across AOIs for all equations in the balancing task.

The four groups did not differ in terms of volatility measures while viewing the different components of unbalanced equations presented during the balancing task (figure 4.20). Mean and standard deviation values: FC = 0.47 (0.09), GS = 0.44 (0.10), UG = 0.41 (0.05), PU = 0.49 (0.10).

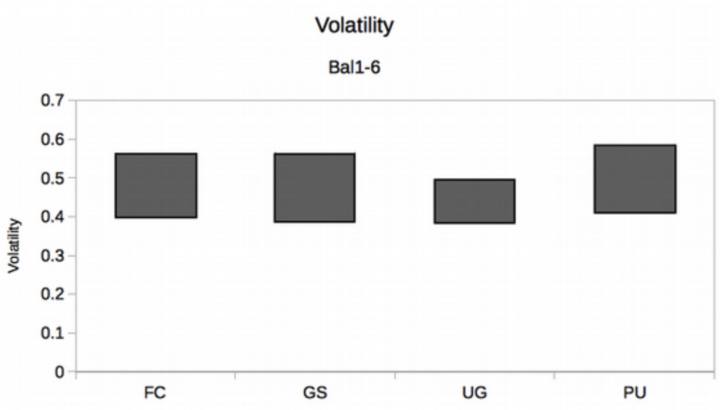


Figure 4.20 Mean volatility values for all equations in the balancing task.

4.3 General discussion

The step 1 (categorization) results indicate a clear developmental pattern in the ER integration abilities across the four cross-sectional groups. Professors exhibit the highest ER integration abilities, followed by graduate students, and then undergraduate students. The pre-university students exhibit the lowest ER integration abilities.

Step 2 analysis (eye-tracking) revealed only indicative patterns of development. The pre-university students often exhibited gaze behavior similar to those of professors; whereas the graduate students exhibited moderate values across parameters, although it was expected from the results of experiment 1 and the categorization trends among the four groups in experiment 2, that the pre-university students would exhibit significantly different gaze parameter values in comparison to the professors. Regardless, the professors, graduate students and pre-university students seem to conform to a weak developmental trend across several gaze behavior parameters. In terms of the nature of gaze transitions, for instance, across graphical representations, professors appear to be interested in deriving meaning from how the dependent variable (curve shape) is responding to the independent variable (Y-axis; process dynamics – RQs 2 and 4) while the doctoral and pre-university students seem to exhibit intermediate behavior, as it is not clear if they are deriving or predicting the behavior of the curve by treating values on the Y-axis independent of those on the X-axis. It could also be that they are corresponding between specific features of the curve shape with specific values on the X-axis. Similar behavioral trends are observed in the volatility values, which indicate that the pre-university students are haphazard in terms of their gaze behavior, perhaps as a result of confusion over mapping the different aspects of a representation onto each other. The extremely high volatility values, i.e. high between-AOI activity for graduate and pre-university students indicate their instability in navigating the graphical representation. This could be possibly be a result of confusion over mapping. The professors, who exhibited moderate volatility values, were a relatively stable group.

This developmental pattern, however, is disrupted when gaze-behavior data of undergraduates is considered. Undergraduates always exhibited significantly different gaze behavior in comparison to the other three groups, and were at one extreme of the continuum. The distribution of their gaze-transitions between the different AOIs of graphical representations, for instance, is exactly opposite to the professors' gaze-transition patterns, and qualitatively different from the graduate and pre-university students. Undergraduates reported the lowest volatility values – they were relatively inactive between-AOIs, though they report a high overall saccade frequency. Majority of saccades for undergraduates thus must be within-AOI eye-movements, indicating that they were either confused or over-confident in integrating the different AOIs in graphical representations. This possible confusion among undergraduates, however, is likely to be different from that of the graduate and pre-university students.

How can such unexpected patterns be explained? The pre-university group had just studied general chemistry, so it was fresh in their minds. While in the case of undergraduates and graduate students, the ER integration system appears unstable and undergoing disruptions because of exposure to a lot of new representations and conceptual knowledge. In the case of professors, experiences with chemical ERs have settled into relatively stable internal models. This is perhaps one reason why pre-university and professors exhibit most stable and less skewed/diverse categorization trends, while pre-university and graduate students are somewhere in between and show strikingly skewed categorization, indicating sharp tendencies towards certain grouping schemes. The development of expertise and ER integration seem to follow a pattern similar to the 'development as a complex dynamic system' model (Smith & Thelen, 2003), which shows that well-learned sensorimotor skills can deteriorate when further skills are learned.

The eye-tracking results as well as the instances of epistemic actions (observed only in the case of experts as reported in experiment 1) suggest that expertise is accompanied by a fine-tuned sensorimotor system, which is: (i) oriented towards picking up maximum information from an external representation, and (ii) involved in the task-specific reorganization of information to facilitate problem solving. Conjecture 1 – stating that the way learners perceptually access ERs would change after significant training in a domain – can thus be said to be broadly supported by the evidence. Once fine-tuned through ER-based training, the sensorimotor system is activated or simulated on encounter with ERs, resulting in distinct sensori-motor behavior (Barsalou et al., 1999), in this case the gaze. This claim is supported by findings from the balancing task, which indicate that the revised sensorimotor pattern is not activated outside the ER integration context – thus refuting the alternative explanation of changes in the sensorimotor system as a general phenomenon.

This indicative evidence in favor of conjecture 1 can be further supported by the general observation that expert chemists make instantaneous decisions and actions while working on chemical processes (such as synthesis) in laboratories. Most such instances are reported as anecdotes of 'intuition' (Kutchukian et al., 2012) and 'tacit knowledge'. Our study is among the first to objectively characterize the sensorimotor changes during training, which eventually support this intuition and tacit knowledge. This approach provides a new perspective while understanding RC and expertise.

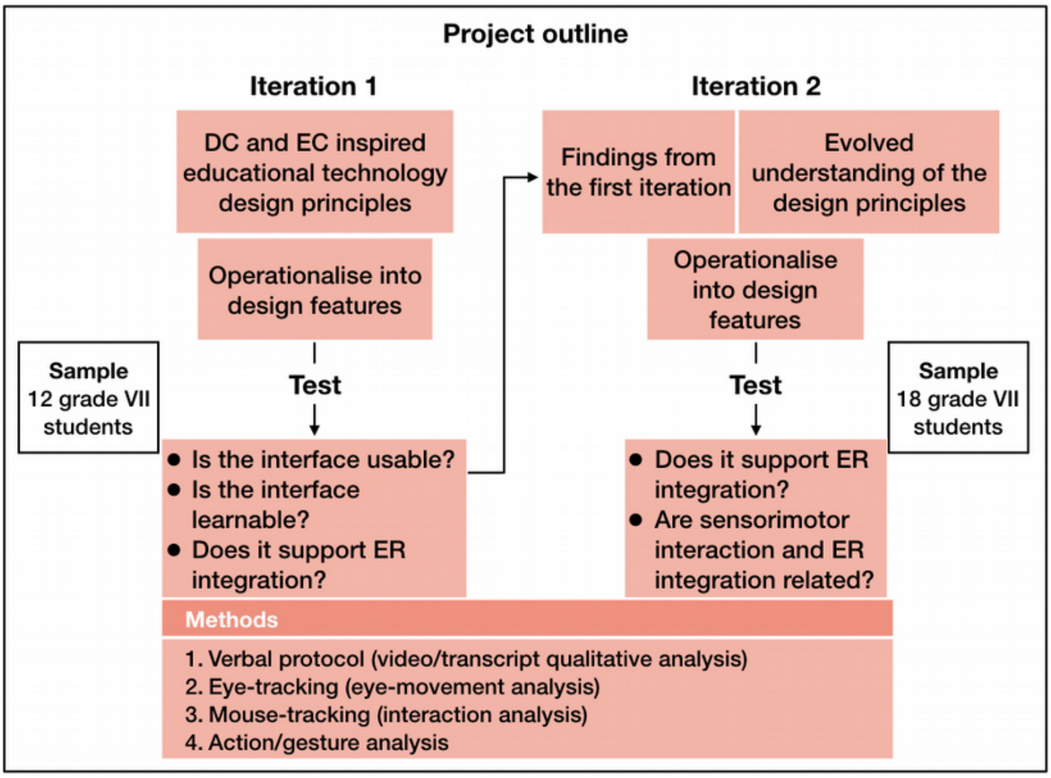
4.4 Limitations

Although the study involves collection and analysis of a huge amount of eye-movement data, the sizes of the participating groups are small, so the results are not confirmative from the point of view of statistical testing. The results are thus only indicative.

Another major methodological issue is that the eye-movement data in relation to dynamic ERs (animations, laboratory videos) were not considered for analysis. This is because the generation of eye-movement data, as well as processing algorithms, are notoriously unreliable when dealing with dynamic stimuli (ERs), and have recently been shown to perform barely above chance (Andersson et al., 2016). Moreover, analysing this type of data is extremely time, effort and computation intensive.

The project attempted to identify sensorimotor markers of the ER integration ability and its development. Similar to most previous experiments, including those replicated in this project, this project does not investigate the influences of (or interferences caused by) conceptual knowledge on ER integration. It is not clear how conceptual knowledge and ER integration are related, and this problem is out of the scope of the aims of this project.

Chapter 5: Can manipulation of enactive ERs lead to integration?: exploring the relationships between interactivity and ER integration using a design-based research model



This chapter discusses a project to test the second conjecture – sensorimotor interaction would support ER integration and its development. The project involved the design, development and testing of an interactive computer interface, with fully manipulable ERs of a physical system, as an intervention to help learners achieve ER integration. The system allowed learners to interact with and control coupled ERs of a phenomenon (oscillation) in their static as well as dynamic states.

The study initially sought to build on the first project, using ERs in chemistry. However, interaction with chemical ERs is complex and counterintuitive, at least for a novice learner (e.g. interaction with real chemicals is often not possible, direct interaction with molecules is impossible; chemical ERs are relatively abstract, and 'acting' on them is a conceptual process). It thus seems difficult to dissociate conceptual understanding and ER integration in the case of chemical representations. Also, it is nearly impossible to find naive subjects who can process representations in chemistry without understanding any of the chemical concepts.

For these reasons, I chose to explore a simple physical system and its representations, where learning the relationship between ERs can happen without requiring conceptual understanding. Oscillation and the ERs related to it were found to be ideal, as one could interact with the ERs physically or virtually at a more everyday world level, in contrast to, say, interacting with molecules. The pendulum is also a physical system with simple dynamics and a trigonometric equation that is relatively easy to understand. The primary motivation for this multi-representational interactive simulation interface is achieving RC or ER integration, and not conceptual understanding, although the possibilities of the latter are not denied.

As also discussed during the review in chapter 2, computer-based interactive learning environments are not new to science, mathematics and engineering education. There exist a vast number of animation and visualization tools, as well as interactive simulations (de Jong & van Joolingen, 1998; Rutten, van Joolingen & van der Veen, 2012) designed and developed primarily to improve conceptual, phenomenon and procedural understanding among students in these fields (e.g. Danish et al., 2015). Many of these interfaces support learners in relating and transforming across ERs using three major types of features: (i) an integrated presentation of ERs on the same screen in order to reduce the split-attention effect (Bodemer et al., 2004), (ii) or in a predetermined sequence to make salient the relation between them (Boucheix et al., 2013; Lowe & Boucheix, 2008), and (ii) dynamic linking of ERs (Stieff & Wilensky, 2003; van der Meij & de Jong, 2006) that automates the task of translation between representations, thus reducing cognitive load. Although these interventions differ significantly from each other in the degree as well as nature of interactivity they support, they can be broadly grouped into two types: 1) Visualizations with no interactivity, or visualizations with minimal interactivity, where operations such as play/pause/rotate/zoom are made possible, and 2) variable manipulation simulations or visualizations with dynamically linked ERs which allow the user to make changes in the given representations. In many systems, these changes are dynamically reflected in the other representations.

The former category of interfaces typically present or display learners with pre-simulated and/or animated ERs of a phenomenon (often not-to-scale) simultaneously or in a sequence, where the learner has either no or minimal control (such as play/pause, zooming in and out of a 2D/3D structure, audio/mute mode, etc.) over the interface. Some examples of such interfaces are 4M:Chem, SMV Chem (Russell &Kozma, 1994; Russell, Kozma, Becker, & Susskind, 2000) and VisChem (Tasker et al., 1996) which are visualization software that simultaneously present on the screen chemical equations, concentration and/or energy graphs, molecular-level animations and laboratory experiment video. However, the design of such interfaces is currently based on the information processing theories of cognition; the central role of the interface is to reduce the learners' cognitive load (specifically working memory load). Ironically, these software interventions, as a result of their visually complex design and minimal interactivity, do not seem to help the learner offload memory and processing to the computer screen.

A second set of computer tools such as PhET (Perkins et al., 2006), NetLogo (Wilensky, 1999), Molecular Workbench (Concord Consortium, 2010), SimQuest (van Joolingen & de Jong, 2003) are remarkably interactive and include features such as integrated and dynamically linked ERs. (i) PhET simulations (Perkins et al., 2006) offer a suite of variable manipulation simulations for various topics, with integrated and dynamically linked multiple representations of a concept, such that changing parameters of the physical phenomenon changes its associated graph. (ii) NetLogo (Wilensky, 1999) is an agent-based modeling language that allows students to create an agent-based model of complex systems and shows the linked behavioral graph. (iii) SimQuest (van Joolingen, Wouter, de Jong, 2003) is an authoring system to generate learning environments that include embedded simulations and adaptive instructions that supports learning through discovery. (iv) GeoGebra, which seeks to connect geometry and algebra, which is a different objective from PhET/Netlogo, which seeks to support the learning of individual concepts. GeoGebra is thus closer to supporting RC, rather than individual concepts (geogebra.org). GeoGebra is an interactive and dynamic mathematics software that allows students to construct geometrical objects with points, vectors, segments, lines, polygons, inequalities, polynomials and functions, thus allowing students to dynamically link diagrams (geometry) and graphs and equations (algebra). These interfaces provide the learner with varying degrees of control over different parameters in one or more representations, where manipulations in that representation may reflect in other representations (such as graph, simulated physical object/system) in real time. For instance, PhET and NetLogo simulations provide varying degrees of control over variable (such as temperature, pressure, speed, etc.) values, where changes in the values reflect in other representations. Geogebra allows dynamic control over geometrical shapes that reflect in the equations. Changes in the equation(s), however, have to be made through an input (code) language. SimQuest learning environments are very flexible in terms of dynamic control over representations, the degree of which may vary from one module to the other. The structural, quantitative as well as dynamic relationships between ERs in these and other such interfaces are, however, often implicit, as the code is designed to simulate representations.

The learning effects of the different types of ER-based computer interfaces (for detailed reviews see Mayer, 2004; McElhaney et al., 2014) have been studied extensively. For instance, interfaces with dynamically linked multiple representations have been found to be not universally useful for learning (McElhaney et al., 2014), although learning is found to improve when students actively manipulate the ERs to produce an integrated format, rather than observing an already integrated representation (McElhaney et al., 2014). Recent reports also suggest that learner interaction and engagement with the interface alone may not be sufficient for conceptual learning and may need further explicit support from the teacher and/or peers (e.g. Danish et al., 2015; Enyedy et al., 2012). In summary, reports on the effectiveness of the current computer interfaces have been mixed, and there is consensus that the expected learning benefits from such computer interfaces have not been fully derived (de Jong & van Joolingen, 1998; Rutten, van Joolingen & van der veen, 2012).

We hypothesize that the mixed results on the effectiveness can be attributed to the following issues related to the current interface designs. Firstly, the existing computer interfaces mostly focus on improving concept learning rather than RC, and research has typically measured conceptual understanding. Intervention designs under such a paradigm often undermine the roles of multiple external representations, treating them merely as ‘tools to learn the concept’, than as a critical component of the target concept itself (the idea of constitutivity). As a result, the effectiveness of such interventions for improving RC and conceptual understanding is not clear. Interfaces specifically targeting the development of RC (e. g. Johri & Lohani, 2011; Stieff, Hegarty & Deslongchamps, 2011; Wilder & Brinkerhoff, 2007) among learners are few and far apart (Pande & Chandrasekharan, 2017). Secondly, as has been found through the theoretical investigation in chapters 2 and 3, most interface designs are broadly based on information processing theories of cognition, and place the emphasis on cognitive load and working memory capacity, rather than the cognitive mechanisms involved in processing and integrating ERs. Thirdly, the designs of these interfaces are largely influenced by the general usability principles from human-computer interaction design (Hutchins, Hollan & Norman, 1985), which are directly applied to the learning problem, without an understanding of how these interactivity principles translate to the learning scenario (Hutchins, Hollan & Norman, 1985).

Our interface design focuses exclusively on helping students with ER integration and RC development, based on the conjecture that ER integration builds on (biological) sensori-motor integration. We consider conceptualization (used synonymously with conceptual integration in this dissertation) is a more complex process (figure 5.1).

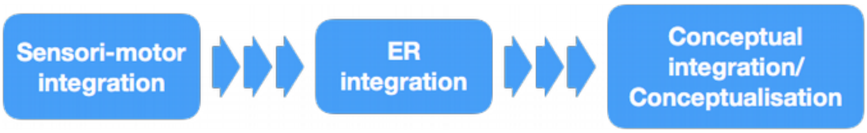


Figure 5.1 Conceptual hierarchy of cognitive processes involved in learning. There may be a continuum of processes with feedback loops in between them.

Based on the conjecture relating the three integration processes with each other, we designed an interface specifically for ER integration, and not conceptual understanding, although conceptual learning may happen as a side effect. Further, our interface design is rooted in the TUF model, as it interconnects the dynamics embedded in the three ERs of a simple pendulum system: a dynamic simple pendulum, its trigonometric equation, and a frequency graph. Thus, unlike simulation models with similar elements such as Net logo (Wilensky, 1999), PhET (Perkins et al., 2006) and SimQuest (van Joolingen & de Jong, 2003), our design is derived from basic research, particularly education research examining ER integration, and our own theoretical account based on recent models of cognition from the distributed and embodied cognition perspectives (Pande & Chandrasekharan, 2017). According to the TUF model, ERs of scientific phenomena and entities typically involve a physical system (such as a pendulum or a moving object) or its description, an equation capturing its behavior, and graphs that display the equations' output for different sets of values. ER integration requires the learner to develop an integrated internal representation of the three ERs – the phenomenon, its equation and the graphs. In order to do this, the learner needs to understand each representation separately, i.e., the dynamic and spatial nature of the phenomenon, the static and numerical nature of the equation, and the static and spatial nature of the graph. Next, the learner needs to relate the representations in pairs (revisit figure 3.1). For instance, the learner must understand that the equation acts as a *controller setting* the behavior of the phenomenon, that is, as the variables take on different values, each set of values causes the physical system to behave in a particular way. This understanding requires the student to "unfreeze" the static elements in the equation into dynamic behavior of the physical system. Conversely, s/he must understand that the dynamic and spatial phenomenon has been "frozen" to a different mode, namely the static and numerical equation. Further s/he must understand how this "freezing" has been done by capturing static states of dynamic ERs into symbolic elements.

In order to start the "unfreezing" of ERs and to generate dynamics from statics, the motor system in the brain would need to be (covertly) activated to run in simulation mode, as this neural system offers the closest approximation for generating movements in imagination. It is plausible that this activation process is difficult for novices to do and (new media-based) educational technology interventions that allow manipulations on the ERs could help trigger this activation and thus begin the "unfreezing" process. This is the key theoretical idea behind the design of our computer interface. I highlight that this approach differs from the designs suggested by the cognitive load account, where *manipulation of ERs is not the central feature of the intervention*. Our approach is also in agreement with the ER "incorporation" idea, developed in recent work in distributed cognition (Chandrasekharan, 2014; Chandrasekharan & Nersessian, 2015), as it suggests that manipulation of the ERs could be a way of promoting the "incorporation" of ERs by the imagination system, thus forming a smooth coupling between internal and external representational components, where changes in one are integrated directly into the other. Moreover, actions done on dynamic ERs would help students understand "freezing" the different ER states and hence improve integration. The central idea is that actions and manipulation require integrating multiple cognitive, perceptual and proprioceptive inputs and feedback loops, and so actions and manipulations performed on ERs in an interface would also trigger the neural networks involved in integration and help in integrating the ERs. This view goes beyond the standard principle that interaction is good, by providing a mechanism explanation for why interactivity provided by new-media technologies might help improve understanding and integration of ERs, and why understanding and integration is limited with static media (Majumdar et al., 2014).

One central feature of the system that is derived from this reasoning based on basic cognition principles and research is the full manipulability of all the ERs in the interface*, including equations.* This is a design requirement emerging from the theoretical model, as the model suggests full manipulabiliy would promote integration of ERs. This design principle is derived from an embodied cognition idea – that actions and manipulation, and feedback based on these, i.e. motor control, requires integrating multiple cognitive and perceptual inputs as well as feedback loops, suggesting that actions and manipulations performed on ERs in an interface would trigger/prime the neural processes involved in integration of inputs, which would in turn help in integrating the ERs as well. Apart from making the equation components manipulable. This theoretical approach also introduces in the system the controller role of the equation, where the full manipulable equation acts as a controller for the states of the other ERs, a feature not seen in standard simulation models mentioned above, which do not present the equation as a manipulable entity fully connected to other manipulable ERs. In this design, students control and 'enact' the equation, and integration is hypothesized to result from this control feature. Testing the development of ER integration based on this design thus also involves testing these hypotheses, and by extension, the cognitive theory that underlies it.

The learning objective of this interface was to support students in developing:

- An enactive understanding of each representation (enaction).
- A dynamic understanding of equations and graphs (‘unfreezing’ static ERs – imagination).
- An ability to capture in imagination static states of dynamic ERs at will (‘freezing’)
- An understanding of equations as controllers.

- An integrated internal representation, consisting of the physical system, equation and graph.

This study employed a design-based research (DBR) approach, involving iterative cycles of design, development, deployment/testing, analysis and redesign (Cobb et al., 2003; Wang & Hannafin, 2005). It emphasizes an iterative research process, where the theories, design principles and (technological) solutions systematically evolve across iterations, 'leading to a better understanding of the process of intervention' (Amiel & Reeves, 2008). The main research goal of this work was to test whether a naïve student can understand the relationships between dynamically linked ERs and integrate them through embodied interactions. Another major objective of this project was methodological, particularly developing an effective strategy (or set of strategies) to analyze student interaction with such an interface, to unearth patterns of behavior, primarily related to gaze and mouse-control, that could be possibly linked to ER integration/RC. This aspect of the project thus seeks to extend the results and methods from the chemistry education project. The DBR project involved two design-testing iterations, where findings, specifically related to interactivity-related design features, from the first testing phase were used to revise the design in second iteration.

The following are specific research questions this project sought to answer.

- After interacting with the interface, can naive learners imagine and describe the dynamic relationship between the following ERs of an oscillation system:
    - simulation of a physical system and its graph,
    - simulation of a physical system and its equation, and
    - equation and graph?
- What patterns of interaction are related to successful ER integration? How are interactivity and ER integration related?

- What kind of interactivity is desirable for ER integration?

This study involved two design-testing iterations.

5.1 Evolution of the simulation design: Iteration One

The first design iteration of the interactive computer simulation interface consisted of three versions (1.0, 1.1 and 1.2). The main design principles were: complete manipulability of all the ERs on the interface, and a sense of control over the ERs and their behavior experienced by the learner (Kirsh, 2010; Kirsh & Maglio, 1994; Chandrasekharan, 2009). Version 1.0 of the interface included simple 2-dimensional representations of a pendulum, a general form of the differential for the motion of a simple pendulum followed by its specific form where the initial angle and length are manipulable (-45 to +45 degrees for angle; 0.1m to 1.5 m for length), and a sine-curve. Each representation was presented in separate panels as shown in figure 5.2. The programming tool Processing was used to develop this version.

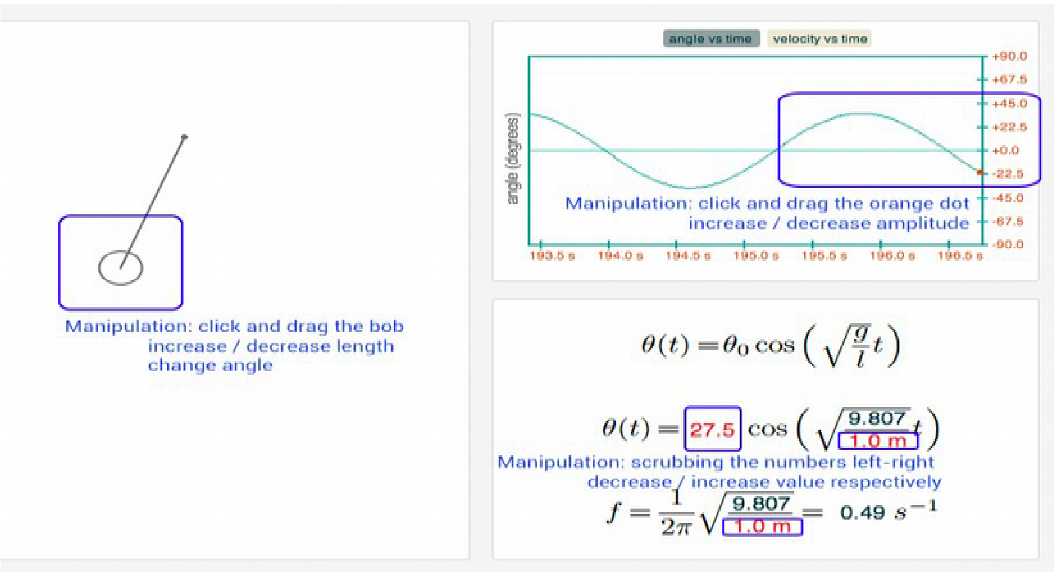


Figure 5.2: Version 1.0 of the interface design. The three components (pendulum, graph and equation) are presented in separate panels on the computer screen. Each of them is manipulable, and the instructions for manipulation are presented inside the respective panel, near the respective representation. To manipulate: (a) the pendulum, click and drag the bob to change the length/angle; (b) the graph, click and drag the orange dot on the curve to change the amplitude; (c) the equation, scrub the numbers highlighted in red to change the values.

In this version, the focus was on the following action sequence a learner could perform in order for the integration of the ERs and embodiment to occur: 1) Manipulate the pendulum to a particular length and release the pendulum at an initial angle to begin the simulation. The equation of oscillation along with its graph is displayed simultaneously. 2) Change the period in the equation of oscillation. The pendulum's length changes and it oscillates at this new period. The graph is updated accordingly. 3) Manipulate the graph to change the frequency of the sinusoid. The change in frequency is translated to a change in length of the pendulum by the simulation which updates the equation and pendulum accordingly. 4) Repeat the same manipulations for initial angle. Table 5.1 outlines the design principles and their operationalization into respective interface design features implemented (Majumdar et al., 2014).

Table 5.1: Design principles and their operationalization into design features.

Principle	**Operationalization**
Multiple representations provide different perspectives about the same phenomenon they represent and are complementary to each other (Ainsworth, 1999 & 2008).	The interface has three representations of the oscillation phenomenon – a simple pendulum, an equation and a graph.
External representations allow processing not possible/ difficult to do in the mind (Kirsh, 2010).	The interface three external representations. The simulation plots the graph of the equation/motion of the pendulum for various lengths and initial angles of the pendulum in real time, thus simulating the corresponding states of a representation into others.
Cognition emerges from ongoing interaction with the world (Brooks, 1991).	The interface is fully manipulable, i.e., the learner can control the pendulum, equation and graph, to see how change in each affects the other elements.
Action patterns can activate concepts, hence actions and manipulations of the representations should be related to existing concepts (O'Malley & Soyer, 2012).	The learner can interact with the pendulum by changing its length and initial angle by clicking and dragging the mouse. The parameters in the equation can be changed by scrubbing the numbers highlighted in red.

In the next version, version 1.1, emphasis was given to improving the nature of interactions: To do this, the previous version 1.0 was examined against the existing literature, and the action sequence was modified, in order to (a) incorporate structured feedback to the learner during manipulation, as feedback is necessary to complete any action, and (b) ensure too much is not happening at the same time (reduce split attention). A separate panel for displaying instructions was added to the top right corner of the screen. The presentation of representations is now distributed across three screens unfolding serially as the student advances through the different screens. A next/back button was added to facilitate navigation between the three screens. The student interacts with the pendulum in the first screen. Then the equation is added to the (second) screen; the student can manipulate the equation and pendulum to see how they are connected. This helps the student understand the relation between the physical system and its equation. Finally the graph for the motion of the pendulum is introduced in the third screen. The manipulations of changing the frequency and initial angle from the graph were removed to reduce the complexity of interaction, rendering the graph non-manipulable. The graph design was simplified; numerical details were reduced to make it more presentable to grade VII students. A play/pause button was introduced just below the pendulum panel on every screen to control the simulation at will (e.g. pause the simulation at anytime, change the desired parameter and resume/play the simulation to observe changes in the representations). The amplitude was restricted to a maximum of 45 degrees, the length could vary between 0.1m to 1.5m, and the simulation would stop at 15 seconds. Another important variation was the introduction of a separate panel for displaying the instructions and information about the ERs. The panels were rearranged to achieve enough and uniform separation between the representations while also keeping in mind that the panels fill the screen well. The modified interface version 1.1 is shown in figure 5.3.

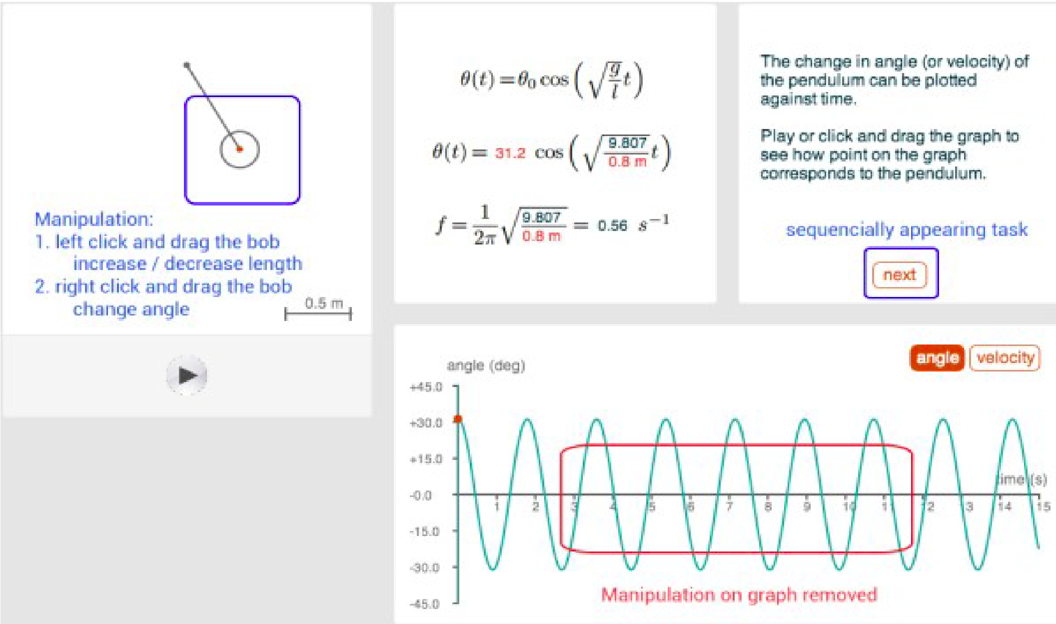


Figure 5.3 Screen 3 of version 1.1 of the first interface design iteration. Incremental introduction of the ERs across three screens; pendulum is introduced in the first screen, the second screen includes pendulum and equation, followed by introduction of the graph. Only pendulum and equation are manipulable. To manipulate: (a) the pendulum, click and drag the bob to change the length/angle; and (b) the equation, scrub the numbers highlighted in red to change the values. Separate panel added for text instructions. A play/pause button was introduced just below the pendulum panel on every screen to control the simulation. A next/back button was added to facilitate navigation between the three screens. The amplitude was restricted at 45 degrees, while the length could vary between 0.1m to 1.5m. The simulation would stop at 15 seconds.

Finally, in version 1.2, the focus was on extrinsic motivation to ensure rich interaction, and making the interaction even more intuitive and aesthetic. To ensure that the students actively interact with all the three representations and to facilitate a comprehensive exploration and ER integration, three learning tasks were introduced. Figure 5.4 shows a screenshot of this new interface version.

