

“Concept” and “Context”: Toward Modelling Understanding in Physics Education Research



Philip Southey

Thesis presented for the Degree of Doctor of Philosophy in Tertiary

Physics Education, University of Cape Town

August 2017

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Abstract

Title: “Concept” and “Context”: Toward Modelling Understanding in Physics Education Research.

Author: Philip Southey

Degree: Doctor of Philosophy in Tertiary Physics Education Research

Supervisor: Saalih Allie

University of Cape Town, 2017

“Context sensitivity” is a core issue in physics education research (PER). Why does student understanding of a concept depend so crucially on the context in which it is embedded? This dissertation attempts to answer this question by using a variety of theoretical tools to model understanding. We conducted three empirical studies which probed context sensitivity of student understanding of (i) Vector Addition; (ii) The FCI (Force Concept Inventory); and (iii) the learning of the concept of a Mathematical Group. (i) Regarding vector addition, we discovered context sensitivities involving the type of physical quantity added (e.g. force or momentum); the textual prompts “total”, “net” and “resultant”; and the object on which a force acts. (ii) In the FCI, we discovered a moderate context sensitivity to unfamiliar words (i.e. when familiar words like “box” were substituted for unfamiliar words like “kist”.) This sensitivity was moderately correlated with the difficulty of the question. (iii) Previous studies have shown that learners exhibit a sensitivity to the concreteness of the learning condition of a Mathematical Group; our study shows that students are engaged in different types of activity in these conditions. A variety of theoretical tools from PER, Cognitive Linguistics, Cognitive Psychology and other areas of Education Research are used to model student understanding in these various studies. Three key insights emerged. (a) The importance of one’s model of “concept” – how it relates to the notion of “context”, and how one chooses an appropriate grain size. (b) The difference between “expert” and “novice” – how this difference influences one’s model of “concept”, and how it influences one’s notion of “sameness” and “difference”. (c) Student reasoning – how a framing of a situation might result in fast, associative, linguistic reasoning on the one hand, or slow, deliberate simulative reasoning on the other. Finally, this thesis is grounded in Wittgensteinian ordinary language philosophy which maintains that notions of “concept”, “context” and “understanding” obtain meaning not by referring to some transcendental “thing”, but by being embedded in our messy form of life. In other words, by modelling understanding we are not approaching the “true meaning” of the term. Instead we are demonstrating how our various models are constitutive of what we *mean* when we say: “My students understand this concept”.

Acknowledgements

Thank you Saalih for all the conversations and arguments.

Thank you to the SASOL Inzalo Team for all the financial and moral support over the past 7 years. It has been a privilege to be part of the Inzalo Community. Thank you to ADP for their additional financial support.

Thank you to the UCT Physics Department. To Steve Peterson, Gregor Leigh and Spencer Wheaton in particular, for allowing me to take up some of your precious lecture time. Thank you to all the staff and postgrad students for creating an environment one looks forward to working in every day. Enkosi ooMama!

Thank you to my friends and colleagues in the USA – Abigail Daane, Vashti Sawtelle, Leslie Atkins Elliot and Angie Little for your input and support. Thank you too to all the inspiring friends and colleagues from AAPT/PERC and FFPER. I feel incredibly fortunate to belong to such an enthusiastic community of extraordinary people.

Thank you to Anna Sfard for your correspondence and support.

And thank you to Rosie, Mom, Dad, Gus. In many ways, large and small.

Contents

Abstract.....	i
Acknowledgements.....	ii
Contents.....	iii
List of Figures	vi
List of Tables	viii
i A Note on Style	ix
ii Thesis Overview	x
1. CHAPTER 1: Introduction	1
1.1. Understanding	1
1.2. Concept.....	6
1.3. Context.....	9
2. CHAPTER 2: Theoretical Background	12
2.1. Resources / Knowledge in Pieces.....	12
2.2. Simulation and Dual Process Theory	13
2.3. Resources and Simulation.....	15
2.4. Reconciling Wittgenstein	16
3. CHAPTER 3: Vector Addition in Different Contexts	18
3.1. Introduction	18
3.2. Literature Review	19
3.3. Overview	22
3.4. Study 1: Vector Addition in the Context of Different Physical Quantities.....	24
3.5. Study 2: Vector Addition in Different Textual Contexts - “Total” and “Net”	27
3.6. Study 3: Vector Addition in Different Textual Contexts – “Total” and “Resultant”	39
3.7. Study 4: Vector Addition in Different Physical Contexts.....	45
3.8. General Discussion and Conclusion	52
3.9. Post Script: Textbooks and Total Force.....	54
4. CHAPTER 4: Context Sensitivity in the FCI	55
4.1. Introduction and Literature Review.....	55
4.2. Overview	56
4.3. Study 1: Pilot Study March 2013.....	57
4.4. Study 2: October 2013	68
4.5. Study 3: February 2014.....	75
4.6. Study 4: “Post Script” on Familiarity	77
4.7. General Discussion and Conclusion	78

INTERMISSION: Prelude to Chapters 5, 6 and 7.....	80
5. CHAPTER 5: A Qualitative Study of the Transfer Experiments of Kaminski et al. and De Bock et al. 82	
5.1. Abstract.....	82
5.2. Introduction	82
5.3. Key Results from Kaminski et al. and De Bock et al.'s Experiments.	83
5.4. Our Study	85
5.5. Instrument	85
5.6. Methodology.....	88
5.7. Quantitative Results.....	89
5.8. Qualitative Results	90
5.9. Discussion.....	94
5.10. Conclusion.....	97
6. CHAPTER 6: A Knowledge in Pieces Perspective of Analogical Transfer in the Studies of Kaminski et al. and De Bock et al.	99
6.1. Abstract.....	99
6.2. Introduction	99
6.3. Theoretical Background	101
6.4. Kaminski et al.'s Central Experiment	105
6.5. De Bock et al.'s Experiment	109
6.6. Southey et al.'s Central Experiment.....	111
6.7. Discussion.....	111
6.8. Conclusion.....	116
7. CHAPTER 7: Instantiations of Concepts – Sameness and Difference.....	119
7.1. Introduction	119
7.2. Three Illustrative Examples.....	120
7.3. Perspectives on “Sameness” and “Difference”	124
7.4. Conclusion.....	131
8. CHAPTER 8: General Discussion and Conclusion	133
Appendix A:.....	135
Three Uses of “Concreteness” in Education Research	135
1. Introduction	135
1.1. Origins and Current Usage	136
1.2. Concrete Objects and Concrete Concepts	136
2. Tangibility.....	137
3. Familiarity	138

3.1. Familiarity and Tangibility	139
3.2. Two Models of Familiarity	141
4. Perceptual Richness	144
4.1. Perceptual Richness and Familiarity	145
4.2. Perceptual Richness as an Explanatory Tool.....	148
5. Conclusion.....	148
Appendix B:	150
Supplementary Material for Vector Addition	150
First Coding Key 2014 Qualitative Data	150
Second Coding Key 2014 Qualitative Data.....	151
Supplementary Material for FCI.....	152
References	154

List of Figures

Figure 1: A Picture Which Seems to Force a Particular Interpretation On Us: A Person Walking a Slope	2
Figure 2: Bloom's Taxonomy of Knowledge Might Be Thought of as a Hierarchy of Different Abilities. (Picture Under Creative Commons License Courtesy of Vanderbilt University Center for Teaching)	4
Figure 3: A "something" occurring in two different contexts	10
Figure 4: Activation Time Curves for Linguistic System and Situated Simulation System. Taken from L. W. Barsalou et al. (2008, p.248)	14
Figure 5: Percentage of students to correctly add vectors in different contexts (n=227).	25
Figure 6: Distribution of answers for the addition of various vector quantities (n=227)	25
Figure 7: Percentage of students to correctly add vectors in different contexts.	28
Figure 8: Comparison of Percentage of Students to Give Correct Answer for Forces Addition Before and After Being Exposed to Scenario F	30
Figure 9: Comparison of Percentage of Students to Give Correct Answer for Momenta Addition Before and After Being Exposed to Scenario M	31
Figure 10: Percentage of students to correctly add vectors in different contexts (n=205).	41
Figure 11: A comparison of students' correct responses for the total force (dark grey) and resultant force (light grey) before and after being exposed to Scenario F.	42
Figure 12: A comparison of students' correct responses for the total momentum of the two balls (dark grey columns) and two ball system (light grey columns) before and after being exposed to Scenario M.	43
Figure 13: Percentage of Students in the PB group to Correctly Calculate the Total Force Acting on the Person and the Net Force Acting on the Ball.	47
Figure 14: Percentage of Students in the BP group to Correctly Calculate the Net Force Acting on the Ball and the Total Force Acting on the Person.	47
Figure 15: Question by Question Comparison of Control Group (n=92) and Experimental Group (n=90)	63
Figure 16: The Difference Between Control and Experimental Groups' Scores. (Control – Experimental).	63
Figure 17: Comparison of Control and Experimental Groups' Overall Scores	64
Figure 18: Comparison of Control and Experimental Group's Scores for Difficult Questions (5,8,9,12,13)	64
Figure 19: Comparison of Control and Experimental Group's Scores for Calculation Questions (1,2,3)	65
Figure 20: Comparison of Control and Experimental Groups' Scores for Conceptual Questions	65
Figure 21: The Difference Between Control and Experimental Scores for Chemical Engineering Students (n=123) in Green, and Life Sciences Students (n=59) in Purple	66
Figure 22: Difference between Control and Experimental n = 220	71
Figure 23: Comparison of Control and Experimental Groups' Overall Scores	72
Figure 24: Comparison of Control and Experimental Groups' Scores for Difficult Questions	72
Figure 25: Comparison of Control and Experimental Groups' Scores for Calculation Questions	72
Figure 26: Correlation Between Difficulty of Question and Difference in Performance Between Control and Experimental Groups. $R = 0.42$	73
Figure 27: Histogram of Number of Students in Bins Associated with the Difference in Their Performance on Control and Experimental Questions	74
Figure 28: Difference between Control and Experimental n = 420	76
Figure 29: Comparison of Control and Experimental Groups' Overall Scores	76
Figure 30: Comparison of Control and Experimental Groups' Scores for Difficult Questions	76
Figure 31: Correlation Between Difficulty of Question and Difference in Performance Between Control and Experimental Groups. $R = 0.45$	77
Figure 32: Two Different Instantiations of the Concept of Vector Addition in Two Different Contexts	80
Figure 33: Two Different Instantiations of the Concept of Force Addition in Two Different Contexts	80
Figure 34: Two Different Instantiations of the Concept of Newton's 2 nd Law in Two Different Contexts	81
Figure 35: Two Different Instantiations of the Concept of a Mathematical Group in Two Different Contexts	81
Figure 36: Schematic Representation of Kaminski et al. and De Bock et al.'s Experiment	83
Figure 37: De Bock et al.'s Extension of the Original Experiment	84

<i>Figure 38: Rules Used in Learning Domain 1. Taken from Supporting Online Material (SOM) (Jennifer A. Kaminski et al., 2009b)</i>	85
<i>Figure 39: Three of the six rules used for Learning Domain 2. Taken from SOM (Jennifer A. Kaminski et al., 2009b)</i>	86
<i>Figure 40: Isomorphic Questions from Learning Domain 1 and Learning Domain 2 Respectively. Taken from SOM</i>	86
<i>Figure 41: The Rules of Combination for the Transfer Domain. Taken from SOM.</i>	87
<i>Figure 42: Rules of Combination for the Transfer Domain in Cohorts AC and CC</i>	87
<i>Figure 43: Objects Used in the Four Domains of Our Experiment</i>	87
<i>Figure 44: Percentage of students in different cohorts whose free writing responses were categorised as exhibiting characteristics of “symbol association” and “counting”.</i>	90
<i>Figure 45: Histogram for “Concrete” Transfer Domain of CC Cohort</i>	91
<i>Figure 46: Mapping of symbols used by student AC34.</i>	92
<i>Figure 47: Percentage of Students in Different Cohorts to Rate the Learning and Transfer Domains as Either Identical, Similar or Different.</i>	94
<i>Figure 48: Rules Used in Learning Domain 1. Taken from SOM (Jennifer A. Kaminski, Sloutsky, & Heckler, 2009b)</i>	106
<i>Figure 49: Three of the six rules used for Learning Domain 2. Taken from SOM (Jennifer A. Kaminski et al., 2009b)</i>	106
<i>Figure 50: The Rules of Combination for the Testing Domain. Taken from SOM</i>	107
<i>Figure 51: Isomorphic Questions from Learning Domain 1 and Learning Domain 2 Respectively. Taken from SOM</i>	107
<i>Figure 52: Objects used in the four domains of de Bock et al.'s experiment</i>	110
<i>Figure 53: Activities involving the posited family of resources of symbolic combination/association</i>	111
<i>Figure 54: Examples of tasks that make use of a “counting” resource.</i>	112
<i>Figure 55: Two Different Instantiations of the Same Concept of Subtraction</i>	121
<i>Figure 56: Two Different Instantiations of the Same Concept of Multiplication</i>	122
<i>Figure 57: Two Different Instantiations of the Same Concept of a Mathematical Group</i>	124
<i>Figure 58: A Picture Depicting the Parable of the Blind Men and the Elephant</i>	128
<i>Figure 59: The “concreteness space” for a Brazilian high school physics student, who is unfamiliar with rugby balls and the idea of a gravitational field.</i>	140
<i>Figure 60: Two models of how familiarity might hinder learning. The black square represents a concept connect to other familiar concepts in a knowledge network. The triangle represents something to be added to the knowledge network.</i>	143
<i>Figure 61: Representation of the results from Petersen and McNeil's study in which student performance on a counting task was found to depend both on familiarity and perceptual richness</i>	146
<i>Figure 62: After learning the concept of competitive specialisation through either a perceptually rich or sparse simulation, students were tested in a transfer test. Their results for the transfer test depended both on the concreteness of the initial simulation (x axis), and on their performance in the initial simulation (y axis).</i>	147

List of Tables

<i>Table 1: A Table Showing Different Models of the Understanding of Subtraction in Different Contexts</i>	x
<i>Table 2: A Table Showing Different Models of the Understanding of Vector Addition in Different Contexts</i>	xii
<i>Table 3: A Table Showing Different Models of the Understanding of Basic Mechanics Concepts in Different Contexts</i>	xiii
<i>Table 4: A Table Showing Different Models of the Understanding of a Mathematical Group in Different Contexts</i>	xiii
<i>Table 5: Examples of Different Nouns used in Control and Experimental Questionnaires</i>	57
<i>Table 6: Relationship Between Questions in Study 1 and Questions in Study 2</i>	69
<i>Table 7: Table Showing Relationship Between Questions in Study 2 and Study 3</i>	75
<i>Table 8: Familiarity Ratings of Five Questions taken from our FCI study (n = 224)</i>	78
<i>Table 9: The Objects Used for Kaminski et al.'s Central Experiment</i>	105
<i>Table 10: Objects Used in Experiment 4</i>	115

i A Note on Style

Please take note of the following unconventional stylistic idiosyncrasies:

- i. The pronoun “we” is often used when referring to the author of this thesis. This is because much of the text began as articles for journal papers, where the authors include both the PhD candidate and his supervisor. (These journal papers are still in the process of being submitted for publication).
- ii. Bar graphs with error bars are presented with concomitant statistical analyses of significance. We found that delaying all statistical analysis to the “Discussion” sections proved irksome to the reader.
- iii. A literature review is not contained in a separate chapter, but is woven into the other chapters of the thesis.
- iv. Chapters 5, 6 and 7 are largely in the style of journal papers, and will contain some repetition.

ii Thesis Overview

“My students understand the concept of subtraction”. An unremarkable statement. But what does it mean? Table 1 shows five different contexts involving the concept of subtraction, and five different ways we might model students’ understanding of subtraction in each of these contexts. (For the sake of simplicity, I have used the generic term “schema” to illustrate what a model of understanding might be in these contexts).

Concept	Context	Possible Models of Understanding
subtraction (for non-negative answers)	$5 - 3$	e.g. a schema involving “take away” or “remove”, such as “5 sheep <i>take away</i> 3 sheep”
subtraction (of whole numbers)	$3 - 5$	e.g. a schema of “moving left along a number line”
subtraction (of fractions)	$\frac{4}{5} - \frac{3}{4}$	e.g. the algorithmic “butterfly method”
subtraction (of fractions (involving estimation))	Estimate: $\frac{19}{20} - \frac{19}{21}$	e.g. any number of embodied schemas such as representing the fractions as “slices of pizza” divided between 20 and 21 people respectively
subtraction (of vectors)	5 N East – 3 N North	e.g. a complex combination of skills and knowledge including the cardinal directions, graphical representation of vectors, Pythagoras’ theorem, etc.

Table 1: A Table Showing Different Models of the Understanding of Subtraction in Different Contexts

This table illustrates the three core areas of interest of this thesis. Firstly, how do we model students’ **understanding** of physical and mathematical concepts? The literature in education, psychology and cognition provides a wealth of modelling tools: cognitive resources, p-prims, image schemas, coordination classes, metaphors, embodied simulation, dual processing, concreteness, working memory, framing, priming, epistemological resources, etc. This thesis predominantly makes use of “cognitive resources” (Hammer, 2000) and “embodied simulation” (Lawrence Barsalou, 2003) and argues that these perspectives are particularly productive in analysing specific areas of physics (and math) education research. Secondly, how do we make sense of the **context** dependency of understanding? In particular, how do we identify one context as being sufficiently *different* from another context? And thirdly, what is a **concept**? How do we identify, or delineate a concept, and

what is its relation to **context** and to our model of **understanding**? For example, in the table above, all contexts deal with *the same* concept of subtraction (which occurs in **boldface**). Yet, in each context, this concept can be *differentiated* at a finer grain size. Note how the terms in nested parentheses define a finer grain size of **concept**, but are *also* related to the **context**. This problematises both the appropriate grain size of a **concept**, as well as the boundary between **concept** and **context**.