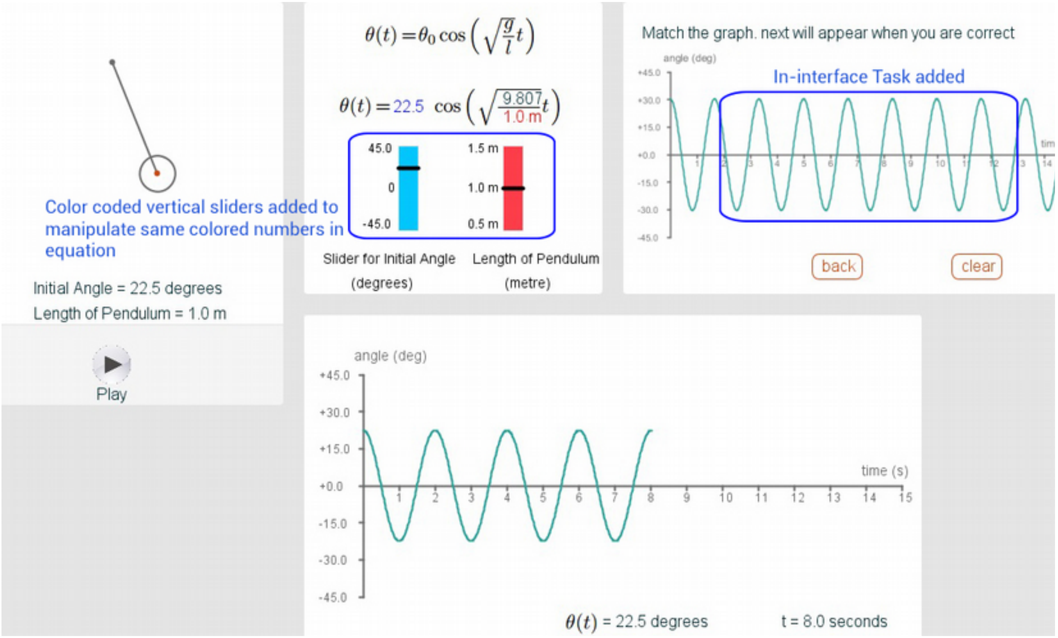


Figure 5.4 Screen 4 of version 1.2 of the computer interface with all the representations and task 1. From screens 1-3, respective instructions would appear in the place of tasks. Note the addition of: (a) vertical sliders below the equation, (b) Initial angle and length values below the pendulum panel, and (c) learning tasks – here in screen 4, task 1 is displayed. The colors of the values in the equation match the colors of sliders to maintain uniformity of meaning.

The interface now has six screens. Screen 1 displayed the manipulable pendulum. Screen 2 showed the manipulable pendulum and equation/sliders, and screen 3 had the manipulable pendulum, equation/sliders and graph. The remaining three screens – 4, 5 and 6 had learning tasks (Appendix 6). In each learning task screen, the instruction panel was replaced by a task panel that displayed a screenshot of a pre-simulated curve (corresponding to different settings or combinations of the length and initial angle of the pendulum), and learners were required to manipulate the equation and/or pendulum to generate a curve that matches the given curve by playing/pausing the simulation as required. The complexity of the curve to be reproduced increased sequentially across the three tasks. For instance, task 1 would require setting the initial angle and length (say at Θ=30 degrees and l=0.7m) only once in the beginning (i.e. at t=0) and press the play button to generate the curve; for tasks 2 and 3, one had to change the parameters more than once – first in the beginning (at t = 0), and then again after the simulation had run for a certain time (say at t = x, where x could be anything between 2-14 seconds).

To change the parameter the second time, one would have to run the simulation with the initial settings, pause it after a while (at t = x seconds, depending on the target curve), change the parameters as required, and then resume the simulation to complete the curve. A mechanical error of +/- 1 seconds in playing/pausing the simulation was allowed in the simulation code. As a result, the student had a 2 second window to press pause to change a parameter in order to achieve the target curve.

Screens 1-3 were the ‘free-exploration’ phase, as the learner is exploring the interface with no specific goal. Screens 4-6 marked the task-specific exploration phase, where the learner explored the interface as s/he solved the learning tasks requiring specific manipulation.

These learning tasks were designed to encourage both actions (manipulating the pendulum and the equation) and imagination or (mental) simulation (anticipating the structure of the curve given a set of states of the pendulum and the equation) in the learner, and to achieve the learning outcomes of RC, which included an understanding of equations as dynamic entities and controllers and an integrated internal representation, consisting of the physical system, equation and graph.

Another major change in this version was that the equation scrubbers (where left and right “scrubbing” actions were required to change the equation parameters) were replaced by vertical sliders. This important change in the design was inspired by a recent finding in numerical cognition, that numbers are grounded by associating small magnitudes with ‘lower’ space and larger magnitudes with ‘upper’ space (Fischer, 2012). These interactions, particularly with equations, distinguish our interface from other variable manipulation simulations (e.g. PhET; Perkins et al., 2006) where the manner in which values are changed is not relevant, whether by slider, input box or multiple options. In fact, a PhET pendulum simulation (Perkins et al, 2006) does not have the equation and graph, and there is only one interaction on the pendulum, while the other variable is manipulated via horizontal sliders. By contrast, our interface is specifically designed to make the learners do certain actions which mimic the behavior of the system so that the system can be 'enacted' -- the learning is through a form of participation with the system.

Table 5.2 summarizes the new design principles and corresponding design features introduced in version 1.2; table 5.3 lists some of the design features that were revised in this version and their corresponding design principles. Table 5.4 presents the highlights of the design feature comparison between the three versions.

Table 5.2 New design principles and respective features introduced in version 1.2

Principle	**Operationalization**
Features of the world are used directly for cognitive operations. Hence the interface features should support integration directly (Landy et al., 2014).	The interface has the physical system, equation and graph, along with different numerical values. The dynamic nature of elements, and their interconnections are made transparent, so that learners can integrate across spatial-numerical and dynamic-static modes.
The active self is critical for integration of features (Reed, 1988).	The interface is introduced with a task-specific exploration phase in which the learner must perform a set of tasks requiring specific manipulation of the interface. It was hypothesized that these tasks were sufficiently complex in order for the learner to actively engage in the problem solving, resulting in comprehensive exploration and manipulation of the interface by the student, so that the three representations are integrated.

Table 5.3 Revision to some design features in version 1.2 and the respective design principles.

Principle	**Operationalization**
Action patterns can activate concepts, hence actions and manipulations of the representations should be related to existing concepts (Fischer, 2012).	Vertical sliders are introduced to manipulate the equation. The values in the equation can now be increased or decreased by clicking and dragging the vertical sliders ‘up and down’ respectively. The interface seeks to make the learners do actions that mimic the behavior of the system, so that the system can be 'enacted' - the learning is thus through a form of participation with the system.
The interface should allow coupling of internal and	The task requires student to match a given graph. Learners change the parameters of the pendulum/equation to generate the graph, and

external representations (Chandrasekharan & Nersessian, 2015).	visually match the task graph to their graph. This develops learner's imagination and coupling between their internal model and the external representation.

Table 5.4 Highlights of the design comparison between the interface versions.

<table>
<tr><th>Version/ Feature category</th><th>Panel/ representation</th><th>1.0</th><th>1.1</th><th>1.2</th></tr>
<tr><td rowspan="4">Manipulation</td><td>Pendulum</td><td colspan="3">Clicking and dragging the pendulum</td></tr>
<tr><td>Equation</td><td colspan="2">Hovering over the numbers highlighted in red activates the scrubbers, clicking and dragging changes the values in the respective numbers</td><td>Vertical sliders added below the equation. Scrubbing function removed from the numbers. Numbers in the equation no more manipulable directly. Values in the equation can now be changed by clicking and dragging the sliders up and down.</td></tr>
<tr><td>Graph</td><td>Clicking and dragging the orange dot on the curve to change the amplitude</td><td colspan="2">Not manipulable</td></tr>
<tr><td>Instruction</td><td colspan="3">No manipulation</td></tr>
<tr><td rowspan="4">Extent of explicit link between the representations/panels</td><td>Pendulum</td><td rowspan="4" colspan="2">Minimal: Symbols, numbers and text not connected. e.g. No explicit link mentioned between Θ and 'initial angle' within or across the representations/panels</td><td>Moderate-to-optimum: Corresponding values for initial angle and length shown, symbols 'Θ' and 'l' not used</td></tr>
<tr><td>Equation</td><td>Phrases 'slider for initial angle (degrees)' and 'length of pendulum (meter)' introduced below respective slider</td></tr>
<tr><td>Graph</td><td>Corresponding Θ and t values shown</td></tr>
<tr><td>Instruction</td><td>No change</td></tr>
<tr><td rowspan="2">Position on the screen</td><td>Pendulum</td><td>Left half</td><td>Top left corner</td><td>Top left corner</td></tr>
<tr><td>Equation</td><td>Bottom half in the right half</td><td>Top centre</td><td>Top centre</td></tr>
</table>

Version/ Feature category	**Panel/ representation**	**1.0**	**1.1**	**1.2**
		of the screen		
	Graph	Top centre half in the right half of the screen	Bottom half of the right half of the screen	
	Instruction	Embedded within the respective panel	Separate panel, top right corner	
ER introduction	Pendulum	Simultaneous	Introduced first (screen 1)	
	Equation		After the pendulum (screen 2)	
	Graph		After both pendulum and equation (screen 3)	
	Instruction		Specific instructions present on each screen	
Active engagement		Learner’s will		Built-in graph-matching/learning tasks

This version of the interface (version 1.2) was considered ready for the first DBR exploration cycle. A two-group controlled pilot study was performed to address the following objectives and research questions.

5.1.1 Pilot study

Iteration 1 focused on the evaluation of usability and learning effects of the system, through a two-group controlled study.

5.1.1.1 Broad objectives and research questions

The pilot was conducted to understand:

- Are instructions (text-based guidance) necessary for manipulation (sensorimotor interaction)? What actions do the manipulation features afford?
- How easy is the interface to use (usability) and/or learn (learnability)? What aspects of the interface do the learners find difficult or problematic?

What aspects of the interface do students appreciate, dislike, and/or overlook? How can knowing this feed into the next iteration?

- How can we characterize the learner's sensorimotor interactions with the interface? What sensorimotor interaction patterns do the learners show, if any? If and how are these interaction patterns related to ER integration and its development? Answering these would particularly help in developing an analysis methodology within the larger embodied and distributed cognition account of RC proposed in chapter 3.

These broad goals led to outlining the following specific research questions (RQs):

4. After interacting with the interface, can naive learners imagine the dynamic relationship between the ERs, in the absence of (physical) manipulation and dynamics?
5. What are the differences in learner exploration of the interface, particularly in terms of manipulation or control, between text-guided and self-guided conditions?
6. What is the difference in learner exploration of the interface between the free-exploration phase (i.e. screens 1-3, before presenting the learning tasks) and the task-specific exploration phase (i.e. screens 4-6, while solving the learning tasks)?
7. How to analyze the sensorimotor interaction data in order to achieve insight into the process and mechanism of ER integration? What differences can be identified in the interaction (particularly eye-tracking and mouse-tracking) behavior between participants who are good at imagining the dynamic relationship between ERs after the interaction and the participants who are not?

5.1.1.2 Sample

12 students (6 female) studying in 7^{th} grade from two urban schools in western India participated in this study. This grade level was chosen because the oscillation concept is not introduced at this level, and the system thus presents only an ER integration problem to these students, and not a concept problem. The interface is designed primarily to cater to the development of RC and not the concept of oscillation.

Parents of the participants and concerned school teachers were informed that each participant would be playing a science game on a laptop using a mouse controller, that the game had stages of increasing complexity similar to any other game, that his/her eye and mouse interactions would be recorded during the playtime, and that s/he would be answering a set of questions related to the game after passing through all the stages. Written consent was obtained from the parents and the students, while the school teachers helped in logistics such as preparing the study time-table in relation to the school timings.

Half the students (text-guided group; 3 female; students code-named L1 through L6) received an interface which had text instructions of how to use the various manipulable features on the interface (e.g. sliders). The remaining students (self-guided group; 3 female; code-named L7 through L12) received an interface without these instructions.

5.1.1.3 Experiment protocol

Each participant sat in-front of a laptop, attached with a Tobii X2-60 portable eye tracker (Tobii Technologies, Sweden, sampling rate of 60Hz.), at a distance of 50-70 cm (figure 5.5). Students who indicated that they were not comfortable with computers were given a few minutes to practice with the mouse before they were introduced to the interface. Once the student was ready, the preset eye-tracker (Tobii X2-60) was calibrated and the interface window was opened for student interaction. The eye-tracker also logged mouse-events.

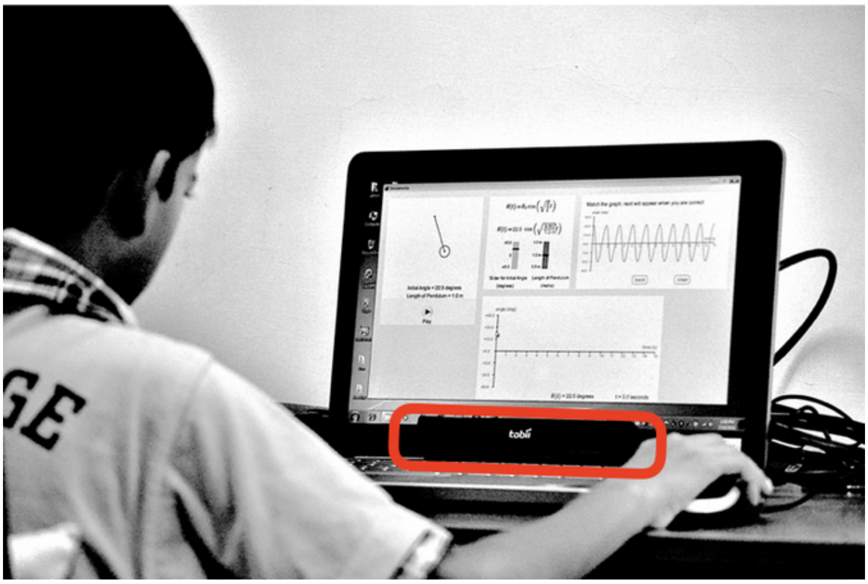


Figure 5.5 Experimental setup: A participant can be seen interacting with the interface on the laptop using a mouse. At the bottom of the screen is fixed a Tobii X2-60 eye-tracker (highlighted in a frame) which records the student's gaze data during the interaction.

Participants in the text-guided group interacted with the interface version 1.2. The self guided group also received the same interface except that all the text instructions from the interface were removed. The latter group was only told that they could manipulate items on the screen with both left and right clicks of the mouse. Each student was allowed to work independently with the interface for as long as s/he wished, advancing through the screens and tasks by clicking the "Next" button. The experimenter only intervened when students had a question, and provided only minimal hints appropriate to their condition namely, text-guided or self-guided.

Once the student indicated that the tasks were completed or that s/he wanted to quit, s/he was interviewed regarding her/his background (general information such as family, area of residence, whether and how often s/he interacts with computers/ mobile phones, favorite subjects in the school and hobbies), and impressions of the interface (e.g. what feature s/he liked or disliked, how the interface could be improved, etc.) The student was then administered some pencil/paper-based tasks to test if s/he imagines or simulates the dynamics of the interface (or her/his sensorimotor interaction) in absence of interaction, and hence answer RQs 1 and 4.

5.1.1.4 ER integration tasks

Stimuli that prompt imagination of phenomena readily activate (existing) mental simulations (e.g., Schwartz & Black, 1999). To prompt student imagination or simulation and answer RQ 1, six pencil/paper-based tasks were developed (see Appendix 7). One of the tasks showed a point on the curve and the student had to indicate the corresponding position of the pendulum i.e. if the pendulum would be: on the right side (of the normal), on the left side (of the normal), in a vertical position, or in a horizontal position. There were three multiple choice questions of this kind. In the remaining three tasks, the student was shown a screenshot of the pendulum in a certain position and was asked to spot/mark point(s) on a curve that would approximately correspond to (the depicted state of) the pendulum. Solving these would ideally require imagining the dynamics of the interface or the ERs with respect to one's sensorimotor interaction with them, and capturing specific state(s) of a representation in relation to a given state of another representation.

As reflected in the RQ1 statement as well as the ER integration task design, equations were deliberately avoided in these tasks due to their complexity in relation to the participants' background and exposure to scientific representations; it was important to keep in mind that the sample comprised of 7th grade students, who probably had no previous 'perceptual' experiences with equations such as the one presented in the interface. I also did not want the students to feel uncomfortable and/or discouraged by facing tasks they may not (be able to) solve.

Student responses to the tasks were subjected to accuracy assessment.

5.1.1.5 Sources of data

(a) Student responses to the ER integration tasks: To answer RQs 1 and 4, students responses to the six ER integration questions were evaluated for accuracy. Good, average and poor performers were identified on the basis of the number of correct answers provided (Correct number of answers to qualify as a: good performer = 5-6; average performer = 3-4; and poor performer = 0-2).

(b) Researcher observations and notes: This included information about the status of completion of the learning tasks by the students, and an unstructured log of student behavior and facial expressions while s/he interacted with the interface (e.g. what features of the interface go overlooked, or are utilized more, etc.), as well as her/his responses to the interview questions.

(c) Sensorimotor interaction data: Student eye movements recorded using a Tobii X2-60 eye-tracker, capturing students' gaze behavior as they explored the interface. The eye-tracker also records mouse-event data (e.g. whether right or left click, location of the click on the screen, etc.). The eye and mouse data were the main sensorimotor data collected. These data could help decipher the dynamic interaction process involved in integration. Both gaze and mouse-click data are available as dynamic screen-activity recordings in the software used to run the eye-tracker. The data can also be extracted as raw data in the form of log-sheets from the software for a customized analysis.

Data of one student (L12) from the self-guided group were not considered for analysis due to gaze-data file corruption which happened during the experiment trial.

5.1.2 Development of interaction analysis

In addition to how the mouse was moved during interaction with the interface, the focus was also on *the task-oriented movements of the eye* (and not how attention was captured by visual elements). Eye movements, in this approach, are treated similar to mouse movements, and are considered as sensorimotor actions that can lead to integration. Based on this view, a novel analysis strategy of eye movements as actions was developed, in order to understand how interactivity is related to learning. This analysis allows studying action patterns that are correlated with ER integration (as measured by ER integration tasks).

Tobii Studio-3.2 (eye-tracking data analysis package from Tobii Technology) was used to extract and process the raw interaction data. The unprocessed or raw data are in the form of dynamic screen-activity recordings (video) and time-stamped logs of gaze as well as mouse activity. For a detailed analysis, these data needed filtration and refinement. The first step was to generate, for each participant, separate segments of interaction with each screen (see section 5.4.2 of Tobii Studio User's Manual Version 3.4.5 for procedure details). For a total of six screens, the segmentation yielded a total of six segments per participant. Segments for the participants of each group (text-guided or self-guided) for each screen were compiled to generate a scene that contained interaction data for all the participants in that group for that particular screen (Section 5.5 of Tobii Studio User's Manual Version 3.4.5). Each scene has two elements: dynamic interaction data for all the participants for a screen, and a static image of the respective screen as a background on which all the compiled gaze data is either superimposed to generate a static visualization (such as a heat map depicting gaze or mouse activity distribution) or dynamically played and visualized as required. This yielded a total of six scenes per group (total 12 scenes; e.g. Scene for screen-1 for the text-guided group had interaction data of all the text-guided participants with the screenshot of screen 1 in the background; scene for screen-2 for the self-guided group had the group's participants' interaction data for screen 2 laid out on the image of screen 2, and so on).

Once the interaction data of the respective participants for each screen were available on a common image of the respective screen, specific areas – known as the areas of interest (AOIs) – could be defined for isolating and capturing the gaze and/or mouse activity happening in those areas.

To perform statistical analysis at a different scale of detail, AOIs were generated at two different layers. First, an overall/general layer of AOIs (figure 5.6) was defined for each of the six scenes to achieve an understanding about the overall distribution of the interaction data across the three different representations (pendulum, equation and graph) for the participants in that group.

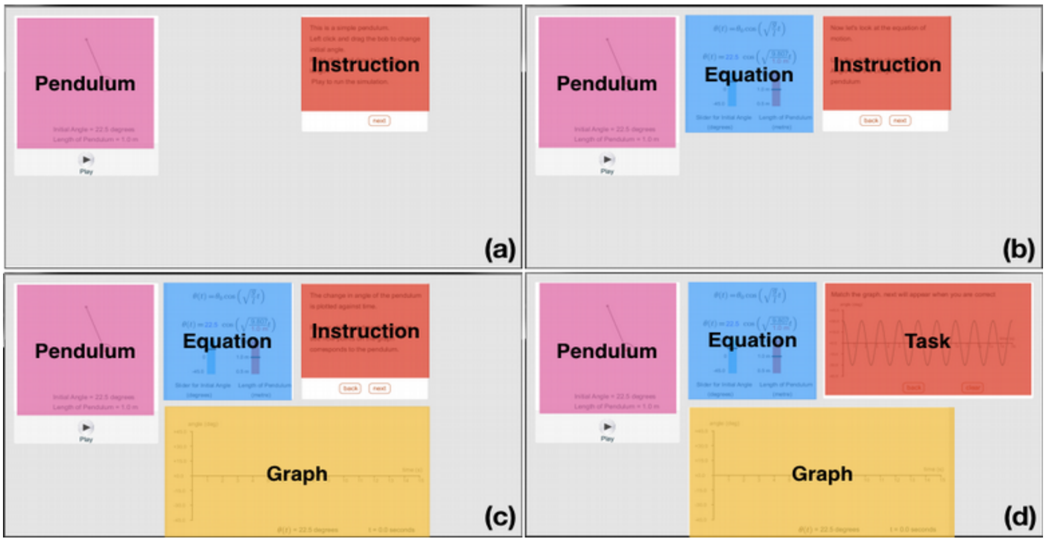


Figure 5.6 General/overall AOIs for: (a) screen 1 has two AOIs – pendulum and instruction, (b) screen 2 has AOIs for pendulum, equation and instruction, (c) screen 3 has four AOIs – pendulum, equation, instruction and graph, and (d) screens 4-6 all have four AOIs each – pendulum, equation, instruction and task. Note that for the self-guided group, the instruction panel was blank for screens 1-3.

The second layer of AOIs, exploring participant interaction in a more detailed and specific manner, was defined based on the nature of information in a representation panel. This included two subsets of AOIs across the ERs: spatial and numerical AOIs (see figure 5.7), formed to explore the integration between spatial and numerical aspects of ERs.

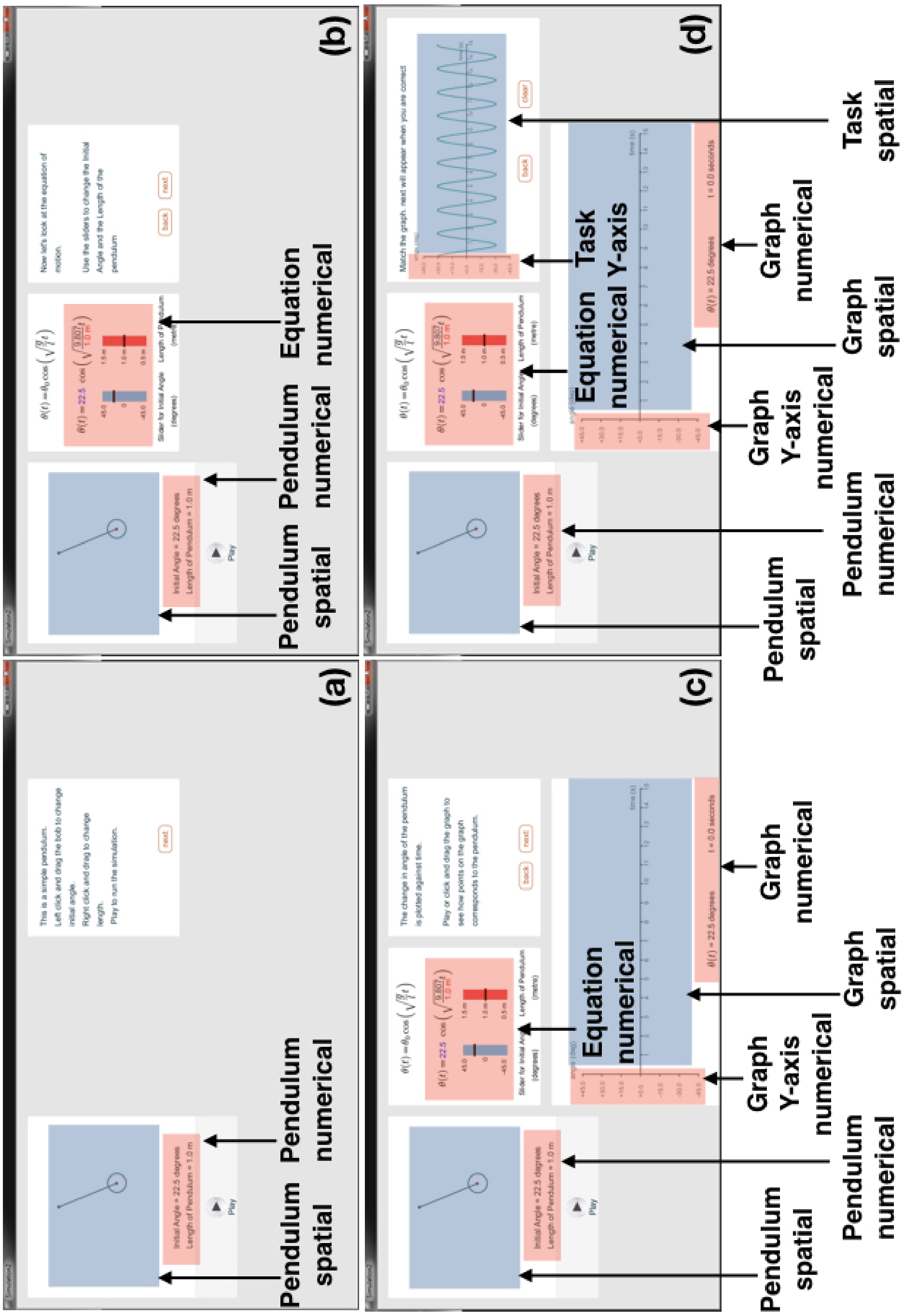


Figure 5.7 Specific AOIs for: (a) screen 1, (b) screen 2, (c) screen 3, and (d) screens 4-6. These specific AOIs are classified into two categories – spatial AOIs (in blue) and numerical AOIs (in red). Spatial AOIs concern the spatial information in a representation/panel while the numerical AOIs are concerned with information of numerical nature in a representation/panel. This division is to capture the relative gaze activity between two epistemologically different aspects of the ERs. Note that for the self-guided group, the instruction panel was blank for screens 1-3.

Using the AOI-based data, four levels of analysis were devised to account for interaction behavior at different depths and extent of abstraction. Figure 5.8 shows a schematic of the levels of interaction data analysis.

At level 1 (frequency distribution analysis), the following statistics were generated for each participant per AOI, per screen: (1) Total visit duration/time spent, (2) visit count, (3) mean fixation duration, (4) fixation count, (5) total number of mouse-clicks, and (6) mean number of mouse-clicks (Tobii Technology, 2014; Appendix 4). Statistics from individual participants for each screen were tabulated into two groups, depending on the nature of participants' exploration (text-guided versus self-guided groups). Combined statistics for all the participants in that group were then used for graphical analysis.

For level 2 (sequence analysis), student interaction with the interface was conceptually divided in two cycles using mouse-clicks as a clustering factor: a perception-action cycle, and an imagination or simulation or thinking cycle. The perception-action cycle comprises of: students manipulating features on the screen (e.g. sliders) and observing changes in other representations. The latter happens when the pendulum simulation is played or paused; students in this cycle attend to the static or dynamic features on the screen (e.g. length/angle values and the graph) in an expectation of how the representation(s) may behave (i.e. imagine or simulate the dynamics) with respect to the manipulation performed. In this level, sequences of fixation events and mouse click events are determined, and grouped as events occurring either in the perception-action cycle, or events occurring in the simulation/imagination cycle.

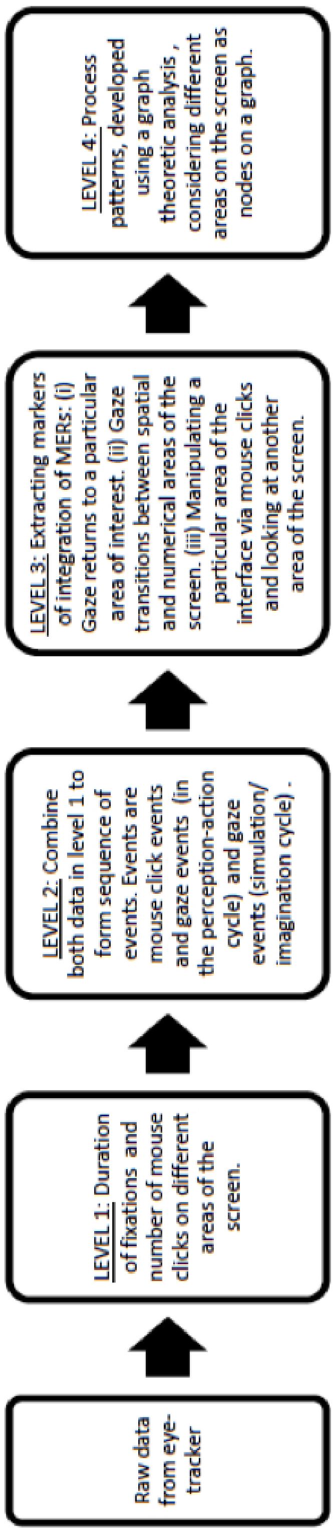


Figure 5.8 Levels of interaction analysis.

At level 3 (local integration marker analysis), local sensorimotor markers that signify integration are defined; the data are processed to achieve more abstraction to identify these markers of integration. An example of a local sensorimotor marker is gaze transitions (discussed in section 4.1.8) between, say a numerical AOI (e.g. the equation) and a spatial AOI (e.g. the pendulum); this specifies integration between numerical and spatial modes. A second example is returns, i.e. a learners' gaze returning to a particular AOI (say 'A') after going elsewhere (say AOI 'B'). Such sequences of events would be characterized as an A-B-A return, where A and B are any two AOIs. This indicates that the learner is retaining a particular feature in memory and returning to it. If there are multiple AOIs involved, the returns would look like A-B-C-A, A-B-C-D-A, and so on, where A, B, C and D are different AOIs. In the pilot, only A-B-A type of returns were explored. The third example of local markers is a consecutive gaze-mouse or mouse-gaze activity, which may include the learner manipulating a representation on the screen (e.g. pendulum) and looking at another representation or AOI (e.g. graph), as this indicates the integration of two representations via the systematic variation offered by control. The local markers cut across the two cycles of interaction defined in level 2. Once these markers were obtained, a goodness measure for these markers is defined by comparing against marker values of experts – learners who perform well on the ER integration tasks.

The final level of abstraction (level 4, global integration marker analysis) in sensorimotor interaction data analysis involves generating global process patterns of how the learners interacted with the interface, using a graph theoretic framework, wherein the AOIs are the nodes and the transitions between the various AOIs are the weights of the branches.

Data analysis level 1 utilizes the general scheme of AOIs (represented in figure 5.6), whereas for levels 2-4 of data analysis, the data related to interaction activity in the specific AOIs (figure 5.7) are used.

5.1.3 Results

5.1.3.1 Performance on ER integration (RQ1) and learning tasks

The text-guided and self-guided groups did not differ in the accuracy of their responses to the learning tasks. Table 5.5 presents the total number of correct and incorrect responses to each task by the two groups. Irrespective of their experiment condition, both the types of tasks seemed equally challenging to the students.

Table 5.5: Between-group comparison: ER integration task accuracy (C=correct, W=wrong). Highlighted in dark grey rows are good performers (L3, L5 and L7), white cells present the accuracy data of average performers (L2, L8 and L11), and students highlighted with light grey cells are poor performers.

	Student	**Task 1**	**Task 2**	**Task 3**	**Task 4**	**Task 5**	**Task 6**	**Total Correct**
Text-guided	**L1**	W	W	W	W	W	C	1
	L2	W	W	C	W	C	C	3
	L3	C	C	C	C	C	C	6
	L4	C	W	W	W	W	W	1
	L5	C	C	C	W	C	C	5
	L6	W	W	W	C	W	W	1
	Total Correct	3	2	3	2	3	4	17
Self-guided	**L7**	C	C	C	W	C	C	5
	L8	C	W	C	W	W	C	3
	L9	W	W	W	W	W	W	0
	L10	W	W	W	W	W	C	1
	L11	W	C	C	W	C	C	4
	Total Correct	2	2	3	0	2	4	13

Some members in both the groups were more accurate than their group-mates. Three students (L3, L5 and L7) were found to be ‘good performers’; two belonged to the text-guided group while one student was from the self-guided group. Three students showed average performance on the tasks (L2 from text-guided group, L8 and L11 from self-guided group). The remaining five students were identified as poor performers (L1, L4 and L6 from the text-guided group, and L9 and L10 from self-guided group).

The students were able imagine the dynamic relationship between pendulum and graph, in the absence of (physical) manipulation and dynamics; however, the data were insufficient to determine whether and to what extent did the (i) student interaction with the interface mediate their answers to the ER integration tasks and hence their ability to imagine the dynamic relationships between ERs, and (ii) how the provision of instructions affect ER integration.

There was no difference between the two groups in performance on the learning tasks. Table 5.6 shows between-group comparisons of: number of students who completed, gave up or did not attempt a learning task, and the average time spent by the students per task.

Table 5.6 Between-group comparison of performance on learning tasks

Learning tasks		**1**		**2**			**3**
Group	**Status of completion**	**No. of students**	**Average time-spent (min)**	**No. of students**	**Average time-spent (min)**	**No. of students**	**Average time-spent (min)**
Text-guided	Completed	5	8.934	3	19.58	3	11.51
	Gave-up	1	30.42	1	5.34	2	8.94
	No attempt/ program error	-	-	2	2.89	1	-
Self-guided	Completed	5	10.51	3	14.36	2	17.26
	Gave-up	0	-	2	16.74	1	13.49
	No attempt/ program error	-	-	-	-	2	-

There was no noticeable correlation between student performance on the ER integration tasks and the status of their completion of the learning tasks; however, this does not mean that there is no correlation between ER integration and student exploration of the interface. This point is further explored in later sections.

5.1.3.2 Level 1: Activity frequency distribution analysis (RQs 2 & 3)

The average time spent looking at each AOI by a group is presented in figure 5.9. A clear pattern, common to both the groups, was that the total time spent looking at a screen increased markedly during the task, compared to before the tasks. Secondly, in free-exploration, the text-guided group spent more time looking at the screen than the self-guided group. Though the fixation duration of the two groups averaged across all tasks are comparable, the text-guided group spent more time in task 1 and successively kept looking lesser, while the self-guided group spent less time in task 1 and successively kept looking more. Further, this data showed that during the tasks in screen 4-6, both the groups spent comparable time looking at the equation. Both groups also spent more time looking at the equation than the pendulum. However, the text-guided group spent more time looking at the pendulum than the self-guided group.

Figure 5.9 Between-group comparison: Average time spent (visit duration) in each AOI screen-wise.

Next, figure 5.10 presents a between-group comparison in the average number of mouse clicks in each AOI.

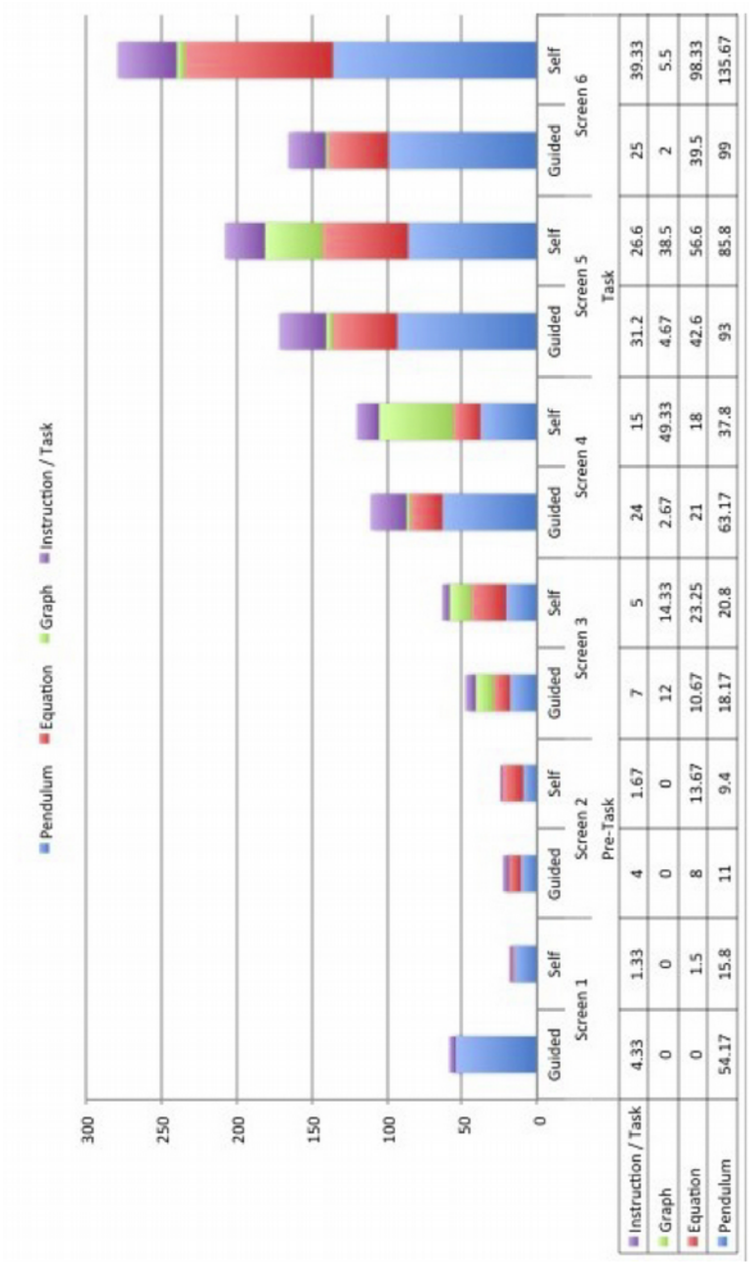


Figure 5.10 Between-group comparison: Average number of mouse clicks in each AOI per screen.

The figure shows that, consistent with increased fixation time, there was a significant increase in the number of mouse clicks as the participants advanced to the tasks from free-exploration. Secondly, starting from screen 2, the self-guided group had a higher number of total clicks on each screen than the text-guided group. Further, except for screen 4 in which both groups were similar, the self-guided group clicked more on the equation AOI (containing the sliders) than the text-guided group. The self-guided group also clicked more on the graph AOI than the text-guided group. Mouse-clicks for the text-guided group were fairly distributed across the different representations in the interface.

The locations of the mouse clicks on the different AOIs of the interface during free-exploration and tasks are shown in figures 5.11 and 5.12 respectively. As seen in figure 5.11, the text-guided group manipulated the pendulum more than the self-guided group, while the self-guided group manipulated the equation more than the text-guided group.

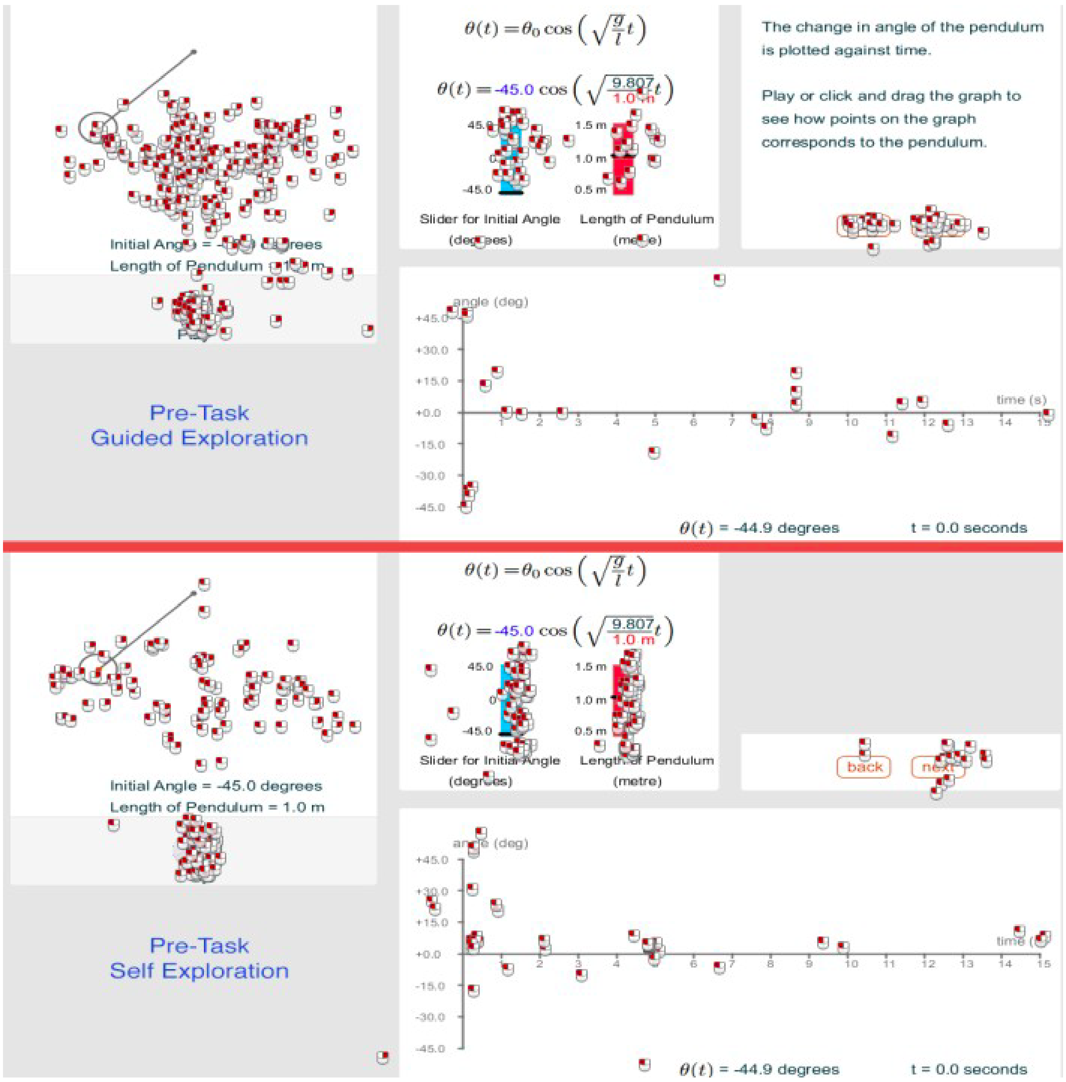


Figure 5.11 Between-group comparison: Average number of mouse clicks on screens 1-3 combined.

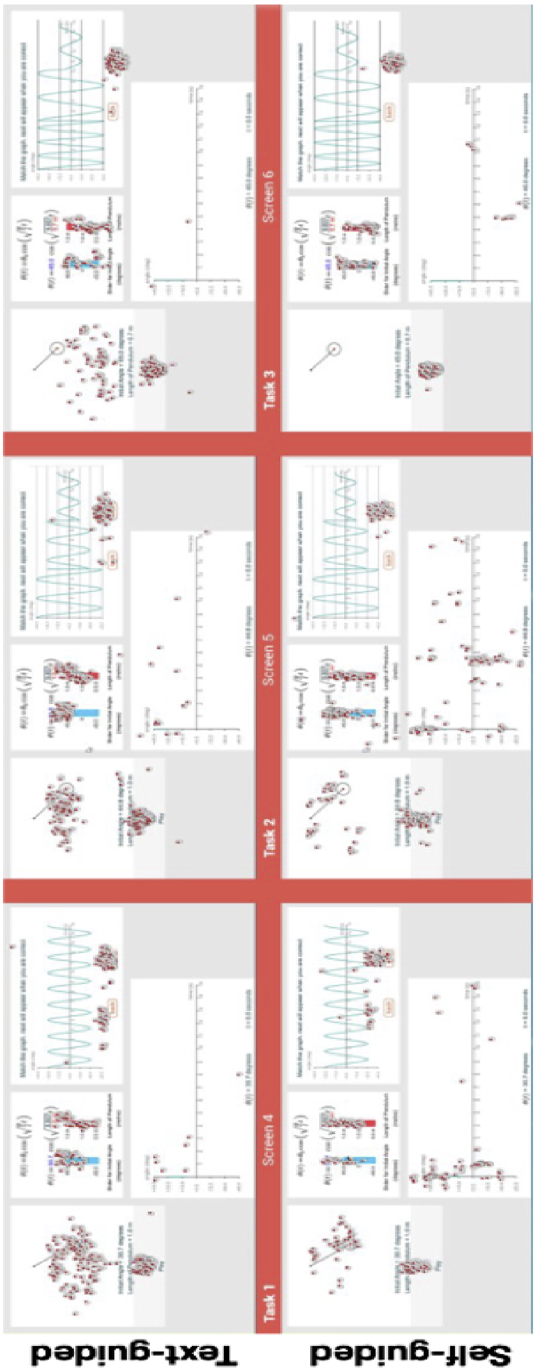


Figure 5.12 Between-group comparison: mean mouse clicks during the learning tasks. Tiny red dots on the screenshots are mouse-clicks. The text-guided group has more and denser interactions with the interface while solving the tasks than the self-guided group on average. However, the self-guided group seems to click on the graph a lot more than the text-guided group. Researcher's notes and analysis of time-spent per screen reveals that students did not interact with the interface during screens 1-3, and as a result, may not have learnt about the manipulations – particularly that they need to hit the play button to generate a curve. These students would keep swinging the pendulum

with mouse instead, wonder why the graph would not appear, and end up clicking on the graph several times, thinking that doing so would generate one.

Figure 5.12 shows that the text-guided group used both pendulum and slider manipulation for all the tasks, however their pendulum usage was considerably lesser than the sliders. They rarely clicked on the graph area. The self-guided group also rarely used the pendulum for manipulation, reaching an extreme case in screen 6, in which they did not use the pendulum at all, and instead, performed the entire task with the help of sliders. This group clicked on the play button more than the other group, especially in screen 6. They also clicked on the graph region several times in screen 4, but when they reached screen 6 they stopped this action. These behavioral differences between the groups are summarized in table 5.7.

Table 5.7 Relative explorations of participants in text and self guided conditions in comparison with an ideal case. Direction of the arrow indicates frequency (up = increased, down = decreased).

Exploration	**Screens 1-3 (free-exploration)**		**Screens 4-6 (learning tasks)**	
Activity	**Look**	**Click**	**Look**	**Click**
Good/ideal	↑	↑	Focused	↓
Text-guided group	↑	↑	↑	↓
Self-guided group	↓	↓	↓	↑

As shown in table 5.7 (RQs 2 and 3), based on the reported behavior of experts in the RC literature, good exploration could be defined as one in which there are more looks and clicks during the free-exploration phase, and focused look and fewer clicks during task-specific exploration phase. Comparing the exploration of the two groups against this canonical exploration, it was found that the text-guided group had more looking and clicking activity during free-exploration. During the task, however, neither group demonstrated 'good' exploration, though, the text-guided group showed exploration closer to the 'good' exploration in terms of less clicking. This indicates that the text-instructions are necessary to help students explore the interface, but not sufficient to achieve good exploration.

To answer RQ 3, which is, "What is the difference in learner exploration of the interface between the free-exploration phase (i.e. screens 1-3, before presenting the learning tasks) and the task-specific exploration phase (i.e. screens 4-6, while solving the learning tasks)?" it was found that in both groups the looking and the clicking increases after the task is presented and hence task-oriented exploration is better than 'naked' exploration as it leads to more manipulation of the interface, which was one of our goals.

5.1.3.3 Levels 2-4: Isolating local markers of ER integration from activity sequences (RQ 4)

In this section, I outline the results obtained through levels 2-4 of analysis. These results are indicative of the work done to devise interaction data analysis strategies. The level 2 and 3 results discussed here present a case of the best performing student (L3) in order to provide a quick glance of the outcomes of analysis at those levels. At level 4, I present an overall comparison of this student with one of the poor performing students (L9) from our sample.

Level 2 data dealt with determining sequences of the sensorimotor interaction events. Figure 5.13 shows a sample sensorimotor event sequence for a good performer (L3) between two consecutive clicks on the play button. Note that the events between play and pause buttons are happening in the perception-action cycle, while the events happening after the pause button is hit are from the simulation or imagination cycle. This sequence shows that the student transitions between spatial and numerical regions during both the perception-action and simulation or imagination cycles.

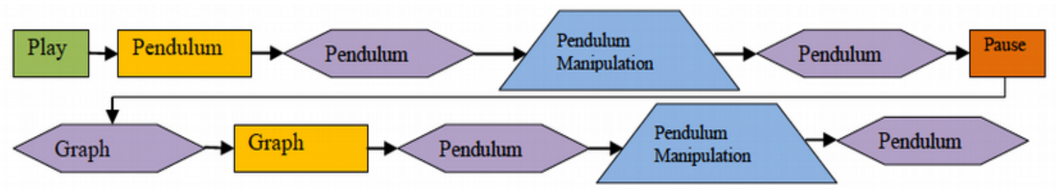


Figure 5.13 An example of events for a good performing student (L9). Rectangle indicates a numerical AOI, hexagon denotes a spatial AOI, while trapezium stands for interface manipulation.

These sequence data were fed into further levels of analysis. Two markers of integration, for instance, were identified at level 3. The first is event transitions between numerical and spatial areas on the screen (Figure 5.14) and the other is the returns between activities or events on different areas of the screen (Figure 5.15).

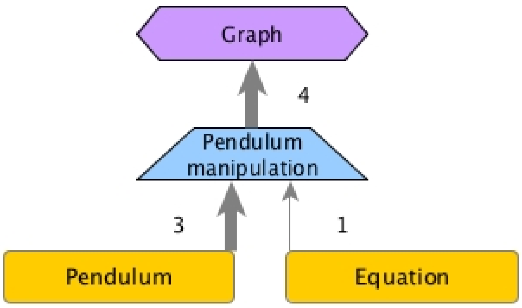


Figure 5.14 Click-gaze transitions for screen 3, student L3; rectangle indicates numerical AOI, hexagon denotes a spatial AOI, while trapezium indicates interface manipulation, arrow indicates the direction of transition, while its thickness symbolizes the number of transitions (normalized).

Figure 5.14 shows that the student transitions from looking at initial angle/length information presented in the pendulum-numerical AOI (imagination) to manipulating the pendulum 3 times – possibly to set the pendulum at desired values, or to understand how the values change after the manipulation (perception-action). The student is also trying to understand the effects of the manipulation on the curve being generated, as indicated by the transitions from the pendulum (after manipulation) to the graph AOI (imagination or simulation).

Figure 5.15 shows that this student looks from the spatial area of the graph to the spatial area of the task and returns 11 times; likely in an attempt to make mappings between the shapes of the curve to be generated and the curve being generated. Interestingly, the student's activity is fairly distributed between the numerical and spatial AOIs; for instance, a significant number of returns for this student occur between the task Y-axis numerical AOI and the task spatial AOI, as well as between graph spatial and graph Y-axis AOIs, pointing to her attempt to compare the initial angle values between the expected and actual curves.

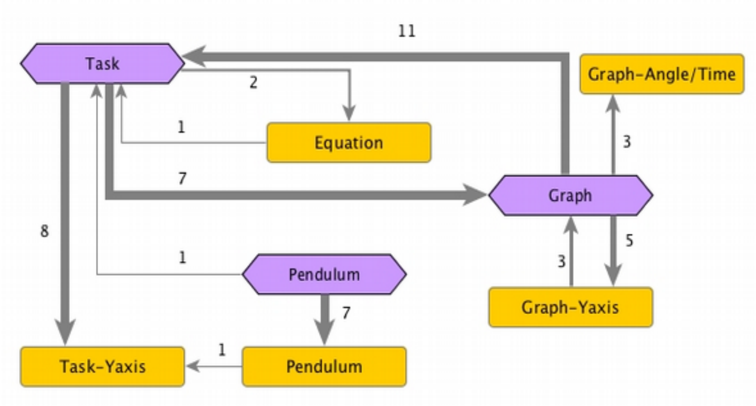


Figure 5.15 Numerical-spatial returns for screen 4, student L3; rectangle indicates numerical AOI, hexagon denotes a spatial AOI, trapezium indicates interface manipulation, arrow indicates the direction of return, while its thickness symbolizes the number of returns (normalized).

Level 4 can be considered to provide a more holistic perspective on the level 3 analysis as it abstracts out several local markers of sensorimotor interaction, to develop global patterns. The transitions defined and identified in level 3, for instance, evolve into transition networks at this level (figures 5.16); whereas the return diagrams evolve into return networks (figures 5.17). Overall, comparison of both the transition and return networks of the students indicates that the best performing student L3 (figures 5.16b and 5.17b) had considerably richer as well as more diverse sensorimotor interactions with the different AOIs of the interface than the poorly performing L9 (figures 5.16a and 5.17a). Student L9 has not only made fewer attempts to manipulate the different features of the interface but also has less gaze activity; and hence, less exploration of the interface in comparison to L3.

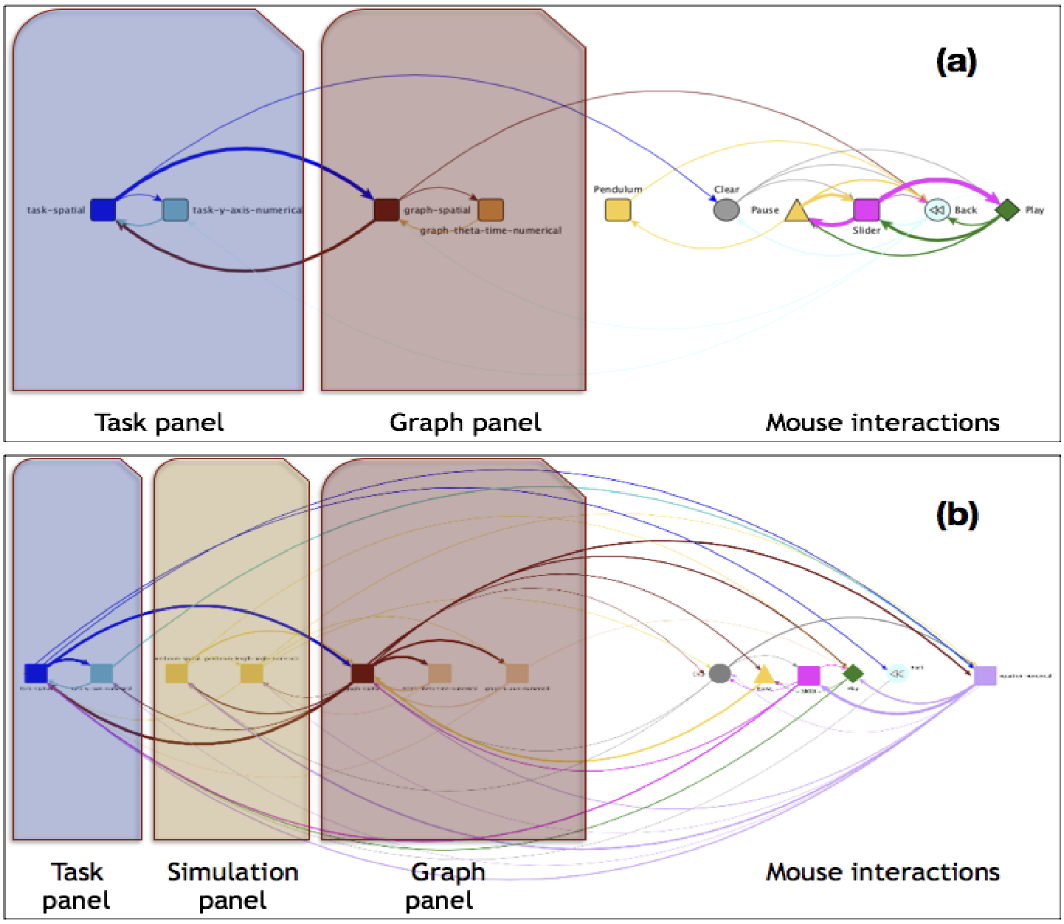


Figure 5.16 Normalized transition networks of interaction during the learning tasks/screens 4-6 of (a) a poor performer L9, and (b) the best performer L3. Direction of arrows indicates the direction of transition, while its width is proportional to percent transitions. Each shape represents an AOI from the specific set of AOIs. On the right side are all mouse interactions performed on the different features (e.g. back, next, play/pause buttons, etc.). On the left are gaze activities.

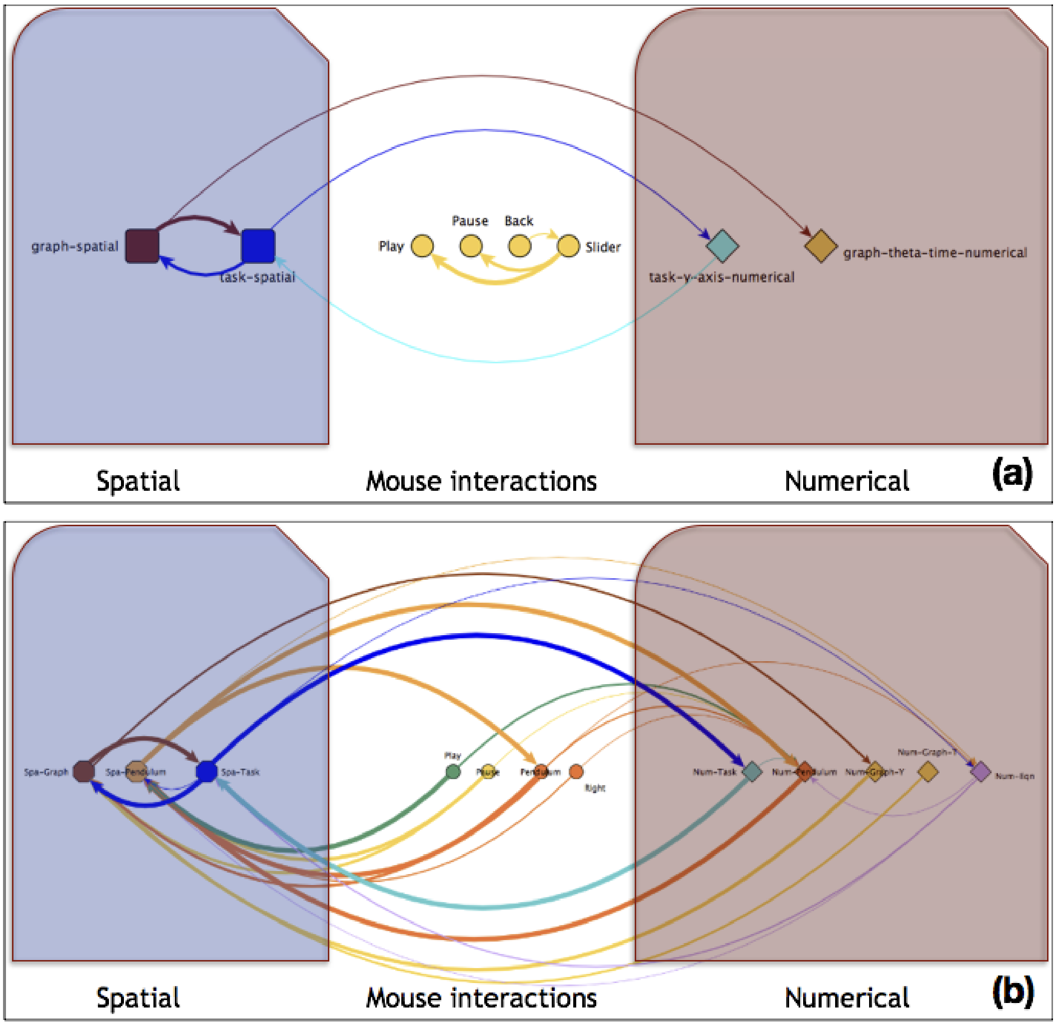


Figure 5.17 Normalized network of returns between spatial AOIs, numerical AOIs and mouse interactions during the learning tasks/screens 4-6 of (a) L9, and (b) L3. Direction of arrows indicates the direction of returns, while its width is proportional to percent returns. Each shape represents an AOI from the specific set of AOIs.

Further, it appears that the richness and diversity of interaction with the interface is related to student performance in the ER integration tasks; thus, the ability to simulate or imagine the dynamic relationship between the ERs. However, the data are only indicative as only one student exhibited qualities similar to an expert, and its associated sensorimotor marker; most students fell into the poor and average performing categories.

5.1.4 Discussion

Preliminary results from the pilot indicate that the students could imagine the dynamic relationships between the ERs with which they interact. Our interface is thus a potential candidate as an educational technology intervention to help students develop RC, particularly in relation to the concept of oscillation (RQ1).

The self-guided group focused on the equation (sliders) and graph, before and during the task. This was perhaps because most students in the self-guided group found it difficult, in the absence of instructions, to discover the various affordances or action possibilities offered by the features of the interface. These students could explore the affordances only by 'playing' with the panels. Students in the text-guided condition, on the other hand, looked at and manipulated all the elements of the interface, in varying degrees, both before and during the tasks. Their exploration of the interface was fairly distributed across the different ERs, and hence was more desirable for ER integration than students in the self-guided condition. In summary, specific text-instructions are thus a necessary feature of the interface as they help learners explore the different affordances of the interface features (RQ 2).

Student sensorimotor interaction with the interface, both in terms of eye and mouse-activity, increased significantly after task presentation in both the groups, hinting that exploration during the task is more desirable than free-exploration. The tasks mediate specific and targeted sensorimotor interaction with the different interface features and are thus integral to the interface design (RQ 3).

Analysis of differences between the best and the poorest performing students across the different levels of sensorimotor interaction show that the devised multi-level analysis approach is a potential candidate for assessing ER integration and its relationship with interactivity (RQ 4).

The interaction data also revealed a few limitations of the interface design, particularly in relation to how the sensorimotor behavior of the self-guided group was possibly skewed towards the sliders reflected as high slider manipulation, particularly during the tasks. However, it is not clear whether the skewed behavior was because the students understood the controller aspect of the equation or because they were familiar with the slider form of interaction. It is also possible that the sliders were more intuitive to the students than other representations. Further studies are required to evaluate the influence of such a familiar UI element.

The ER integration tasks were insufficient to capture student imagination. In the next iteration, these tasks need to be complemented with qualitative think-aloud or verbal reasoning data for richer analysis of the student thinking processes. Further, these tasks could be integrated into the interface for easy deployment without altering the actual experiment settings. This would also help log student responses digitally, and capture student eye-behavior as they imagine while answering the tasks, particularly to see if the eye movement patterns during imagination match those exhibited during interface exploration (Thomas & Lleras, 2009).

Following sections describe the second iteration of this DBR.

5.2 Evolution of the simulation design: Iteration Two (main study)

5.2.1 Modifications to the computer interface design

It was realized from the results of the first iteration that, in order to help students achieve a more desirable (expert-like) exploration of the multi-representational interactive simulation interface, the following design changes were needed:

- The interface must have instructions regarding manipulation affordances of all the representational elements: *The instruction panel was moved to the top of the screen as a white elongated strip. The back, next and clear buttons were also presented along the same strip in the top right corner of the screen.*

- The learner must be given specific tasks which would require her/him to manipulate all the representational elements: *The three learning tasks embedded in the interface were retained in this iteration. The screen-order was also maintained as before.*

- All the interaction affordances must be equally familiar (or unfamiliar) to the students so that they don't gravitate towards using one: *Sliders in the equation were changed to bevel buttons over each variable which change on click and drag.*

- From the researchers notes taken during student interaction in iteration one, it came to light that owing to the difference in the scales of the task graph and the dynamically generated representation graph, students had difficulty in doing the task of graph matching as they could not often 'see' the similarities. This led to frustration, which would have affected exploration. In order to make this process easier, *a grey colored sine wave (a "ghost" graph) which is always displayed on the screen, corresponding to the current pendulum parameters (length and initial angle), was introduced. This graph could be manipulated during screens 1-3 (change maximum amplitude by left clicking and frequency by right clicking) leading to changes in the length and initial angle in both the pendulum and equation. This rendered the interface fully manipulable again (see version 1.0). When the learner clicked play, a blue graph would be generated, as the pendulum oscillates, over this grey 'ghost' graph reflecting the pendulum dynamics. After the learning tasks appeared, while the grey colored "ghost" graph would remain in the background, it would no longer be manipulable. This was to ensure that learners don't end up using only the graph for accomplishing the tasks.*

Figure 5.18 below presents all these features highlighted using red-colored borders on a screenshot of the new interface version 2.1 (The most recent version of the system is available here: http://bit.ly/pendulum_old).

Figure 5.18 Screenshot of the computer interface version 2.1 with all 3 representational modes, a task and the "ghost" graph feature. Inside the red rectangular border on the top are instructions and back/next and clear buttons. In the centre of the screenshot inside the two different square-like red borders are the bevel buttons or bulged scrubbers (as opposed to sliders from the earlier interface version). Clicking and dragging the bevels would change respective values in the equation. Highlighted with a rectangular border below the back and clear buttons is a learning task. In the graph, blue-colored part of the curve is the curve being generated, while grey colored curve in the background is 'ghost' graph on which the blue colored curve is generated once the simulation is played. Only for screens 1-3, the ghost graph is manipulable by clicking and dragging either vertically (to change initial angle) or horizontally (to alter length). During the learning tasks, the graph cannot be manipulated.