Chapter 1 provides the philosophical underpinning of the thesis. In his book “Saussure and Wittgenstein: How to play games with words”, Roy Harris (1990) outlines a particular view of language which he calls “nomenclaturist”. This is the idea that words get their meaning by naming things “out there” in the world. Obvious examples include words such as “table” and “tree”. He then argues that Saussure and Wittgenstein’s insights into the workings of language are fundamentally anti-nomenclaturist. Chapter 1 applies these insights to the core concepts of this thesis: “understanding”, “concept” and “context”. It argues that before we can begin using these words, we must better appreciate how they come to have meaning: their meaning does not come from naming objects, but from having a particular use in the “games” we play with them.

Chapter 2 provides some theoretical background. While most theoretical concepts are introduced on a “need to know” basis during the thesis, this chapter outlines two of the major theoretical frameworks used in this thesis: (i) Resources/Knowledge in Pieces, and (ii) Simulation and Dual Process Theory.

Chapters 3 – 7 report on three empirical studies.

(i) Chapter 3: Context Dependency and Embodied Simulation in Vector Addition

Three series of questionnaires and interviews were carried out with freshman concerning their understanding of basic 2-d vector addition. We found a striking context dependency; students easily add *displacements* and *forces*, but struggle to add *momenta*. Furthermore, we found that students display a context sensitivity to the textual prompts “*net*” and “*total*”. This latter context sensitivity is hypothesised to result from two competing embodied simulations. This hypothesis was tested in a further (fourth) pilot study.

<u>Concept</u>	<u>Context</u>	<u>Possible Models of Understanding</u>
Vector Addition (of forces and displacements (in 2d))	Forces and Displacements	Mostly algorithmic/fast processing; some embodied simulation
Vector Addition (of momenta (in 2d))	Momentum	Cognitive resources activated are significantly different to those activated when adding “forces” and “displacements”
Vector Addition (involving total (forces (in 2d)))	Total Force	some embodied simulation of felt experience
Vector Addition (involving net (forces (in 2d)))	Net Force	some embodied simulation of resultant motion
Vector Addition (of forces (on person (in 1d)))	Forces on Person	Embodied simulation of force experienced
Vector Addition (of forces (on hard ball (in 1d)))	Forces on Hard Ball	Embodied simulation of resultant motion

Table 2: A Table Showing Different Models of the Understanding of Vector Addition in Different Contexts

(ii) **Chapter 4: Context Dependency and Embodied Simulation in the FCI**

The phenomenon of context dependency in the FCI has been well documented (Stewart, Griffin, & Stewart, 2007). For example a change of object (e.g. a book supported by a spring vs. a book supported by a table (See Palmer (2001)) can have a major effect on student performance. We asked the question: what if the modified context contained *unfamiliar words*? In other words, what would happen if we replaced common words such as “box”, “car” and “man”, with unfamiliar words such as “kist”, “jalopy” and “Australopithecus”. We hypothesised that students make sense of FCI questions through embodied simulation (e.g. they actually simulate a man pushing a box when making sense of the question). Our hypothesis predicts that the presence of unfamiliar words would frustrate and hinder embodied simulation, thereby adversely affecting the performance of students in this modified version of the FCI. We compared the performance of students in this experimental context (questions with unfamiliar words), with that of a control group (questions with the original, familiar words). Two experiments were performed, sometimes yielding conflicting results. A larger, third study demonstrated that performance is moderately better in the “familiar word” context, and that the effect of the unfamiliar words is moderately correlated with the difficulty of the questions.

Theoretical frameworks of Language and Situated Simulation) (L. W. Barsalou, Santos, Simmons, & Wilson, 2008) and working memory (Baddeley, 2007) were used in the analysis.

<u>Concept</u>	<u>Context</u>	<u>Possible Models of Understanding</u>
Basic Mechanics Concepts (with familiar words)	Familiar Words	Evidence of shallow linguistic processing, and possible embodied simulation.
Basic Mechanics Concepts (with unfamiliar words)	Unfamiliar Words	

Table 3: A Table Showing Different Models of the Understanding of Basic Mechanics Concepts in Different Contexts

(iii) **Chapters 5, 6 and 7: A Resources Perspective of the Studies of Kaminski et al: Context Dependency and Embodied Simulation in Transfer.**

Kaminski, Sloutsky, & Heckler (2008) published a series of studies arguing that teaching math by means of abstract examples results in better transfer than teaching by means of concrete examples. Their studies involved a questionnaire which tested student transfer of the concept of “mathematical group of modulo-3”. Students were introduced to this concept in one of two contexts: by means of concrete examples, or by means of abstract examples. Kaminski et al. found that students who learned by means of abstract examples outperformed those students who learned the concept by means of concrete examples. This thesis argues that these conclusions are confused due to a lack of interrogation of the notions of “context” and “concept”. We replicate and extend Kaminski et al.’s study, and demonstrate that a Resources perspective provides a clearer and more productive lens through which to interpret the results of these and similar studies. In particular, we collect and analyse qualitative data to show that students are engaged in different *activities* in the different contexts.



<u>Concept</u>	<u>Context</u>	<u>Model of Understanding</u>
Mathematical Group (involving of abstract symbols)		Resource activation involving “symbol association”.
Mathematical Group (involving containers filled with discrete levels)		Resource activation involving “counting”.

Table 4: A Table Showing Different Models of the Understanding of a Mathematical Group in Different Contexts

Note that there is a prelude to chapters 5, 6 and 7 which better explains how these chapters are related to each other.

Chapter 8 provides a final discussion and conclusion.

CHAPTER 1: Introduction

Introducing the three main ideas in the title of the thesis.

1.1. Understanding

Understanding is not a “thing”. What do I mean by this. “Modelling understanding” is profoundly different to, for example, “modelling planetary motion”. When we model planetary motion, we aim to truthfully/accurately represent a particular phenomena which exists independently of our modelling. We develop better and better models in our quest for this truth/accuracy. (Compare epicycles, to Kepler’s Laws, to GR). The concept of “understanding” however, is so deeply embedded in our human form of life, that our models of understanding partly constitute what understanding *is*. (And here I include “lay-models” of ordinary usage). Like the artichoke, if we peel away all our models of understanding, we are not left with the “something at the centre” which we are trying to model. This section fleshes out this insight.

1.1.1. Grammatical Beginnings

“Understanding” is not a mental or physical state; it is not a “something” that obtains when certain necessary and sufficient conditions are met.

“What really comes before our mind when we *understand* a word?” (Wittgenstein, 1958, §139)

It is common to picture “understanding” as having a source; as emanating from *something*. We might picture the source of our understanding to be a mental image, a schema, or a formula. For example, the mathematics education researcher, Shlomo Vinner, claims that “to understand... means to have a concept image.” (Vinner, 2002, p.69). The philosopher, Ludwig Wittgenstein, was well aware of this tendency, and devoted much of his *Philosophical Investigations* (Wittgenstein, 1958) to illustrating how this tendency results in a variety of philosophical (i.e. conceptual) confusions. Wittgenstein shows that “understanding” is not a physical or a mental state. It is not a *something* that can be discovered. This is a grammatical fact, not an empirical claim. A vast literature elucidates Wittgenstein’s insight. I will illustrate one small facet regarding “mental pictures”, but it can be extended to other mental phenomena such as “conceptual images” etc. Suppose that the source of understanding is a mental picture, akin to the graphic shown in Figure 1 below:



Figure 1: A Picture Which Seems to Force a Particular Interpretation On Us: A Person Walking a Slope

The meaning of this picture is “a person walking up a slope”. However, it could also be understood as “a person sliding down a slope (in a somewhat peculiar position)”. The point is that the picture does not force a particular meaning on us; a picture (or a concept image, or a schema, etc.) always requires a further interpretation. In his interpretation of Wittgenstein’s famous “rule following arguments”, Kripke (1982) provides a similar analysis of the addition sign, “+”. He demonstrates that it can always be interpreted in a deviant manner, and concludes that there is no symbol, rule, or formula that can be said to be the source of our understanding of addition. (See Southey’s (2012) master’s thesis for further details regarding this argument). To re-iterate: “understanding” does not get its meaning by referring to a particular “thing”. Instead, it has a particular use in our language games¹; in our forms of life². This is the bedrock of its “true meaning”. This is a crucial point. It provides us with an anchor; a beginning to our investigation. If we realise that a mission to discover the *actual* nature of *actual* understanding is as fruitless as a mission to discover the East Pole, we can focus our efforts on a more productive project: modelling understanding.

1.1.2. Criteria

Our criteria for ascriptions of understanding are always behavioural. i.e. “inner” mental events are not candidates for criteria of understanding.

Another facet of the Wittgensteinian argument against positing “understanding” as a type of mental state is the “Beetle in the Box” argument” (Wittgenstein (1958), §293), also known as the “Idle Wheel Argument” (Schroeder (2006)). The essential idea is that even if concomitant mental states

¹ The phrase “language game” was coined by Wittgenstein (1958), and is similar to the notion of “discourse”. Different language games govern different aspects of life (e.g. a language game for commerce, for romance, for academia, etc.) and it is *within* these language games that word meaning gains traction.

² “form of life” is another Wittgensteinian term, which is similar to the notion of a “sociocultural context”; that complex web of activities, expectations, intentions, norms, etc. that constitute a particular form of life.

accompanied understanding, they play no role in our ascription of understanding.³ We cannot literally look into people's minds for evidence of understanding; any ascription of understanding involves behavioural criteria.⁴ This behaviour might be a written answer to exam questions (understanding a theoretical concept), whistling a melody (understanding a musical theme), prodding someone's stomach in the appropriate places (understanding human physiology), or identifying physical features of a building (understanding architecture). This is not to say that these behaviours *are* the understanding.⁵ Just as "understanding" is not to be *equated* with a particular mental state or brain process, so it is not to be *equated* with particular behaviour. Nevertheless, just as a well-functioning brain is necessary for understanding, so appropriate behaviour play an essential role in our ascription of understanding.

In summary, mental states or objects (such as mental pictures, schemas, etc.) do no work in our ascription of understanding; they are idle wheels. Thus, in using such mental objects in our models of understanding, we must be aware that these entities do not obtain meaning by corresponding to "things" in the mind/brain which can act as criteria in our ascriptions of understanding. Instead, these mental entities gain traction through the immensely complex web of interconnected behaviours, expectations, intentions, cultural norms, etc. In short, these mental entities play a role in our language games which occur in our particular forms of life.

1.1.3. Understanding and Ability

It is better to think of understanding as being more akin to an ability (or potentiality) than an occurrent mental state. However, understanding is not equivalent to ability. The relationship between understanding and ability is complex, and underpins (i) the notion of different types (or levels) of understanding and (ii) the notion of actual understanding and imitated understanding.

If understanding is not to be equated with mental entities, or with behaviour, how should we think of understanding? Expanding on Wittgenstein's heuristic of "meaning as use", we can think of understanding as being akin to ability; "To know something is... more akin to a power or potentiality than a state or actuality" (Bennett & Hacker, 2003, p.148). However, as any teacher will ratify, being able to do something does not entail that one truly understands what one is doing. e.g. "my students

³ Note that the asymmetry between first and third person ascription is a major theme in Wittgenstein's philosophy (see for example (Wright, 1998)). This dissertation will not consider first person ascriptions of understanding, but acknowledges that students' perception of their own understanding is both interesting and valuable.

⁴ Of course, various technologies such as fMRI do offer us a peek inside the brain. However, these technologies do not provide us with *criteria* for ascriptions of understanding.

⁵ An identification of a mental phenomenon, such as understanding, with behaviour is indicative of classic behaviourism. Wittgenstein's position is demonstrably distinct from behaviourism. See Hacker (1993).

are able to perform integrals, but they do not understand what integration really is.” The relationship between understanding and ability is important for at least two reasons: (i) It underlies the notion of “levels” of understanding; (ii) It underlies the much-used distinction between *actual* understanding (or sense-making) and *imitated* understanding (or answer-making).

Understanding is not a dichotomous state: we don’t either have it, or not. In this sense, it is akin to an ability; as something which develops over time. (Compare the view of understanding as the possession of a particular mental image or schema. This is more likely to lead to a dichotomous view: If you have the mental image, then you have the understanding; If you don’t have the mental image, then you don’t have the understanding). The idea of understanding as something which develops over time is perhaps best represented in Bloom’s Taxonomy of knowledge shown in Figure 2 below. Note that this taxonomy also highlights the complex relationship between the concepts of understanding and ability. Each level is characterised by particular abilities. One might ask: (a) Is it the possession of understanding that leads to the exercise of these abilities? Or (b) Is it the exercise of these abilities that *mean* the particular level of understanding? These questions fall prey to the types of conceptual confusions this chapter seeks to dispel. To re-iterate: (a) Understanding is not the possession of a “thing” (such as a mental image or a schema) from which our abilities might flow; and (b) While behavioural criteria are essential for an ascription of understanding, this does not make understanding *identical* to the behaviour (or the exercise of particular abilities). To get clearer on the distinction between “ability” and “understanding”, we will look next at how ability is related to *imitated* understanding and *actual* understanding.

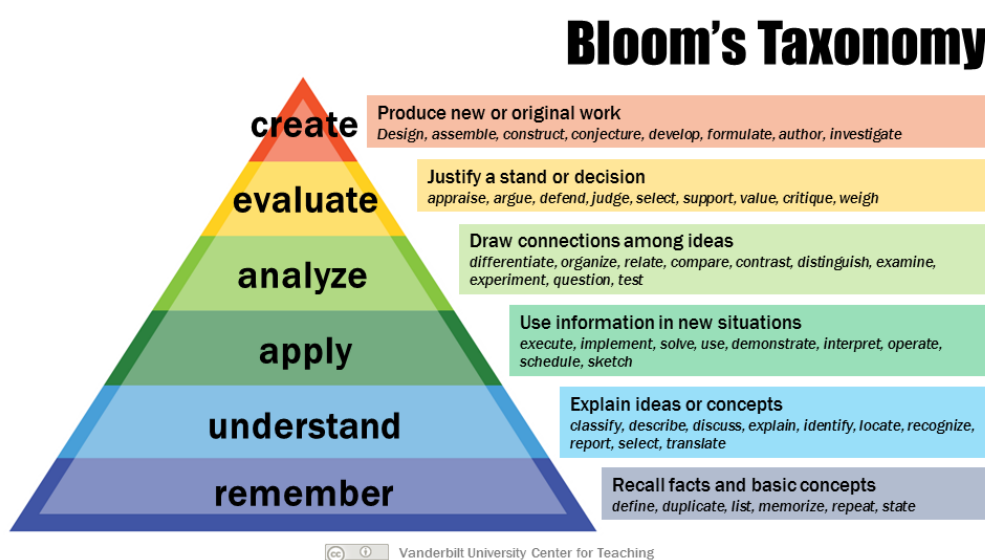


Figure 2: Bloom's Taxonomy of Knowledge Might Be Thought of as a Hierarchy of Different Abilities.
(Picture Under Creative Commons License Courtesy of Vanderbilt University Center for Teaching)

1.1.4. Knowing How and Knowing That

The Knowing How/Knowing That distinction underpins our intuitive ideas of two different levels of understanding – shallow understanding and deep understanding.

Gilbert Ryle (1945) distinguished between two types of knowledge: knowing how and knowing that. He believed that knowing how to do something is always and essentially different from knowing that something is thus and so. This distinction is well illustrated in a famous thought experiment by John Searle (1980) called the “Chinese Room Argument”. Searle imagines himself alone in a locked room. Chinese symbols are slipped under the door. Searle understands nothing of Chinese and only recognises the Chinese symbols by their shape. He possesses a set of rules, written in English, which instruct him how to generate symbols in response to the symbols he receives. Unbeknownst to him, the symbols slipped under the door are questions, and the symbols he generates, with the help of his set of rules, are answers to those questions. He imagines his set of rules to be sufficiently sophisticated that his answers are indistinguishable from that of a native Chinese speaker. People outside the room are therefore fooled into believing that he understands Chinese, but Searle insists that although he has developed an ability to respond to written Chinese, he has only *imitated* understanding and not achieved *actual* understanding of the language. He *knows how* to generate symbols, but does not *know that* symbol X means Y. The thought experiment was devised in order to question the attribution of (actual) understanding to computers and the development of AI.

However, its relevance for education research is self-evident. We often attribute something like *imitated* understanding (or knowing *how*) to students. In the education literature we use terms such as plug-and-chug, pattern matching and algorithmic manipulation (Kuo, 2013); rote learning (Elby, 1999); shallow processing (F Marton & Säljö, 1976); answer making (Turpen & Finkelstein, 2010); memorization and regurgitation, etc. Just as in Searle’s thought experiment, we might attribute a particular ability to students in these cases, but we do not attribute *actual* understanding to them. Conversely, we also use terms to attribute *actual* understanding to students, such as deep processing (F Marton & Säljö, 1976) and sense-making (Dervin, 1983). Sense-making emphasises the connective and constructive aspects of knowledge generation; the integration of new information with past experience. For example Spillane, Reiser, & Reimer (2002) define sense-making as “how prior knowledge, beliefs, and experiences influence construction of new understandings” (Spillane et al., 2002, p.388), while Kuo (2013) states that authentic sensemaking involves seeking “coherence between intuitive ideas and formal representations.” (Kuo, 2013, p.95). Quinn (2009) states that “sense-making involves looking for patterns in one’s experience to make plausible judgments about future experiences” (Quinn, 2009, p.64).

Subsequent discussion on Ryle's distinction (see for example Bennett & Hacker (2003) and White (1982)) points to the fact that the distinction between "knowing how" and "knowing that" is not as rigorous as Ryle would have believed; indeed, these authors show that "knowing that" often entails "knowing how" to do something. In other words, the knowing how/knowing that distinction cannot ultimately be used to define two different levels knowledge. Rather, the relationship between "behavioural abilities" (knowing how) and "cognitive abilities" (knowing that) is complex, and forms a continuum of levels of understanding.

To summarise: This subsection speaks to the complex relationship between understanding and ability. The "knowing how" and "knowing that" distinction speaks to an intuitive sense we have of the relationship between understanding and ability. i.e. that actual understanding is *more* than basic behavioural ability. However, this does not negate the idea that actual understanding is still akin to ability.