- However, it was hypothesized that this "ghost" graph could also become a crutch which the learner would use to complete the tasks without employing their own imagination to generate the given graph (hypothesis 1): *Another version of the interface was created where the "ghost" graph was not available; this was to examine the role of the "ghost" graph in imagination-based integration.* A screenshot of the interface version 2.2 after incorporating these changes in shown in Figure 5.19.

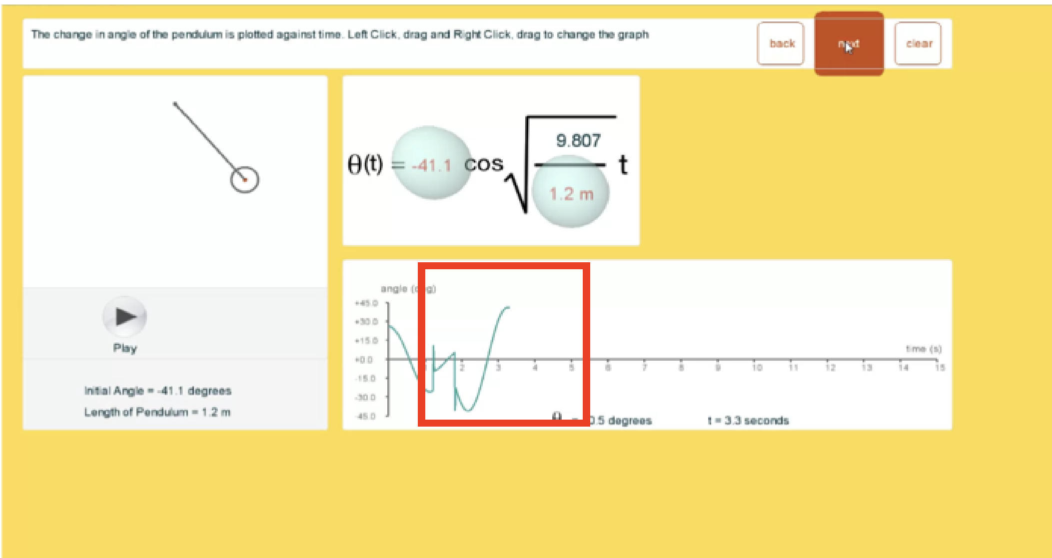


Figure 5.19 Screenshot of version 2.2, without the "ghost" graph. It is still possible to click and drag inside the graph panel to change parameters, except that the resultant curve after manipulation is no longer available. Any change would appear only in the pendulum and equation panels. The resultant graph would be available only after resuming/running the simulation by clicking 'play'.

Note that the two versions 2.1 and 2.2 differed only in the ghost-graph feature.

- The interface must offer a more holistic assessment of the ER integration ability: *The 6 questions presented in the ER integration tasks from the previous iteration were complemented with 8 new questions.* Total 14 questions (see Appendix 8) now together catered to the learning objectives presented in table 5.8 below.

Table 5.8 Learning objectives and question categories. For specific questions, see Appendix 8.

Learning objectives: The student will be able to -	**Question category**
Map phenomenon and graph	Check whether learner can relate points on graph to phenomenon and vice versa (6 questions)
	Check whether given a word problem, a learner can imagine the phenomenon and its graph? - Oscillatory graphs (1 question)
	Non-oscillatory graphs, non-sinusoidal movement in time (2 questions)
Map phenomenon and equation	Describe damped pendulum and ask what is the equation (1 question)
	How to modify behavior of pendulum (1 question)
	Modify equation, ask about behavior (1 question)

Map equation and graph	Show underdamped pendulum graph and ask what is the equation? (1 question)
	Modify equation, ask about graph (1 question)

Questions of the last two categories involving equations were anticipated to be difficult, particularly for the student participants in this study who were studying in 7th grade at the time, and were naive to complex equations and mathematical forms such as the pendulum equation embedded in the interface. Nevertheless, I wanted see if the students could still develop an implicit sense of the different components of the equations.

The ER integration tasks were included within the interface and were introduced after the learning tasks. Figure 5.20 presents a screenshot of ER integration question 1.

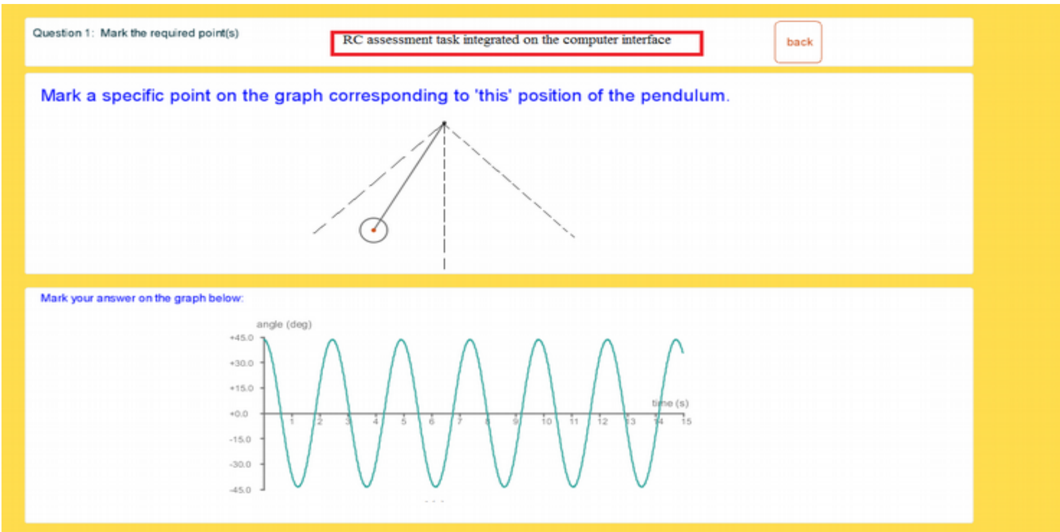


Figure 5.20 Screenshot of the computer interface with a RC assessment task. Students could indicate their choice on the screen using radio buttons. A ‘next’ button would appear only after an answer is marked or chosen.

A summary of the final design principles and corresponding design features, as they evolved from iteration one to two, are shown in Table 5.9 below, with the changes highlighted in bold.

Table 5.9 Comparison of the design principles and respective design features in the two iterations.

Principle	**Design feature in Iteration one**	**Design Feature in Iteration two**
External representations allow processing not possible or difficult to do in the mind (Kirsh, 2010).	The interface plots the graph of the equation/motion of the pendulum for various lengths and initial angles of the pendulum.	The interface plots the graph of the equation/motion of the pendulum for various lengths and initial angles of the pendulum
Cognition emerges from ongoing interaction with the world (Brooks, 1991).	The interface is fully controllable, i.e., the learner can control the pendulum and equation, to see how change in these affects the other element and the graph.	**The interface is fully manipulable, i.e., the learner can control the pendulum, equation and graph, to see how a change in these affects the other two elements.**
The features of the world are used directly for cognitive operations, hence the interface should have all the features needed for integration of representations (Landy et al., 2014).	The interface has the physical system, equation and graph, along with different numerical values. The dynamic nature of elements, and their interconnections, are made transparent, so that learners can integrate across spatial-numerical and dynamic-static modes.	The interface has the physical system, equation and graph, along with different numerical values. The dynamic nature of elements, and their interconnections, are made transparent, so that learners can integrate across spatial-numerical and dynamic-static modes.
The active self is critical for integration of features (Reed, 1988).	The exploration on the interface is guided by tasks which the learner must do.	The exploration on the interface is guided by tasks which the learner must do.
Action patterns can activate concepts, hence actions and manipulations of the representations should be related to existing concepts (O'Malley & Soyer, 2012).	The learner can interact with the pendulum by changing its length and initial angle by clicking and dragging the mouse. This interaction is meant to mimic the interaction with a real pendulum. The parameters in the equation can be changed using vertical sliders – moving up	**The learner can interact with the pendulum by changing its length and initial angle by clicking and dragging. This interaction is meant to mimic the interaction with a real pendulum. The parameters in the equation can be changed using bevel buttons which are placed over the variables in the equation and can be changed by clicking and dragging left and right to decrease and increase values. This interaction highlights the role of the equation as a controller. The**

	indicates increase in parameter, moving down indicates decrease.	**parameters of the graph can be changed by clicking and dragging up and down to change amplitude and left and right to change frequency.**
The interface should allow coupling of internal and external representations (Chandrasekharan & Nersessian, 2015).	The learning task requires the learner to match a given graph. Learners change the parameters of the pendulum/ equation to generate the graph and visually match the task graph to their graph. This develops learner's imagination and coupling between their internal and the external representation.	**The learning task requires the learner to match a given graph. Learners change the parameters of the pendulum/equation to generate the graph and visually match the task graph to their graph. This develops learner's imagination and coupling between their internal and the external representation. A "ghost" graph may or may not be present on the interface during the task which aids the matching process, but may adversely affect the development of the learners' imagination.**

The new design versions 2.1 and 2.2 were evaluated using a lab study described in the next section.

5.2.2 Methods

The methods from the first iteration were adapted with some changes.

5.2.2.1 Sample

18 students (9 female, age range ~11-13 years) studying in 7th grade from an urban school in western India volunteered to participate in the study. I wanted to test our interface with naïve participants who were not formally introduced to the concept of oscillation, simple pendulum, time period, etc. as well as multiple representations of these concepts such as diagrams, graphs and equation. However, I did not explicitly control the familiarity variable.

Our participants belonged to socio-economically underprivileged communities with most of them residing in densely populated areas. Initial interactions with the students revealed that only one of them had a computer system (desktop) at home. Others reported to have some experience interacting with a desktop/laptop (specifically the 'paint/drawing' application in MS Windows operating system), mostly in the school. Every child, however, frequently interacted with touch-screen cellular phones and had at least one such phone at home.

None of the participants spoke English. Their English literacy, vocabulary and understanding was limited to simple 5-7 letter words. These students could barely read, pronounce or write words such as ‘graph’, ‘pendulum’ and ‘equation’. All the participants fluently spoke Marathi and Hindi (major regional languages in this part of India), and preferred Marathi as the language of communication during the experiment.

Before they agreed to participate, each student and his/her parents were informed that s/he would be playing a science game on a laptop using a mouse controller, that the game had stages of increasing complexity similar to any other game, that his/her eye movements would be recorded during the playtime, and that s/he would be answering a set of questions related to the game after passing through all the stages. On expressing willingness to participate, a written consent was obtained from at least one parent of each child.

5.2.2.2 Experimental setup and protocol

Before commencing the experiment, students who expressed unfamiliarity or discomfort using laptop and mouse control were allowed to familiarize themselves with the mouse for 10-20 minutes by practicing with applications such as the Microsoft Paint. The student was asked to verbally indicate when s/he was ready to start the game.

Below is the schematic of the experiment sequence:

Start >> Introduction to the setup >> Eye-tracker calibration >> Interaction with the interface >> Relax >> ER integration questions >> Relax >> Interview >> End.

The overall setup and experiment procedure was similar to the first iteration (revisit section 5.1.1.3).

For each screen, a Marathi translation of the instructions was read to the participant. The researcher repeated the instructions as and whenever the student needed. The researcher also provided appropriate hints when students had a question, but carefully avoided providing any cues related to the functional relationships between the ERs in the interface and solutions to the learning tasks. After the student completed or quit the learning tasks, s/he was given a small break (ranging from 1-3 minutes) to relax. S/he then proceeded to attempt the ER integration questions; the student was told that s/he could skip questions or quit at any time. A Marathi translation of each question was read to the participant when and as frequently as the student required. To ensure that students understand each question, the interviewer rephrased the translated version of the question, if required, without using any representations or gestures. After completion of the questions, students could relax for about 5-10 minutes.

Finally, each participant was interviewed about his/her (i) overall experience with the interface, (ii) assessment of the interface in terms of usability, learnability and interactivity, (iii) observation of own actions performed during the interaction, effects of those actions and their own thoughts about why they performed those actions, and (iv) strategies used or thinking process employed while answering the questions. During (iv) the researcher walked the participant through each question on the interface as well as reminded the answer provided by the student, while the student reasoned about why s/he chose that answer. The interview session was video recorded using a Sony camcorder (DCR SR40).

Interestingly, a post-facto analysis (not included in this dissertation) indicated a pedagogical advantage of this interview session as it allowed students to reflect on their own actions, their effects on the ERs (Danish et al., 2015; Sengupta, Krinks and Clark, 2015) and the relationships between those ERs, thus making the imagination richer with a strong possible reasoning component.

Each student took 45-70 minutes for interacting with the interface (inclusive of the time spent on ER integration questions). The interview with each participant took 10-17 minutes.

11 students (code named with letter 'G') received the interface version 2.1 (with ghost-graph) – hence the ghost graph condition, while 7 students (code named with letter 'N') interacted with interface version 2.2 which did not have a ghost graph – hence 'no ghost graph' condition.

5.2.2.3 Data sources

Eye Tracker: Eye and mouse activities were recorded using a Tobii X2-60 portable eye-tracker.

Mouse tracker: Code was designed and embedded within the interface to record logs of mouse movements and clicks as the participant interacted with the interface. This data could then be synchronized with the eye-tracker data for a more holistic interaction analysis.

Researcher Observations: The researcher kept an unstructured log of student behaviors and facial expressions while they interacted with the interface, as well as their responses during the interview.

ER integration questions: These questions attempted to evaluate the extent to which students are able to imagine and simulate the movement(s) they observed on the interface. There were a total of 14 multiple choice questions (Appendix 8). The option chosen by each student was automatically recorded into a log after the interaction was completed. Accuracy data for all the students were captured from the logs, and tabulated.

Verbal responses and video recording: The interview session was video recorded and transcribed. The transcripts, tagged with student gesture data, served as a source of data on reasoning process. These data were coded and then correlated with the accuracy data. This combined analysis could help identify students who better imagine or simulate the pendulum behavior and oscillation phenomenon, and therefore, integrate the ERs.

5.2.2.4 Data analysis strategies

The following data related to student performance were obtained from the logs created by the interface code: (a) time spent on each screen, (b) time spent on each learning task, (c) status of completion of the learning task (e. g. successful, did not attempt, etc.), and (d) accuracy on the ER integration questions (Q1-14). These data were compiled from all the students and tabulated.

The transcripts (statements tagged with gesture/action data) from the interviews were coded for accuracy as well as qualitative patterns of reasoning. Table 5.10 presents in detail the coding scheme, with examples, that emerged out of the transcript analysis, in relation to the learning objectives of the interface. The categories are hierarchical; for instance, a student is deemed 'successful' on providing a justification related to the dynamic relationship between ERs; 'less successful' if the student's explanation about relationships between ERs is based on physical features or numbers in ERs or a combination of both; and 'least successful' on providing an explanation that does not involve either of the first two types of reasoning.

Table 5.10 Reasoning categories and verbal/non-verbal behavior pointers.

Category/ code	**Explanation**	**Example behavior (verbal + non-verbal)**	**Color code**
Mapping dynamics between (M)ERs (simulation)	1. Explicit description of effects of the act of changing one representation on the other (involving self as the cause of change). 2. No explicit mention of the dynamic cause-effect link, but reasoning includes description of coupling between multiple states of two (more) representations. 3. Description of	- When I move/change 'this', 'that' changes in a certain way; often accompanied by relevant gestures such as pointing to specific feature(s) in representation(s). - When 'this' goes here, 'that' goes there; complemented with gestures such as pointing on the screen to the feature(s) referred to. - 'This' moves like 'this'	

	covariance between physical appearance of a representations, say shape of the curve (such as up and down, straight line) and numbers/quantity/magnitude (numbers increasing and decreasing, identical numbers in two representations). Such an explanation is also often accompanied by gestures such as waving hands.) 4. Use of contexts/meaning or function of features of representations such as the '-' (minus) sign with decrease in quantity, to map equation or graph with phenomenon. Such explanations often tend to be incorrect, although they involve some imagination of the dynamics.	(accompanied by a waving-hand gesture, etc.) - When 'this' is 'up', 'that' is on the right side (indicated by a waving gesture or pointing to the features referred to) - When 'that' is 'less' (pointing to a number), this is like 'this'. - Because there is a minus sign, the curve will be less/down.	
Physical feature or number mapping between the ERs	5. Descriptions of mapping between physical features of representations (e.g. straight-line indicating car direction and a plateau curve) accompanied with relevant gestures (such as moving hand in a straight line, pointing along a curve on the screen). 6. Explicit description solely based on mapping between numbers in the representations – typically between graph and equation. Student would indicate specific number(s) in ERs (by pointing or explicit mention or a combination of both)	- The car is in a straight line (pointing to the car direction) and this (pointing to a straight line curve from the options) is also a straight line so this is the answer. - The length is 10 in this equation.. and here.. it is 5 in the second equation. 5 is half less than 10.. so it'll take less time than the original.. Maybe half of it..	
Performing spatial	7. This involves explicit descriptions and/or gestures	- I thought that there is this mid-point here in the pendulum, so if	

operations on ERs. Three kinds: (a) ER super-imposition, (b) Extension of feature(s) of a representation till it intersects with feature(s) in another representation, (c) Spatial division of a representation into parts that are further mapped with other ERs	related to the student imagining a superimposition of images, as if s/ he were taking the two images in hand and laying them over each other to see points of coincidence(s), typically between the curve and the pendulum. The superimposition may also be an attempt to match the 'point' features present in pendulum and curve images. 8. Descriptions of imagining an extension of, say the pendulum length, so that it intersects with, say graph or curve, to establish some relation between the ERs. The student either reports this verbally or using gestures or a combination of the two. 9. Explicit descriptions of mental operations on ERs accompanied by appropriate gestures. For instance, splitting the screen or a representation into two or more parts to reason about the direction of pendulum or curve. This could be accompanied sometimes with physical feature or number mapping between the ERs.	we put this on the graph, there will be a midpoint here too.. - Gesture: imaginary drawing pendulum image on graph - The bob is here, if we pull it more till here (gesture indicating an extension of the pendulum length till it intersects the curve), the curve will be here (at the point of intersection). - I imagined cutting the pendulum image into right half and left half and compared it to the graph cut into two halves to point the answer on the curve. - Gesture: chopping figures with hand(s), indicating the parts	

An inter-rater reliability test using 33.33% of the sample was done by a researcher other than the experimenter. This test showed 100% agreement on the coding scheme.

The sensorimotor interaction data included eye as well as mouse data. Since our analysis treats eye movements similar to mouse movements – actions that can potentially lead to integration -- the eye and mouse tracking data logs were collected from respective sources, synchronized and compiled in a single file. The interaction analysis strategies developed during iteration 1 were employed to analyze these data (see section 5.1.2 for details; also Kothiyal et al., 2014 &

Majumdar et al., 2014). Figures 5.21a to 5.21d depict the specific AOIs generated for the current interface versions 2.1 and 2.2.

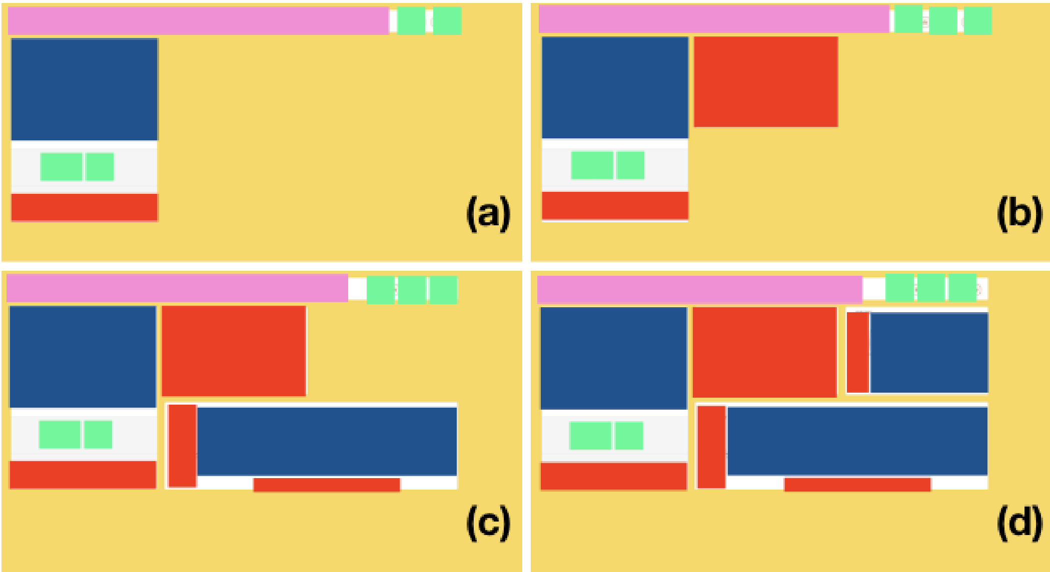


Figure 5.21 AOIs for (a) Screen 1, (b) Screen 2, (c) Screen 3, and (d) Screens 4-6. Color codes: Pink = instructions, green = buttons (play and pause below the pendulum; back, next and clear next to instructions in the top right corner), blue = spatial (pendulum in screens 1-6, graph in screens 3-6 and task in screens 4-6), and red = numerical (pendulum in screens 1-6, equations in screens 2-6, graph Y-axis and graph numerical in screens 3-6, and task Y-axis in screens 4-6).

In iteration 1, I discussed the following four hierarchical levels of interaction data analysis varying in the degree of abstraction:

(i) Level 1 analysis provides data on spread of attention (e. g. number of fixations or mouse clicks per AOI, time spent on each AOI, etc.).

(ii) Level 2 analysis concerns movement of participants from one AOI to the other; sequences of fixation events and mouse click events are determined here (e.g. figure 5.13) and classified into two cycles – perception-action cycle and simulation or imagination cycle.

(iii) At level 3, level 2 data are fed in to define and compute markers that signify integration. An example of a marker is returns – identified as an A-B-A movement of the eye, where A and B are two different representations (AOIs).

Such a movement indicates that the learner is retaining a particular feature in memory and returning to it.

(iv) Finally at the fourth level, process patterns of how the participants interacted with the interface are generated from level 3 data using a graph theoretic framework such as a transition diagram, wherein the AOIs are the nodes and the transitions between the various AOIs are the weights of the branches.

Apart from generating the transition and return diagrams, level 4 in the present analysis also focuses on defining and computing interaction parameters specific to the simulation or imagination cycle of interaction. This is primarily because it is postulated that, during this cycle, the participant would expect an outcome using a forward model (Schubotz, 2007; Rahaman et al., 2017) of the action (i.e. mouse click) performed during the interaction as an active effort to understand system behavior. This expectation is, in a way, a simulation of the system mediated by the interaction with ERs (Pande & Chandrasekharan, 2017).

In the following sections, data analysis is reported at levels 3 and 4 only as it directly addresses an important objective of this iteration – i.e. understanding (inter)action patterns, and not attention. Table 5.11 presents definitions of the parameters computed in this iteration to characterize the interaction of participants. See Appendix 9 for details on the data analysis steps taken.

Table 5.11 Definitions of interaction parameters calculated in this iteration.

Parameter	**Definition**
Gaze Transitions (Level 3)	Eye movements between two consecutive fixations (e. g. A-B, where A and B are two different AOIs)
Gaze Returns (Level 3)	Eye movements between two or more AOIs of the nature A-B-A, A-B-C-A, A-B-C-D-A, and so on, where A, B, C, D are different AOIs. Returns can be thought of consisting multiple transitions, for instance, the return A-B-A has an A-B transition and then a B-A transition. Similarly, A-B-C-A- consists of three transitions, A-B, B-C and returning from C to the AOI A i.e. a C-A transition.
Useful A-B-A returns	Returns of the nature A-B-A between two successive mouse clicks.

(Level 4)	
Useful A-B-C-A returns (Level 4)	Returns of the nature A-B-C-A between two successive mouse clicks.
Unique AOIs count between mouse clicks (Level 4)	The number of AOIs visited between two successive mouse clicks, where even multiple visits to an AOI are counted as a single entry (e. g. if a participant visits AOI A twice, AOI B four times and AOI C just once between two successive mouse clicks, his/her unique AOI count will be 3 irrespective of the number of times s/he visited each of the AOIs).
AOIs count between mouse clicks (Level 4)	Total count of AOI visits, where multiple visits are counted separately (in the above example, the total AOI count between the two mouse clicks will be recorded as 2+4+1 = 7 counts).
Average spread (Level 4)	The average number of occurrences of different AOIs between mouse clicks. Spread = Average number of AOIs visited between mouse clicks.
Elasticity (Level 4)	The weighted sum of the average number of useful returns of the nature ABA and ABCA. Elasticity = 1*(average number of ABA returns) + 2*(average number of ABCA returns). Elasticity also shows how elastic or fluent a person is transitioning between AOIs. It can be understood in contrast to a general meaning of 'inertia' which usually signifies rigidity. Elasticity would thus indicate how easily does a person navigate between the different parts of a stimulus.

5.2.3 Results

5.2.3.1 ER integration

Table 5.12 below presents the top-level data: learner time spent on screens and learning tasks, accuracy on the ER integration questions, and the qualitative category of reasoning in the ER integration questions.

Table 5.12: Group level data (I: Incorrect answer, C: Correct answer, DNA: Did not attempt). Color coding of the cells indicates category of reasoning (Dark grey: Dynamic mapping; light grey: Feature/number mapping; white: Spatial operations or no mapping or no response).

Student		Time on screen (seconds)						Performance on ER integration questions													
		Screen 1	Screen 2	Screen 3	Task 1	Task 2	Task 3	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
With ghost graph	G1	89.49	158.09	70.15	775.46	870.13	195.66	C	I	I	I	C	I	C	C	I	C	I	C	I	C
	G2	145.88	78.23	76.11	1370.39	496.92	0.49	I	C	I	I	I	I	C	C	I	I	I	I	I	I
	G3	160.73	92.01	66.44	695.12	569.85	10.42	I	I	I	I	I	I	C	C	I	I	C	C	C	I
	G4	334.40	408.52	283.04	458.89	1114.57	268.44	I	I	I	I	I	I	I	C	I	I	I	C	I	C
	G5	118.42	48.99	66.25	588.45	261.22	27.23	I	I	C	I	I	I	C	C	I	I	I	I	I	I
	G6	140.77	165.11	123.71	1031.85	116.58	8.18	I	I	I	I	C	I	I	C	I	C	I	C	I	C
	G7	628.41	571.44	1148.60	409.76	2609.81	852.35	I	I	C	C	C	C	C	I	C	I	I	C	I	I
	G8	331.20	240.98	242.90	518.67	314.31	8.50	I	I	C	I	I	I	C	C	I	C	I	I	C	I
	G9	198.56	176.20	201.00	1749.97	1151.69	10.35.5	I	I	I	I	I	I	C	C	I	I	I	I	C	I
	G10	157.95	119.61	1771.84	874.76	528.02	6.80	I	I	I	C	I	I	I	I	I	DNA	DNA	DNA	DNA	DNA
	G11	234.07	83.86	189.52	947.01	173.98	7.96	I	I	I	I	I	I	C	C	I	C	C	I	I	I
Without "ghost" graph	N12	270.68	89.63	94.71	2180.59	505.31	11.46	I	I	I	C	I	C	C	I	I	I	I	C	I	I
	N13	127.08	104.96	53.44	466.31	396.55	112.52	C	C	C	I	I	C	C	C	I	I	I	C	I	I
	N14	164.93	96.03	181.33	156.81	653.64	141.61	I	I	I	I	C	C	C	C	I	I	I	I	I	C
	N15	208.35	229.03	407.51	466.95	758.63	356.77	I	I	I	C	I	C	C	C	I	I	I	I	C	C
	N16	240.15	247.21	119.54	900.67	1057.23	659.01	I	I	I	I	I	I	C	C	I	C	I	C	I	C

Student	Time on screen (seconds)						Performance on ER integration questions													
	Screen 1	Screen 2	Screen 3	Task 1	Task 2	Task 3	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
N17	268.16	139.50	239.40	451.97	303.05	141.75	I	I	I	I	I	C	I	I	C	C	I	I	C	I
N18	898.63	328.74	783.62	822.61	352.12	91.41	I	I	I	I	I	I	C	C	C	I	I	C	I	I
Average (with ghost graph)	230.90	194.82	385.41	856.39	746.10	138.60	9.09	9.09	27.27	18.18	27.27	9.09	72.72	81.81	9.09	36.36	18.18	45.45	27.27	27.27
Average (without ghost graph)	311.14	176.44	268.51	777.99	575.22	216.36	14.29	14.29	14.29	28.57	14.29	71.43	85.71	71.43	28.57	28.57	0	57.15	28.57	42.86
	Time on screen (seconds)						Percent correct students/responses													

There were no noticeable differences between the ghost-graph and no-ghost-graph conditions in terms of both time spent on screens and accuracy on the ER integration questions, except that the participants from the ghost-graph condition took longer to solve learning tasks 1 and 2 on an average (hypothesis 1). The researcher's records from the interaction session note that several students from the ghost graph condition spent the initial few minutes constantly clicking at different points in the graph panel during the learning tasks; few of these students eventually got frustrated as the curve would not generate. The researcher had to intervene by reminding the students to try some other features in the interface (such as the 'play' button, without explicitly mentioning it). Students in the no-ghost-graph condition also scored well above average, similar to students in the ghost-graph condition, in questions 6 and 12; the former requires one to map the movement of pendulum with that of the curve, while the latter requires mapping pendulum to its equation. There was no difference between the conditions in qualitative reasoning.

There was no correlation between the time spent per screen and the accuracy on the ER integration questions. The accuracy as well as reasoning patterns are similar among students in both the conditions. Highest accuracy among both the conditions is recorded on questions 7 and 8 (72.72% and 81.81% for ghost-graph condition, and 71.43% and 85.71% for the no-ghost-graph condition respectively). Question 7 tested if the students transfer the learning or achievement of ER integration (in relation to the phenomenon of oscillation) from simple pendulum-based ERs to a real-life situation (mood swings). Question 8 (and 9) did not involve the oscillation situation. Most students, including the poor performers, described behavior of the curve in relation to the event referred to in the question, while answering these questions. One student (G7), for instance, said the following in response to question 8:

> My friend's mood is good in the morning (<u>pointing on the answer graph</u> at the beginning of the curve), <u>then it decreases as the time goes (moves finger along the curve</u>), then it gets better again and decreases as the time goes, so this is the answer (<u>Moving fingers along the curve throughout</u>)... G7 while explaining her answer to question 7.

G7's response shows that she not only related the 'up' and 'down' states of mood with respective features in the curve (crest and trough respectively), which would have been coded as a feature-based reasoning, but also indicated an understanding of the dynamic change or process of the phenomenon of mood change in relation to the static curve.

On a question testing if the students transfer the understanding of oscillation dynamics to other representations or situations, N15 exhibits her understanding that not every curve represents change in a parameter of oscillatory nature.

> Car runs at 60km.. they said.. and <u>it does not go up-down</u>.. Speed is <u>same</u>… (Researcher: What does not go up and down?).. Its <u>speed remains the same</u>.. so.. <u>straight</u>... N15 on question 8.

Further, based on the accuracy and quality of reasoning, the students were classified into three performance categories irrespective of the condition they received: Good integrators, intermediate or partial integrators, and poor integrators. Below is the student distribution across the categories.

1) Good: G7, N13, N15

2) Intermediate/partial: G2, G6, G11, N18

3) Poor: G1, G3, G4, G5, G8, G9, G10, N12, N14, N16, N17

The good integrators (G7, N13 and N15) extensively reported the dynamics of ERs when mapping between the pendulum and the graph (Q1-9), but not in the questions requiring mapping equation with either graph or pendulum (Q10-14). These students exhibited a good understanding of the various graphs shown during the different ER integration questions, including those not related to oscillation. In a few cases, phenomenon dynamics was reported upon observing various ERs even when this was not required to answer the question successfully. For instance, N13, while explaining her answer to Q5 (which presents a static picture of the sine-wave curve with a point highlighted on x-axis and requires one to mark position(s) of the pendulum corresponding to that point) said the following:

Because when it (pointing to the highlighted point on the curve) is in the middle, <u>the ball (pendulum) goes to the left</u>… N13

To answer this question, it is sufficient to correspond the static states of the two representations without imagining their dynamic behavior. The phrase, "goes to the left" in N13's response is a clear description of the dynamic to-and-fro movement of the pendulum. Note that the question (or the correct answer) does not concern the direction of pendulum movement.