1.1.5. Modelling Understanding

This section on "Understanding" is intended to emphasise that ascriptions of understanding belie immensely complex processes. A sentence such as "Student A has an understanding of subtraction" is radically different from "Student A has dark hair". Understanding is not a "thing" one can point to, nor is its possession dichotomous. We can only ever observe a small subset of the relevant criteria for the ascription of understanding. And, what counts as "relevant criteria" depends on our use of "understanding" in a particular context. i.e. our model of understanding. For example, if we model understanding of "subtraction of fractions" to be the ability to apply an algorithmic procedure, we will look for criteria of finding a common denominator etc. However, if we model the understanding of "subtraction of fractions" as the ability to make an estimation based on physical intuition, our criteria will be very different. Here "model of understanding" need not refer only to a robust theoretical model such as "embodied simulation". Rather, our model of understanding is akin to *what understanding means in a particular context*. And thus, the meaning of "understanding" is partly constituted by the models we use.

Next, we look at how our models of understanding depend crucially both on the *context* in which they are to be used, and the *concepts* which they take as their objects.

1.2. Concept

This section covers similar ground to Section 1.1, but instead of considering "understanding" we consider the object of understanding; a concept.

In the context of education, “understand” is a transitive verb; it always takes an object. If one has achieved understanding, we may always ask: An understanding *of what*? Therefore, to speak of a particular model of understanding, is also to speak of its object; typically termed a concept. In physics education a typical object of understanding is a physics concept such as “force”. However, as argued by diSessa (2002) there is very little theorising as to what a concept actually *is*. Just as there are many different models of understanding, there may be very different types of concept. In fact, each model of understanding should have its own model of a concept. An illustrative juxtaposition is the location of “concept” in the broad models of cognitivism and sociocultural theory. A cognitivist might locate the notion of “concept” completely *within* the mind of the subject (“concepts are neural structures” (Lakoff & Johnson, 1999, p.19)). However, in sociocultural theory, our knowledge is understood to be situated within a particular sociocultural context, and the notion of “concept” would be located somewhere *between* the subject and their surroundings. Indeed, “concept” is associated with numerous terms in the education and philosophy literature: prototypes (Rosch, 1999); conception/misconception (Gilbert & Watts, 1983); resource (Hammer, 2000); image schema (Hampe, 2005); concept image (Tall & Vinner, 1981); metaphor (Lakoff & Johnson, 1980); primary metaphor (Grady, 2005); blend (Fauconnier & Turner, 2008); mental models; semantic networks (Lakoff, 1987); neural networks, etc. Each of these terms is associated with a different theory (or model) of understanding. And each these theories will mean something different by the sentence: “Student X understands concept Y”.

1.2.1. Having a Concept and Having an Ability

Our ordinary discourse of concepts makes use of the idea (the metaphor) that concepts are unitary, stable objects. Picture the following commonly used phrases: “I’m hoping to get the concept across to my students”; “This concept builds on previously learned concepts”; “They’ve successfully transferred the concept to a new context”. The model of “concept as object” is particularly apparent when we speak of an “instantiation” or “representation” of a concept. Much theorising about concepts has followed a similar trend: a concept has been modelled as an abstract object (Margolis and Laurence 2007), a mental image or idea (Locke, 1841), a schema (Hampe, 2005) or even a particular configuration of neurons (Lakoff & Johnson, 1999). This understanding of a concept is beset by what Wittgenstein called “one of the greatest sources of philosophical bewilderment: a substantive (i.e. a noun) makes us look for a thing that corresponds to it” (Wittgenstein 1958). While the model of “concept as object” might be useful in particular circumstances, a large number of philosophers and educational theorists have argued that “possession of a concept” is less like having a particular thing (mental image, schema) and more like having a particular ability: “To have a concept is to have mastered the use of a word (or phrase), so concept possession is a complex ability

or array of interconnected abilities.” (Bennett & Hacker, 2003, p.340). In developing their coordination class model of concepts, diSessa & Sherin (1998) define having a concept as the *ability* to “see the information that defines the concept in an appropriate range of relevant situations”. Sfard (2008) defines a concept as “a symbol together with its uses” (Sfard, 2008, p.111). This emphasis on ability or activity is fundamental to the discourse of situated learning, which speaks not of “concept acquisition” but of “legitimate peripheral participation”. (Lave & Wenger, 1991).

Sfard provides a compelling account of how abilities or activities come to be labelled as concepts.⁶ Sfard argues, that via the human ability of *objectification*, we replace “sentences about processes and actions with propositions about states and objects” (p.44). For example: “In the majority of school tests and tasks dealing with function she regularly *did well* and *attained* above average scores”, becomes “She *has acquired a conception* of function”. Thus the ability of scoring well in a test, becomes labelled as the acquisition of an *object* – the conception of a function. Bennet and Hacker give a similar account of the objectification of ability: “A concept is the abstraction from the use of the word.” (p.339).

This shift in focus from *concept as object* to *concept possession as ability possession* is obviously linked to a shift in modelling understanding as a *state*, to modelling understanding as an *ability*. A concept is therefore not to be understood as a unitary, stable object, but rather as a collection of various interconnected ideas and activities which have been objectified and subsumed under a single label.

1.2.2. Grain Size

Consider the first column of Table 1, Table 2, Table 3 and Table 4. Note how the text in bold refers to the overarching concept at issue, while the text in nested parenthesis refer to what might be termed “sub-concepts”. (Note that we would normally refer to the text in nested parenthesis as referring to the particular sub-*contexts* in which the concept is embedded). For example, in Table 2 the concept at issue is “vector addition”. It would be very strange for a lecturer to say: “My class understand the concept ‘vector addition involving the net force in 2d’”. And yet, if it can be shown that students understand “vector addition involving the **net** force in 2d” *differently* to “vector addition involving the **total** force in 2d” shouldn’t these be considered as two *different* concepts? Why do we speak about the concept of “vector addition” in the *contexts* of “net force” and “total force”, as though there is this “thing” called “vector addition”, which is embedded in two different environments?

⁶ For the purposes of this paper, the possession of an *ability* is to be thought of as proficiency in a variety of interconnected activities, and acquaintance with a variety of associated ideas (which themselves might be objectifications of other abilities).

Why do we not speak of two “smaller things”: (i) “vector addition involving the **total** force” and (ii) vector addition involving the **net** force”. Indeed, why is the grain size of the concept “vector addition” not drawn at a larger grain size? Couldn’t one say that “vector addition” is simply the concept of “vector operations” in the context of addition? (Which students understand differently to the concept of “vector operations” in the context of multiplication). Perhaps there is no “natural” grain size for a concept. There are only pragmatic reasons for choosing one grain size over another in order to delineate between “the concept” and “the context”.

1.3. Context

1.3.1. Definitions and Models of Context

Taken from Merriam-Webster Dictionary (2016):

Context

- 1: the parts of a discourse that surround a word or passage and can throw light on its meaning
- 2: the interrelated conditions in which something exists or occurs: environment, setting

These two definitions from the Merriman Webster mirror the two models of context used by Finkelstein (2005) in his paper “Learning Physics in Context”. Citing Cole (1996), Rogoff (1992) and McDermott (2001), Finkelstein (2005) identifies two broad models of context:

(i) that which surrounds, and (ii) that which weaves together, or mutually constitute each other.

The central issue is that of *separateness* or *difference*. In education, if we speak of the “context dependency” of a concept, the grammar (or logic) of the term “context” implies that there is *something separate* from the context, that occurs *within* the context. For example, “the concept of subtraction *within* the context of vectors”. This logic corresponds to the first definition/model mentioned above – context as surrounding *something*. This logic is depicted in Figure 3.

Interestingly, Figure 3 might also represent a basic model of transfer: a concept (the small blue square) occurs *within* Context A, but is separate from Context A. Thus, it is possible to transfer this

concept to Context B.

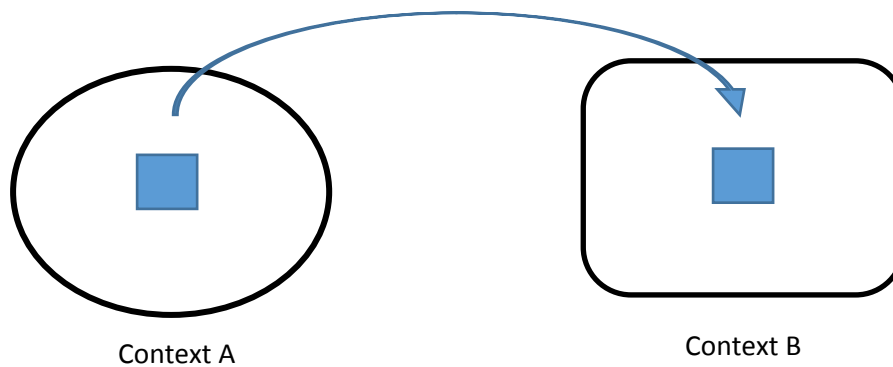


Figure 3: A "something" occurring in two different contexts

An obvious critique of this model of context (which can be extrapolated to a critique of the notion of "transfer" (Lave & Wenger, 1991)) is that it is not always clear how to *separate* the surrounding *context* and the *something* within the context. A striking example of this is wave/particle duality, where the very nature of a photon depends on its context. i.e. the photon is not simply a *something* which occurs, unchanged, in different contexts. This notion of "context" as "that which surrounds" is further critiqued by van Oers in his paper "The Fallacy of Decontextualisation" (Van Oers, 1998). Van Oers addresses particular accounts of abstraction which imagine that one can peel away layers of context (or situatedness, or concreteness) until one is left with a pure, abstract object (cf. the blue square in Figure 3). Van Oers argues that if context is understood as providing *meaning* (see definition above (Merriam-Webster Dictionary, 2016)), then the end result of decontextualisation is meaningless-ness.

In the second model used by Finkelstein (2005), context is described as "the connected whole that includes constituent elements and the relations among them." (p.1191). In this model "the task and the context are thought of as *mutually constitutive*" (ibid).⁷ This model does not seem very helpful for our purposes, as it does not provide a sufficient "figure/ground" contrast for considering *context dependency*; concepts *within* a particular context. Thus, this thesis will largely adopt the first model of context which is thought to be more relevant for our purposes. However, a central theme of the thesis will be the limitations and pitfalls of this particular model of context. For example, instead of thinking of "context" as a passive "surrounding" we can think of it in terms of a locus for meaningful actions.

⁷ See Barad (2007) for an thorough account of an ontological system based on this notion of mutual constitutiveness, inspired by quantum physics.

Finally, cognitive linguists are insistent that “context” is a mental phenomenon. (See Ungerer & Schmid (2013) and (Langacker, 2008)). This is also stressed by Redish (2004): “In the mind of a particular individual, context is the state of activation of each of the neurons in the individual’s brain at a particular instant” (p.16). The introduction of a thinking subject will be important when considering differences between novice and experts’ understanding of a concept within a particular context.

1.3.2. Context and Understanding

To re-iterate the moral of section 1.1.1: we are not looking for a universal set of necessary and sufficient conditions that completely capture the meaning of “understanding”. Instead, by adopting Wittgenstein’s notion of “meaning as use” (see for example Baker & Hacker (2008)) the meaning of “understanding” is determined by its use in a particular context. We may understand a picture (§526)⁸, a language (§199), a musical theme (§527), a sentence (§60) or a poem (§531). It is clear that our criteria for understanding, our model of understanding, and our *meaning* of understanding is very different in these different contexts.

In STEM education the context-dependency of understanding is manifold. Consider different representational contexts: is our understanding of a concept represented in an equation similar to our understanding of the same concept represented as a graph? Is our understanding of a concrete representation of a concept similar to our understanding of an abstract representation of the same concept? Consider different epistemological contexts: is our understanding of a concept learned empirically in a laboratory investigation, similar to our understanding of the same concept when learned from a textbook? Consider different disciplinary (or discourse) contexts: is a sociologist’s understanding of the concept of “measurement” the same as a physicists’ understanding of “measurement”? Consider different cognitive developmental contexts. Is a novice’s understanding of a concept similar to an expert’s understanding? Consider the contexts of different concepts. Is our understanding of “force” similar to our understanding of “subtraction”, or our understanding of “dog”? And consider different physical contexts. Is our understanding of force, when exerted by a table, the same as our understanding of force when exerted by a spring?

What does it mean for understanding to be *different* in different contexts? It means that the criteria/model/meaning of understanding is different. An aim of this thesis is to determine what criteria/models/meanings of understanding might be most productive in specific contexts in physics education research.

⁸ References are to paragraph numbers in the Philosophical Investigations (Wittgenstein, 1958)

CHAPTER 2: Theoretical Background

As mentioned in the Introduction, there are numerous theoretical tools that can be brought to bear when modelling “understanding”. For example: embodiment, perceptual symbol systems, knowledge in pieces, misconceptions, image schemas, metaphor theory, blending theory, coordination classes, symbolic forms, cognitive resources, p-prims, image schemas, coordination classes, metaphors, embodied simulation, dual processing, concreteness, working memory, framing, priming, epistemological resources, etc. Most of these terms will be introduced and defined during the thesis. Indeed, most of the required theoretical background is woven into the thesis on a “need-to-know” basis. Here, I will only introduce two of the main frameworks used: Resources/Knowledge in Pieces and Embodied Simulation. These frameworks will also be revisited and refined during the thesis.

2.1. Resources / Knowledge in Pieces

Our ordinary discourse of concepts makes use of the idea (the metaphor) that concepts are unitary, stable objects. This ontological view of concepts as robust units has had a major impact on ideas of conceptual change and transfer, which continue to hold sway. In particular, it has been the foundation of a vast literature on “misconceptions”. These are generally thought to be stable, coherent concepts (or quasi-theories) that students incorrectly hold about the world. (See DiSessa (2006) for a review). Some education researchers have challenged this view, and advocate that knowledge should not be viewed as occurring in coherent cognitive chunks, but that “knowing” should be thought of as the activation of multiple fine grained cognitive elements in a particular context. These have been dubbed “cognitive resources”, or simply “resources” (Hammer, 2000). Unlike a unitary view of concepts, these elements are fragmented; perhaps loosely associated with other elements, but with limited coherence. They are also highly context sensitive.

A powerful example of the usefulness of this perspective is students’ descriptions of the forces acting on a ball after it is tossed vertically in the air. (See diSessa (1993) and Hammer et al. (2005)). The only force acting on the ball in the “ball toss” problem, is the constant downward force of gravity, causing the ball to decelerate to a stop, and then accelerate back to earth. However, during its upward trajectory, it is common for students to incorrectly posit a gradually diminishing upward force. From a “unitary” perspective of concepts, one could attribute the misconception “force is proportional to velocity” to the students; as the ball slows down, so the erroneous upward force becomes less. However, at the top of ball’s trajectory, it is common for students to posit two equal and opposite forces: (i) gravity, and (ii) an equal and opposite upward force. This phenomenon is

difficult to explain from a unitary, misconceptions perspective. The velocity of the ball at the top of its trajectory is zero, therefore the presence of the “balancing force” directly contradicts the idea that students hold the robust misconception that “force is proportional to velocity”. However, students seem to have no qualms in positing both the diminishing upward force and the balancing force in their description of the ball’s trajectory. This scenario is better explained from a “knowledge in pieces” or “resources” perspective. During the ball’s upward flight, the cognitive resource of “force implies motion” is activated. However, at the top of its trajectory, the cognitive resource “balancing” is activated.

Importantly resources are not “correct” or “incorrect” in themselves, but they can be correctly or incorrectly *applied*. In other words, a resource only becomes correct or incorrect when used in a particular context. For example “force implies motion” is a perfectly good resource which can be usefully applied in a variety of contexts, but it has been misapplied in the above example.

Another example of Knowledge in Pieces comes from Southey’s investigation into the concept of density. Southey found that students make sense of *density* in a variety of different ways: packing; floating and sinking; heaviness; unchanging intensive property of a material; amount of stuff in a designated volume; the equation $\rho = \frac{m}{v}$. These might be thought of as different resources students draw on in order to make sense of the overarching concept of density. Questionnaires testing students’ understanding of density (see for example (Kohn, 1993)) tend to use a variety of questions which trigger these different resources. Southey argues that instead of interpreting these questionnaires as testing for students’ understanding of the concept “density”, it is more accurate to think of the questionnaires as testing students’ ability to apply a variety of resources. In other words, instead of thinking of students’ knowledge of density as occurring in one large chunk, one should think of students’ knowledge of density as occurring in pieces.

2.2. Simulation and Dual Process Theory

A key question in cognitivist theories of knowledge is: how is knowledge represented? Some assume a single amodal system; a type of “mentalese” in which our knowledge is coded. The most famous proponent of which is probably Jerry Fodor (Fodor, 1975). Other theories assume a multimodal model. These theories assume knowledge is distributed in a variety of systems. For example Barsalou’s (1999) theory of perceptual symbol systems posits a “concept-like” entity, a “perceptual symbol”, to underlie knowledge representation. A perceptual symbol is record of all the neural states that underlie perception. It is not a *perfect* recording, but a type of schematic representation. A perceptual symbol contains multi-modal information, including sight, taste, touch, sound, and smell. Therefore, according to Barsalou’s (2003) account, using one’s knowledge entails re-activation

of the multi-modal systems. This re-activation is the basis of the notion of “simulation”. (See also Barsalou (2009)).

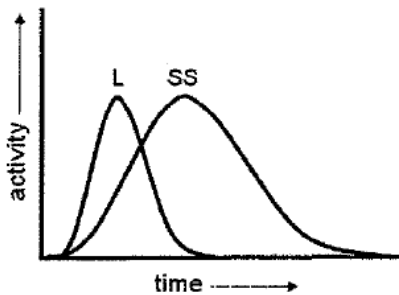


Figure 4: Activation Time Curves for Linguistic System and Situated Simulation System. Taken from L. W. Barsalou et al. (2008, p.248)

In their paper *Language and Simulation in Conceptual Processing*, L. W. Barsalou et al. (2008) contrast this type of simulative processing, with linguistic processing. Like Paivio (1990) they posit a type of dual processing theory of knowledge. Their theory of *Language and Situated Simulation* (LASS) posits (i) a linguistic system, and (ii) a conceptual system. (i) The linguistic deals with the processing of linguistic units such as words, and includes processes such as basic word association. The linguistic system might be described as “shallow” and “fast”: “We

assume that these linguistic strategies are relatively superficial. Rather than providing deep conceptual information, these strategies provide shallow heuristics that make correct performance easily possible.” (L. W. Barsalou et al., 2008, p.249). (ii) The conceptual system deals with simulation. It is associated with “deeper” knowledge, which is grounded in our perceptual systems, rather than our language systems. Although both systems are activated simultaneously, Figure 4 shows how the activity of the linguistic system is thought to peak first. In other words, if a problem can be solved using the shallower, associative processing of the linguistic system, there will be no need to engage the conceptual system in situated simulation. “...executive processing can focus on the linguistic system as its primary source of information for at least several seconds, before simulations begin to have effects on behaviour... When fast responses are possible, information from the linguistic system dominates, and information from the simulation system has less effect.” (L. W. Barsalou et al., 2008, p.250).

There seems to be an obvious analogy to be made here with the two systems of Kahneman's (2011) Dual Process Theory: the fast thinking “System 1” (which Kahneman calls “the associative machine”), and the slow thinking “System 2” (which Kahneman calls the “lazy controller”), which requires substantial mental effort. When confronted with a problem, System 2 is only engaged if System 1 does not provide a satisfactory solution. This is analogous to the temporal sequencing of the linguistic system (involving linguistic association) and the conceptual system (involving simulation) of Barsalou et al.

The analogy is not complete, however. In the opening to his chapter describing System 1, Kahneman asks the reader to look at the following words: “banana” and “vomit”.

A lot happened to you in the last second or two. You experienced some unpleasant images and memories. Your face twisted slightly in an expression of disgust, and you have may have pushed this book imperceptibly farther away. Your heart rate increased, the hair on your arms rose a little, and your sweat glands were activated. In short, you responded to the disgusting word with an attenuated version of how you would react to the actual event. (Kahneman, 2011, p.50)

In short, Kahneman demonstrates that *simulation* can also be associated with System 1. Indeed, in physics education research, Dual Process Theory is predominantly used when researching students' metacognition. (See for example Kryjevskiaia, Stetzer, & Grosz (2014)). That is, students' ability to "slow down, take a step back", and assess their thinking process. A helpful theoretical term to introduce here is "backstage cognition" (Turner, 2000). These are "mental operations that typically operate below the horizon of observation". Thus, Barsalou et al.'s linguistic system is comparable to Kahneman's System 1, and both deal with backstage cognition; cognition below the horizon of conscious awareness. Kahneman's System 2 is, by definition, "frontstage cognition", or slow, effortful thinking. However, while Barsalou et al.'s conceptual system, is typically associated with slower, conscious, simulative sensemaking, we can also identify elements of as backstage cognition as being of a simulative character.

Finally, the simulative character of conceptual processing lies on a spectrum. For example, describing scoring the winning goal of a soccer match, including all emotional and physical detail, is a paragon of simulative processing. However, briefly recalling that one lefts one's keys on the coffee table would be a shallower example of simulation; involving mainly visualization. (Compare recalling how one placed the keys on the table, including the feel of the keys in one's hand, and the texture of the surface of the table). In this thesis we hope to use the theoretical notion of Simulation as a tool to model student sensemaking. Identifying the simulative character of students' approaches to a question from their written response involved coding their response as lying sufficiently "high-up" on the spectrum of simulation. The methodology is explored in section 3.5.4.

2.3. Resources and Simulation

As these two theoretical perspectives are repeatedly drawn upon in this thesis, it is instructive to make a few comments on how they might be related. They arise in two very different contexts. As described above, Resources arose in opposition to a misconceptions framework, and offers a new description of elements of student cognition. Instead of thinking of elements of student cognition as occurring in large, coherent, pre-exisiting "chunks" (i.e. "conceptions" or "misconceptions"),

resources were posited as fine grained elements of cognition whose activation is heavily context dependent. On the other hand, Simulation arose as a reaction against amodal theories of conceptual processing. The emphasis in this thesis is contrasting simulative processing with linguistic processing (as described above). Thus while both theories are concerned with conceptual processing, Resources focuses specifically on the size of the conceptual elements themselves, while Simulation Theory focuses more on the representation, and processing of those elements. Thus, one could say that resources are cognitive elements that can be used in simulative processing.

2.4. Reconciling Wittgenstein

There might seem to be an irreconcilable tension between the previous chapter and the current chapter. For example, is a resource not a *thing* that one has, which implies the having of a particular kind of knowledge?

To re-iterate, the first chapter is not intended as a positive theory of “what knowledge *truly is*”. (For example, it is not intended to place a moratorium on expressions involving having or possessing knowledge, nor on modeling knowledge as a type of *thing*.) Instead it is intended to provide a framework for what an investigation of *understanding* or *knowledge* might look like. In particular, it attempts to demonstrate that the foundation (the bedrock) of such an investigation is a grammatical foundation, *not* an empirical foundation. In other words, if we wish to look for a type of “source” of meaning for these terms, we are to look to the role they have in our grammar, the use they have in our particular forms of life. It is our distinctly human, messy forms of life that gives substance and traction to these terms, *not* unseen mental entities.⁹

Nevertheless, as scientists we are drawn to model “knowledge”. However, more pertinently, as educators we wish to model the mind of the student. Given a particular set of initial conditions, we wish to predict the behaviour of the system (Reif, 1986). This has been shown to be a productive exercise, as various models of knowledge used in PER have lead to improvements in teaching and learning. However, the emphasis of the first chapter is to warn against interpreting these successes as demonstrating that a “solution” for the notion of “knowledge” or “understanding” has been found. All one can say is that in this particular context (and here we must operationally define “context”), we have found a *useful model* of student understanding/knowledge (and here we must operationally define “understanding”/“knowledge”) of this particular concept (and here we must operationally define “concept”). In other words we have to accept that the meanings of the terms

⁹ Although a well-functioning cognitive system is required for possessing knowledge, that does not mean that knowledge is identical to that system, nor somehow literally contained within that system.

“understanding”, “knowledge”, “concept” and “context” themselves depend on the “concepts”, “contexts” and types of “understanding” which we are interested in.

CHAPTER 3: Vector Addition in Different Contexts

<u>Concept</u>	<u>Context</u>	<u>Model of Understanding</u>
Vector Addition (of forces and displacements (in 2d))	Forces and Displacements	Mostly algorithmic; some embodied simulation
Vector Addition (of momenta (in 2d))	Momentum	Activates resources significantly different to “forces” and “displacements”
Vector Addition (involving total (forces (in 2d)))	Total Force	mostly algorithmic; some embodied simulation of felt experience
Vector Addition (involving net (forces (in 2d)))	Net Force	mostly algorithmic; some embodied simulation of resultant motion
Vector Addition (of forces (on hard ball (in 1d)))	Forces on Hard Ball	Simulation of resultant movement
Vector Addition (of forces (on person (in 1d)))	Forces on Person	Simulation of force experienced

3.1. Introduction

“Vector addition” is a concept. We might also think of it as a type of skill or ability. In either case, it is at the grain size of “vector addition” that we most often ascribe this particular skill or concept to students. It is common to say: “my students can add vectors”. It is not common to speak at a finer grain size, and say: “my students can add force vectors, but they haven’t yet added momentum vectors”. To an expert, the latter sentence might sound something like: “my students can add even numbers, but they have not yet added odd numbers”. Vector addition is typically understood as a unit of knowledge; a concept; a transferable skill. If one knows that momentum is a vector, then one should simply add momenta in the same way one adds forces. Quite Easily Done. This is not what we found in our study. Students who readily admit that momentum is a vector, and who capably add force and displacement vectors, do *not* necessarily transfer this skill to the addition of momentum vectors.

Physics education practice and research are guided by overarching theoretical frameworks. One of the more important frameworks is that of conceptual change. In his paper “A History of Conceptual Change” diSessa (DiSessa, 2006) argues that a major fault line in the conceptual change literature is that of coherence/fragmentation: to what extent are student ideas coherent and integrated, or quasi-independent and fragmented. Toward one end of this axis is the “theory theory”: the idea that students’ concepts cohere in robust (albeit naïve) theories. Toward the other end of the axis is the Knowledge in Pieces or Resources framework developed by diSessa (1993) and Hammer (2000) among others. The theory theory is more compatible with traditional notions of abstraction and transfer, while the knowledge in pieces framework foregrounds the role of context. Indeed, Disessa (2002) argues that context should be a central focus of education research.

This chapter looks at students’ ability to add vectors in various contexts. It begins with a study whose findings might be viewed as classic example of transfer failure, and goes on to examine the effect of further contextual changes on student performance. The results presented here are most productively explained by means of a Knowledge in Pieces or Resources framework.¹⁰

3.2. Literature Review

Students have difficulty with two-dimensional vector addition. After administering a vector knowledge test which included problems involving two dimensional algebraic and graphical vector addition, Knight (1995) concluded that of 286 freshmen only about a third had sufficient vector knowledge to proceed with mechanics. Nguyen & Meltzer (2003) administered a test to 2000 students which focused solely on two-dimensional graphical addition of vectors. More than one quarter of the students in the calculus based course, and more than half of the students in the algebra based course were unable to add vectors satisfactorily. Flores, Kanim, & Kautz (2004) developed a conceptual test of vectors which included one question on two dimensional, graphical vector addition. Only 30% of students added the vectors satisfactorily.

One of the difficulties students face is that vectors appear in a variety of physical and non-physical contexts.¹¹ For example, one may add different physical quantities, such as force and displacement, and the addition may be framed in a graphical or algebraic context. Indeed, all of the above-

¹⁰ It is acknowledged that not all PER researchers support a Resources framework. Nevertheless, I believe the fine-grained nature of the investigation, and the quantitative and qualitative results will support the use of this theoretical framework for these series of studies.

¹¹ Note that although some authors refer to a “pure” mathematical context as “contextless”, I believe this to be misleading. A mathematical context provides meaning and use for the concept of vector addition, in a similar way that a physical context might.

mentioned research flags “physical context” as a salient factor. Knight (1995) observes: “By the far the most common partially correct response [to the question ‘define a vector in your own words’] equated vectors with specific physical quantities (e.g. ‘A vector is a force’)” (p.76). Nguyen & Meltzer (2003) suggest that the displacement vector might be the best physical context in which to teach vector addition. However, in a paper which presents his cumulative knowledge of teaching vectors to students, Roche (1997) observes that “the displacement vector is easiest to explain, perhaps too easy. It is very different from other vectors such as force.” When discussing implications for developing tutorial curricula, Shaffer & McDermott (2005) cite context as being ‘critically important’. These comments can be understood as suggesting that “vector addition” is not a homogeneous concept, and that different physical contexts allow for the activation of different resources, some of which might be a help or a hindrance.

There has been a variety of research comparing vector addition in different contexts. Allie & Buffler (1998) found that students were sensitive to the particular vector quantity being added. Heckler & Scaife (2015) found that students perform better when adding vectors in i, j, k representation than in graphical arrow representation. Shaffer & McDermott (2005) show that students are better able to subtract vectors when no physical context is involved. Van Deventer's (2008) master's thesis compares student understanding of vectors in physics contexts and math contexts. 270 students were given a vector test embedded in either a math context or a physics context. Like Shaffer and McDermott, van Deventer also found that students were significantly better at subtracting vectors in a math context than a physics context. However, he did not find any significant difference when comparing students' two-dimensional addition of vectors in math and physics contexts, although the pattern of incorrect answers did depend on context. (The physics context involved two forces acting on an object). van Deventer did not find any significant difference between student overall performance in the math context and the physics context. An important point to note is that van Deventer's questionnaire involving a “physics context” includes not only different vector *operations* (such as addition, subtraction, and multiplication), it also includes a variety of physics contexts involving vectors. These physics contexts included force (tension, normal, and torque), magnetic field, electric field, velocity, acceleration, and displacement. It should be clear by the end of this literature review that student performance is sensitive both to the particular vector operation, and to the particular physical context. Thus, we would argue that van Deventer's notion of “physics context” occurs at an overlarge grain size. Barniol & Zavala (2010, 2014a, 2014b) conducted a similar series of studies comparing student performance on vector operations in math and physics contexts. Their questionnaire involved ten different activities, including “finding the magnitude”, “finding the direction”, vector addition, subtraction and multiplication. They also found no overall significant

difference between student performance in the math and physics contexts. However, like van Deventer (2008), they did find significant differences on individual items. For example, for the question involving the dot product of vectors, students performed significantly better in the physics context. They argue that this is because in the context of work (the physics context) students have an additional resource on which to draw, which helps them in answering the question. Regarding two dimensional vector addition, their findings were identical to Van Deventer (2008); there was no significant difference when comparing math and physics contexts, although the pattern of incorrect answers did depend on context.

The literature does not include much discussion of hypotheses that might describe or predict systematic sensitivities to context. Almost all of the above research includes categories of common errors. For example, in their question on vector addition Barniol & Zavala (2014) cite four common errors: “tip-to-tip”; “closing the loop”; “horizontal bisector” and “general bisector”. It is also common for researchers to discuss the misapplication of algorithms, such as the Pythagorean method, or the “head to tail” method. In comparing student performance in math and physics contexts, Barniol & Zavala (2014a) do venture a theory for why the physical context of work ($W = \vec{F} \cdot \vec{d}$) aided students in their calculation of the dot product: “The work context helps the students in the conceptual problems, since in this context students have more resources; i.e., work allows students to have an “extra” physical interpretation that is ‘close’ to them.” (p.13). In order to explain why this “extra physical interpretation” is not observed to benefit students in the physical contexts of velocity and force, they argue that “in these contexts [of velocity and force] the physical and mathematical interpretations are so closely related that students do not receive any extra physical interpretation”. (p.14).

In summary: Previous research demonstrates that when investigating the context sensitivity of students’ understanding of vectors, it is not helpful to work at the overlarge grain size of “vectors”, “physics” and “math”. Researchers should be cognizant of at least three important ways in which the research topic can be divided up: (i) The umbrella term “vectors” can include a variety of different activities, and these different activities seem to activate very different cognitive resources. Compare finding the direction of a cross-product using the right hand rule, to performing two-dimensional vector subtraction. (ii) The “physics context” includes a variety of ways in which vectors can be represented (as force, velocity, etc.) which also seem to activate different cognitive resources. Compare the addition of two forces acting on an object, to walking two successive displacements and determining the resultant displacement. (iii) Vectors can be represented in algebraic or graphical form. These representations may be used in the question or requested in the answer.

Our focus in this study is two dimensional vector addition. Van Deventer et al. (2007), Van Deventer (2008) and Barniol & Zavala (2010, 2014a, 2014b) concluded that student performance in two dimensional addition of vectors did not differ between math contexts and physics contexts. The physics context used by all researchers was that of “force”. In our study we investigate the difference between student performance in vector addition, in the physical contexts of force, displacement and momentum.

3.3. Overview

This chapter on students’ understanding of vector addition consists of four separate studies. (i) The first considers the three different physical contexts of force, displacement and momentum. (ii) The second looks at the different textual contexts of “total force” vs. “net force” and “total momentum” vs. “net momentum”. (iii) The third considers the textual contexts of “total force” vs. “resultant force”, and “momentum of the *two balls*” vs. “momentum of the *two ball system*”. (iv) The fourth compares the contexts of forces acting on a person (“squishy” animate object), and forces acting on a marble (hard, inanimate object). I shall briefly give an overview of the main findings from the studies, and how they followed on from each other. All studies were conducted on the new intake of about 220 freshmen medical students who take a semester course in physics. All studies involve a questionnaire, and in two of the studies we conducted follow-up interviews. The questionnaires included forced choice responses (multiple choice) for quantitative statistical analysis, and free writing responses which were qualitatively analysed by a method suggested by grounded theory (Strauss & Corbin, 2007). The vector addition problems involved two vectors acting at right angles to each other, and were isomorphic in each context. (i) The first questionnaire asked students to add forces, displacements and momenta, and involved only forced choice responses. The results clearly showed that students performed significantly worse when adding momenta than when adding forces and displacements. Note that students only encounter two-dimensional addition of forces and displacements (but not momentum) at high school. In follow up interviews, students who had incorrectly answered the question on momenta, indicated that even though they knew force and momentum were both vectors, they understood momentum to be very different from force, and therefore used a different strategy of addition. One might say that “momentum” activated a different set of resources for these students. The interviews also revealed that students seemed to be sensitive to the textual prompts of “total” and “net”, as well as “add the momenta of the *two balls*” and “add the momenta of the *two ball system*”. This context sensitivity seemed grounded in an embodied simulation of the scenario – “total” caused students to simulate the total force experienced (e.g. the total “punch” experienced) while “net” caused students to simulate of the net

effect on the motion of the object. (ii) Our second study pursued the possible sensitivity to the textual contexts of “total” and “net”. Half the class were asked to determine the net vector quantity (i.e. net force/displacement/momentum) and the other half were asked to determine the total vector quantity. Only in the context of momentum (i.e. in comparing those students who were asked to determine “net momentum” with those asked to determine “total momentum”) did we observe a significant difference. However, we included a question which explicitly asked students whether they thought “net force” was different to “total force”. 74% of students claimed there is a difference. We then posed a free response question in which they were asked to explain why they were different. About two thirds of the respondents used some type of fairly shallow algorithmic reasoning (e.g. citing pythagoras’ theorem), while one third reasoned using a type of embodied simulation. The latter group equated “total force” with the total “felt” force (or pressure), while “net force” was associated with the resultant direction of movement. Follow up interviews were conducted in which students were asked to simulate the questions using a “demo kit”. A demo kit consists of small objects such as bottle tops and elastic bands. Similar reasoning to that found in the Free Writing Responses emerged. (iii) The third study also split the class in half – one half were given the textual prompts: “total force” and “total momentum of the *two balls*”. The other half were the textual prompts: “resultant force” and “total momentum of the *two ball system*”. No significant differences were observed. However, we then explicitly asked students whether they believed these two textual prompts referred to different scenarios. 74% of students claimed that “resultant force” was different to “total force” and 65% of students claimed that “total momentum of the *two balls*” was different to “total momentum of the *two ball system*”. (iv) Our previous three studies indicated that a significant proportion of students performed different embodied simulations for the notions of “total” and “net” force. We conducted one final study to see whether we could explicitly prompt these simulations. Students were given isomorphic questions involving two forces acting on a marble, and two forces acting on them (a human body). They were asked to find the total force on the body, and the net force on the marble, and then asked to explain why they chose the same, or different answers to the questions. We found a significant difference between their calculation of the force on the body and on the marble, and their free written responses supported the previous studies’ findings. Students who added the vectors in a scalar manner tended to mention the felt experience of the forces, while students who added forces vectorially, tended to mention the resultant motion of the object.