Nothing conclusive could be said about student performance on the questions involving equations (Q10-14), as all the participants uniformly found them difficult; most students relied on number and feature mapping, including the good integrators. Specifically, it was observed that:

a. For Q10, which requires changing the equation by mapping the problem statement in imagination to static equation components, none of the participants could provide any acceptable justification.

b. For Q11, which requires imagination of pendulum movement, change in one of its parameters, and its effect on that movement, correct answer accompanied by acceptable explanation about the dynamics was given only by three students (G3, G7, G11).

c. For Q12, which requires reading equations and mapping the differences in those equations to the differences in the behavior of pendulum, correct answer and reasoning was given by 8 students (G1, G3, G4, G6, G7, N12, N13, N16). However, most provided a number mapping-based answer though they could not map these numerical components with the pendulum behavior.

d. For Q13, which requires mapping the equation and its components with an unfamiliar, damped graph, 5 students (less than half) 'guessed' the answer correctly (G3, G8, G9, N15, N17) but failed to provide intelligible reasoning about mapping between ERs.

e. For Q14, which requires mapping equation and its components with a familiar graph, 6 students answered correctly (G1, G4, G6, N14, N15, N16). However, analysis of their transcripts revealed that they did so by mapping numbers between the equations and physical features on the graphs without transforming between the equation and the graph.

5.2.3.2 Interaction Patterns

The table below (Table 5.13) shows the interactivity parameters defined earlier in table 5.10 for all participants, except for G5 whose interaction data was not included due to technical problems.

Table 5.13: Interactivity parameters for screens 2 and 3, and tasks 1 and 2.

Screen 2																
		Useful Returns-ABA – between mouse clicks			**Useful Returns-ABCA – between mouse clicks**			**Unique AOIs count between mouse clicks**			**AOIs count between mouse clicks**			**Average Spread + Average Elasticity**		
Performance	**Student**	**Individual**	**Average**	**Standard Deviation**	**Individual**	**Average**	**Standard Deviation**	**Individual**	**Average**	**Standard Deviation**	**Individual**	**Average**	**Standard Deviation**	**Individual**	**Average**	**Standard Deviation**
	G7	6.09		8.94	0.00	0.00	0.00	0.98	1.18	0.28	3.73	4.45	1.35	9.82		

	Performance	Average				Good		
	Student	G6	G2	N18	G11	N15	N13	G7
Useful Returns-ABA – between mouse clicks	Individual	2.70	0.89	2.72	3.50	5.53	18.50	3.99
	Average	2.45				9.34		
	Standard Deviation	1.10 0.24				7.97		
Useful Returns-ABCA – between mouse clicks	Individual	0.50	0.00	0.36	0.50	0.83	2.50	0.60
	Average	0.34				1.31		
	Standard Deviation	0.24				1.04		
Unique AOIs count between mouse clicks	Individual	0.50	0.89	0.55	0.83	1.69	2.00	1.23
	Average	0.69				1.64		
	Standard Deviation	0.20				0.39		
AOIs count between mouse clicks	Individual	1.60	2.33	1.74	2.00	5.19	3.00	3.04
	Average	1.92				3.74		
	Standard Deviation	0.32				1.25		
Average Spread + Average Elasticity	Individual	5.30	3.22	5.18	6.50	12.38	26.50	8.23
	Average	5.05				15.70		
	Standard Deviation	1.36				9.58		

Screen 3

		Poor									Average				Good	
		N17	G1	N12	G9	G10	G8	G3	G4	N16	G6	G2	N18	G11	N15	N13
Useful Returns-ABA – between mouse clicks	Individual	6.65	11.83	2.50	1.57	1.95	5.48	4.93	7.11	35.50	6.06	3.86	4.95	1.93	4.57	20.75
	Average	8.61									4.2				10.47	
	Standard Deviation	10.57									1.76					
Useful Returns-ABCA – between mouse clicks	Individual	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Average	0.00									0.00					
	Standard Deviation	0.00									0.00					
Unique AOIs count between mouse clicks	Individual	0.05	0.06	1.00	0.74	0.40	0.48	0.19	0.78	1.50	0.39	0.86	0.30	0.65	1.07	1.50
	Average	0.58									0.55					
	Standard Deviation	0.48									0.25					
AOIs count between mouse clicks	Individual	0.80	1.06	3.75	2.11	1.80	1.73	1.30	2.31	6.00	1.42	3.86	1.25	2.39	3.61	6.00
	Average	2.32									2.23					
	Standard Deviation	1.63									1.20					
Average Spread + Average Elasticity	Individual	7.45	12.89	6.25	3.68	3.75	7.21	6.23	9.42	41.50	7.48	7.72	6.20	4.32	8.18	26.75
	Average	10.93									6.43				14.92	
	Standard Deviation	11.80									1.56				10.28	

Poor						Average				Good			Performance			Poor								
G9	G10	G8	G3	G4	N16	G6	G2	N18	G11	N15	N13	G7	Student		Task 1	N17	G1	N12	G9	G10	G8	G3	G4	N16
9.32	4.51	5.91	6.43	2.26	7.22	9.03	5.72	7.59	6.33	0.95	7.93	0.92	Individual	Useful Returns-ABA – between mouse clicks		1.28	5.29	2.56	4.65	2.72	6.71	3.50	7.29	7.29
5.63						7.17				3.27			Average			4.59								
3.14						1.47				4.04			Standard Deviation			2.22								
1.06	0.62	0.84	0.85	0.40	0.96	1.07	0.60	1.02	0.85	0.15	1.33	0.28	Individual	Useful Returns-ABCA – between mouse clicks		0.15	0.71	0.56	0.50	0.36	0.79	0.50	0.86	0.86
0.75						0.89				0.59			Average			0.59								
0.40						0.21				0.65			Standard Deviation			0.24								
1.27	0.64	1.39	1.39	1.21	0.86	1.15	0.91	0.34	1.14	0.90	2.04	1.24	Individual	Unique AOIs count between mouse clicks		0.05	0.00	2.44	1.20	NaN	0.36	1.00	1.07	NaN
1.12						0.89				1.39			Average			0.87								
0.32						0.38				0.59			Standard Deviation			0.85								
3.48	1.71	3.05	2.85	2.95	2.56	2.90	2.32	1.05	2.57	2.08	7.04	2.88	Individual	AOIs count between mouse clicks		0.89	1.00	5.00	4.65	NaN	1.14	2.40	2.57	NaN
2.85						2.21				4.00			Average			2.52								
0.79						0.81				2.66			Standard Deviation			1.71								
14.92	7.46	10.64	10.98	6.01	11.70	14.07	9.24	10.68	10.60	3.33	17.63	4.36	Individual	Average Spread + Average Elasticity		2.47	7.71	8.68	10.30	3.44	9.43	6.90	11.58	
9.98						11.15				8.44			Average			7.56								
4.49						2.06				7.98			Standard Deviation			3.20								

	N12	4.33			0.50			1.34			3.40			8.73		
	G1	11.00			1.54			1.14			3.84			17.92		
	N17	1.26			0.22			0.58			1.49			3.19		
Task 2																
		Useful Returns- ABA – between mouse clicks			**Useful Returns- ABCA – between mouse clicks**			**Unique AOIs count between mouse clicks**			**AOIs count between mouse clicks**			**Average Spread + Average Elasticity**		
Performance	**Student**	**Individual**	**Average**	**Standard Deviation**	**Individual**	**Average**	**Standard Deviation**	**Individual**	**Average**	**Standard Deviation**	**Individual**	**Average**	**Standard Deviation**	**Individual**	**Average**	**Standard Deviation**
Good	**G7**	7.35			0.69			1.55			3.48			12.21		
	N13	6.72	5.80	2.16	0.44	0.45	0.24	1.28	1.22	0.37	3.60	3.07	0.81	11.20	9.77	3.40
	N15	3.33			0.21			0.82			2.14			5.89		
Average	**G11**	3.76			0.24			1.28			2.60			6.84		
	N18	3.21	2.64	1.53	0.14	0.13	0.10	NaN	0.84	0.38	NaN	2.07	0.48	3.49	4.45	2.11
	G2	3.21			0.14			0.60			1.93			5.42		
	G6	0.37			0.00			0.63			1.68			2.05		
Poor	**N16**	7.24			0.71			1.25			3.27			11.93		
	G4	8.64			0.82			1.43			3.84			14.12		
	G3	3.97			0.30			0.78			2.19			6.76		
	G8	2.37			0.11			1.47			4.00			6.59		
	G10	3.00	4.91	3.02	0.22	0.44	0.33	0.31	1.04	0.52	0.96	2.80	1.30	4.40	8.59	4.58
	G9	9.38			0.93			1.26			3.44			14.68		
	N12	1.71			0.15			0.85			2.07			4.08		
	G1	6.25			0.60			1.73			4.47			11.92		
	N17	1.67			0.11			0.27			0.96			2.85		

Statistical significance could not be established in any case as the category sizes are very small and the standard deviation values within each category are considerably high. Overall, good integrators had the highest average values on all interactivity parameters except for task 1, where they record the lowest values. Poor performers had the second highest average values across all interactivity parameters for all the screens and tasks; although, not significantly different from the values for good integrators. Intermediate performers had the lowest average values across all the interactivity parameters for all screens and tasks. The interactivity values suggest that good integrators had the richest sensorimotor interaction with the interface, followed by poor integrators whose values are slightly lower; while the average integrators had the least diverse interaction with the interface.

Next, figure 5.22 presents how the overall ‘spread+elasticity’ patterns vary for the three performance categories, as students’ interface exploration progressed through the different screens and tasks. Average spread and elasticity provide a more holistic picture of the interaction patterns (described in table 5.10) as they are abstracted from other level 3 and 4 parameters.

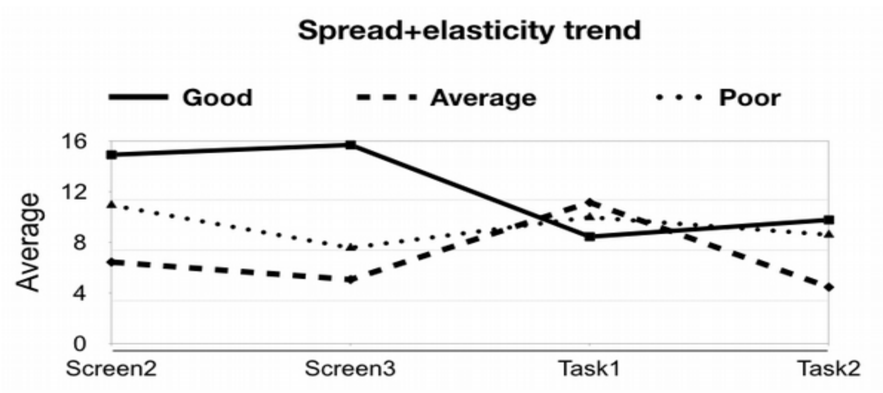


Figure 5.22 Trend across screens and tasks for average spread+average elasticity values of good, intermediate and poor integrators. Notes: (1) Screen 1 interactivity data are not calculated as only one ER (pendulum) is available for interaction at that stage; ER integration-related interactivity thus is neglected for screen 1. (2) Task 3 data are not considered as most participants did not attempt this task.

Across the screens and tasks, good integrators exhibit a strong trend with a drop in their interaction with the interface during the learning tasks as compared to screens 2 and 3. They have the lowest values on all interactivity parameters for task 1. The values in task 2, although more than those in task 1, are almost half the values recorded for screens 2 and 3. For average and poor integrators, although the overall nature of their interaction with the interface does not seem to vary significantly as they transit from screen 2 through task2, a weak pattern exactly opposite to that for the good integrators is noticeable with an increase in activity after transit from screen 3 to task 1.

These trends match the ideal (expert-like) eye-behavior patterns described previously in table 5.7, suggesting that good integrators had already explored the interface and the dynamic relationships between the ERs while in screens 2 and 3, and that they were able to do the tasks with a more focused approach, where they relied more on imagination than sensorimotor interactivity.

To dig more into the correlation between patterns of reasoning and patterns of interaction, four participants (two each from the good and poor performance categories), who were able to articulate the reasons for their answers clearly and elaborately, were selected for detailed interaction analysis at level 4. This was because such computer interfaces are expected to develop learners' implicit understanding of the target domain and I wanted to examine learners who were able to make their implicit knowledge explicit and describe what they had learned from the interface.

The four participants selected were G7, N13, G9 and N16. Of these G7 and N13 (good integrators) developed correct understanding of the dynamic relationship between the ERs and were able to imagine it accurately later on in the absence of interactivity, while the other two, G9 and N16 (poor integrators), developed incorrect understanding of the relationship between the ERs (for example, physical feature or number based) but were able to imagine this relationship clearly later on.

Figure 5.23 compares the transition networks of these participants for screens 2 and 3 and task 1. Figure 5.24 presents a comparison of the plots of the returns.

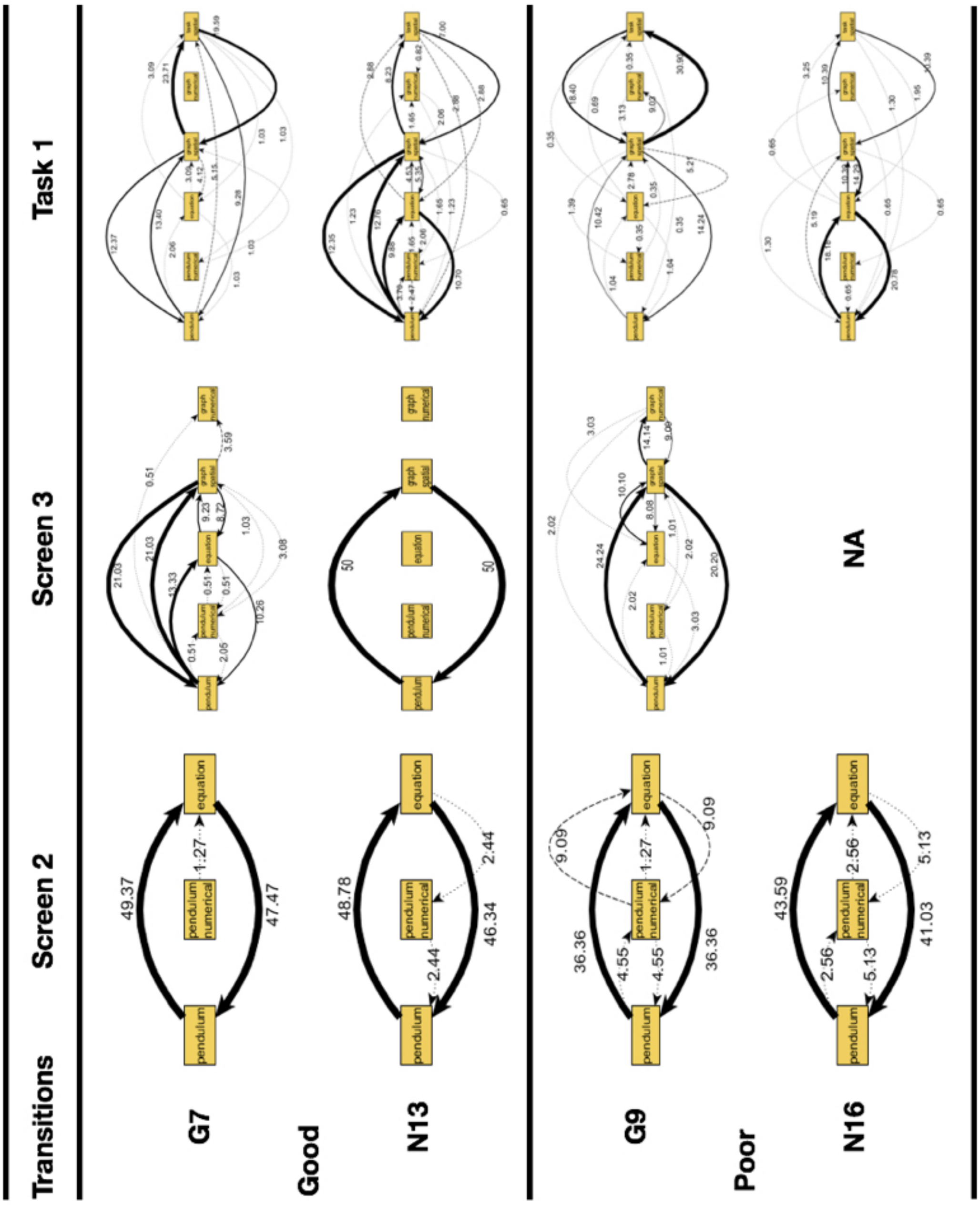


Figure 5.23 Comparison of patterns of transitions for screens 2 and 3 and task 1. Rectangular boxes with a yellow fill are AOIs and the arrows connecting them are transitions. Direction of arrow signifies direction of transition. The width of an arrow is proportional to relative number (tag) of transitions in that direction.

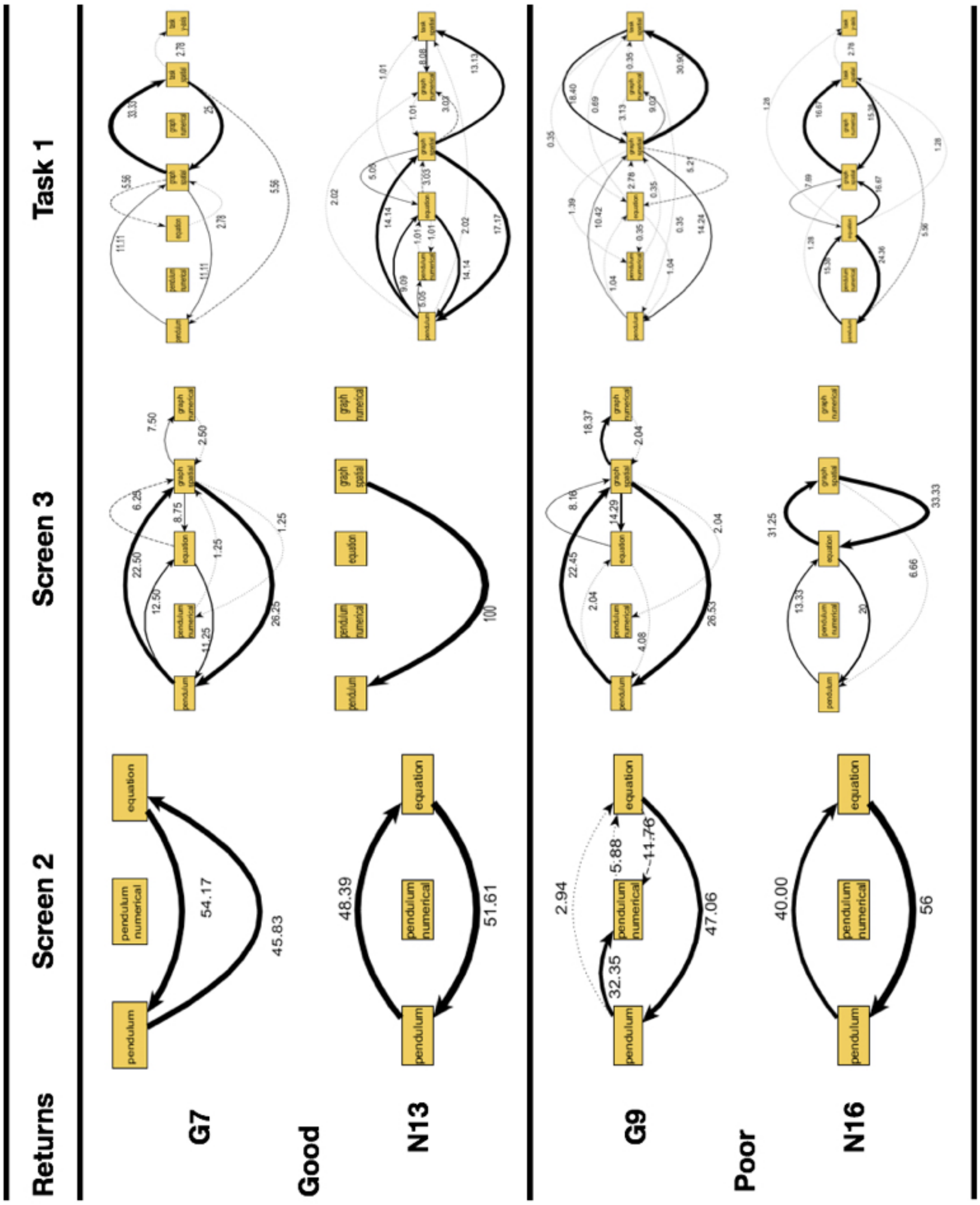


Figure 5.24 Comparison of returns for screens 2 and 3 and task 1. Rectangular boxes with a yellow fill are AOIs and the arrows connecting them are returns. Direction of arrow signifies direction of return. The width of an arrow is proportional to relative number of returns (tag) in that direction.

The transition networks for the good and poor integrators during screen 2 interaction are more or less similar, while there are clear qualitative differences between the students for interaction during screen 3 and learning task 1. Particularly for task 1, the transition networks for the good integrators are qualitatively richer than those of the poor integrators; the former group not only transited more frequently between the AOIs but also exhibited activity between diverse AOIs. However, these observations are only indicative as this difference could not be established quantitatively. Moreover, no such patterns were observed in the return diagrams where both low and high performers exhibit similar interaction patterns (e.g. transition networks of all the four students for screen 2, return networks of N13 and N16 during screen 2; figure 5.24). In summary, participants who could not integrate ERs may have interacted with the interface in ways qualitatively similar to those employed by good integrators.

An overall comparison between the transition and return plots of these four students shows that their interaction patterns do vary qualitatively in terms of emphasis laid on the different representations and the sequence of looking and clicking. Importantly, these are individual variations and not just variations between the performance categories. In fact, these strong individual differences within groups indicates that there are multiple patterns of interaction among good integrators as well as poor integrators. This suggests that not only there are multiple patterns supporting integration but also that the same interaction pattern can lead to different integration performance (good or poor).

5.2.4 Discussion

In this section, I discuss how findings presented in the previous section inform us about the relationship between interactivity, imagination and ER integration.

5.2.4.1 Hypothesis 1: Ghost graph may be a hindrance to ER integration and imagination

Considering that the ghost graph changes in real time in accordance with the manipulation of any of the parameters, having a ghost graph in the background should have reduced the time taken to solve the problem significantly, as both the target curve and the curve that would be generated as a result of manipulation would be available for direct visual comparison. Yet, students in the ghost-graph condition took longer to solve the learning tasks on an average than students in the no-ghost-graph condition. These students also performed considerably more 'extra' clicks in the graph region during the tasks than students in the other condition. Why?

Before I seek to answer this question, I must remind the reader of the nature of the learning tasks. Completing task 1 required changing only one setting of parameters right in the beginning (i.e. at $t = 0$). Once the correct pendulum length and initial angle values were set, playing the simulation would generate the curve and complete the task. For solving tasks 2 and 3, one had to change the parameters more than once – first in the beginning (at $t = 0$), and then again after the simulation has run for a certain time (say at $t = x$, where x could be anything between 2-14 seconds, depending on the task). To change the parameter the second time, one needed to run the simulation with the initial settings, pause it after a while (at $t = x$ seconds, depending on the target curve), change the parameters as required, and then resume the simulation to complete the curve.

Considering that the average time that was taken by students in both conditions to achieve precision at playing/pausing the simulation within the mechanical error window, was more or less similar, the following are a few possible mutually non-exclusive events that may have led to this inverse result:

(a) In task 1, because the students in the ghost graph condition could visualize in the graph the effects of changing a parameter without any delay, they did not feel it necessary to play/run the simulation to cross check their settings with the task image. However, it took them longer to realize that they had to generate a curve over the ghost graph by playing the simulation so that the task could be complete. It is possible that the students deemed it unnecessary to play/run the simulation much earlier, during screen 3 interaction. The no-ghost-graph condition students, on the other hand, realized the need to play/run the simulation quickly as they had no option (right from their interaction in screen 3) but to complete the simulation in order to observe effects of their manipulation and check if their initial settings were correct.

(b) In tasks 2 and 3, playing/pausing the simulation was critical to a successful completion. Students from the ghost graph condition who did not learn to play the simulation in task 1 either failed in task 2 for obvious reasons or took more time to engage in the play/pause cycles.

(c) In tasks 2 and 3, at the first pause after a certain curve had been generated, changing a parameter would update the ghost graph only from the end-point of the previously generated partial curve, and continue till the cycle ends. The graph panel now showed a 'hybrid' or combined curve (that had a partial curve generated till the pause with initial settings + ghost graph for new settings which originated at end of the first pause). Figure 5.25 shows a screenshot of such a hybrid curve.

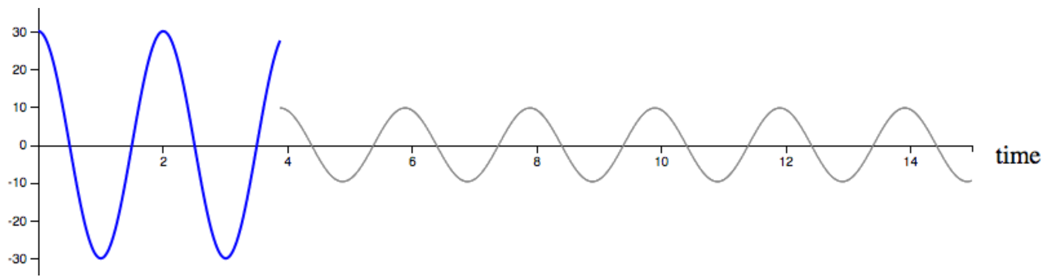


Figure 5.25 Example of a hybrid curve displayed in the simulation interface. The curve with the larger amplitude (from t = 0 to t = 4) is generated with initial settings, the curve with smaller amplitude is a ghost graph displayed after pausing the simulation and changing the initial angle setting, as a result, obtaining a hybrid of a generated curve and a resultant ghost curve.

This hybrid curve confused students as it would look different from the target curve. Getting out of this confusion needed a significant leap of thought – that the simulation could be run and paused to change the parameters again, and that the ghost graph shows only a possible curve every time a setting is changed. The more the manipulations involved in each play/pause cycle, the easier it would be to achieve this leap; and hence, more time. With no ghost graph available, students in the other condition did not face this problem.

In all the learning tasks, the ghost graph turned out to be a crutch instead of support, working as a means to offload the imagination of parameter change and curve generation dynamics, thus confirming hypothesis 1.

5.2.4.2 Hypothesis 2: Integration is correlated with interactivity

The result that distinct patterns of interaction were identified with each performance category shows that interactivity and ER integration are related. The consistently higher values for interaction parameters during screens 2 and 3 indicate a positive relationship between interactivity and ER integration, suggesting that interactivity is necessary for integration.

5.2.4.3 Hypothesis 3: Interactivity does not guarantee integration

High interactivity did not always lead to integration. Despite exhibiting values as high as the good integrators across various interactivity parameters, poor integrators failed to imagine the dynamic relationship between the ERs. Intermediate or partial integrators, who could develop partial imagination-based integration of ERs, reported the lowest interactivity with the interface, indicating that low eye and mouse interactivity leads to partial, but incomplete integration.

The good integrators, however, exhibited a strong downward trend of spread+elasticity values (figure 5.23) as they moved from screen 2 through the learning tasks, consistent with the ideal interactivity behavior (see table 5.7). This means that the participants interacted and explored the dynamic ER relationships more during the screens 2 and 3. A decrease in their activity during the learning tasks indicates that they focused more on solving the problems as, perhaps, they had already explored the relationships between ERs. The interactivity (spread+elasticity values) for average as well as poor integrators, on the other hand, does not seem to change across the different phases in the interface; although the average integrators show slightly increased activity as they move from screen 3 to the learning task 1. These trends in interactivity across the different phases of the interaction may have some role to play in ER integration. While it can be tempting to conclude from these results that it may be good to explore the interface freely (i.e. when there are no tasks or specific goals to accomplish) in order for integration to happen, results from our pilot study reported earlier suggest otherwise.

Secondly, good integrators imagined the phenomenon dynamics through ERs even during situations where it was not necessary to do so. N13's response to Q5, as discussed in the results section shows this clearly. Note that the question does not concern the direction of pendulum's movement as it could have been going either ways (left to right or right to left) when the picture was captured. A possible explanation to this behavior could be that N13 got drawn into the dynamics of the pendulum-graph relationship. It is possible that her interaction with the simulation interface and the three representation was intense/effective enough so that just the perception of a state of the static representation readily triggered an imagination of the movement of the pendulum and graph dynamics, consistent with what she had experienced during the interaction. This maybe a case of automatic simulation, which is one aspect of RC as per our model presented in chapter 3. The other important aspect of RC is to gain explicit control over the dynamics in order to capture its static states at will (Pande & Chandrasekharan, 2017). However, despite exhibiting gaze and click interactivity as high as the good integrators, not even a single such incident was observed in case of poor integrators. The poor integrators, in fact, failed to imagine the dynamics almost always.

While good and poor integrators shared interaction patterns to some extent, the strong individual differences observed within each performance group in the nature of interaction leads to the final finding discussed below.

5.2.4.4 Integration has no unique pattern of interactivity

This statement is a corollary of the previous finding that interactivity does not guarantee integration. It is supported by the individual differences between transition and return networks of the good and poor integrators. The transition and return patterns of these four students showed that their interaction patterns varied qualitatively in terms of emphasis laid on the different representations and the sequence of looking and clicking.

Importantly, these are individual variations and not just variations between the performance categories. These strong individual differences within groups suggest that there are multiple patterns of interaction among good integrators as well as poor integrators. This suggests that not only are there multiple patterns supporting integration, but also that the same interaction pattern can lead to different integration performance (good or poor). The results imply that there is no unique or "ideal" interaction pattern that can guarantee integration.

In summary, these findings indicate that the relationship between sensorimotor interaction and ER integration is more complex than assumed in conjecture 2, and may involve other factors such as facilitation by a teacher, context, etc. In summary, extensive interaction is necessary for ER integration, but it is not sufficient.

5.3 Limitations

1. Students in this study experienced a simulation intervention for about ten minutes with minimal instruction. Future studies are needed to investigate whether (a) a longer exposure to the interface would help ER integration more, and (b) would it be possible for a teacher to smoothly connect this computer interface with the existing classroom dynamics, possibly using QR codes (Borar et al., 2017) linking the simulation to the textbook, to scaffold ER integration.

2. The intervention and assessment modules are presented separately in the current interface. Students interact with dynamic ERs but are presented with static ERs during the assessment (ER integration questions). This design allows investigating how students imagine based on the static ERs, but does not provide information on how they would use interactions in the simulation itself to solve the problems. An ongoing revision of the design seeks to include both static image based tasks and simulation based tasks to address this question.

Chapter 6: Concluding remarks

6.1 Summary

The theoretical and empirical work reported in this dissertation focuses on ER integration, which is central to RC – a critical skill in learning science, mathematics and engineering. I argue that a theoretical account of ER integration, based on recent developments in distributed cognition (DC) and embodied cognition (EC), taking into account the constitutive character of ERs, is needed, particularly to (a) understand the cognitive mechanisms underlying ER integration, and (b) develop design guidelines for developing enactive new media interventions. As a first step to develop such an account, I reviewed the theoretical frameworks proposed for ER integration as well as RC development, and related studies within and across the STEM domains (chemistry, biology, physics, mathematics, engineering). The review revealed that existing accounts and approaches to ER integration are primarily rooted in classical information processing theories of cognition, particularly cognitive load-based models. Such accounts make the development of ER integration appear mysterious, as they do not seek to unravel the underlying cognitive mechanisms. Further, the computer-based interventions derived from such frameworks consider ERs merely as tools to achieve conceptual understanding, and ironically end up helping offload some of the learner's cognitive processes to the computer screen.

To address the need for a state-of-the-art understanding of the cognitive mechanisms that support ER integration, I outline a theoretical account extending the idea of constitutivity. This model (the TUF model) focuses on the interaction between internal cognitive processes and external representations, applying and extending recent advances in distributed and embodied cognition theory. The account illustrates how learners incorporate ERs, by interacting with them using sensori-motor mechanisms. ERs thus gradually become part of, and thus extend, the cognitive system, as well as form and extend the internal model of the scientific phenomena they represent. Imagination based on these internal models can thus connect and integrate external models smoothly. Further, activations of the sensori-motor system during interactions with these ERs, as well as mental simulations based on the internal traces of these sensorimotor interactions, facilitate 'freezing' and 'unfreezing' the different states of ERs in imagination.

This theoretical approach suggests that:

(1) The development of the ER integration ability (expertise) would result in a reorganization of the cognitive system, particularly the sensorimotor system. This suggests that the way learners perceptually access ERs would change after significant training in a domain.

(2) Sensorimotor interaction would support ER integration and its development.

To test these predictions, I developed two empirical projects. The first explored behavioral markers of sensori-motor mechanisms associated with ER integration. The second developed a novel interaction-based learning environment, and tested it extensively to understand the role of interaction in ER integration. Both empirical investigations treat eye-movements (or gaze) as actions similar to hand movements.

Project 1 concentrated on identifying gaze and other behavioral markers across various expert and novice populations, to understand the development of ER integration in chemistry. The results confirmed that, among the multiple variables at work, a sensorimotor change is critically associated with the development of ER integration. This sensorimotor component, in our sample, was identified as a tuning of the perceptual system, in the process of novices turning into experts (marked by changes in eye movements and gaze patterns while viewing ERs). This tuning helps in quickly and effectively picking up relevant information from the ERs. Interestingly, experts also appeared to 'simulate' the chemical phenomenon dynamics during their context-based encounters with chemical ERs, suggesting that expertise is supported by a close coupling between perceptual and imagination systems, thus confirming the first conjecture. This study is among the first to objectively characterize the sensorimotor changes facilitated by training in a discipline.