3.4. Study 1: Vector Addition in the Context of Different Physical Quantities

3.4.1. Sample

The sample comprised 227 freshmen enrolled for a medical degree at the University of Cape Town (UCT). The students would have encountered the two dimensional addition of forces and displacements at high school, but would have been exposed to adding momenta in only one dimension. The questionnaire was given at the beginning of the university year, prior to instruction.

3.4.2. Methodology for Questionnaire

This exploratory multiple choice questionnaire was administered in February 2013. The questionnaire was answered under exam conditions but without time pressure. The first question asked students to categorize various physical quantities (force, displacement and momentum) as being vectors or scalars. The remaining three questions involved quantitative vector addition. Students were able to select a number from 0 – 11 as their answer. For the sake of brevity, these multiple choice answers are only included for question 4 below. The following instructions appeared at the top of the sheet:

In each of the problems below circle the answer that you think is correct. Where the answer is a vector you need only select the magnitude.

2. *Two forces are applied to a box. Force 1 has a magnitude of 3 Newtons and acts East while force 2 has a magnitude of 4 N and acts North. The total force on the box is:*

3. *A student walks 2 m West, then 5 m East and finally 4 m North. The final displacement of student is:*

4. *Ball 1 has a mass of 1.5 kg and a velocity of 2ms^{-1} North while ball 2 has a mass of 1 kg and a velocity of 4ms^{-1} East . The total momentum of the two balls is:*

0 1 2 3 4 5 6 7 8 9 10 11 kgms^{-1}

Note that questions 2, 3 and 4 all require the same calculation of the hypotenuse of a 3-4-5 triangle to give a correct answer of 5.

3.4.3. Findings

The vast majority of students correctly identified force, displacement and momentum as vector quantities (94%, 93% and 88% respectively). However, they were not as successful when adding these various quantities.

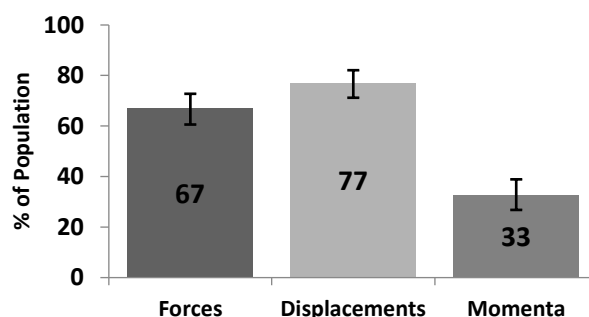


Figure 5: Percentage of students to correctly add vectors in different contexts ($n=227$).

Figure 5 shows the results for questions 2 – 4. Even though students correctly identified force, displacement and momentum as vector quantities, their success at vector addition is sensitive to the context of the particular vector quantity being added ($p < 0.001$, Cochran's q test). Figure 6 shows that students' responses clustered around the multiple choice values of 5 and 7.

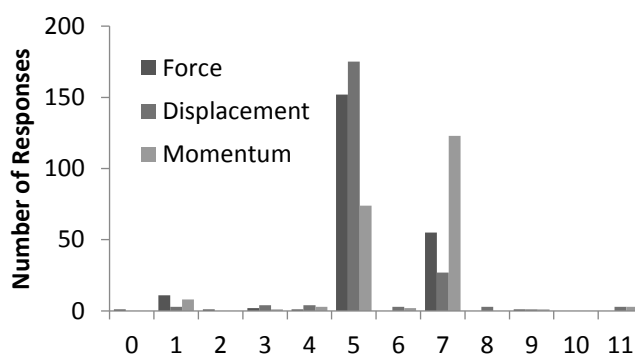


Figure 6: Distribution of answers for the addition of various vector quantities ($n=227$)

3.4.4. Methodology for Follow Up Interviews

Follow up interviews were conducted with eight students. These students were selected for having drawn the correct vector diagram (i.e. a 3-4-5 triangle) on their answer sheets but choosing the incorrect answer of 7 N and/or 7 kgms^{-1} . The interviews occurred 8 months after they had answered the questionnaire during which time they had completed one semester of physics 1 instruction. The

students were given their previous answer sheet, asked to look over their previous answers, and to explain why they might agree or disagree with any of answers they had given.

The interview data was analyzed using an approach suggested by Grounded Theory (Strauss & Corbin, 2007). General patterns were identified within students' responses. Three of these general patterns are described below.

3.4.5. Findings

We summarize our three main findings below.

1. Adding momenta is different to adding forces because momentum is different to force.

Students often cited their lack of knowledge of momentum as a reason for their difficulty with adding momenta. However, from an expert's point of view, they had all the relevant information they required, as they had all correctly indicated that momentum was a vector quantity. Three of the students acknowledged that they had initially seen the momentum question as being similar to the force question, but that some aspect of the question had subsequently led them to believe that different addition strategies were required: "...basically these two questions [q2 and q4] are the same, like to me they look the same... but when I see total momentum and I see total force, my thinking changes."

2. Students have different conceptions of "total" and "net".

Note that the questions asked for the *total* force and the *total* momentum. Some students claimed that it was the word "total" that caused them to sum the magnitudes algebraically and that they would have answered differently if they had been asked for the net quantity. On questioning them further, four of the students explicitly identified a difference between a *total* vector quantity and a *net* or *resultant* vector quantity. On trying to disentangle the notions of "total" and "net" some students gave a procedural distinction: "whenever I hear "net" I just automatically go for the calculation where you use the triangle..." Other students provided a more robust conceptual distinction: "I do agree that there is 7 newtons applied to the box, but, it's a 5 newton... result. ... Just think, if I was getting punched with 3 newtons and 4 newtons [gestures throwing punches] I'd have 7 newtons pressed on me... in total!"

3. The notion of a system: students have robust ideas of separateness and togetherness

While correctly revising their answers for the momentum question, three students changed their minds and reverted back to their incorrect answer because of the notion of "separateness":

"...I think what I did was took momentum as a scalar quantity and just added the two... whereas I think what should happen is... [inaudible]... no...I think its fine...because the two balls are unrelated

and moving in different directions... you can't sort of have the resultant momentum between the two."

When prompted, all students admitted that if the question had read: "Find the momentum of the two ball *system*" instead of "Find the momentum of the two balls", they would have given the correct answer of 5 kgms^{-1} .

3.5. Study 2: Vector Addition in Different Textual Contexts - "Total" and "Net"

In response to the results from our interviews we conducted a further study to test whether students are sensitive to the textual contexts of "total" and "net" when finding a **total** vector quantity and a **net** vector quantity.

3.5.1. Sample

The sample of students comprised 209 freshmen enrolled for the same medical degree at the University of Cape Town (UCT) as the cohort from Study 1. The students would have encountered the two-dimensional addition of forces and displacements at high school, but would have been exposed to adding momenta in only one dimension. The questionnaire was given at the beginning of the university year, prior to instruction.

3.5.2. Methodology: Part 1 (Forced Choice Responses)

This revised questionnaire was given to the new cohort of pre-medical freshmen in February 2014. Approximately half of the class received the identical multiple choice questions as the 2013 cohort, which is provided below for convenience:

In each of the problems below circle the answer that you think is correct. Where the answer is a vector you need only select the magnitude.

2. *Two forces are applied to a box. Force 1 has a magnitude of 3 Newtons and acts East while force 2 has a magnitude of 4 N and acts North. The total force on the box is:*

3. *A student walks 2 m West, then 5 m East and finally 4 m North. The final displacement of student is:*

4. *Ball 1 has a mass of 1.5 kg and a velocity of 2 ms^{-1} North while ball 2 has a mass of 1 kg and a velocity of 4 ms^{-1} East. The total momentum of the two balls is:*

0 1 2 3 4 5 6 7 8 9 10 11 kgms^{-1}

The other half received similar questions, but with three small changes: Instead of being asked for the **total** force, **total** momentum and **total** displacement, they were asked for the **net** force, **net** momentum and **net** displacement.

3.5.3. Findings

The light grey columns in Figure 7 represent the proportion of correct answers for “total force”, “total displacement” and “total momentum”. The dark grey columns are the proportion of correct answers for the equivalent net vector quantities. As in Figure 5, Figure 7 demonstrates that students perform differently depending on the type of vector quantity being added. ($p < 0.001$, Cochran’s q test for both “total” and “net” groups.¹²)

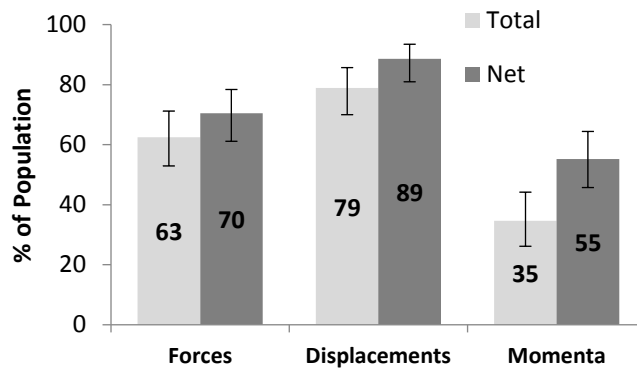


Figure 7: Percentage of students to correctly add vectors in different contexts.

Regarding the effect of the textual prompts “total” “net”, three χ^2 tests of association were performed for forces, displacements and momenta (i.e. comparing the light and dark grey bars) and yielded p-values of $p = 0.22$, $p = 0.57$ and $p = 0.003$ respectively. After adjusting for a Bonferoni correction at the 95% confident level ($\alpha = 0.017$), we determined that the textual prompts only resulted in a significant difference in the case of adding momenta.

¹² Note that McNemar’s test for paired data (with Bonferoni correction), yielded a significant difference between Displacements&Momenta, and Forces&Momenta, but not between Forces&Displacements, for both dark grey (total) and light grey (net) columns.

3.5.4. Methodology: Part 2 (Forced Choice and Free Writing Responses)

After the initial multiple choice questions, the questionnaire presented all students with the following scenario (hereafter known as *Scenario F*):

As in question 1 above: Two forces are applied to a box. Force 1 has a magnitude of 3 N and acts East while force 2 has a magnitude of 4 N and acts North. Three students have the following conversation:

Student A: "The magnitude of the total force on the box is 7 N, and the magnitude of the net force on the box is 7 N."

Student B: "I agree that the magnitude of the total force on the box is 7 N, but the magnitude of the net force on the box is 5 N."

Student C: "The magnitude of the total force on the box is 5 N and the magnitude of the net force on the box is 5 N."

(i) With whom do you most closely agree? (Circle your choice). A B C

(ii) Explain why you agree with that student:

They were then given an identical scenario where three posited students are discussing possible values for *total momentum* and *net momentum*. (Hereafter known as *Scenario M*).

Using an approach suggestion by Grounded Theory (Strauss & Corbin, 2007), broad patterns of response were identified in students' reasoning for distinguishing between "total force" and "net force". We initially coded for over 20 different categories and subcategories (see Appendix B). We then refined our coding procedure to six categories of interest. These are given below.

One of the recurrent patterns of response was the use of visualisation and simulation. This type of response was eventually coded under the category "simulation". As mentioned in section 2.3, identifying the simulative character of students' approaches to a question from their written response involved coding their response as lying sufficiently "high-up" on the spectrum of simulation. There are no "hard and fast" rules for coding a response as being simulative in character. However, as one can see from the examples provided below, evidence of visualization or simulation is readily identifiable, and contrasts markedly with responses coded as "Procedural".

3.5.5. Findings

After being presented with Scenario F, 82% of students claimed that “total force” was different to “net force” (i.e. they agreed with Student B). 12% correctly identified “total force” and “net force” as having the same value of 5 N (i.e. agreed with Student C). And 6 % agreed with Student A, and incorrectly thought that both “total force” and “net force” had a value of 7 N. This led to a radical change in the distribution of correct answers. Figure 8 shows that 40 of the 65¹³ students in the “total” cohort who had added forces correctly, altered their answer to the incorrect value of 7 N after being exposed to scenario F. This left only 15 students (14%) with the correct answer of 5 N. (Only one student altered his answer from the incorrect value of 7N, to the correct value of 5 N). On the other hand, after being exposed to scenario F, 24 of the 30 students in the “net” cohort who had incorrectly answered the question on forces, changed their answer to the correct value of 5N. Thus, post scenario F, there is an obvious significant difference in students’ performance in the “net” and “total” cohorts. ($p < 0.001, \chi^2$ test).

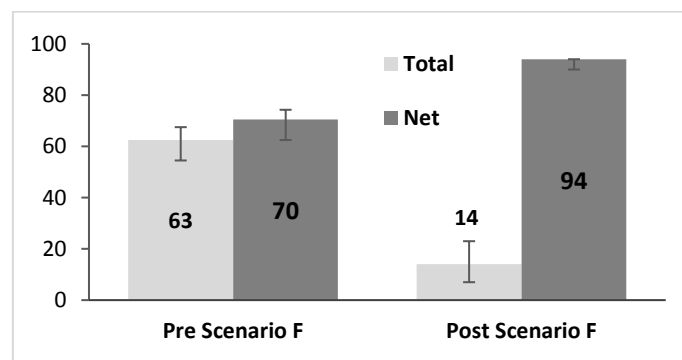


Figure 8: Comparison of Percentage of Students to Give Correct Answer for Forces Addition Before and After Being Exposed to Scenario F

A similar effect was observed post Scenario M. 55% of students thought that “total momentum” and “net momentum” were different (i.e. they agreed with student B). Only 10% correctly identified “total momentum” and “net momentum” as having a value of 5 $kgms^{-1}$. An unexpectedly large proportion of 30% incorrectly affirmed that both quantities have a value of 7 $kgms^{-1}$. (5% of students left this question blank). As shown in Figure 9, 40% of students in the “total” cohort who had answered the questions correctly, altered their answer to the incorrect value of 7 $kgms^{-1}$. This led to a drop of 25 percentage points (35% to 10%) in those who answered correctly. Interesting,

¹³ The 65 students are the 63% seen in Figure 8.

the number of students in the “net cohort” to alter their answer lead to only a 9 percentage point increase in the choice of correct answer.

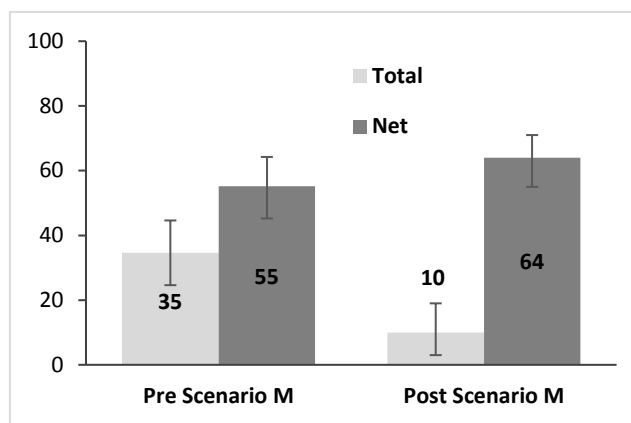


Figure 9: Comparison of Percentage of Students to Give Correct Answer for Momenta Addition Before and After Being Exposed to Scenario M

Now we turn to the reasons students gave for making such radical changes to their forced choice responses. The findings from the free writing responses are summarized below. They are organised into six categories, including the number of student responses to fall into each category, and a few exemplars from that category.

Simulation: Effect, Result, Movement and Experience [26% of Students (i.e. 54 students)]

Just over a quarter of students were categorized as using a type of simulation in their explanation of why total force differed from net force. Examples include:

- The total force being applied to the box is 7N because it is experiencing 3N from an easterly direction and 4N from a Northerly direction. (123T)¹⁴
- I also realized that the box is experiencing a total force of 7N (the 4N north and 3N east). (112N)¹⁵
- The box experiences a total force of 7N because this is the sum of the 2 forces acting on it. (164N)
- The amount of force the box experiences or that are acting on it add up to 7N. (196T)
- The magnitude of the total force is 7N because there is a 3N and 4N forces acting on the box, however the net force is 5N because the 2 forces acting on the box pushes it in a north east direction (123T)

¹⁴ This is anonymized student 123 from the Total Cohort

¹⁵ This is anonymized student 112 from the Net Cohort

- *The net force is 5N because the box is being pushed/pulled 3N east and 4N north, this will change the direction of the ball to approximately North-North-East by that time the ball experiences 5N (167N)*
- *the net force indicates the force resulting in the movement of the box... (119N)*
- *The net force is the resultant force on the box that determines the correct magnitude and direction the box will actually move. (136N)*
- *The net force is an indication of the effect the combined forces will have on the box (176T)*

Ontology of Net Force [17% of students (i.e. 39 Students)]

Some students drew the distinction between “total” and “net” by positing the existence of a “new”, “overall” force.

- *because they act in different directions they cannot be added - instead they must be combined to calculate the force at an angle, a "new" overall force. (121T)*
- *the two forces act concurrently to create one net force. (116T)*
- *In order to get the net force, the diagonal force applied needs to be calculated which equals to 5N. (173N)*
- *The term “total” refers to the sum of the forces acting on the box, whereas the magnitude of the 'net force' refers to the magnitude of the one force that could replace all other forces acting on the box and yet produce the same result (184N)*

Explicit Procedural [30% of students (i.e. 63 students)]

About a third of students drew the distinction by the procedure used in calculating “total” and “net”. This often included mentioning Pythagoras’ Theorem. Note that we do *not* believe this category to be equivalent to *shallow (algorithmic) reasoning*. See Discussion below.

- *Force 1 and Force 2 add up to 7, so the magnitude of the total force on the box is 7N. To find the magnitude of the net force you would have to use the theorem of Pythagoras:
 $3^2 + 4^2 = 5^2$. Since the forces make a right angled triangle. (108N)*
- *Total force can be calculated by simple addition of all unbalanced forces. i.e. $3+4=7$. Net force is calculated differently and amounts for 5N (126N)*
- *The total force is equal to the sum of the two forces, independent of direction. however, the net force depends on the direction in which the forces act and will be calculated as follows:
 $x^2 + y^2 = r^2$ (138N)*

Direction [40% of students (i.e. 84 students)]

Approximately 40% of students stated that the difference between total force and net force resides in the fact that net force accounts for direction, while total force does not. Examples include:

- *The total force exerted on the box or any object doesn't consider in which direction they are exerted. (176)*
- *The total amount of force, disregarding direction, is 7N (119)*
- *However the net force is still 5 N as this is the smallest combination of the forces and their direction is taken into account (125)*
- *Force 1 and force 2 are acting on the box ∴ total F on box is 7N. but they are in different directions so resultant force (using pythag) is 5N (127)*
- *The total force on box is 7N, because of the 3N and 4N force, but because they're in different directions the magnitude of the net force is only 5N. (135)*

Total is "Just Adding", Net is Something Else [62% of students (i.e. 130 students)]

Over half the students explicitly mentioned that in order to find the total force you simply add the two forces together.

- *The total force is 7N as you just add the 2 forces being applied. (137N)*
- *I agree with student B because technically there is 7N total force acting on the object the 4N force and the 3N force. If one had to simply total the forces.*
- *The total force is 7N is a simple process of adding together all the forces acting on the object to calculate their magnitude.*

A handful of students also fell under other interesting categories such as:

Force is Lost: *"7N however is not the net force on the box due to some forces being lost since there are two directions involved. (146T)"*

Some Mention of Opposition: *"While force of 3N and 4N are acting on the box, neither force is opposing the other, thus the total force is 7N and the net force is 7N."*

3.5.6. Methodology: Part 3 (Follow Up Interviews)

Follow up interviews were conducted with eight students. The interviews occurred after they had completed 3 months of physics instruction in mechanics. When being probed about the question involving the two forces being exerted on the box, a small demo kit was used. A red bottle cap represented the box, and two stones, one white and one green, represented the 3N force and 4N force respectively. A brown stone represented the resultant 5N force. Using the demo kit, we represented two scenarios: Scenario A, in which the white and green stones (i.e. the 3N and 4N forces) acted at right angles on the bottle cap (i.e. the box). And Scenario B, in which the brown stone (the equivalent 5N resultant force) acted on the bottle cap at an appropriate angle.

3.5.7. Findings

The students agreed that these two scenarios (Scenario A and Scenario B) were essentially the same: the bottle cap (i.e. the box) would have the same acceleration in both conditions. However, we then asked two students if Scenario A and Scenario B would be equivalent, if the bottle cap represented *them*, instead of the box. A small portion of the interview transcripts are included below. The first excerpt is an interview with Gareth¹⁶. Note how Gareth's focus switches from the experience of the resultant motion, to the tactile experience (pressure) of the forces being exerted on him.

Interviewer: *What if we put Gareth here. Would you say it was the same or different [in Scenario A compared with Scenario B]?*

Gareth: *I suppose it would feel the same, you would move the same amount. It would feel the same I suppose.*

Interviewer: *... how would it feel the same?*

Gareth: *You would still move the same...*

[After discovering that Gareth plays soccer, I give him the same scenario, but the white and green stones are represented by soccer players that are performing 'body checks' on him].

Interviewer: *Would that [having two players body check you] feel the same if one guy body checked you with 5 N.*

Gareth: *no it wouldn't. Because, this one there's two guys putting force on you... and there only one guy.*

Interviewer: *... what were you focusing on [before]*

Gareth: *I was focusing on the movement of this [points at the bottle cap, which represents him]...*

Interviewer: *When I talk about the body check example, why does your thinking change? ...*

Gareth: *Then I'm thinking about what it feels like to have a force pressed on me... whereas the first time I was thinking about the result of the forces.*

The second excerpt is from an interview with Pimtsi. Note that when she considers being pushed by the two forces (Scenario A), she thinks this will produce a different resultant motion than if she was just pushed by the one (resultant) force (Scenario B).

¹⁶ All names have been changed for anonymity.

KP: 'Cause the thing is, with these two forces [gestures at white stone (3N force) and green stone (4N force)], wouldn't it happen like, somehow they'd restrict me from moving because this one is causing me to move this way, that one is causing me to move that way [gesturing]... so I'd go that way... and... ja... I don't think I'd move the same distance as this one [gestures to the brown stone (resultant 5 N force)] ... this one [the 5N force] keeps heading in one direction...

Interviewer: *Very cool Pimtsi. So my last question is going to be... if there was a box, would the box move differently or the same [in Condition A versus Condition B]*

PK: It would move the same.

Interviewer: *But you're saying that you would move differently [in Condition A versus Condition B]?*

KP: [Very Long Pause]... Ja, ja, I'd move differently I think.

Finally, three other students, Buhle, Jason, and Serje were asked what use the concept of "total force" might have in daily life. These were their answers:

Buhle: *"...total force, let's see. I guess you can calculate... a lot of things... like how much damage that is being done to a car"*

Jason: *"the total force I'd use maybe to measure... how much force it takes to compress a can."*

Serje: *"the total force tells you how much compression an object can withstand."*

3.5.8. Discussion

This study confirmed the 2013 findings that students exhibit context sensitivity when adding vectors in the contexts of forces, displacements, and momenta. In particular, students perform significantly worse when adding momenta than when adding either forces or displacements.

A new finding from this study is the observation of a context sensitivity to the textual prompts of "total" and "net". This sensitivity was observed in the spontaneous response to the questions on "total momentum" and "net momentum". Students answering the question on "net momentum" performed significantly better than those answering the question on "total momentum". A possible reason for why this spontaneous sensitivity was not observed for the questions involving forces and displacements is because students are already familiar with these contexts from high school. Research in cognitive psychology on topics such as perceptual habituation (Johnston & Hawley, 1994) and registration without learning (D. Hintzman, Curran, & Oppy, 1992) suggest that "once the cognitive system knows enough about an object to deal with it effectively, the system may resist wasting resources by analysing the object further." (D. L. Hintzman & Curran, 1995, p.223). In other

words, the mind is biased toward the top-down processing (or conceptual processing) of familiar situations. Note that top-down processing is the foundation of the theory of framing (Whitney, 2002). Thus, when students encountered the vector addition problems involving forces and displacements, they are likely to have skimmed over the details of the question, and framed the questions as requiring a similar strategy to the many other similar questions they had previously encountered at high school. Furthermore, a large body of education research suggests that children (and adults) often resist changing their well-established knowledge of objects, concepts, or procedures (e.g., Duncker, 1945; Mack, 1995; McNeil & Alibali, 2005). In other words, students would be reluctant to change their established strategies for adding forces and displacements in the face of an unfamiliar textual prompt such as “total”. To re-iterate – our hypothesis for why a spontaneous sensitivity to the textual prompts of “total” and “net” was only observed in the vector addition of momentum is because students had not previously encountered such a scenario. They did not have a well-established procedure for this type of scenario. This is likely to have forced them to consider the details of the question, making the textual prompts of “total” and “net” more conspicuous.

Another striking finding of the study is that, after prompting, 82% of students agreed that “total force” is different to “net force”. The fact that this was only observed after explicitly suggesting that such a distinction might exist, suggests that this may be another framing effect. Recall that Scenario F involves the explicit introduction of the terms “total force” and “net force”, as well as the suggestion that these might be different concepts. If the students framed the questionnaire as a typical physics test, they might think it strange to include a discussion question about two concepts, if there was no difference between them. On the other hand, as suggested above, students’ familiarity with the contexts of force and displacement means that they might not have engaged very deeply with a possible distinction between “total force” and “net force” without being explicitly prompted to do so. One might therefore argue that students held a *latent* belief in the distinction between “total” and “net”, but required an opportunity for deep reflection on the matter, in order to bring this belief to light. Whatever reasoning one might ascribe to students distinguishing between “total force” and “net force”, one thing is clear: The fact that students could easily *justify* a distinction between “total force” and “net force” indicates that students have resources readily available to make this distinction.

The set of resources we are particularly interested in are those that involve embodied simulation. However, it is important to note that the majority of students did *not* reason using any type of embodied simulation. Many students referred to a procedural distinction: Approximately two thirds of students reasoned that total force is calculated using simple addition ($3\text{ N} + 4\text{ N} = 7\text{ N}$) while

one third of students explicitly mentioned Pythagoras' theorem as a means for calculating the net force. The use of "simple addition" seems like an instance of the *availability heuristic* (Kahneman, 2011) - answering a question by accessing whatever is readily available. The simple addition of $3\text{ N} + 4\text{ N} = 7\text{ N}$ is a readily available method of combining these numbers to produce a result different from the net force of 5 N . Furthermore, the textual prompt "total" arguably activates ideas of "totaling" or "adding together". It is tempting to identify students who pointed only to procedural differences, or who used only mathematical reasoning, as "shallow reasoners". For example: *"Total force is measured by adding the forces together. 3+4=7. Total force = 7N." (144N)*. However, the more important question might be *why* students used this type of reasoning. Kuo's (2013) thesis explores the notion of epistemological framing (Hammer et al., 2005). i.e. how students answer the question: "what kinds of knowledge or reasoning are appropriate here?". Kuo (2013) demonstrates that students not only have access to a wide range of epistemological resources (Hammer & Elby, 2002), but that their mathematical reasoning can also be associated with deep processing. (See Sherin's (2001) notion of *symbolic forms*). In other words, it is possible that students who used shallow reasoning did so because they framed the situation as requiring a response involving shallow reasoning. Therefore it is not the *students* who are necessarily "shallow reasoners", but the situation (e.g. the particular wording of the question) which prompted shallow reasoning. For example, three students who had largely used mathematical reasoning in their written answers in the questionnaire, reasoned using embodied simulation in their interviews. Furthermore, in Study 4: Vector Addition in Different Physical Contexts below, we use an alternative framing of the "addition of forces" problem which prompted less mathematical and procedural reasoning, and more embodied simulation.

Finally, the interviews also revealed that students may be sensitive to the contexts of "forces exerted on a box" and "forces exerted on a person". A possible reason for this sensitivity is hinted at in Gareth's interview, where his focus shifts from the resultant motion of the body, to the felt experience of the forces. As mentioned above, 27% of responses were coded as representing embodied simulation. After re-analysing these responses, it became quite clear that students were simulating either the resultant motion of the box, or the "felt experience" of the box. Compare these two groups of responses:

Group 1

"The net force is 5N because the box is being pushed/pulled 3N east and 4N north, this will change the direction of the ball to approximately North-North-East by that time the ball experiences 5N." (142T)

"...the net force refers to the force experienced in the direction of motion. The box is being pulled north and eastwards simultaneously so it will be displaced in a northeasterly direction." (211T)

"Therefore the box will experience a net force in one direction only." (213N)

Group 2

"The total force being applied to the box is 7N because it is experiencing 3N from an easterly direction and 4N from a Northerly direction." (201N)

"I also realized that the box is experiencing a total force of 7N (the 4N north and 3N east)" (104T)

"The box experiences a total force of 7N because this is the sum of the 2 forces acting on it." (167T)

This distinction between simulating the felt experience, and the resultant motion squares well with cognitive science research which indicates that "when we hear or read about objects, we mentally simulate them from the perspective of someone actually experiencing the scene – not from a God's-eye view..." (Bergen, 2012, p.71). In his paper on model development, Hestenes states: "The object description requires a decision as to the type of model to be developed. For example, a given solid object could be modelled as a *material particle*, a *rigid body*, or an *elastic solid*." (David Hestenes, 1987) The data suggest that "net force" might prompt students to model the box as a point particle, while "total force" might prompt students to model the box as an elastic solid, capable of experiencing pressure. Study 4 in Section 3.7 explores this in more detail.

3.5.9. Conclusion

The data presented here present a strong case that when students perform the algebraic, two dimensional addition of vector quantities they are sensitive both to (a) the type of vector quantity being added (force, displacement, momentum) and (b) to the textual prompts of "total" and "net". Furthermore, the latter sensitivity is particularly apparent if the students are presented with (i) a vector quantity with which they are not very familiar (i.e. momentum) or (ii) if it is suggested that there might be a difference between the notions of "total" and "net". Students can readily provide reasons for distinguishing between "total force" and "net force". Most of this reasoning was found to be procedural and/or mathematical. However a significant proportion of students reasoned using embodied simulation. Their simulations could be categorized as focusing either on (i) the resultant motion of the box, or (ii) the "felt experience" of the box.

3.6. Study 3: Vector Addition in Different Textual Contexts – “Total” and “Resultant”.

3.6.1. Introduction

Our previous studies have shown that students are sensitive both to the type of vector quantity being added (force, displacement, momentum) and to the textual prompts “total” and “net”. “Total” tended to prompt an arithmetic sum, and “net” a vector sum. The present study explores additional aspects of this sensitivity to textual context, through the use of multiple choice questions, and free writing. The findings of the present study indicate that students are also sensitive to the textual prompts, “resultant” and “system”. We suggest that embodied cognition may be a helpful lens for describing why students exhibit these context sensitivities.

3.6.2. Sample

The sample of students comprised 205 freshmen enrolled for the same medical degree at the University of Cape Town (UCT) as in Studies 1 and 2. The high school curriculum of these students includes the two dimensional addition of forces and displacements, but only one dimensional addition of momenta. The questionnaire was administered at the beginning of the university year, prior to instruction.

3.6.3. Methodology: Forced Choice Responses Part I

A multiple choice questionnaire (the same questionnaire used in study 2, but with minor changes) was given to 205 pre-medical freshmen. The first question asked students to categorize various physical quantities (e.g. force, displacement and momentum) as being vectors or scalars. The remaining three questions involved quantitative vector addition. Approximately half the class received the questions listed below. This group was designated the control group. The other half of the class received similar questions, but with three small changes: Instead of being asked for the **total** force and **total** displacement, they were asked for the **resultant** force, and **resultant** displacement. In addition, instead of being asked for the total momentum of the **two balls**, they were asked for the total momentum of **the two ball system**. This second cohort constituted the experimental group. Students were able to select a number from 0 – 11 as their answer. For convenience, the questions for the control group are shown below:

In each of the problems below circle the answer that you think is correct. Where the answer is a vector you need only select the magnitude.

2. Two forces are applied to a box. Force 1 has a magnitude of 3 Newtons and acts East while force 2 has a magnitude of 4 N and acts North. The total force on the box is:

3. A student walks 2 m West, then 5 m East and finally 4 m North. The total displacement of student is:

4. Ball 1 has a momentum of 3kgms^{-1} North while ball 2 has a momentum of 4kgms^{-1} East . The total momentum of the two balls is:

0 1 2 3 4 5 6 7 8 9 10 11 [kgms^{-1}]

Note that questions 2, 3 and 4 all require the same calculation of the hypotenuse of a 3-4-5 triangle to give a correct answer of 5.

After answering the above questions, the questionnaire presented the students with the following scenario (hereafter, *Scenario F*). This is the same as the *Scenario F* in Study 2, except that “net force” has been replaced with “resultant force”.

As in question 2 above: Two forces are applied to a box. Force 1 has a magnitude of 3 N and acts East while force 2 has a magnitude of 4 N and acts North. Three students have the following conversation:

Student A: “The magnitude of the total force on the box is 7 N, and the magnitude of the resultant force on the box is 7 N.”

Student B: “I agree that the magnitude of the total force on the box is 7 N, but the magnitude of the resultant force on the box is 5 N.”

Student C: “The magnitude of the total force on the box is 5 N and the magnitude of the resultant force on the box is 5 N.”

(i) With whom do you most closely agree? (Circle your choice). A B C

(ii) Explain why you agree with that student:

And, finally, the students were presented with this scenario (hereafter, *Scenario M*).

As in question 4 above: Ball 1 has a momentum of 3kgms^{-1} North while ball 2 has a momentum of 4kgms^{-1} East. Three students have the following conversation:

Student A: “The magnitude of the total momentum of the two balls is 7kgms^{-1} and the magnitude of the total momentum of the two ball system is 7kgms^{-1} ”.

Student B: “I agree that the magnitude of the total momentum of the two balls is 7 kgms^{-1} , but the magnitude of the total momentum of the two ball system is 5 kgms^{-1} ”.

Student C: “The magnitude of the total momentum of the two balls is 5 kgms^{-1} and the magnitude of the total momentum of the two ball system is 5 kgms^{-1} ”.

(i) With whom do you most closely agree? (Circle your choice). A B C

(ii) Explain why you agree with that student:

Students were not permitted to “turn back” and change their answers to previous questions.

3.6.4. Findings

Almost all students correctly identified force, displacement and momentum as being vector quantities (98%, 97% and 93% respectively). However, their vector addition of these quantities did not exhibit the same regularity.

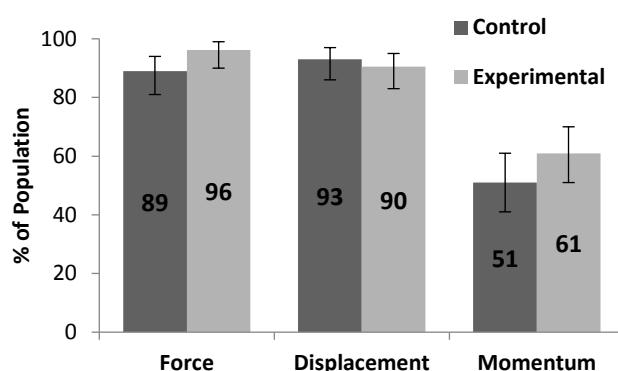


Figure 10: Percentage of students to correctly add vectors in different contexts (n=205).