The DBR project reported two design-development and testing iterations of a fully manipulable, interactive multi-representational computer interface. It was conceptualized to support integration of ERs at the middle-school level, based on the concept of oscillation. Results revealed that although sensorimotor interaction in general facilitates ER integration, high interactivity does not always lead to integration. As a corollary, there is no unique interaction pattern leading up to ER integration. This indicates that the relationship between sensorimotor interaction and ER integration is more complex than assumed in conjecture 2.

In summary, the similarity in eye movements in the way experts (study 1) and good integrators (study 2) interacted with the ERs together support the idea of sensorimotor markers of ER integration. These data are only indicative given the small samples, but it suggests the possibility of developing assessment models based on sensorimotor markers. However, such sensorimotor change is just one trackable outcome of ER integration. Generating such change does not guarantee ER integration.

6.2 Educational implications

This research, particularly its unique perspectives on the problem of science learning, has many different implications. I highlight a few important ones below.

Firstly, the conjecture that concepts are constituted by interaction with many ERs, and the converging results corroborating this model based on theoretical and empirical work, indicate that sensorimotor interaction supports ER integration. This suggests a focus on manipulative-based pedagogies, particularly those utilizing the potential of computer technologies and new-media as they make possible manipulation of ERs and observing the effects in real time. They also allow coupling static and dynamic states of ERs at will.

Secondly, the proposed model provides a theoretical justification for action-based learning. Recent research argues that (embodied) interactivity leads to learning (Abrahamson & Sánchez-García, 2016; Borar et al., 2017), particularly manipulation based on new-media. But it is not clear how manipulation contributes to learning. One proposal is that the process of interaction associates the self with perception and memory, and this leads to better cognition (Hung et al., 2014). A second approach argues for constructionism (Papert & Harel, 1991), which is considered as a new epistemology, where the central role of interactivity is the support it provides for collaborative building, of mathematical objects (using Logo) and complex systems (using NetLogo), based on manipulation-based programming. A recent third approach considers gestures in new computational media as similar to the process of gestures during the mathematical discovery process, which are hypothesized to be part of the mechanism that helps shift body-based intuitions (about possible results) into external symbolic proofs built using known and accepted mathematical structures (De Freitas & Sinclair, 2014; Sfard, 1991 & 2000). There is also a recent (fourth) approach which, through empirical research, indicates that (conceptual) learning through interaction (with ERs) is further strengthened when learners reflect on their interaction explicitly (Danish et al., 2015; Sengupta et al., 2015). The constitution view argued for in this dissertation suggests that actions done on manipulatives help in learning because actions are inherently integrative in nature. Every action requires a complex integration process, bringing together objects, forward models and feedback from various channels (visual, tactile, proprioception). This integration process would be primed when manipulatives are used to interact with symbolic entities, and this priming would help integrate different symbolic components in imagination.

Next, while this research suggests that interactivity is necessary for ER integration, it also shows that interactivity is not sufficient. The role of the teacher is crucial in directing learner attention to important parts of the interface, to facilitate the ER integration quickly and optimally. Apart from such cultural support needed to support interactivity, the system we developed has a more content-related limitation. Our computer interface was very useful for learners in understanding the 'controller' role of equations, where the equation is used to set the initial value of the variables. Once the oscillation starts, the equation works as a 'descriptor', because the variable values change as the simulation progresses, and this change is captured by the graph. However, the general equation embeds a third aspect, where it describes an idealized system that is true of all natural number values of the variables. This idealization, and the process by which it is derived using modeling and deductive thinking, are not supported by our interactive system. This is because the system only presents an instantiation of the general description provided by the equation, and this simulation of the general system only illustrates the oscillation behavior for a range of values. These illustrative cases may help in understanding the general case, but possibly not its derivation. The illustration of oscillation behavior can be considered similar to the way teacher demonstrates examples, by embodying and simulating the dynamic behavior using the blackboard and gestures. This process may allow the student to extend the specific cases the teacher illustrates, to reach a general case. However, there is another process the teacher illustrates, where she derives the oscillation equation. The inductive extension process does not work in understanding the derivation case, as derivations are based on model-based reasoning. This derivation process is not supported by current interactive systems. It is an open question whether the model-based reasoning involved in this process can be supported by interactive media, as the reasoning here proceeds using uninstantiated variables and general principles. These are integrated by the imagination process, to arrive at the model system. This integration capacity based on variables may well be a unique affordance of the imagination process, and manipulatives may not contribute much to this process.

Finally, extending the above reasoning, it is possible that a lot of interactivity may hinder or suppress imagination. If every manipulation is made possible externally, students may end up over-relying on the external world, and not feel the need to develop the internal-external coupling. This is not a problem if the external affordances are always available, as in the case of scientific calculators in advanced classes. But it would definitely be a problem if the external action possibilities are unavailable – which is most often the case.

6.3 Other contributions of the dissertation to the field

1. The comprehensive review of literature around ER integration and RC development provided in chapter 2 is one of its kind, as it brings together extensive and highly diverse theoretical as well as experimental work from different disciplines, among which no common threads are readily apparent.

2. The empirical projects reported in the dissertation are among the first to objectively characterize the sensorimotor changes facilitated by training in a domain. Findings from these projects and their conceptual background provide a fresh perspective towards theories of ER integration and expertise.

3. Our fully manipulable interface is one of the first theoretically motivated interventions targeting ER integration, by using interaction features emerging from DC and EC theories. It is also among the few DBR projects studying the development of RC using eye and mouse tracking.

4. The idea of making equations as an interface element, making them manipulable, and using equations as controllers, is first proposed and developed in this work, based on theoretical considerations. All other existing simulation systems either hide equations in code, only allowing discrete parameter changes, or display equations in a haphazard way. Our theory-based approach to using equations as control elements in the interface leads to the fundamental insight that formal systems are best understood as dynamic systems, capturing dynamic real-world behavior continuously.

5. This work is the first to systematically examine the relationships between interactivity, ER integration and learning. In contrast, most existing computer interventions assume interactivity is good, based on design principles coming from usability paradigms in HCI (human-computer interaction) and educational

technology design. The analysis presented here shows that usability design principles cannot be applied directly to the problem of learning complex representations and conceptual content.

6. The empirical work reported here led to the development of novel interaction-based methods to study problem-solving, using gaze and (inter)action tracking. The interaction analysis methods described in the dissertation are state-of-the-art, and emerged from dedicated collaborative work over the years with contributions from cognitive scientists, educators, computer scientists, teachers and students.

6.4 Limitations and future work

The work reported here only provides indicative data to support the conjectures related to the theoretical model, as the studies have the following set of limitations.

A major limitation of the empirical work reported in this dissertation is related to the use of eye-tracking methodology. When using eye-tracking technology, firstly, it cannot be known for sure if looking at something equals (consciously) processing it. Given that it is possible for humans (and perhaps animals as well) to physically have eyes pointed to an object in the external world and simultaneously attend to something else in imagination, the gaze behavior captured by eye-tracking may not always be related to the participant's cognitive processes. Related to the above, changes in the cognitive processes may not always reflect in changes in gaze behavior.

Secondly, the statistical outputs of eye-tracking often contain systematic errors to a certain degree, arising out of individual differences in calibration accuracy and precision. Particularly when dealing with stimuli displayed on computer screens, the issues of staggering of stimulus or task windows (for each individual as well as across participants) can further increase the errors. Moreover, there is often loss of gaze data points due to several unavoidable factors such as blinks, proximity to the laptop screen, rapid head-movements, or moving out of the eye-tracking zone, etc. However, considering the state-of-the-art of the technology, such errors are permitted within a certain range, depending on the context of the experiment.

Next, part of the work reported is based on the conjecture that ER integration is cognitively more fundamental as well as simpler than conceptualization. This conjecture grants that the process of conceptualization, in relation to sensorimotor and ER integration, may involve feedback loops, and that there may be many different phases in-between. Future research is needed to test this conjecture systematically from the point of view of designing and sequencing instructions specifically for conceptualization, and understanding the complex relation between sensorimotor interaction and conceptualization. One interesting question here is: should instructional tools provide perceptual experiences related to learning content (such as simulations), before introducing concepts (such as physical laws), to help students have a concrete cognitive base for better comprehension of the concepts?

In the second project, students experienced a simulation intervention that provided perceptual bases for about ten minutes, instruction was provided for about five minutes. If students spent more time interacting with the simulation, and more instruction was provided, it is possible that results with larger effect sizes would have emerged. Generally, it takes a considerable amount of time before complex embodied experiences are internalized, to the point that the internalized traces can be used to run imagination events. Even with a carefully designed instructional process to provide embodied experience, more intervention time (long enough for students to fully embody their perceptual experiences) would make the imagination effects stronger. It would also be useful to investigate whether different instruction types activate the multimodal representation differently.

Further, the result that the same patterns of interactivity lead to different understandings about the relationships between ERs suggests that these patterns of sensorimotor interaction need to be supported by instructions, constantly modulated based on contexts and the nature of students, which only a good teacher can provide. Future studies are needed to investigate (a) how the interface would help a teacher in scaffolding ER integration, and (b) how the teacher would address the need to constantly modulate the complex classroom dynamics based on interactions between technology, conceptual content, materials and students. Future work in this direction may also include opportunities for students to reflect on their own actions, effects of those actions on the ERs, the relationships between the ERs, etc. A post-facto analysis (not included in this dissertation) indicated that the interviews conducted in Iteration 2 may have had further effects on student ER integration abilities. This has been reported previously in the case of conceptual learning through agent-based modelling environments (e.g. Sengupta et al., 2015) and computer-supported collaborative learning environments (e.g. Danish et al., 2015).

From the technology point of view, the interface design discussed in this work allowed student to see in real-time on screen the changes that resulted from interaction or manipulation. However, with the emergence of more enactive, embodied and immersive (new-media) technology platforms, the instruction design principles and features outlined in this dissertation could be fused with gesture-based control (Kinect, Wii, LeapMotion) and/or haptic devices, which allow experiences that imitate kinaesthetic movements associated with interaction with physical objects. Such experiences would allow students to have a richer (multimodal) experience of formal systems. Future studies are needed to compare the advantages and limitations of such different new-media interventions.

References

Acevedo Nistal, A., Van Dooren, W., Clarebout, G., Elen, J., & Verschaffel, L. (2010). Representational flexibility in linear-function problems: A choice/no-choice study. In L. Verschaffel, E. D. Corte, J. Elen, & T. D. Jong (Eds.), *Use of representations in reasoning and problem solving: Analysis and improvement* (pp. 74–93). UK: Routledge.

Acevedo Nistal, A., Van Dooren, W., & Verschaffel, L. (2012). What counts as a flexible representational choice? An evaluation of students' representational choices to solve linear function problems. *Instructional Science*, *40*(6), 999–1019.

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*(2), 131–152.

Ainsworth, S. (2008). The educational value of multiple-representations when learning complex scientific concepts. *Visualization: Theory and Practice in Science Education*. Retrieved from http://link.springer.com/content/pdf/10.1007/978-1-4020-5267-5.pdf#page=194

Amiel, T., & Reeves, T. C. (2008). Design-Based Research and Educational Technology: Rethinking Technology and the Research Agenda. *Journal of Educational Technology & Society*, *11*(4), 29–40.

Anastasiadou, S., & Chadjipantelis, T. (2008). The role of representations in the understanding of probabilities in tertiary education. In *Eleventh International Congress on Mathematical Education*. Monterrey, Mexico.

Anzai, Y. (1991). Learning and use of representations for physics expertise. Towards a general theory of expertise.

Anzai, Y., & Yokoyama, T. (1984). Internal Models in Physics Problem Solving. *Cognition and instruction*, *1*(4), 397–450.

Aurigemma, J., & Chandrasekharan, S. (2013). Turning experiments into objects: The cognitive processes involved in the design of a lab‐on‐a‐chip device. *Journal of*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/jee.20003/full

Barsalou, L. W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences*, *22*(4), 577–609; discussion 610–60.

Battista, M. T. (2003). Understanding students'. In *thinking about area and volume measurement. Learning and teaching measurement* (pp. 122–142).

Battista, M. T., & Clements, D. H. (1996). Students' Understanding of Three-Dimensional Rectangular Arrays of Cubes. *Journal for Research in Mathematics Education*, *27*(3), 258–292.

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1987). Students' visualization of a chemical reaction. *Education in Chemistry*, *24*(4), 117–120.

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1988). Theories, principles and laws. *Education in Chemistry*.

Biber, A. Ç. (2014). Mathematics teacher candidates' skills of using multiple representations for division of fractions. *Educational Research and Evaluation: An International Journal on Theory and Practice*, *9*(8), 237–244.

Billings, E., & Klanderman, D. (2000). Graphical Representations of Speed: Obstacles Preservice K−8 Teachers Experience. *School Science and Mathematics*, *100*(8), 440–450.

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and Instruction*, *14*(3), 325–341.

Bodner, G. M., & Domin, D. S. (2000). Mental models: The role of representations in problem solving in chemistry. *University Chemistry Education*, *4*(1).

Borar, P., Karnam, D., Agrawal, H., & Chandrasekharan, S. (2017, September). Augmenting the Textbook for Enaction: Designing Media for Participatory Learning in Classrooms. In *IFIP Conference on Human-Computer Interaction* (pp. 336-339). Springer, Cham.

Boucheix, J.-M., Lowe, R. K., Putri, D. K., & Groff, J. (2013). Cueing animations: Dynamic signalling aids information extraction and comprehension. *Learning and Instruction*, *25*(Supplement C), 71–84.

Braithwaite, D. W., Goldstone, R. L., van der Maas, H. L. J., & Landy, D. H. (2016). Non-formal mechanisms in mathematical cognitive development: The case of arithmetic. *Cognition*, *149*, 40–55.

Brenner, M. E., Herman, S., Ho, H.-Z., & Zimmer, J. M. (1999). Cross-National Comparison of Representational Competence. *Journal for Research in Mathematics Education*, *30*(5), 541–557.

Briggs, M., & Bodner, G. (2005). A model of molecular visualization. *Visualization in Science Education*. Retrieved from http://link.springer.com/content/pdf/10.1007/1-4020-3613-2.pdf#page=97

Brooks, R. A. (1991). Intelligence without representation. *Artificial intelligence*, 47(1-3), 139-159.

Bub, D. N., & Masson, M. E. J. (2012). On the dynamics of action representations evoked by names of manipulable objects. *Journal of Experimental Psychology. General*, *141*(3), 502–517.

Cardella, M. E., Atman, C. J., & Adams, R. S. (2006). Mapping between design activities and external representations for engineering student designers. *Design Studies*, *27*(1), 5–24.

Çelik, D., & Sağlam Arslan, A. (2012). The Analysis of Teacher Candidates' Translating Skills in Multiple Representations. *Elementary Education Online*, *11*(1), 239–250.

Chandrasekharan, S. (2009). Building to discover: a common coding model. *Cognitive Science*, *33*(6), 1059–1086.

Chandrasekharan, S. (2013). The cognitive science of Feynmen. *Metascience*, *22*(3), 647–652.

Chandrasekharan, S. (2014). Becoming knowledge: Cognitive and neural mechanisms that support scientific intuition. *Rational Intuition: Philosophical Roots, Scientific Investigations*, 307–337.

Chandrasekharan, S. (2016). Beyond Telling: Where New Computational Media is Taking Model-Based Reasoning. In L. Magnani & C. Casadio (Eds.), *Model-Based Reasoning in Science and Technology* (pp. 471–487). Springer International Publishing.

Chandrasekharan, S., & Nersessian, N. J. (2015). Building Cognition: The Construction of Computational Representations for Scientific Discovery. *Cognitive Science*, *39*(8), 1727–1763.

Chandrasekharan, S., Nersessian, N.J. (2018). Rethinking correspondence: how the process of constructing models leads to discoveries and transfer in the bioengineering sciences.. Synthese, DOI 10.1007/s11229-017-1463-3

Chandrasekharan, S., & Osbeck, L. (2010). Rethinking Situatedness. *Theory & Psychology*, *20*(2), 171–207.

Chandrasekharan, S., & Tovey, M. (2012). Sum, quorum, tether: Design principles underlying external representations that promote sustainability. Pragmatics & Cognition, 20(3), 447-482.

Chen, F., Hu, Z., Zhao, X., Wang, R., Yang, Z., Wang, X., & Tang, X. (2006). Neural correlates of serial abacus mental calculation in children: A functional MRI study. *Neuroscience Letters*, 403(1), 46–51.

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.

Çikla, O. A., & Çakiroglu, E. (2006). Seventh grade students' use of multiple representations in pattern related algebra tasks. In *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi* (pp. 31–31).

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

Cole, M., & Engeström, Y. (1993). A cultural-historical approach to distributed cognition. *Distributed Cognitions: Psychological and Educational Considerations*, 1–46.

Consortium, C. (2004). Molecular workbench. *Java Simulations and Modeling tools), (2004–2013). A*.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, *90*(6), 1073–1091.

Cook, M., Wiebe, E. N., & Carter, G. (2008). The influence of prior knowledge on viewing and interpreting graphics with macroscopic and molecular representations. *Science Education*, *92*(5), 848–867.

Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction*, *9*(4), 343–363.

Danish, J. A., Enyedy, N., Saleh, A., Lee, C., Andrade, A. (2015). Science through Technology Enhanced Play: Designing to support reflection through play and embodiment. In Lindwall, O., Häkkinen, P., Koschman, T. Tchounikine, P. & Ludvigsen, S. (Eds.) *"Exploring the material conditions of learning: opportunities and challenges for CSCL," the Proceedings of the Computer Supported Collaborative Learning (CSCL) Conference,* (Vol. 1, pp. 332 – 339). Gothenburg, Sweden: The International Society of the Learning Sciences.

Davidowitz, B., Chittleborough, G., & Murray, E. (2010). Student-generated submicro diagrams: a useful tool for teaching and learning chemical equations and stoichiometry. *Chemistry Education Research and Practice*, *11*(3), 154–164.

De Jong, T., & Van Joolingen, W. R. (1998). Scientific Discovery Learning with Computer Simulations of Conceptual Domains. *Review of Educational Research*, *68*(2), 179–201.

DeWolf, M., Son, J. Y., Bassok, M., & Holyoak, K. J. (2017). Relational Priming Based on a Multiplicative Schema for Whole Numbers and Fractions. *Cognitive Science*. https://doi.org/10.1111/cogs.12468

Dickes, A. C., Sengupta, P., Farris, A. V., & Basu, S. (2016). Development of Mechanistic Reasoning and Multilevel Explanations of Ecology in Third Grade Using Agent−Based Models. *Science Education*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/sce.21217/full

Diefes-Dux, H. A., Moore, T., Zawojewski, J., Imbrie, P. K., & Follman, D. (2004). A framework for posing open-ended engineering problems: model-eliciting activities. In *34th Annual Frontiers in Education, 2004. FIE 2004.* (Vol. (pp., pp. F1A–3–8 Vol. 2). IEEE.

Diezmann, C. M., & English, L. D. (2001). Promoting the use of diagrams as tools for thinking. In A. A. Cuoco (Ed.) (pp. 77–89). National Council of Teachers of Mathematics.

diSessa, A. A. (2004). Metarepresentation: Native Competence and Targets for Instruction. *Cognition and Instruction*, *22*(3), 293–331.

diSessa, A. A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *The Journal of Mathematical Behavior*, *10*, 117–160.

Disessa, A., & Sherin, B. (2000/0). Meta-representation: an introduction. *The Journal of Mathematical Behavior*, *19*(4), 385–398.

Domin, D., & Bodner, G. (2012). Using students' representations constructed during problem solving to infer conceptual understanding. *Journal of Chemical Education*, *89*(7), 837–843.

Dreyfus, A., & Jungwirth, E. (1990). Macro and micro about the living cell: Which explains what. Relating macroscopic phenomena to microscopic particles.

Dreyfus, T. (2002). Advanced Mathematical Thinking Processes. In D. Tall *Advanced Mathematical Thinking* (pp. 25–41). Springer, Dordrecht.

Duval, R. (2006). A Cognitive Analysis of Problems of Comprehension in a Learning of Mathematics. *Educational Studies in Mathematics*, *61*(1-2), 103–131.

Elia, I., Panaoura, A., Eracleous, A., & Gagatsis, A. (2007). Relations between secondary pupils' conceptions about functions and problem solving in different representations. *International Journal of Science and Mathematics Education*, *5*(3), 533–556.

Enyedy, N., Danish, J. A., Delacruz, G., & Kumar, M. (2012). Learning physics through play in an augmented reality environment. *International Journal of Computer-Supported Collaborative Learning,* 7(3), 347- 378.

Enyedy, N., Danish, J. A., DeLiema, D. (2015). Constructing liminal blends in a collaborative augmented-reality learning environment. *International Journal of Computer Supported Collaborative Learning,* 10(1), 7-34.

Fischer, M. H. (2012). A hierarchical view of grounded, embodied, and situated numerical cognition. *Cognitive Processing*, *13*(1), 161–164.

Garforth, F. M., Johnstone, A. H., & Lazonby, J. N. (1976). Ionic Equations: Difficulties in Understanding and Use. *Education in Chemistry*, *13*, 72–75.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise Differences in the Comprehension of Visualizations: a Meta-Analysis of Eye-Tracking Research in Professional Domains. *Educational Psychology Review*, *23*(4), 523–552.

Giere, R. N., & Moffatt, B. (2003). Distributed cognition: Where the cognitive and the social merge. *Social Studies of Science*, *33*(2), 301–310.

Gilbert, J. (2005). Visualization: A metacognitive skill in science and science education. *Visualization in Science Education*. Retrieved from http://link.springer.com/content/pdf/10.1007/1-4020-3613-2.pdf#page=46

Goel, V. (1995). *Sketches of Thought*. MIT Press.

Goldin, G. A., & Kaput, J. J. (1996). A joint perspective on the idea of representation in learning and doing mathematics. *Theories of Mathematical Learning*. Retrieved from https://www.researchgate.net/profile/Gerald_Goldin/publication/269407907_A_joint_perspective_on_the_idea_of_representation_in_learning_and_doing_mathematics/links/55ecacac08ae65b6389f3d5b.pdf

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585–612.

Gooding, D. C. (2006). From Phenomenology to Field Theory: Faraday's Visual Reasoning. *Perspectives on Science: Historical, Philosophical, Social*, *14*(1), 40–65.

Grove, N. P., Cooper, M. M., & Cox, E. L. (2012). Does Mechanistic Thinking Improve Student Success in Organic Chemistry? *Journal of chemical education*, *89*(7), 850–853.

Hanakawa, T., Honda, M., Okada, T., Fukuyama, H., & Shibasaki, H. (2003). Neural correlates underlying mental calculation in abacus experts: A functional magnetic resonance imaging study. *NeuroImage*, 19(2), 296–307.

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, *8*(6), 280–285.

Henderson, J. M., & Ferreira, F. (2004). Scene Perception for Psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 1–58). New York: Psychology Press.

Herbert, D. M. (1988). Study Drawings in Architectural Design: Their Properties as a Graphic Medium. *Journal of Architectural Education*, *41*(2), 26–38.

Herron, J. D., & Greenbowe, T. J. (1986). What can we do about Sue: A case study of competence. *Journal of Chemical Education*. Retrieved from http://pubs.acs.org/doi/pdf/10.1021/ed063p528

Hinton, M. E., & Nakhleh, M. B. (1999). Students' microscopic, macroscopic, and symbolic representations of chemical reactions. *Chemical Educator*. Retrieved from http://link.springer.com/article/10.1007/s00897990325a

Hitt, F. (1998). Difficulties in the articulation of different representations linked to the concept of function. *The Journal of Mathematical Behavior*, *17*(1), 123–134.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed Cognition: Toward a New Foundation for Human-computer Interaction Research. *ACM Trans. Comput. -Hum. Interact.*, *7*(2), 174–196.

Holsanova, J. (2014). Reception of multimodality: Applying eye tracking methodology in multimodal research. *Routledge Handbook of Multimodal Analysis*. Retrieved from https://www.researchgate.net/profile/Jana_Holsanova/publication/259501157_Reception_of_multimodality_Applying_eye_tracking_methodology_in_multimodal_research/links/546e31990cf2b5fc17606f88.pdf

Hutchins, E. (1995a). *Cognition in the Wild*. MIT Press.

Hutchins, E. (1995b). How a Cockpit Remembers Its Speeds. *Cognitive Science*, *19*(3), 265–288.

Hutchins, E. (2014). The cultural ecosystem of human cognition. *Philosophical Psychology*, *27*(1), 34–49.

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1985). Direct Manipulation Interfaces. *Human–Computer Interaction*, *1*(4), 311–338.

Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. *The Interface of Language, Vision, and Action: Eye Movements and the Visual World*, *217*, 105–133.

Izsák, A. (2011). Representational Competence and Algebraic Modeling. In *Early Algebraization* (pp. 239–258). Springer, Berlin, Heidelberg.

Jacobson, M. J. (2001). Problem solving, cognition, and complex systems: Differences between experts and novices. *Complexity*, *6*(3), 41–49.

Jensen, W. B. (1998). Logic, History, and the Chemistry Textbook: I. Does Chemistry Have a Logical Structure? *Journal of Chemical Education*, *75*(6), 679.

Jmol. (2014). an open-source Java viewer for chemical structures in 3D. Retrieved from http://www.jmol.org/

Johnstone, A. H. (1982). Macro and microchemistry. *School Science Review*, *64*(227), 377–379.

Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning*, *7*(2), 75–83.

Johnstone, A. H. (2000). TEACHING OF CHEMISTRY - LOGICAL OR PSYCHOLOGICAL? *Chemistry Education Research and Practice*, *1*(1), 9–15.

Johri, A., & Lohani, V. K. (2011). Framework for improving engineering representational literacy by using pen-based computing. *International Journal of Engineering Education*. Retrieved from http://mason.gmu.edu/~johri/publications/Johri %20Lohani%20IJEE%20Representational%20Learning%20Draft.pdf

Johri, A., Roth, W. M., & Olds, B. M. (2013). The role of representations in engineering practices: Taking a turn towards inscriptions. *Journal of Engineering Education*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/jee.20005/full

Kamii, C., Kirkland, L., & Lewis, B. A. (2001). Representation and abstraction in young children's numerical reasoning. The roles of representation in school mathematics.

Kamii, C., & Kysh, J. (2006). The difficulty of "length× width": Is a square the unit of measurement? *The Journal of Mathematical Behavior*, *25*(2), 105–115.

Kandel, S., Orliaguet, J. P., & Viviani, P. (2000). Perceptual anticipation in handwriting: the role of implicit motor competence. *Perception & Psychophysics*, *62*(4), 706–716.

Kapteijn, M. (1990). The functions of organizational levels in biology for describing and planning biology education. *Relating Macroscopic Phenomena to*.

Kaput, J. J. (1995). A Research Base Supporting Long Term Algebra Reform?

Karnam, D., Agrawal, H., Mishra, D., & Chandrasekharan, S. (2016). Interactive vectors for model-based reasoning. In W. Chen, T. Supnithi, A. F. Mohd Ayub, M. Mavinkurve, T. Kojiri, W. Chen, J. Yang, S. Murthy, S. L. Wong & S. Iyer. (Eds.), *Proceedings of the 24th international conference on computers in education* (pp. 401–406). Mumbai, India: Asia-Pacific Society for Computers in Education.

Keller, B. A., & Hirsch, C. R. (1998). Student preferences for representations of functions. *International Journal of Mathematical Education in Science and Technology*, *29*(1), 1–17.

Kellman, P. J., & Garrigan, P. (2009). Perceptual learning and human expertise. *Physics of Life Reviews*, *6*(2), 53–84.

Kelly, R. M., & Jones, L. L. (2008). Investigating Students' Ability To Transfer Ideas Learned from Molecular Animations of the Dissolution Process. *Journal of Chemical Education*, *85*(2), 303.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics*, *12*(3), 317–326.

Kindfield, A. C. H. (1994). Biology Diagrams: Tools to Think With. *Journal of the Learning Sciences*, *3*(1), 1–36.

Kirsh, D. (1996). Adapting the Environment Instead of Oneself. *Adaptive Behavior*, *4*(3-4), 415–452.

Kirsh, D. (2001). The Context of Work. *Human–Computer Interaction*, *16*(2-4), 305–322.

Kirsh, D. (2010). Thinking with external representations. *AI & SOCIETY*, *25*(4), 441–454.

Kirsh, D. (2013). Embodied cognition and the magical future of interaction design. *ACM Transactions on Computer-Human Interaction: A Publication of the Association for Computing Machinery*, *20*(1), 1–30.

Kirsh, D., & Maglio, P. (1994). On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science*, *18*(4), 513–549.

Kirshner, D., & Awtry, T. (2004). Visual Salience of Algebraic Transformations. *Journal for Research in Mathematics Education*, *35*(4), 224–257.

Knuth, E. J. (2000). Student Understanding of the Cartesian Connection: An Exploratory Study. *Journal for Research in Mathematics Education*, *31*(4), 500–507.

Koedinger, K. R., Alibali, M. W., & Nathan, M. J. (2008). Trade‐Offs between grounded and abstract representations: Evidence from algebra problem Solving. *Cognitive*

Science. Retrieved from http://onlinelibrary.wiley.com/doi/10.1080/03640210701863933/full

Kohl, P. B., & Finkelstein, N. D. (2008). Patterns of multiple representation use by experts and novices during physics problem solving. *Physical Review Special Topics - Physics Education Research*, *4*(1), 010111.

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, *13*(2), 205–226.

Kozma, R. B., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*. Retrieved from http://www.academia.edu/download/31906904/kozma_russell_jrst_article.pdf

Kozma, R., Chin, E., Russell, J., & Marx, N. (2000). The Roles of Representations and Tools in the Chemistry Laboratory and Their Implications for Chemistry Learning. *Journal of the Learning Sciences*, *9*(2), 105–143.

Kozma, R., & Russell, J. (2005). Students Becoming Chemists: Developing Representationl Competence. In *Visualization in Science Education* (pp. 121–145). Springer, Dordrecht.

Kukkonen, J. E., Kärkkäinen, S., Dillon, P., & Keinonen, T. (2014). The Effects of Scaffolded Simulation-Based Inquiry Learning on Fifth-Graders' Representations of the Greenhouse Effect. *International Journal of Science Education*, *36*(3), 406–424.

Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic component of image perception in mammogram interpretation: gaze-tracking study. *Radiology*, *242*(2), 396–402.

Landy, D., Allen, C., & Zednik, C. (2014). A perceptual account of symbolic reasoning. *Frontiers in Psychology*, *5*, 275.

Landy, D., & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(4), 720–733.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*(1), 65–100.

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

Latour, B. (1993). Pasteur on lactic acid yeast: a partial semiotic analysis. *Configurations*, *1*(1), 129–145.

Lehrer, R., & Schauble, L. (2006). *Cultivating Model-Based Reasoning in Science Education*. Cambridge University Press.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, Graphs, and Graphing: Tasks, Learning, and Teaching. *Review of Educational Research*, *60*(1), 1–64.

Lesh, R., & Doerr, H. (2003/May). Foundations of a model and modeling perspective on mathematics teaching, learning, and problem solving. In R. Lesh & H. Doerr (Eds.), *Beyond Constructivism: Models and Modeling Perspectives on Mathematics Problem Solving, Learning, and Teaching* (pp. 3–33). Lawrence Erlbaum Associates.

Lesh, R., Post, T., & Behr, M. (1987). Representations and translations among representations in mathematics learning and problem solving. *Problems of Representation in the Teaching and*.

Levy, S. T., & Wilensky, U. (2009). Crossing Levels and Representations: The Connected Chemistry (CC1) Curriculum. *Journal of Science Education and Technology*, *18*(3), 224–242.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by Embodiment: Six Precepts for Research on Embodied Learning and Mixed Reality. *Educational Researcher,* 42 (8), 445-452. doi: 10.3102/0013189x13511661

Longden, B. (1982). Genetics—are there inherent learning difficulties? *Journal of Biological Education*, *16*(2), 135–140.

Lowe, R., & Boucheix, J. M. (2008). Learning from animated diagrams: How are mental models built?. Diagrammatic representation and inference.

Madden, S. P., Jones, L. L., & Rahm, J. (2011). The role of multiple representations in the understanding of ideal gas problems. *Chemistry Education Research and Practice*, *12*(3), 283–293.

Majumdar, R., Kothiyal, A., Ranka, A., Pande, P., Murthy, S., Agarwal, H., & Chandrasekharan, S. (n.d.). The Enactive equation: Exploring How Multiple External Representations are Integrated, Using a Fully Controllable Interface and Eye-Tracking. In *2014 IEEE Sixth International Conference on Technology for Education* (pp. 233–240). Amritapuri, Clappana P. O., India: IEEE.

Mammino, L. (2008). Teaching Chemistry with and without External Representations in Professional Environments with Limited Resources. In J. K. Gilbert (Ed.), *Visualization: Theory and Practice in Science Education* (pp. 155–185). Springer.

Manovich, L. (2001). *The language of new media*. MIT press.

Marshall, P. (2007). Do tangible interfaces enhance learning? In *Proceedings of the 1st international conference on Tangible and embedded interaction* (pp. 163–170). ACM.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science*, *29*(4), 587–625.

Mathai, S., & Ramadas, J. (2007). Visualising structure and function of the digestive system. A poster presented at. In *the Gordon Research Conference on Visualization in Science and Education*. Bryant University, RI, USA.

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, *41*, 85–139.

Mayer, R. E. (2005). Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity and Temporal Contiguity Principles. *The Cambridge Handbook of Multimedia Learning.* Retrieved from http://psycnet.apa.org/record/2006-00633-012

McCracken, W. M., & Newstetter, W. C. (2001). Text to diagram to symbol: representational transformations in problem-solving. In *31st Annual Frontiers in Education Conference. Impact on Engineering and Science Education. Conference Proceedings (Cat. No.01CH37193)* (Vol. 2, pp. F2G–13–17 vol.2).

McElhaney, K. W., Chang, H.-Y., Chiu, J. L., & Linn, M. C. (2015). Evidence for effective uses of dynamic visualisations in science curriculum materials. *Studies in Science Education*, *51*(1), 49–85.

Moore, T. J., Miller, R. L., Lesh, R. A., Stohlmann, M. S., & Kim, Y. R. (2013). Modeling in Engineering: The Role of Representational Fluency in Students' Conceptual Understanding. *Journal of Engineering Education*, *102*(1), 141–178.

Nakhleh, M. B., & Krajcik, J. S. (1994). Influence of levels of information as presented by different technologies on students' understanding of acid, base, and ph concepts. *Journal of Research in Science Teaching*, *31*(10), 1077–1096.

Nakhleh, M. B., & Postek, B. (2008). Learning chemistry using multiple external representations. *Visualization: Theory and Practice in Science*. Retrieved from http://link.springer.com/content/pdf/10.1007/978-1-4020-5267-5.pdf#page=212

Nersessian, N. J. (2010). *Creating Scientific Concepts*. MIT Press.

Nitz, S., Nerdel, C., & Prechtl, H. (2012). Modeling the Relationship between Representational Competence and Domain Knowledge. In E. de Vries & K. Scheiter (Eds.), *of Representational Technologies in Education and Training?* (pp. 160–162).

Nitz, S., & Tippett, C. D. (2012). Measuring Representational Competence in Science. In E. de Vries & K. Scheiter (Eds.), *of Representational Technologies in Education and Training?* (pp. 163–165).

Nurrenbern, S. C., & Pickering, M. (1987). Concept learning versus problem solving: Is there a difference? *Journal of Chemical Education*, *64*(6), 508.

O'Malley, M. A., & Soyer, O. S. (2012). The roles of integration in molecular systems biology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*(1), 58–68.

Ozogul, G., Johnson, A. M., Moreno, R., & Reisslein, M. (2012). Technological Literacy Learning With Cumulative and Stepwise Integration of Equations Into Electrical Circuit Diagrams. *IEEE Transactions on Education*, *55*(4), 480–487.

Padalkar, S., & Hegarty, M. (2015). Models as feedback: Developing representational competence in chemistry. *Journal of Educational Psychology*, *107*(2), 451.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*. Retrieved from http://doi.apa.org/journals/cep/45/3/255.pdf

Pande, P., & Chandrasekharan, S. (2014). Categorization of multiple external representations by chemistry undergrads: An eye-tracking study. In *Tan (Eds.), Proceedings of the International Science Education Conference* (pp. 1393–1415).

Pande, P., & Chandrasekharan, S. (2017). Representational competence: towards a distributed and embodied cognition account. *Studies in Science Education*, *0*(0), 1–43.

Pande, P., & Ramadas, J. (2013). Students' Measurement Experiences and Responses to Length Comparison, Seriation and Proportioning Tasks. In N. G., A. Jamakhandi, & S. Ebie (Eds.), *Proceedings of epiSTEME -* (pp. 5–145). Mumbai: HBCSE, TIFR.

Pande, P., Shah, P., & Chandrasekharan, S. (2015). How do experts and novices navigate chemistry representations: An eye-tracking investigation, to appear. In *the Proceedings of episteme-6 the Sixth International Conference to review research on Science Technology and Mathematics Education*.

Pape, S. J., & Tchoshanov, M. A. (2001). The Role of Representation(s) in Developing Mathematical Understanding. *Theory into Practice*, *40*(2), 118–127.

Pea, R. D. (1993). Practices of distributed intelligence and designs for education. *Distributed Cognitions: Psychological and Educational Considerations*, *11*, 47–87.

Perkins, K., Adams, W., Dubson, M., Finkelstein, N., Reid, S., Wieman, C., & LeMaster, R. (2006). PhET: Interactive Simulations for Teaching and Learning Physics. *Physics Teacher*, *44*(1), 18–23.

Prain, V., & Tytler, R. (2012). Learning Through Constructing Representations in Science: A framework of representational construction affordances. *International Journal of Science Education*, *34*(17), 2751–2773.

Purcell, A. T., & Gero, J. S. (1998). Drawings and the design process: A review of protocol studies in design and other disciplines and related research in cognitive psychology. *Design Studies*, *19*(4), 389–430.

Rahaman, J., Agrawal, H., Srivastava, N., & Chandrasekharan, S. (2017). Recombinant Enaction: Manipulatives Generate New Procedures in the Imagination, by Extending and Recombining Action Spaces. *Cognitive Science*. https://doi.org/10.1111/cogs.12518

Rau, A. K., Moll, K., Snowling, M. J., & Landerl, K. (2015). Effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently. *Journal of Experimental Child Psychology*, *130*, 92–105.

Reed, E. S. (1988). James J. Gibson and the psychology of perception. Yale University Press.

Reisslein, M., Moreno, R., & Ozogul, G. (2010). Pre−college electrical engineering instruction: The impact of abstract vs. contextualized representation and practice on learning. *Journal of Engineering*, *99*(3), 225–235.

Rivera, J., & Garrigan, P. (2016). Persistent perceptual grouping effects in the evaluation of simple arithmetic expressions. *Memory & Cognition*, *44*(5), 750–761.

Roth, W., & Tobin, K. (1997). Cascades of inscriptions and the re−presentation of nature: how numbers, tables, graphs, and money come to re−present a rolling ball. *International Journal of Science Education*, *19*(9), 1075–1091.

Russell, J., Kozma, R., Zohdy, M., Susskind, T., Becker, D., & Russell, C. (2000). SMV: Chem (Simultaneous Multiple Representations in Chemistry). New York: John Wiley.

Russell, J. W., & Kozma, R. B. (1994). 4M:CHEM-Multimedia and Mental Models in Chemistry. *Journal of Chemical Education*, *71*(8), 669.

Rutten, N., van Joolingen, W. R., & van der Veen, J. T. (2012). The learning effects of computer simulations in science education. *Computers & Education*, *58*(1), 136–153.

Salomon, G. (1997). *Distributed Cognitions: Psychological and Educational Considerations*. Cambridge University Press.

Sanger, M. J. (2005). Evaluating students' conceptual understanding of balanced equations and stoichiometric ratios using a particulate drawing. *Journal of Chemical Education*, *82*(1), 131–134.

Santos-Trigo, M. (1996). An exploration of strategies used by students to solve problems with multiple ways of solution. *The Journal of Mathematical Behavior*, *15*(3), 263–284.

Savoy, L. G. (1988). Balancing Chemical Equations. *The School Science Review*, *69*(249), 713–720.

Sengupta, P., Krinks, K. D., & Clark, D. B. (2015). Learning to deflect: Conceptual change in physics during digital game play. *Journal of the Learning Sciences*, *24*(4), 638-674.

Scanlon, T. (1998). *What We Owe to Each Other*. Harvard University Press.

Schnepp, M., & Nemirovsky, R. (2001). Constructing a foundation for the fundamental theorem of calculus. The roles. In *of representation in school mathematics* (pp. 90–102).

Schnotz, W. (2002). Commentary: Towards an Integrated View of Learning from Text and Visual Displays. *Educational Psychology Review*, *14*(1), 101–120.

Schönborn, K. J., & Anderson, T. R. (2009). A model of factors determining students' ability to interpret external representations in biochemistry. *International Journal of Science*. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/09500690701670535

Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, *11*(5), 211–218.

Schwartz, D. L., & Black, J. B. (1996). Analog Imagery in Mental Model Reasoning: Depictive Models. *Cognitive Psychology*, *30*(2), 154–219.

Schwartz, D. L., & Black, J. B. (1996). Shuttling Between Depictive Models and Abstract Rules: Induction and Fallback. *Cognitive Science*, *20*(4), 457–497.

Schwartz, D. L., & Black, T. (1999). Inferences through imagined actions: Knowing by simulated doing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *25*(1), 116.

Sevian, H., Bernholt, S., Szteinberg, G. A., Auguste, S., & Pérez, L. C. (2015). Use of representation mapping to capture abstraction in problem solving in different courses in chemistry. *Chemistry Education Research and Practice*, *16*(3), 429–446.

Shubbar, K. E. (1990). Learning the Visualisation of Rotations in Diagrams of Three Dimensional Structures. *Research in Science & Technological Education*, *8*(2), 145–154.

Stewart, J. (1983). Student problem solving in high school genetics. *Science education*, *67*(4), 523–540.

Stewart, J. H. (1982). Difficulties Experienced by High School Students When Learning Basic Mendelian Genetics. *The American Biology Teacher*, *44*(2), 80–89.

Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying Representational Competence With Multi-Representational Displays. *Cognition and Instruction*, *29*(1), 123–145.

Stieff, M., & McCombs, M. (2006). Increasing representational fluency with visualization tools. In *Proceedings of the 7th international conference on Learning sciences, 730-736. International Society of the Learning Sciences*.

Stieff, M., & Wilensky, U. (2003). Connected Chemistry—Incorporating Interactive Simulations into the Chemistry Classroom. *Journal of Science Education and Technology*, *12*(3), 285–302.

Subramaniam, K., & Padalkar, S. (2009). Visualisation and Reasoning in Explaining the Phases of the Moon. *International journal of science education*, *31*(3), 395–417.

Sweller, J., & Chandler, P. (1991). Evidence for Cognitive Load Theory. *Cognition and Instruction*, *8*(4), 351–362.

Talanquer, V. (2013). How Do Students Reason About Chemical Substances and Reactions? In *Concepts of Matter in Science Education* (pp. 331–346). Springer, Dordrecht.

Tasker, R. F., Chia, W., Bucat, R. B., & Sleet, R. (1996). The VisChem Project-visualising chemistry with multimedia. *CHEMISTRY IN NEW ZEALAND*.

Taylor, J. E. T., Witt, J. K., & Grimaldi, P. J. (2012). Uncovering the connection between artist and audience: viewing painted brushstrokes evokes corresponding action representations in the observer. *Cognition*, *125*(1), 26–36.

Thomas, L. E., & Lleras, A. (2009). Covert shifts of attention function as an implicit aid to insight. *Cognition*, *111*(2), 168–174.

Thompson, P. W., & Sfard, A. (1994). Problems of reification: Representations and mathematical objects. In *Proceedings of the annual meeting of the International Group for the Psychology of Mathematics Education—North America, plenary sessions* (Vol. 1, pp. 1–32).

Tsui, C.-Y., & Treagust, D. F. (2003). Genetics Reasoning with Multiple External Representations. *Research in Science Education*, *33*(1), 111–135.

Tsui, C.-Y., & Treagust, D. F. (2013). Introduction to Multiple Representations: Their Importance in Biology and Biological Education. In *Multiple Representations in Biological Education* (pp. 3–18). Springer, Dordrecht.

Tytler, R., & Prain, V. (2010). A Framework for Re‐thinking Learning in Science from Recent Cognitive Science Perspectives. *International Journal of Science Education*, *32*(15), 2055–2078.

Tytler, R., Prain, V., Hubber, P., & Haslam, F. (2013). Reasoning in Science through Representation. In *Constructing Representations to Learn in Science* (pp. 83–107). SensePublishers, Rotterdam.

Tytler, R., Prain, V., Hubber, P., & Waldrip, B. (2013). *Constructing Representations to Learn in Science*. Springer Science & Business Media.

van der Meij, J., & de Jong, T. (2006). Supporting students' learning with multiple representations in a dynamic simulation-based learning environment. *Learning and Instruction*, *16*(3), 199–212.

Van Joolingen, W. R., & De Jong, T. (2003). Simquest. In *Authoring Tools for Advanced Technology Learning Environments* (pp. 1–31). Springer, Dordrecht.

van Someren, M. W., Reimann, P., Boshuizen, H., & de Jong, T. (1998). *with Multiple Representations. Advances in Learning and Instruction Series* (pp. 10160–0757). New York, NY: Elsevier Science, Inc., PO Box 945, Madison Square Station.

Viviani, P. (2002). Motor competence in the perception of dynamic events: a tutorial. *Common Mechanisms in Perception and Action*, *19*, 406–442.

Viviani, P., & Stucchi, N. (1989). The effect of movement velocity on form perception: geometric illusions in dynamic displays. *Perception & Psychophysics*, *46*(3), 266–274.

Viviani, P., & Stucchi, N. (1992). Motor-perceptual interactions. Retrieved from http://psycnet.apa.org/psycinfo/1993-97008-011

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, *53*(4), 5–23.

White, T., & Pea, R. (2011). Distributed by Design: On the Promises and Pitfalls of Collaborative Learning with Multiple Representations. *Journal of the Learning Sciences*, *20*(3), 489–547.

Wilder, A., & Brinkerhoff, J. (2007). Supporting representational competence in high school biology with computer-based biomolecular visualizations. *The Journal of Computers in*. Retrieved from http://search.proquest.com/openview/423910f2efc9cc0190731b5b8a8cb0f0/1?pq-origsite=gscholar&cbl=23124

Wilensky, U. (1999a). GasLab: An extensible modeling toolkit for connecting micro-and macro-properties of gases. Modeling and simulation. In *science and mathematics education* (pp. 151–178).

Wilensky, U. (1999b). {NetLogo}. Northwestern University. Retrieved from http://www.citeulike.org/group/2050/article/1283125

Won, M., Yoon, H., & Treagust, D. F. (2014). Students' learning strategies with multiple representations: Explanations of the human breathing mechanism. *Science Education*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/sce.21128/full

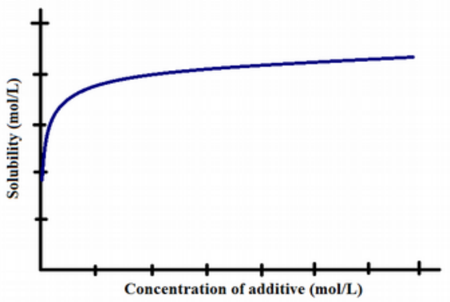
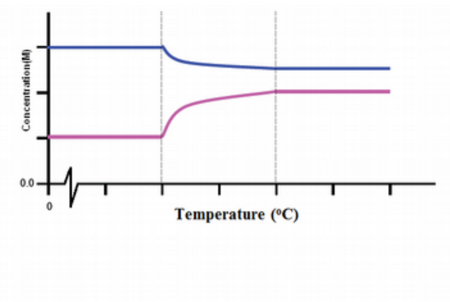
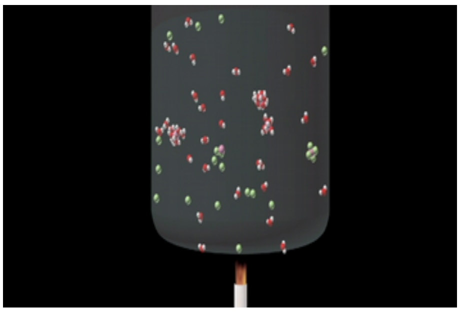
Wu, H.-K., Krajcik, J. S., & Soloway, E. (2001). Promoting understanding of chemical representations: Students' use of a visualization tool in the classroom. *Journal of Research in Science Teaching*, *38*(7), 821–842.

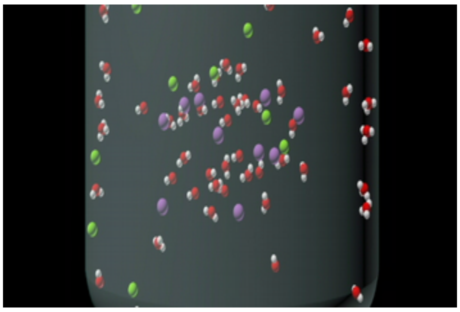
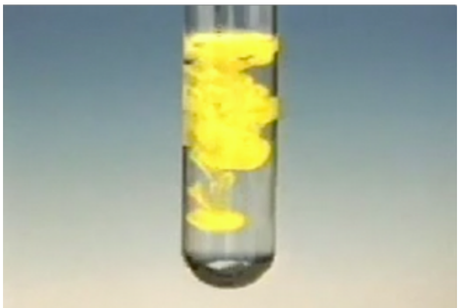
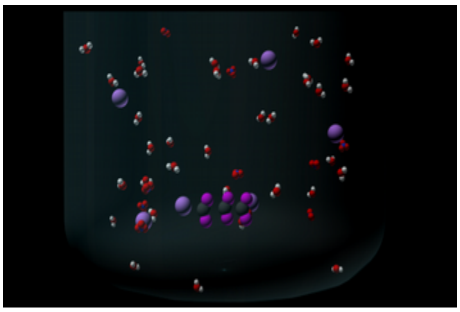
Wu, H. K., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education*, *88*(3), 465–492.

Yarroch, W. L. (1985). Student understanding of chemical equation balancing. *Journal of Research in Science Teaching*, *22*(5), 449–459.

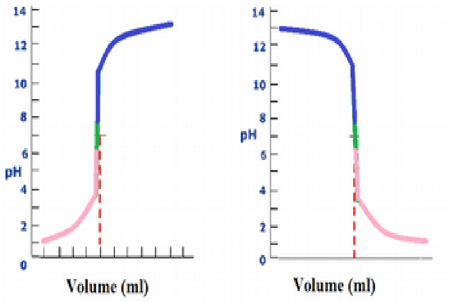
Appendix 1 (Chapter 4)

Representations used in ER categorization (screenshot images)

Representation	**Image**	**Description of the representation**
A		A solubility vs. concentration curve governing the dissoluti silver chloride in relation to the concentration of ammonia.
B		Demonstration of a neutralization reaction between sodium hydroxide and hydrochloric acid. The different colors are due to the addition of an indicator. Corresponding video shows how mixing certain quantities of pink colored acid with blue colored base results in formation of a neutral products – salt and water.
C		Temperature vs. concentration curves related to an equilibrium reaction.
D		3D molecular animation depicting effects of heating on an equilibrium reaction between water and dissolved cobalt chloride.

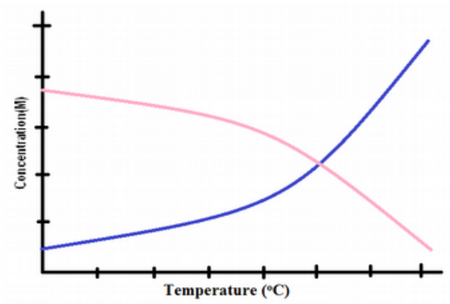
E	$$NaOH_{(aq)} + HCl_{(aq)} \Downarrow NaCl_{(aq)} + H_2O_{(l)}$$	Representative equation of a neutralization reaction between strong base and strong acid.
F		3D molecular animation representing the above neutralization reaction.
G	$$N_2O_{4(g)} \rightleftharpoons 2NO_{2(g)}$$	Equation of the NO2-N2O4 gas equilibrium reaction.
H		Demonstration video of the precipitation reaction between potassium iodide and lead nitrate.
I		3D molecular animation depicting the dynamics of the above reaction.

J



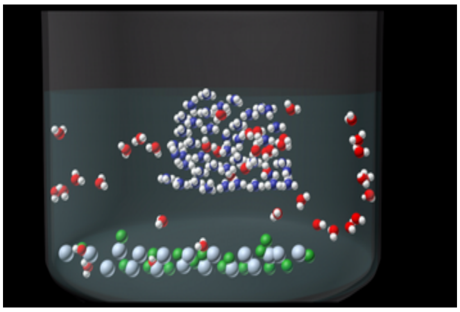
Titration curves of the strong acid-strong base neutralization reaction. One depicts the addition of base in acid while the other captures the opposite.

K



Temperature vs. concentration curves related to an equilibrium reaction.

L



3D molecular animation capturing dynamics of the equilibrium reaction between solid silver chloride and ammonia.

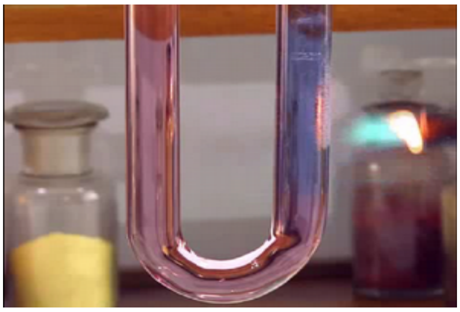
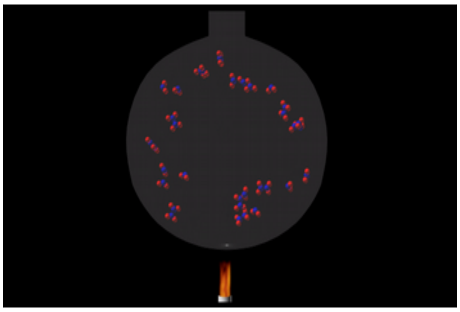
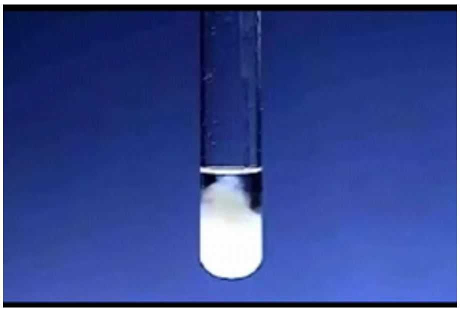
M

A demonstration video showing effects of different temperature conditions on the NO2-N2O4 gas equilibrium reaction filled in closed tubes.

N

$$AgCl_{(s)} + 2NH_{3(aq)} \rightleftharpoons Ag(NH_3)_2^+{}_{(aq)} + Cl^-{}_{(aq)}$$

Equation of the equilibrium reaction between silver chloride and excess of ammonia.

O		Demonstration video showing effects of heating on an equilibrium reaction between water and dissolved cobalt chloride.
P	$$2KI_{(aq)} + Pb(NO_3)_{2(aq)} \Downarrow 2KNO_{3(aq)} + PbI_{2(s)}$$	Chemical equation capturing the precipitation of lead iodide as a result of a reaction between potassium iodide and lead nitrate.
Q		3D molecular animation showing effects of different temperature conditions on the NO2-N2O4 gas equilibrium reaction.
R	$$Co(H_2O)_6^{2+}{}_{(aq)} + 4Cl^-_{(aq)} \rightleftharpoons CoCl_4^{2-}{}_{(aq)} + 6H_2O_{(l)}$$	Equation of the equilibrium reaction between water and dissolved cobalt chloride.
S		Demonstration video showing the dissolution of solid silver chloride on addition of excess liquid ammonia – equilibrium reaction.

Sample animations and demonstration videos can be found at:

http://lsr.hbcse.tifr.res.in/chem/reactions_render/

Appendix 2 (Chapter 4)

Chemical equation balancing task problems:

Question	**Image**
Bal1	$Al(OH)_3 + H_3O^+_{(aq)} \rightleftharpoons Al^{3+} + H_2O$
Bal2	$N_{2(g)} + O_2 \Rightarrow N_2O_5$
Bal3	$C_2H_5I + NH_3 \Rightarrow C_2H_5NH_3^+ + I^-$

Bal4

$$Ag + S_8 \longrightarrow Ag_2S_{(s)}$$

Bal5

$$Cr + O_2 \longrightarrow Cr_2O_3$$

Bal6

$$C_4H_9Br \longrightarrow C_4H_9OH + H_3O^+Br$$

Appendix 3 (Chapter 4)

Instructions given to each participant during the Balancing task:

You will be seeing a simple unbalanced chemical equation and your task is to balance it. No paper and pencil are available as you are expected to do it mentally. There is no time limit, so you can take as long as you want. You can also proceed to the next equation if you find the current one difficult but remember that you will not be allowed to return to the equation you skip.

Instructions for the ER categorization task:

Now we begin the second task. Here, I will be showing you one by one, a number of representations such as chemical equations, graphs, 3D animations and laboratory demonstration videos on the laptop screen. When I show you each representation, I will be handing over to you a card with the corresponding representation printed on it. In case of animations and demonstration videos, the card will have a screenshot of some moment in the movie. Attend to each representation on the screen carefully as you will not be allowed to return to it after you have proceeded to the next one. You can take as much time as you want to view each image, and watch each movie as many times as you want before proceeding to the next. Once you have seen all the representations on the laptop and collected all the corresponding cards with you, I will tell you what to do with them.

Appendix 4 (Chapters 4 & 5)

Definitions of gaze parameters.

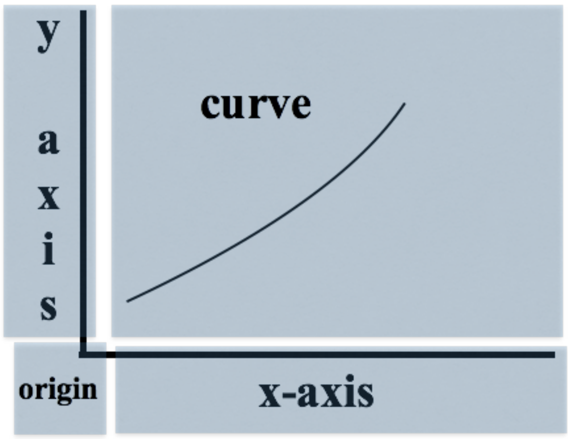
Fixation point	Point (location) on the stimulus where the eye is fixated.
Fixation index	Represents the order in which a fixation event was recorded. The index is an auto-increment number starting with 1 (first gaze event detected).
Visit duration	The duration of each individual visit within an AOI.
Visit count	The number of visits within an AOI.
Fixation duration	The duration of each individual fixation for a participant within an AOI.
Fixation count	This metric measures the number of times the participant fixates on an AOI or an AOI group.
Saccade	Movement of the eye between fixation points.
Gaze Transitions	Eye movements between two consecutive fixations (e. g. A-B, where A and B are two different AOIs)
Inertia	The number of transitions made to the same AOI/total number of transitions.
Volatility	1 – inertia.
Gaze Returns	Eye movements between two or more AOIs of the nature A-B-A, A-B-C-A, A-B-C-D-A, and so on, where A, B, C, D are different AOIs. Returns can be thought of consisting multiple transitions, for instance, the return A-B-A has an A-B transition and then a B-A transition. Similarly, A-B-C-A- consists of three transitions, A-B, B-C and returning from C to the AOI A i.e. a C-A transition.
Useful A-B-A returns	Returns of the nature A-B-A between two successive mouse clicks.
Useful A-B-C-A returns	Returns of the nature A-B-C-A between two successive mouse clicks.
Unique AOIs count between	The number of AOIs visited between two successive mouse clicks, where even multiple visits to an AOI are counted as a single entry (e. g. if a participant visits

mouse clicks	AOI A twice, AOI B four times and AOI C just once between two successive mouse clicks, his/her unique AOI count will be 3 irrespective of the number of times s/he visited each of the AOIs).
AOIs count between mouse clicks	Total count of AOI visits, where multiple visits are counted separately (in the above example, the total AOI count between the two mouse clicks will be recorded as 2+4+1 = 7 counts).
Average spread	The average number of occurrences of different AOIs between mouse clicks. Spread = Average number of AOIs visited between mouse clicks.
Elasticity	The weighted sum of the average number of useful returns of the nature ABA and ABCA. Elasticity = 1*(average number of ABA returns) + 2*(average number of ABCA returns). Elasticity also shows how elastic or fluent a person is transitioning between AOIs. It can be understood in contrast to a general meaning of 'inertia' which usually signifies rigidity. Elasticity would thus indicate how easily does a person navigate between the different parts of a stimulus.

Appendix 5 (Chapters 4 & 5)

Steps taken to process raw eye-tracking data in order to obtain transition matrices:

Step 1: In the raw data sequence generated from Tobii Studio, as there are separate columns for different areas of interest (AOI), there is unique AOI hit (1 denotes a hit) under respective column headers (each AOI). So the AOI columns were merged, using 'CONCATENATE' function in excel, so that a number represents a specific AOI hit. For example, consider the stimulus figure below.



While generating the (gaze or mouse) activity sequences for this stimulus, the number 1000 denotes AOI hit for CURVE, 0100 denotes ORIGIN hit, 0010 marks hit in Y-AXIS area, 0001 for X-AXIS hit. 0000 denotes a non AOI hit, i.e. no fixation was recorded in any of the AOIs. The fixation in this case may lie outside all the AOIs. So from the raw data sequence, the fixation hit sequence was retained in the form of these numbers, stored under a separate column.

Step 2: To this concatenated column, the filter (from 'Sort & Filter" option) was applied wherein 0000 box was unchecked to remove all the non AOI hit entries from the sequence. Then a different column stores this filtered sequence. In the adjacent column, difference between two consecutive cells, (A2-A3, A3-A4 and so on) is calculated. A new filter is applied to this column to filter the cells with the value '0' (zero). In this way consecutive duplicate AOI entries are deleted, to get only the unique AOI hits. For instance, if four consecutive fixations happen in AOI curve while the next three happen in AOI X-axis, after applying our algorithm, these will be filtered as only two AOI events – one happening in AOI curve and the very next in AOI X-axis. This would then be counted as one transition from AOI curve to AOI X-axis. In this way, the entire sequence was analyzed to count transitions between the different AOIs.

Step 3: Subsequently, the transition diagrams were plotted according to the transition sequence using a graph theoretic framework.

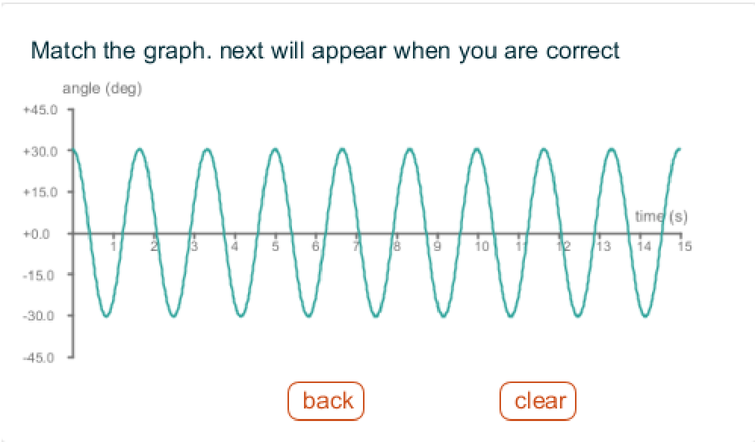
Step 4: For returns, the event or transition sequence obtained in step (2) is further filtered to identify transitions of the A-B-A or A-B-C-A kind, by applying a similar logic.

The tables for fixation duration, fixation count, and saccades were generated from TOBII and the data tables were created to calculate the required parameters across participants and groups and graphs were plotted accordingly.

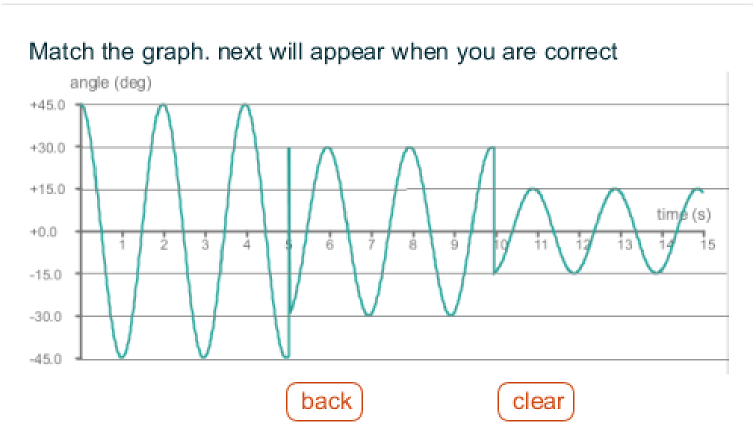
Appendix 6 (Chapter 5)

Learning tasks (images): The task required students to generate a curve in the graph panel similar to that depicted in the respective task image by manipulating the parameters accordingly and playing the simulation.