It is clear from Figure 10 that there are no significant differences between the control and experimental groups for any of the vector quantities being added. However, the momentum cluster is significantly different to both the force and displacement clusters (Cochran’s Q, $p < 0.001$ for both control and experimental groups. McNemar’s test, $p < 0.001$ for both comparisons of “experimental momentum” with “experimental displacement”, and “control momentum” with “control force”, with Bonferoni correction, $\alpha = 0.005$).

For Scenario F, 74% of students agreed with Student B. i.e. 74% of students declared that the resultant force was 5N and that the total force was 7N. Only 2 students (1%) agreed with Student A, and 24% of students agreed with Student C.

For Scenario M, 65% of students agreed with Student B. i.e. 65% of students declared that the total momentum of the two balls is different from the total momentum of the two ball system. 24% of students agreed with Student A and 17% of students agreed with Student C.

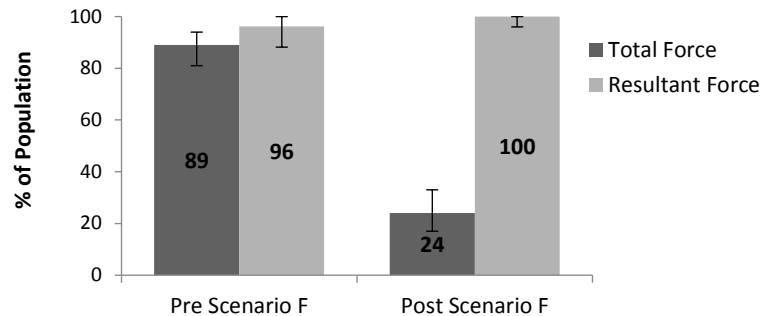


Figure 11: A comparison of students' correct responses for the total force (dark grey) and resultant force (light grey) before and after being exposed to Scenario F.

Comparing the dark grey columns of Figure 3 we can see a large 65 percentage point decrease of correct respondents for “total force”. In other words, after it was suggested that total force might differ from resultant force, 73% of students in the control group who had initially chosen the correct value of 5 N, changed their answer to the incorrect value of 7 N.

Comparing the dark grey columns in Figure 12, we see that the percentage of correct respondents to the total momentum of the two-ball system fell from 51% to 21% after the students had been exposed to Scenario M. In other words, about 60% of students who initially gave the correct answer of 5 kgms^{-1} for the momentum of the two balls, changed their answer to the incorrect value of 7 kgms^{-1} . Comparing the light grey columns, we see that, after being exposed to scenario M, a number of students changed their incorrect answer of 7 kgms^{-1} , to the correct value of 5 kgms^{-1} for the two ball system, leading to an increase of 19 percentage points.

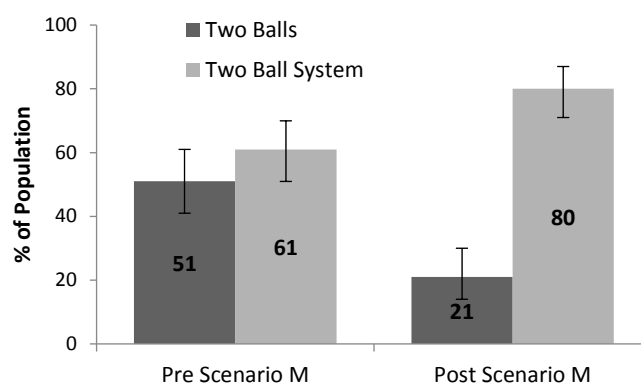


Figure 12: A comparison of students' correct responses for the total momentum of the two balls (dark grey columns) and two ball system (light grey columns) before and after being exposed to Scenario M.

3.6.5. Methodology: Free Writing Responses

As in Study 2, students were invited to motivate their choice of answer for Scenario F in a free response format. The data were analysed using the codes developed in Study 2. Other patterns of response were also identified. For brevity, only two categories are mentioned in the Findings below.

3.6.6. Findings

As in Study 2, the majority of students referred to a type of procedural distinction between “total force” and “net force”. Nevertheless, the main category of interest, “Simulation”, accounted for a significant proportion of students.

Simulation: Effect, Result, Movement and Experience [26% of students (i.e. 54 Students¹⁷)]

In Study 2, 26% of students used some type of simulation in their response. An identical proportion of students were found to use some type of simulation in their responses in Study 3. Some responses included both a simulation involving the “felt experience” of the box *and* a simulation involving the resultant movement of the box:

- *The total force that the box experiences is the addition of the two forces applied to the box, but the resultant force that the box experiences is 5N because it moves in a North Easterly direction as a result of both forces. (179N)*
- *The box, not considering direction, experiences both these forces and the sum of the forces is 7N. However, the resultant force is 5N as when directions are taken into account the box will only move in a North Easterly direction whilst experiencing 5N of force. (113N)*

¹⁷ This is not a “copy&paste” error! Both Study 2 and Study 3 had 54 students categorised as using simulation.

Ontology of Resultant Force as a New, Single, “Independent” Force (7% of students (i.e. 15 students))

An interesting new category to emerge within the superordinate category of “Ontology” was the notion of the resultant force as a type of independent entity. The word “resultant” might have triggered these particular responses, as we did not observe similar responses in Study 2 for “net force”.

- *the resultant force is the combination of individual forces into a single concentrated force that acts on the box (198T)*
- *The resultant force is a single force which has the same effect as both original forces acting at the same time (117R)*
- *The resultant force is the effecting force, or the final force as a result of all the other forces. (133R)*
- *It is the one force that could replace all the others and have the same effect (158R)*
- *while the resultant force speaks of one force that 'represents' all other forces. (199R)*

Note that the qualitative results from the momenta data have yet to be fully coded, although a first pass indicates that the hypothesised notions of “separateness vs. togetherness” do indeed underlie students’ reasoning for making the distinction between the “two balls” and the “two ball system”. For example:

- *The two balls are two different objects and we add their respective momenta to obtain the total momentum [of 7 kgms^{-1}]. The two ball system is now one object so to obtain the total momentum we must find the resultant momentum since momentum is a vector. (175R)*

3.6.7. Discussion

The results from the quantitative analysis section of the paper are consistent with the findings in both Study 1 and Study 2. Figure 10 confirms that students add vectors with different rates of success in different contexts. Figure 11 and Figure 12 show that students are sensitive to the textual prompts. The quantitative analysis also shows that although students may not distinguish between “total force” and “resultant force” at first glance, they readily distinguish between these two notions, if such a difference is suggested to them. We saw an identical result in Study 1 (see also Southey & Allie, 2014) for the notions of “net force” and “total force”. A similar degeneracy is lifted between the notions of (i) the total momentum of two balls, and (ii) the total momentum of a two ball system. Figure 2 and Figure 3 illustrate students’ willingness to make these distinctions. (Chi squared tests, $p < 0.02$, Bonferoni correction, $\alpha = 0.025$).

The qualitative analysis suggests that the majority of student distinguish between “total force” and “net force” using shallow procedural reasoning, such as “arithmetic sum” or “Pythagoras’ theorem”.

However, a significant number of students appeared to run a *simulation*, as described by Barsalou (2003). These students seem to simulate what an experience of the two forces (3N East and 4N North) would be like. Similarly to Study 2, their responses can be grouped according to whether they simulate a tactile experience of pressure, or whether they simulate the resultant motion. These simulations of deformation (which correspond to the above-mentioned simulations of the tactile experience of pressure) depend on how the students model the object in question. As in Study 2, the data suggest that “net force” might prompt students to model the box as a point particle, while “total force” might prompt students to model the box as an elastic solid, capable of experiencing pressure.

3.6.8. Conclusion

The present study is the third such study to suggest that when students perform two dimensional vector addition (in a non-graphical context) they are sensitive to both (a) the type of vector quantity being added, and (b) textual prompts. This study shows that the majority of students distinguish between the terms “total force” and “resultant force”, as well as “momentum of the two balls” and “momentum of the two ball system”. The qualitative data suggest that there might be systematic, experiential reasons for students distinguishing between “total force” and “net force”. For example, their reasoning may be understood in terms of a simulation involving either the tactile experience of an object (total force) or the resultant movement of an object (net force). These distinct simulations would depend on how the students model (cf. Hestenes) the object.

3.7. Study 4: Vector Addition in Different Physical Contexts

3.7.1. Introduction

This final study was conducted in order to further investigate the type of simulation students might be using in order to make sense of the notions of “total force” and “net force”. Evidence from Study 2 and Study 3 suggest that when asked to calculate the total or net (or resultant) force, students may simulate either the “felt experience” of the force on the object, or the resultant motion of the object. We therefore designed a questionnaire that might prompt these particular simulations. The questionnaire consisted of two questions. Both questions explicitly ask students to picture particular scenarios. The first scenario involves a small, hard metal ball on a smooth surface, in which students were asked to find the **net force**. This scenario was intended to prompt the “resultant motion” simulation. The second scenario involves picturing oneself on a soccer field, and having two players simultaneously collide with one. For this scenario, students were asked to find the **total force**. This scenario was intended to prompt the “felt experience” simulation. Importantly, we believed the

order of the questions to be a potential influential factor. We therefore reversed the order of the questions on half the questionnaires. Students were asked to explain their choice of answer as if explaining their reasoning to a child in junior school. We chose this particular audience to further prompt students to make sense via simulation. By deterring students from using higher order concepts such as “vector sum”, or “tip-to-tail”, we hypothesized that they would be less inclined to use shallow, procedural reasoning.

3.7.2. Sample

The sample comprised 293 freshman taking a physics course designed for students majoring in Life Sciences. These students would have taken physics at high school, and would be familiar with the vector addition of forces.

3.7.3. Methodology: Forced Choice Responses

The questionnaire consisted of two questions which are given below. Roughly half the class ($n = 156$) were given the “ball scenario” first, and the “person scenario” second. This group will hereafter be called the “ball-person” or “BP” group. The other half ($n = 138$) were given the questions in the opposite order. This group will be known as the “person-ball” group or “PB” group. The questions were answered under exam conditions, and students were given ample time to complete the task.

1. Picture the following scenario: A small, hard metal ball is at rest on a smooth surface. It is then simultaneously struck by two objects. The first object exerts a force on the metal ball of 4 N East, and the second object exerts a force on the metal ball of 4 N West.

(a) What is the magnitude of the net force experienced by the metal ball? Circle the correct answer.

0 1 2 3 4 5 6 7 8 9 10 11 [N]

(b) Carefully explain your choice of answer to a child in junior school.

2. Picture the following scenario: You are playing soccer on the school playground. You are standing still, having a rest. Suddenly, two large players simultaneously collide with you. The first player exerts a force on you of 4 N West, while the second player exerts a force on you of 4 N East.

(a) What is the magnitude of the total force you experience? Circle the correct answer:

0 1 2 3 4 5 6 7 8 9 10 11 [N]

(b) Carefully explain your choice of answer to a child in junior school.

3. Did you notice any similarities or differences between question 1 and 2? Explain.

3.7.4. Findings

The two main findings were: (i) the scenarios do have a significant impact on student performance; and (ii) the order of the questions matters.

Figure 13 shows that in the PB group, students performed significantly better on the *ball scenario* than the *person scenario* ($p < 0.001$, McNemar's Test for differences between two correlated proportions)

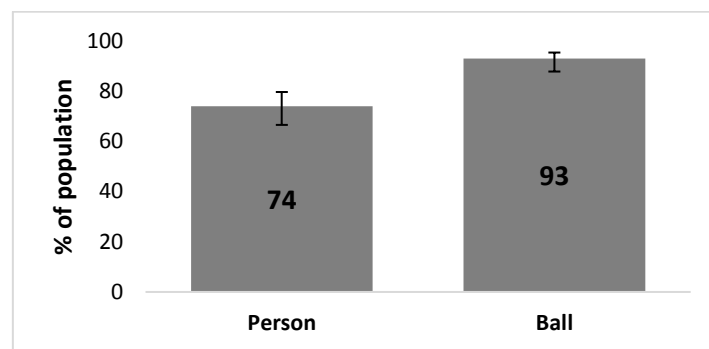


Figure 13: Percentage of Students in the **PB group** to Correctly Calculate the **Total Force Acting on the Person and the Net Force Acting on the Ball**.

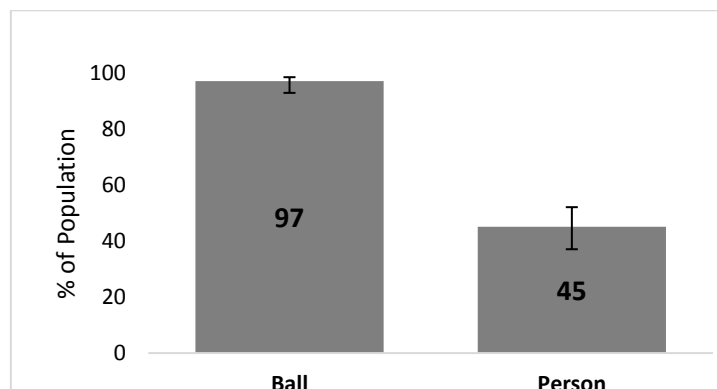


Figure 14: Percentage of Students in the **BP group** to Correctly Calculate the **Net Force Acting on the Ball and the Total Force Acting on the Person**.

Figure 14 shows that in the BP group only 45% of students gave correct answers in the *person scenario*, which was significantly less to the 97% who correctly answered the *ball scenario*. ($p < 0.001$, McNemar's Test for differences between two correlated proportions).

A χ^2 test of association, revealed that the performance on the *Person Scenario* was significantly worse in the BP group than the PB group ($p < 0.0001$). i.e. If students are first given the Ball Scenario, they perform worse in the Person Scenario than if they hadn't received the Ball Scenario. A further χ^2 test revealed no significant difference in the performance on the *Ball Scenario* between PB and BP groups ($p = 0.1$).

Finally, a χ^2 test revealed a significant difference ($p < 0.0001$) between those answering the "Person" question in the PB group, and those answering the "Ball" question in the BP group. (This is comparing the first columns (the 74% and the 97%) of Figure 13 and Figure 14.) This is perhaps the crucial comparison, as it compares students' spontaneous performance in the two scenarios. i.e. it shows that student performance in the *Person Scenario* and the *Ball Scenario* differs, even when the scenarios are not received one after the other.

3.7.5. Methodology: Free Writing Responses

As shown in "Methodology: Forced Choice Responses" above, students were asked to explain their choice of answer to question 1 and question 2 to a child in junior school. Question 3 asked the students to comment on any perceived similarities or differences between question 1 and question 2.

The coding scheme used for Study 2 and Study 3 was used to categorize students' responses according to whether they exhibit features of "simulation". A full qualitative analysis is still ongoing.

3.7.6. Findings

There is evidence to suggest that a simulation of both scenarios provided an underlying basis for student reasoning. In the *Ball Scenario* many students explicitly simulated the "resultant motion of the ball" in their responses. They claimed that, because the ball does not move, the net force on the ball is zero. Examples of responses from students who chose "0 N" for the net force on the ball are given below:

- "The forces cancel each other out and therefore the ball doesn't move." (121BP)
- "The ball is not moving... This means that in the end the total force is zero." (132BP)
- "... the force applied is equal and this causes the ball to not move at all. The ball will stay still and nothing will cause it to move" (153BP)
- "When the ball is hit by the other 2 objects, the ball will not move. So it will stay in the same position". (119PB)
- "So they cancel each other out and the ball would stay at the same place" (146PB)
- "The two forces cancel each other out and the ball would stand still." (144PB)

In their free writing responses to the *Person Scenario* many students explicitly simulated “the felt experience of the person”. They justified their chosen answer of 8 N, by claiming that this is the total force experienced by the person. Examples include:

- *“Each player hits you with a force of 4 N, and you will experience each of these forces (it might hurt), so you will experience a total force of 8 N.” (121BP)*
- *“When two people run into you, you feel pain. This is caused by the force they ran into you with. The amount of force on you is the one players force on you plus the other players force on you.” (153BP)*
- *“When two players bump into you, you get squashed in the middle. You won’t move, but you still get hurt by both players.” (119PB)*
- *“That is a total of 8N force on your body” (144PB)*

However, a particular scenario (*Ball* or *Person*) did not dictate a particular simulation. Rather, the simulation seemed to be correlated with the particular choice of numerical answer. “*Resultant Motion*” was correlated with the answer of 0 N and “*Felt Experience*” with an answer of 8 N. Here are three examples of students who chose 0 N as their answer after applying a “*Resultant Motion*” simulation to the “*Person Scenario*”:

- *... you won’t move from your position” 118 (PB)*
- *“If someone pushes you from the right and another person from the left at the same tie you won’t move...” (109PB)*
- *“In both cases the two objects are struck by equal forces which are oppositely directed and the objects...do not move...” (111PB)*

There are also examples of students performing both “*Resultant Motion*” and “*Felt Experience*” simulations for a particular scenario. Here is an example of a student who performed both simulations for the “*Person Scenario*”, ultimately justifying a choice of “0 N” as her answer:

“If a person with a force of 4 N hits you from one side and a person with the same amount force hits you from the other side, those forces ‘cancel’ each other out. Well, you will still feel the force but in terms of movement, you will not show that a force has been experienced” (123PB)

And she provides similar reasoning in her answer to the *ball scenario*:

“Same reasoning applies. Although it might feel to the ball that 8 N of force is experienced, the net force is 0” (123PB).