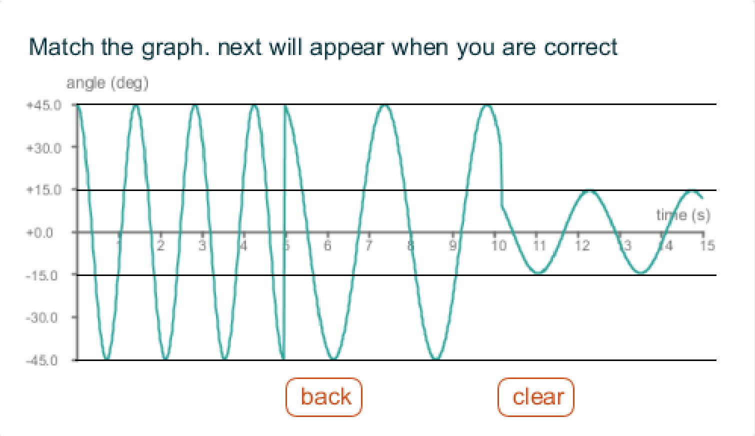
Task 1



Task 2



Task 3

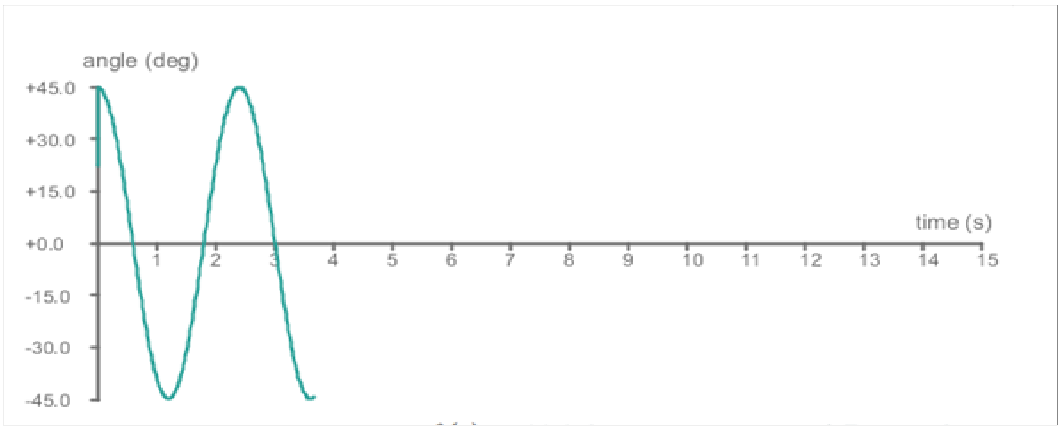


Match the graph. next will appear when you are correct
angle (deg)
+45.0
+30.0
+15.0
+0.0
-15.0
-30.0
-45.0
time (s)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
back
clear

Appendix 7 (Chapter 5)

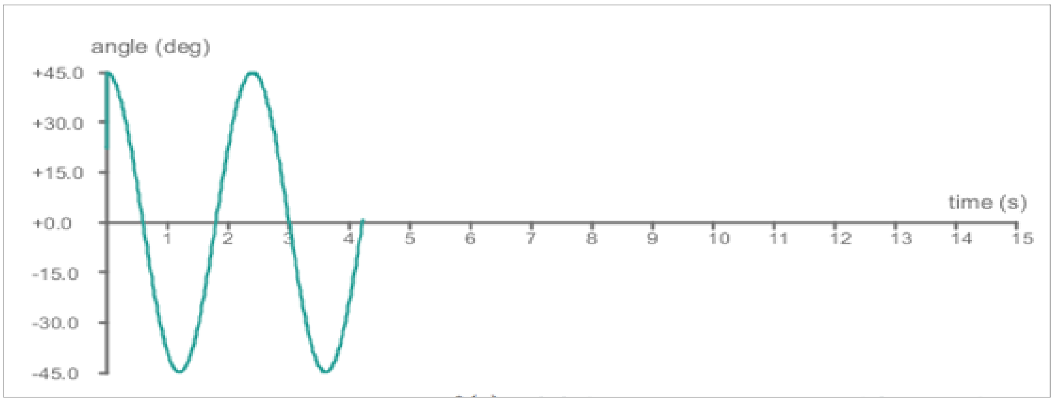
ER integration questions in iteration 1 (printed sheet was presented to the student).

Question 1: On which side will be the pendulum (the moving object) when the end of the graph is negative, as below?



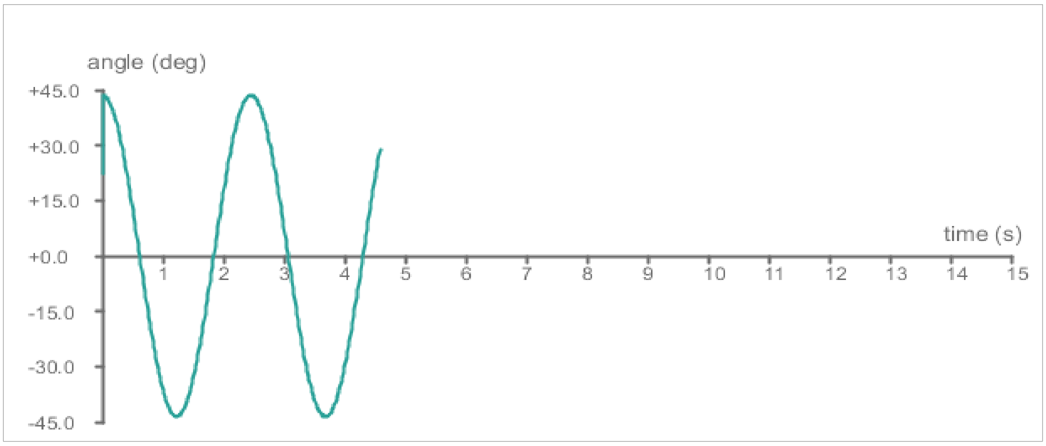
(a) Right (b) Left (c) Exactly vertical (d) Exactly horizontal

Question 2: Where will be the pendulum (the moving object) when the end of the graph is on the x-axis, as below?



(a) Right (b) Left (c) Exactly vertical (d) Exactly horizontal

Question 3: Where will be the pendulum (the moving object) when the end of the graph is as below?



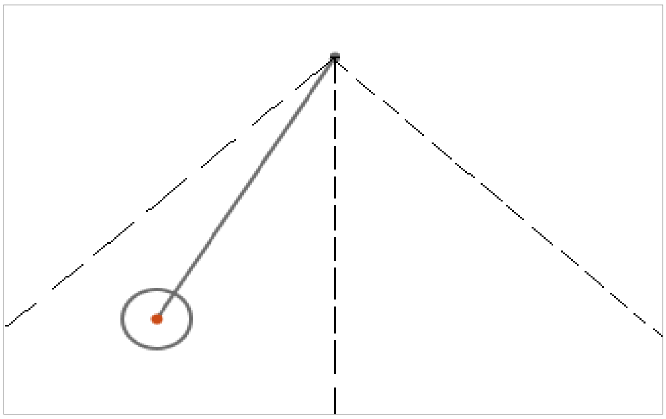
(a) 30 degrees left

(b) 30 degrees right

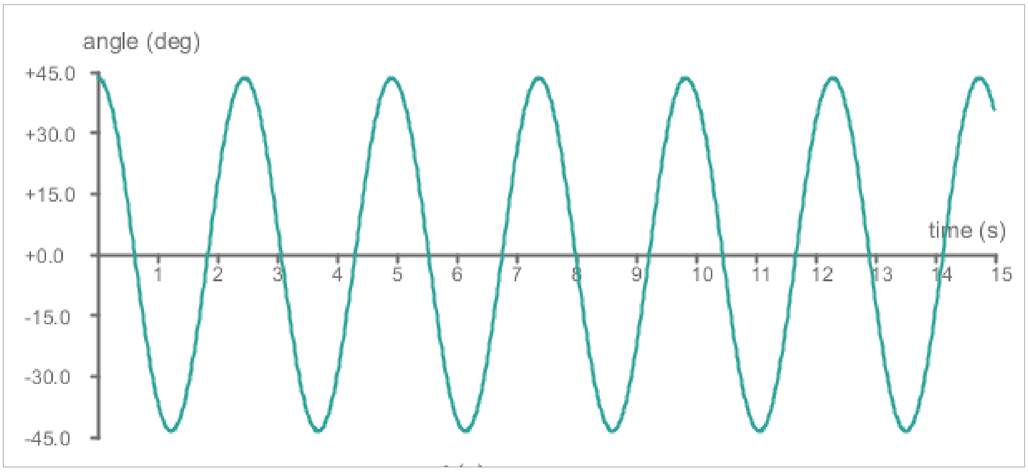
(c) 30 degrees vertical

(d) Exactly horizontal

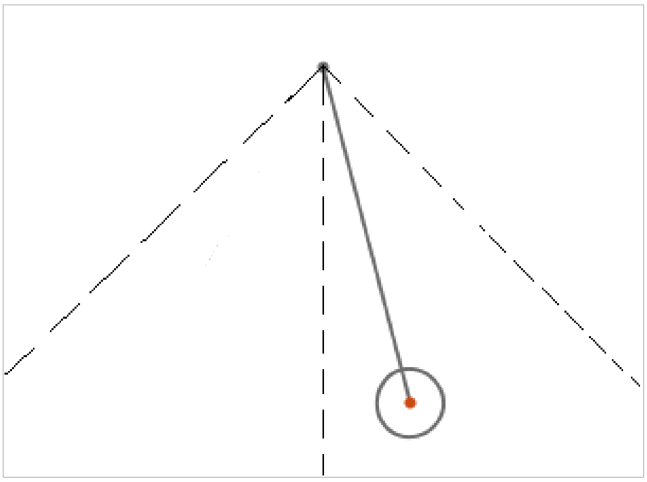
Question 4: The pendulum (the moving object) is at the point shown in the figure below. Where is this point on the graph?



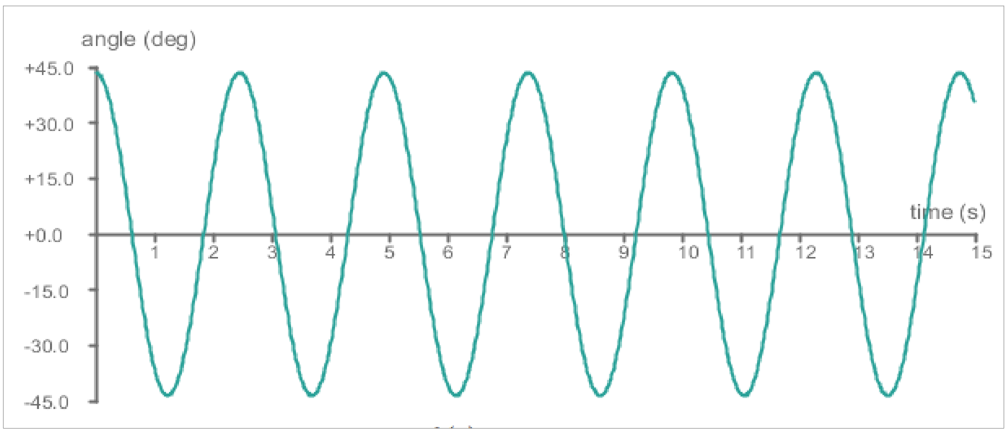
You can mark the point(s) on the graph below:



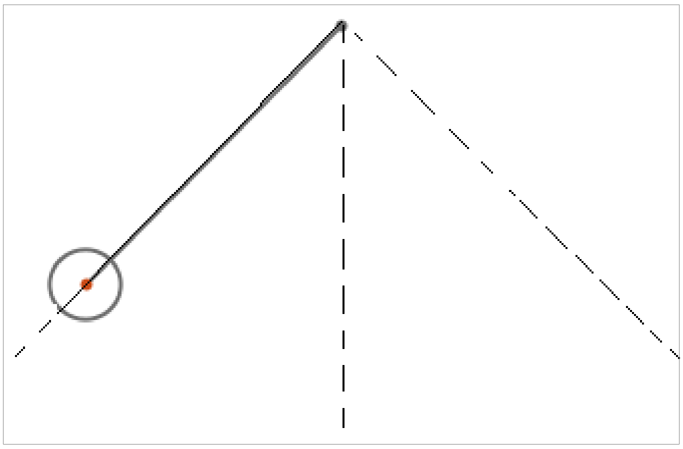
Question 5: The pendulum (the moving object) is at the point shown in the figure below. Where is this point on the graph?



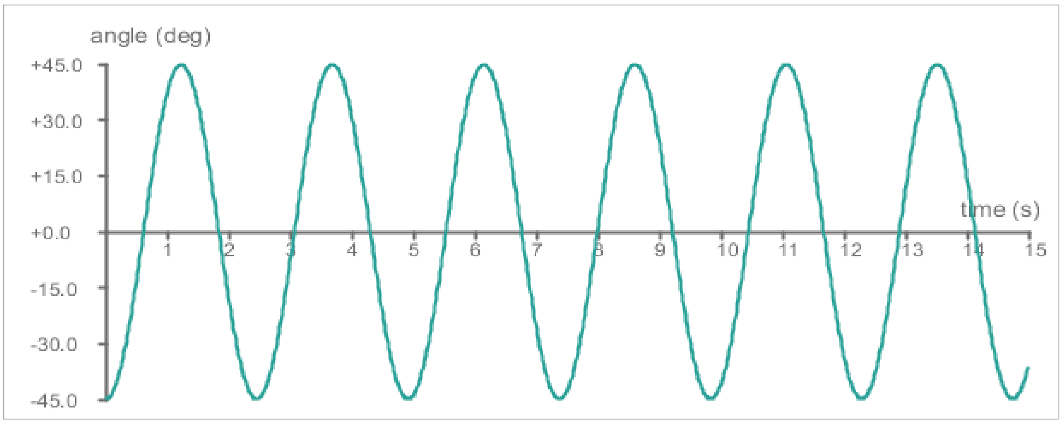
You can mark the point(s) on the graph below:



Question 6: The pendulum (the moving object) is at the point shown in the figure below. Where is this point on the graph?



You can mark the point(s) on the graph below:



Appendix 8 (Chapter 5)

Learning objectives and ER integration questions (general) in iteration 2:

Learning objectives: The student will be able to -	**Question category**	**Multiple-choice question with four options**
Map phenomenon and graph	Check whether learner can relate points on graph to phenomenon and vice versa	Given a particular position on the graph, identify the corresponding position of the pendulum.
		Given a diagram of a pendulum at a certain position, mark the corresponding position(s) on the graph.
	Check whether given a word problem, a learner can imagine the phenomenon and its graph? - Oscillatory graphs	Imagine you have a friend whose mood swings regularly. She starts the day with a very happy mood, but as the day progresses, she gets upset. At night she is very sad. The next morning she is very happy again, and her mood deteriorates as the day progresses. This continues every day. Which graph among the following best represents your friend's behavior? Options (with figures) y = constant; x = constant; parabola; sinusoid.
	Non-oscillatory graphs (non-sinusoidal movement in time):	A car is moving along a road at a constant speed of 60 kph (a line diagram of car). The graph that best represents its speed of movement is:

		Options (with figures) y = constant; x = constant; y=x; sinusoid.
		I threw a ball in the air towards my friend 50 m away. It went towards him, rising initially and then falling down as it approached him (line diagram of two people). The graph that best describes the path of the ball is: Options (with figures) parabola; x = constant; sinusoid; y = x.
Map system and equation	Describe damped pendulum and ask what is the equation.	Consider the simple pendulum you worked with earlier. When you move the bob to a particular angle and release it, the pendulum keeps moving back and forth around the vertical position without stopping. The equation of the angle of the pendulum at any point of time is given by $\theta(t) = \theta_0 \cos \sqrt{\frac{g}{l}}$ t). T is the time taken for one cycle, i.e. starting from one point and returning to the same point. Now suppose there was air drag. When the pendulum is released from a particular angle, the drag will slow it down and after moving back and forth about the vertical position a few times, the pendulum will stop at the vertical position. What would the

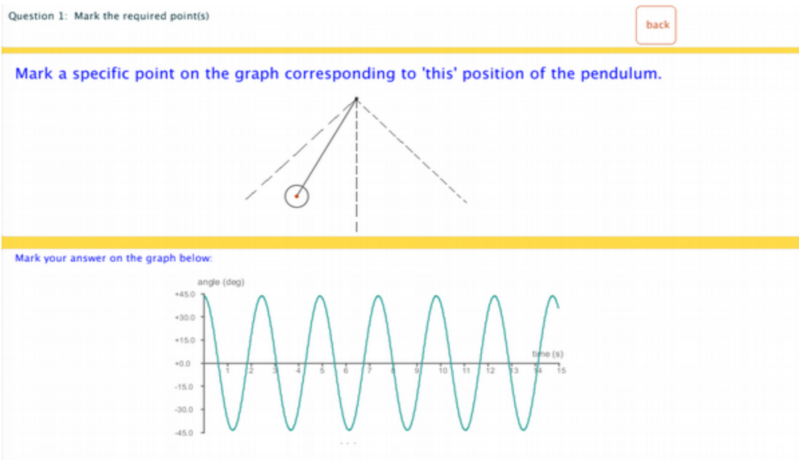
		equation for the approximate angle of such a pendulum at any point of time be? Options: 2*$\theta(t)$, $\theta(t)$/2, $\theta(t)$-2, (1-t/2)
	How to modify behavior of pendulum	Consider the simple pendulum you worked with earlier. The equation of the angle of the pendulum at any point of time is given by $\sqrt{\frac{g}{l}}$ $\theta(t)=\theta_0 \cos$ t). T is the time taken for one cycle, i.e. starting from one point and returning to the same point. Consider a pendulum which completes one cycle in 1.2 sec. Now suppose if you want to make the pendulum go faster and complete one cycle in 1 sec. What would you do? Options: Increase l, Decrease g, Decrease l, Let the pendulum go from a greater angle initially
	Modify equation, ask about behavior	Consider the simple pendulum you worked with earlier. The equation of the angle of the pendulum at any point of time is given by This pendulum completes one cycle in 2 sec, i.e. starts from one point and returns to the same point in 2 seconds. Now consider a pendulum with equation given by. How long

		does this pendulum take to complete one cycle? Options: Less than 2 sec; More than 2 sec; Equal to 2 sec; Half the time
Map equation to graph	Show underdamped pendulum graph and ask what is the equation?	Look at the graphs below (diagrams of sinusoid and damped sinusoid shown). The graph on the left has the equation. What is the approximate equation of the graph on the right? Options: 2*$\theta(t)$, $\theta(t)$/2, $\theta(t)$-2, (1-t/2).
	Modify equation, ask about graph	Look at the graphs below (diagrams of two sinusoids with different amplitudes). The graph on the left has the equation. What is the graph of the equation? $\theta(t)=20\cos\sqrt{\frac{g}{l}}t$ Options: Graph with same frequency and amplitude 20; graph with same frequency and amplitude 40; graph with different frequency and same amplitude; graph with different amplitude and frequency.

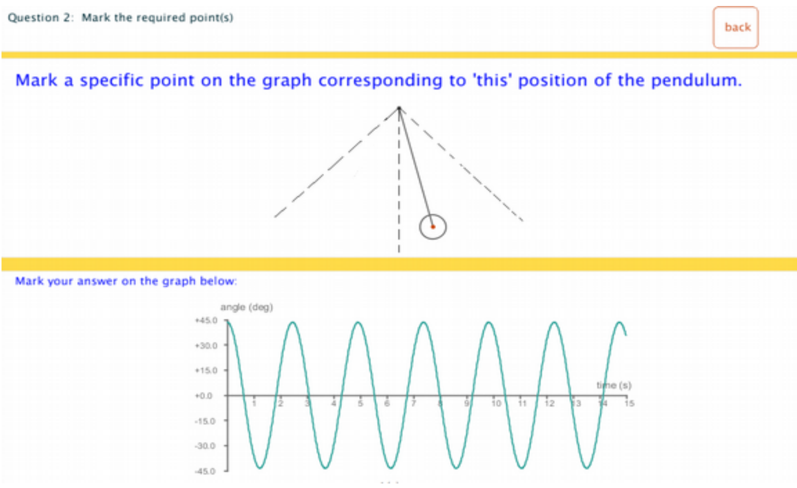
ER integration questions (screenshot images)

Question Screenshot

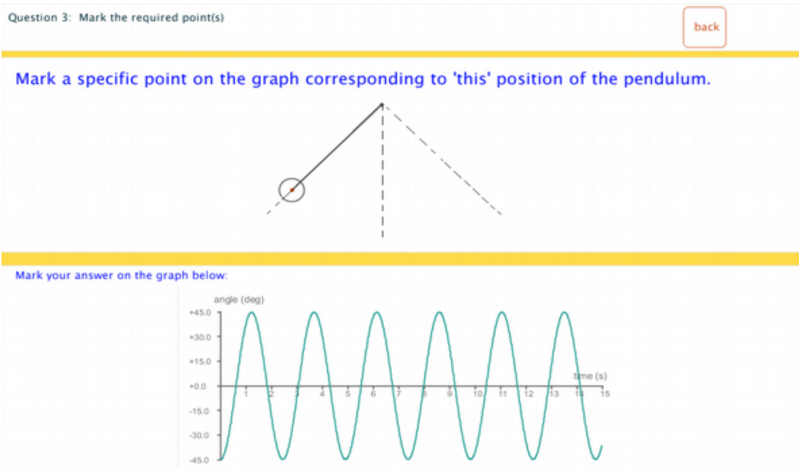
1



2



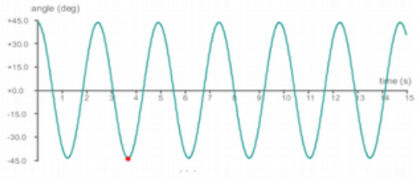
3



4 Question 4: Choose the correct option

back

On which side will be the pendulum (the hanging object) when graph is negative (as depicted)?



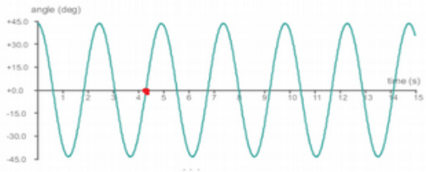
Options

- Right
- Exactly Vertical
- Left
- Exactly Horizontal

5 Question 5: Choose the correct option

back

Where will be the pendulum when graph is on x-axis (as depicted)?



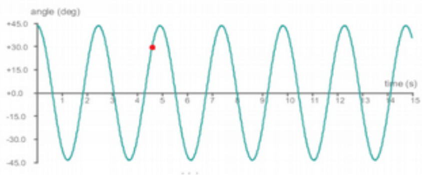
Options

- Right
- Exactly Vertical
- Left
- Exactly Horizontal

6 Question 6: Choose the correct option

back

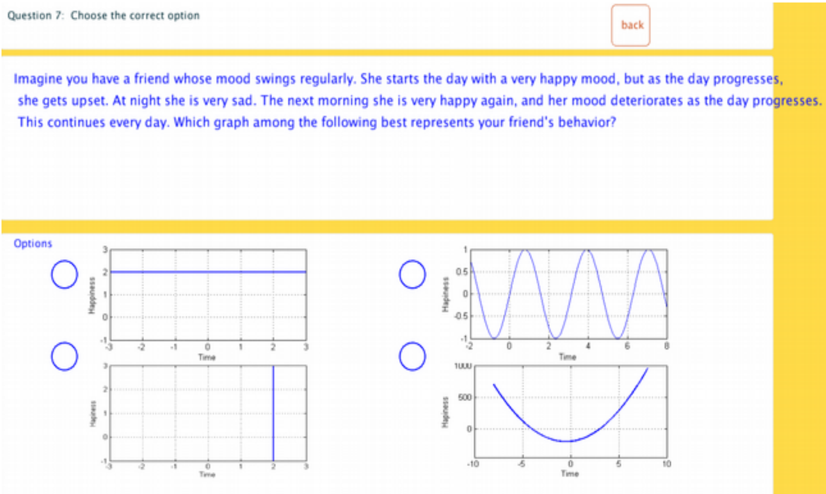
Where will be the pendulum (the hanging object) according to this graph?



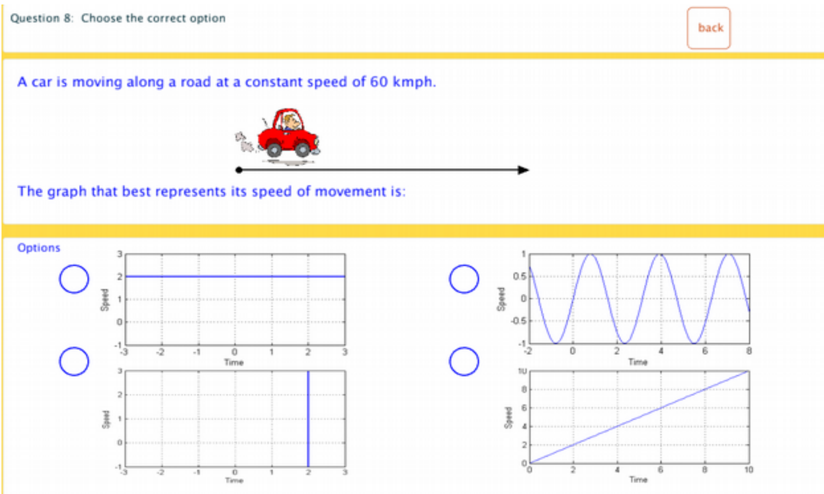
Options

- 30 Degree Left
- 30 Degree Vertical
- 30 Degree Right
- Exactly Horizontal

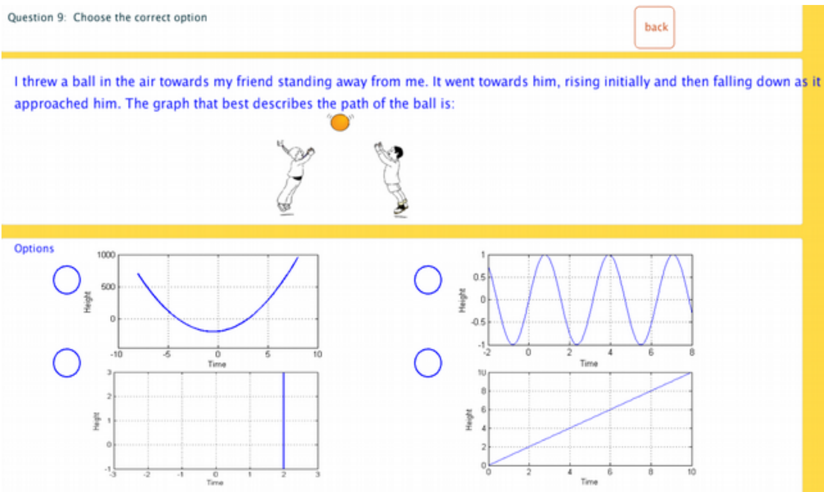
7


Question 7: Choose the correct option
back
Imagine you have a friend whose mood swings regularly. She starts the day with a very happy mood, but as the day progresses, she gets upset. At night she is very sad. The next morning she is very happy again, and her mood deteriorates as the day progresses. This continues every day. Which graph among the following best represents your friend's behavior?
Options
Happiness
Time

8


Question 8: Choose the correct option
back
A car is moving along a road at a constant speed of 60 kmph.
The graph that best represents its speed of movement is:
Options
Speed
Time

9


Question 9: Choose the correct option
back
I threw a ball in the air towards my friend standing away from me. It went towards him, rising initially and then falling down as it approached him. The graph that best describes the path of the ball is:
Options
Height
Time

10 Question 10b: Choose the correct option

back

Now suppose there was air drag. When the pendulum is released from a particular angle, the air drag will slow it down and after moving back and forth about the vertical position a few times, the pendulum will stop at the vertical position. What would the equation for the approximate angle of such a pendulum at any point of time be?

Options

- $2 * \theta(t)$
- $\theta(t) - 2$
- $\theta(t) / 2$
- $\theta(t) * (1-t/A)$ where A is a constant number

11 Question 11: Choose the correct option

back

Consider the simple pendulum you worked earlier with, its equation is

$$\theta(t) = \theta_0 \cos \sqrt{\frac{9.807}{L}}\, t$$

What will you do to make the pendulum go faster?

Options

- Increase length
- Decrease length
- Decrease g
- Let the pendulum go from a greater initial angle

12 Question 12: Choose the correct option

back

When the pendulum returns to its starting position (the point you released it from), it is said to have completed one cycle. Consider a simple pendulum with equation

$$\theta(t) = \theta_0 \cos \sqrt{\frac{10}{L}}\, t$$

which takes 2 seconds to return to its starting position (i.e complete one cycle). Now consider another pendulum with equation

$$\theta(t) = \theta_0 \cos \sqrt{\frac{5}{L}}\, t$$

How long will this other pendulum take to complete one cycle (return to its starting position)?

Options

- less than 2 sec
- equal to 2 sec
- greater than 2 sec
- Half the time

13

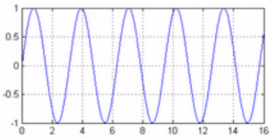
Question 13: Choose the correct option

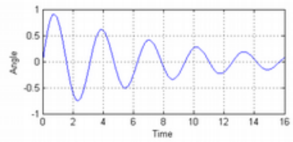
back

Look at the graphs below. The graph on the left has a equation

$$\theta(t) = \theta_0 \cos\sqrt{\frac{9.807}{L}}\,t$$

What is the approximate equation of the graph on the right?





Options

- 2 * θ(t)
- θ(t) - 2
- θ(t) /2
- θ(t) * (1-t/A) where A is a constant number

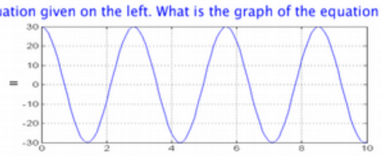
14

Question 14: Choose the correct option

back

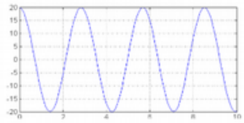
The graph below has equation given on the left. What is the graph of the equation given on the right?

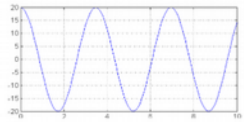
$$\theta(t) = 30 \cos\sqrt{\frac{4.903}{L}}\,t =$$

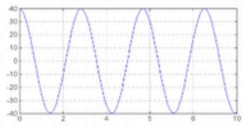


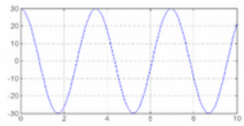
$$\theta(t) = 20 \cos\sqrt{\frac{4.903}{L}}\,t = ?$$

Options









Appendix 9 (Chapter 5)

Data processing steps taken (Chapter 6). This algorithm is based on a logic similar to the gaze data analysis explained in appendix 5. Here I briefly mention the steps from a program we developed to automate the data processing.

Step 1: The raw data is obtained from Tobii.

Step 2: Assign Event Type to mouse and fixation events.

Step 3: Remove all the non-mouse and non-fixation entries (e.g. saccades, microsaccades and other classified events)

Step 4: Assign AOI to fixation (from the list of AOI hits) and mouse events (from the pixel values of the mouse click and the AOI rectangles)

Step 5: Remove all the non-AOI data The timestamp value is corrected, i.e., the gaps in time due to removing data are re-arranged such that all the resulting is continuous in time. Also each data entry is assigned a time duration of (1000/60) milliseconds because of the 60-Hz data entry by Tobii Pupil Data is extracted and the average of left and right eye data is stored separately along with the corresponding AOI data. Next, average and Standard Deviation of the pupil data is calculated for the entire timespan and for individual AOIs.

Step 6: Consolidate the entries with the same Fixation Index.

Step 7: Make 3 data arrays: only fixation events, only with mouse events, with both mouse and fixation events.

Step 8: Calculate the time duration for each fixation event.

Step 9: Club fixations happening in the same AOI into one event entry.

Step 10: Calculate the total time and the time spent in each AOI. Note: The time data in the sequence sheet is the corrected time data. The mouse click could have fixation, saccade or unclassified gaze event. Tobii does not provide AOI-hit information for the mouse click event entries in the data Tobii does not provide the mouse down time, ie, we don’t have the drag data. The mouse click given by Tobii happens when the mouse button is pressed or released.

i This chapter contributed to the following journal publication: Pande, P., & Chandrasekharan, S. (2017). Representational competence: towards a distributed and embodied cognition account. *Studies in Science Education*, *0*(0), 1–43.

ii This chapter contributed to the following journal publication: Pande, P., & Chandrasekharan, S. (2017). Representational competence: towards a distributed and embodied cognition account. *Studies in Science Education*, *0*(0), 1–43.

iii The classroom activities are one way of interacting with the dynamics embedded in ERs, new-media technologies is another. I do not consider educational technologies as stand-alone resources, to be used independent of classroom activity.