Here is another example of a student performing both simulations:

“I think the poor kid (me) would be crushed. The force I feel would not cancel out as I would have loads of bruises! But I need to add that when the collision happens we would not move east or west. Therefore the force would balance out the direction of movement.” (110PB)

In the above answers we can see mention of “pain”, “crushed”, “bruises” etc. The objects (ball and person) involved in each scenario seem to influence the particular simulation. This is well exemplified in the following examples :

- *“I thought in terms injuries the net force would crush the kid. But in terms of a metal ball it can’t really get hurt or crushed and the ball would not move – it will feel the force , but there is no net force.” (110PB)*
- *“ The difference was the object. In question 1, the metal ball was inanimate and could not experience pain, and it was easier to dismiss the ball. The human experienced pain and could not have done so unless there was force”. (153BP)*

Finally, a handful of students articulated the tension between making sense of the scenario with their physical intuition, and providing a “physics answer”. This is well exemplified in the following student’s response:

The forces exerted on me by the two players will cancel each other out since they are equal in magnitude but in opposite directions which means it would be as if no force was exerted on me although it may seem impossible when you think of it. I think I lost understanding of the concept somehow because I thought about it in a physics brain-mannered way...” (122PB)

3.7.7. Discussion

In Study 2 and Study 3 we distinguished between: (i) students spontaneously differentiating between “total force” and “net force” (or “resultant force”), and (ii) students differentiating between “total force” and “net force” after explicit prompting. In both cases we attributed this differentiation to students’ use of different reasoning when dealing with “total force” and “net force”. This difference in reasoning was hypothesized to be grounded in two different simulations – “*resultant motion*” and “*felt experience*”. We hypothesized that the latter differentiation (“after

explicit prompting”) also included a framing effect. By framing the questionnaire as a typical test or exam, students might reason: “It is silly to propose a possible difference between two concepts, if there is *no* difference!” We hypothesize that similar effects were at play in Study 4. It is clear from the qualitative analysis above that two different simulations, “*resultant motion*” and “*felt experience*”, ground students’ reasoning for differentiating between the two scenarios. However, the asymmetry in performance associated with the order of the questions, is hypothesized to be due to a framing effect. If students framed the questionnaire as a typical test or exam, they would think it silly to include two identical questions.

Consider the BP group: After calculating the **net force** as 0 N in the *Ball Scenario*, students’ framing of the questionnaire as a typical test/exam would have primed them to “be on the lookout” for a difference in the second question involving the *Person Scenario*. This difference was easily located in the new object used – an animate, squishy person that feels pain – and in the term “**total force**”. This new object (and its corresponding scenario) allowed for the simulation of “*felt experience*”, which, together with the term “**total**” and its connotations, provided readily available grounds for reasoning to an answer of 8 N .

Now consider the PB group. The quantitative data suggest that the *Ball Scenario* prompts students to reason to an answer of 0 N ; the object (the small, inanimate, metal ball) and the term “**net force**” seem to unambiguously guide students to choose this answer. Now, 74% of students in the PB group chose 0 N as their answer to the first question – the *Person Scenario*. Even if these students were primed to “be on the lookout” for a possibility of arriving at a different answer in the second question (*Ball Scenario*), the unambiguous nature of this question does not allow any room for reasoning to an answer other than 0 N . This hypothesis explains the asymmetry in the performance on the *Person Scenario* question in the PB and BP groups (74% and 45% respectively). Furthermore, an *additional* framing effect might be at play in the PB group. This is due to the effect of familiarity, as outlined in the Discussion section of Study 2. As the students are familiar with addition of forces from high school, they “may resist wasting resources by analysing the object further.” (D. L. Hintzman & Curran, 1995, p.223). At school, the correct answer to a vector addition question involving two equal and opposite forces would be 0 N . Thus, by framing it as a “typical vector addition question”, students would be less sensitive to the prompts of “**total**” and the animate, squishy person, and less likely to give an answer of “ 8 N ”.

3.7.8. Conclusion

This study, the fourth and final of a series of studies on vector addition, provides compelling evidence to suggest that students are adept at grounding their reasoning in simulation (L. W.

Barsalou et al., 2008). Students' responses which exhibited evidence of simulation could be categorized as being simulations of "*resultant motion*" or "*felt experience*". A variety of factors influenced which of these simulations would obtain.

(i) Framing. As argued in the studies above, framing can have at least two different types of effect:

(a) If a question is framed, or categorized, as belonging to a particular question type (e.g. vector addition) which one has encountered on numerous occasions, and for which one has developed a type of algorithm, then this might prompt a type of "automaticity" that prevents one from noticing the detail of the question. (b) If one frames a questionnaire as being a typical test or exam, and encounters two very similar questions, one is likely to "be on the look out" for differences between these two questions. It is these two framing effects that are hypothesized as being responsible for the asymmetry in performance in the *Person Scenario* in the PB and BP groups.

(ii) the words "**total**" and "**net**" and (iii) the given scenario; in particular, the object used in the scenario. Since these factors were conflated in this study, we cannot conclude which might be a greater contributor in prompting a particular simulation

Finally, while the *resultant motion* simulation was generally associated with the *Ball Scenario* and the *felt experience* simulation generally associated with the *Person Scenario*, these two simulations were most tightly correlated with the answers of 0 N and 8 N respectively. This is exemplified in responses which include *both* simulations, but weight one simulation as being more salient than the other. In these cases, the chosen answer is aligned with the more salient simulation.

3.8. General Discussion and Conclusion

"My students understand vector addition". What does this mean? There are a number of different models/meanings we could use:

- i. We might model the concept of "vector addition" as a type of "thing" that we acquire, much like we acquire a tool in our (cognitive) toolbox. Students demonstrate their acquisition of this concept/tool by successfully adding forces and displacements. From this perspective we would expect students to unproblematically transfer this concept to the context of adding momenta, especially because they knew all the relevant information (i.e. the fact that momentum is a vector quantity). This chapter is perhaps first and foremost a challenge to this traditional "acquisition and transfer" model of understanding, as it is applied to the concept of "vector addition".

- ii. We might model our understanding of the concept “vector addition” as a type of ability; the possession of this concept being akin to a skill. Like any skill/ability, we might expect some difficulties in applying it in a new context. For example, if we learn to ride a bicycle on a flat tar surface, we would expect some difficulty in transferring this skill to a rough “off road” context. The crucial question for education research is: how can we know if a new context will impede the smooth exercise of a particular skill/ability? In other words, how can we explain that “adding momenta” is a “rough off road” context in comparison to the “smooth tar road” contexts of “adding forces” and “adding displacements”?
- iii. A resources perspective is helpful in this regard. When seeing “momentum” a different set of resources are activated in the mind of the student. As one student said in her interview: “On seeing the word ‘momentum’, my thinking changed.” This different cognitive activity is perhaps an underlying explanation for why the context of “adding momenta” is understood as different to the context of forces and displacements. The activation of a different set of resources *is* the difference in context.
- iv. The coordination class model further deepens our understanding of understanding of vector addition. “Vector addition” is a sufficiently sophisticated concept to be modelled as a coordination class. In this model the emphasis is why students couldn’t see the momentum context as the same as the force and displacement contexts. We could look for evidence of what occurred in students’ readout strategies or their causal nets, such that information related to “vector addition” could not be straightforwardly coordinated in the context of momentum.
- v. Both the Resources perspective and the coordination class perspective prompt for a finer-grained analysis: what are the smaller “things” underlying the context sensitivity. In the case of the context sensitivity to “total” and “net”, these could be identified as two different type of mathematical resource: “simple addition: $3 + 4 = 7$ ”, and “Pythagorean: $3^2 + 4^2 = 5^2$ ”. For most students, the “cognitive buck” stops here. But perhaps we can go even further; is there a further physical intuition (a type of resource) that underlies the association of “total” with “simple addition” and “net” with “Pythagoras”?
- vi. And here embodied simulation seems to be a helpful model. Some students associate “total force” with the felt experience of a force, and “net force” with resultant movement. Students can be prompted to simulate what two force would *feel* like, or prompted to simulate how two forces cause a resultant motion, and incorporate these simulations into their understanding.
- vii. Finally, we can introduce further models which focus more on the process of reasoning and metacognition. Why might students not spontaneously distinguish between “total force” and “net force”, but justify a distinction between these terms when prompted to do so? Perhaps

because they have particular expectations regarding physics questionnaires; they frame the activity in such a way that it feels expectant upon them to provide a distinction between “total force” and “net force”.

- viii. Or perhaps, due to functional fixedness, they overlooked a possible distinction between “net force” and “total force”. It is only when explicitly asked if there is such a distinction, that they engage slower, more considered reasoning. This would also explain why students are more likely to make a spontaneous distinction between the unfamiliar contexts involving “net momentum” and “total momentum”. The unfamiliarity of the contexts meant that there is no precedent; no readily available answer from their associative machine.

“My students understand vector addition”. This is not a simple sentence. What it means is constituted by our various models of understanding. This chapter has shown that there are various types of data, that can count as criteria in various models, that can productively give meaning to the concept of “understanding” in the context of “vector addition”.

3.9. Post Script: Textbooks and Total Force

Introductory textbooks ordinarily do not bring attention to the fact that “total” and “net” are synonymous in the context of vector addition. While most first year textbooks use “resultant force”, there are examples of first year textbooks, such as Reese’s “University Physics” (Reese, 2000), that use “total force” throughout. Serway & Jewett (2013) and (Cutnell & Johnson, 1995) use the term “net force” but “total momentum”. Older textbooks such as Adair (1969), Abbott (1969) and Beiser (1973) all refer to “total momentum”. Crummett & Western (1994) use “resultant displacement” and “resultant force” when teaching about vector addition. However, some “back of the textbook” questions refer to the “total displacement”, “net displacement”, “total force” and “net force”.

CHAPTER 4: Context Sensitivity in the FCI

<u>Concept</u>	<u>Context</u>	<u>Model of Understanding</u>
Basic Mechanics Concepts (with familiar words)	Familiar Words	Evidence for possible embodied simulation and shallow linguistic processing.
Basic Mechanics Concepts (with unfamiliar words)	Unfamiliar Words	Theoretical models used: LASS (Linguistic and Situated Simulation); Dual Processing; Working Memory.

4.1. Introduction and Literature Review

A particularly useful tool for probing context sensitivity in physics education research, is the Force Concept Inventory (FCI) (David Hestenes, Wells, & Swackhamer, 1992). The FCI is a standardized test, consisting of 30 questions, and is thought to provide an accurate indicator of students' understanding of basic mechanics. Its precursor was the Mechanics Diagnostics Test (Halloun & Hestenes, 1985). There have been various challenges to the comprehensibility of the FCI as a test of students' understanding of mechanics, and these have largely centered around the notion of context sensitivity. Context sensitivity in the FCI has been studied by a variety of researchers. Their findings are collated in the paper *Context Sensitivity in the Force Concept Inventory* (Stewart et al., 2007). Stewart et al. identify eight different types of contextual change. For each contextual change they calculated a percentage point shift in score between the control and the experimental groups. For example, the three largest shifts in score were observed for a change of context involving (i) the removal of a figure; (ii) changing one concrete system to another concrete system; and (iii) removing questions from a group. The most relevant of these change of contexts for our purposes is the changing of one concrete system for another. It has been well established that substituting one noun for another in a physics problem, can have a large impact on student performance. One example involves asking students to consider the normal force exerted on a book when resting on a table. If one substitutes "table" for "spring", a considerable difference in student reasoning is observed (D. E. Brown & Clement, 1989). One way of modeling this context sensitivity is that "table" and "spring" activate different resources (Hammer, 2000), or different semantic networks (Lakoff, 1987). The table activates ideas of rigidity, while the spring activates ideas of flexibility. Another model might be that these words prompt different simulations (Lawrence Barsalou, 2003).

A context that has not been explored is that of *unfamiliar words*. The questions asked in this series of studies is: What if a word is completely unfamiliar, and activates very few productive resources? What if a word does not belong to any existing semantic network, and has no associated mental image or simulation? For example, what happens if we substitute “table” for an unfamiliar word, such as “kist”, or for a nonsense word such as “dinnish”.

The context of “unfamiliar” words is not only interesting from a purely theoretical perspective. It may also shed light on differences between 1st year physics instruction, and instruction in the upper division. Many concepts in first year, such as “force”, “mass”, “acceleration”, “density”, “current”, and “energy”, are all encountered at high school, and also have lay-uses. In other words, these words are already associated with established semantic networks. However, concepts in the upper division such as “superposition”, “Hilbert Space”, and “Hermitian” are encountered for the first time in a physics class. This may point to a significant difference in approach required in teaching first year and upper division content. In first year, educators need to be aware that most of the concepts will activate multiple resources in the minds of first years, and the challenge is to marshal these resources (to suppress certain resources and to activate other resources) into coherent concepts. In contrast, in the upper division, educators need to be aware that they will be the first point of contact for most core concepts. Instead of negotiating established knowledge networks associated with first year concepts, educators will lay the first semantic connections for upper division concepts. Hence the importance of the first analogies used and the first contexts introduced. These will be the foundation from which students begin to integrate these unfamiliar concepts into their existing knowledge networks.

This chapter looks at how we might model students’ understanding of unfamiliar words in physics education.

4.2. Overview

We conducted a series of three studies comparing two groups of students answering different versions of FCI-type questions. The control group received “normal” FCI-type questions, while the experimental group were given questions in which some of the nouns had been replaced with unfamiliar words. Our hypothesis was that the control group would outperform the experimental group. The results of the studies are inconclusive. While students’ overall performance in the control and experimental groups was not consistently significantly different, students’ performance on the more difficult questions does seem to be significantly better for the control group.

A variety of possible influential factors were explored. This chapter focuses on two factors: (i) the difficulty of the questions, and (ii) whether the questions are oriented toward calculation, or conceptual understanding.

4.3. Study 1: Pilot Study March 2013

4.3.1. Sample

The sample consisted of 182 freshmen who were enrolled in either Engineering or Life Sciences. All of these students took physics at high school.

4.3.2. Instrument

The control questionnaire consists of 13 questions which are based on questions from the FCI and the Mechanics Baseline Test (see Hestenes & Wells (1992), Hestenes, Wells, & Swackhamer (1992) and Hake (1998)). The experimental questionnaire was created by substituting familiar nouns in the control questionnaire, with unfamiliar nouns. The full questionnaire is provided below. The experimental version of the questionnaire is in *italics*. Table 5 shows some of the more common noun substitutions.

<u>Control</u>	<u>Experimental</u>
Car	Humvee, Xytherion, Jalopy
Man	Australopithecus, Sambian
Ball	Prolate spehroid, Candelabra
Box	Kist, Skurrit

Table 5: Examples of Different Nouns used in Control and Experimental Questionnaires

1. A car is on a highway and slows down from a speed of 9 m.s^{-1} to a speed of 5 m.s^{-1} in half a second. What is the average acceleration of the car during this time?
1. *An Australopithecus is on a travelator and slows down from a speed of 9 m.s^{-1} to a speed of 5 m.s^{-1} in half a second. What is the average acceleration of the Australopithecus during this time?*

(A) 2 m.s^{-2} (B) 3 m.s^{-2} (C) 4 m.s^{-2} (D) 6 m.s^{-2} (E) 8 m.s^{-2}
2. You exert a force of 200 N on a box of mass 20 kg as you push it over a rough surface. The frictional force on the box is 50 N. What is the acceleration of the block?
2. *You exert a force of 200 N on a kist of mass 20 kg as you push it over a linoleum surface. The frictional force on the kist is 50 N. What is the acceleration of the kist?*

(A) 10 m.s^{-2} (B) 2.5 m.s^{-2} (C) 7.5 m.s^{-2} (D) 2.0 m.s^{-2} (E) 4.0 m.s^{-2}

3. A car has a maximum acceleration of 3.0 m.s^{-1} . What would its maximum acceleration be while towing a second car twice its mass?

3. *A xytherion has a maximum acceleration of 3.0 m.s^{-1} . What would its maximum acceleration be while towing a krypterion, twice its mass?*

(A) 2.5 m/s (B) 2.0 m/s (C) 1.5 m/s (D) 1.0 m/s (E) 0.5 m/s

4. Two metal balls are the same size but one weighs twice as much as the other. The balls are dropped from the roof of a single story building at the same instant of time. The time it takes the balls to reach the ground below will be:

4. *Two prolate spheroids are the same size but one weighs twice as much as the other. The spheroids are dropped from the awning of a single story building at the same instant of time. The time it takes the prolate spheroids to reach the ground below will be:*

(A) about half as long for the heavier spheroid as for the lighter one.

(B) about half as long for the lighter spheroid as for the heavier one.

(C) about the same for both spheroids.

(D) considerably less for the heavier spheroid, but not necessarily half as long.

(E) considerably less for the lighter spheroid, but not necessarily half as long.

5. The two metal balls of the previous problem roll off a horizontal table with the same speed. In this situation:

5. *The two prolate spheroids of the previous problem roll off a horizontal sideboard with the same speed. In this situation:*

(A) both spheroids hit the floor at approximately the same horizontal distance from the base of the sideboard.

(B) the heavier spheroid hits the floor at about half the horizontal distance from the base of the sideboard than does the lighter spheroid.

(C) the lighter spheroid hits the floor at about half the horizontal distance from the base of the sideboard than does the heavier spheroid.

- (D) the heavier spheroid hits the floor considerably closer to the base of the sideboard than the lighter spheroid, but not necessarily at half the horizontal distance.
- (E) the lighter spheroid hits the floor considerably closer to the base of the sideboard than the heavier spheroid, but not necessarily at half the horizontal distance.
6. A stone dropped from the roof of a single story building to the surface of the earth:
6. *A loose shingle falls from the awning of a single story mansion to the surface of the earth:*
- (A) The shingle reaches a maximum speed quite soon after it falls from the awning and then falls at a constant speed thereafter.
- (B) The shingle speeds up as it falls because the gravitational attraction gets considerably stronger as the shingle gets closer to the earth.
- (C) The shingle speeds up because of an almost constant force of gravity acting upon it.
- (D) The shingle falls because of the natural tendency of all objects to rest on the surface of the earth.
- (E) The shingle falls because of the combined effects of the force of gravity pushing it downward and the force of the air pushing it downward.
7. A large truck collides head-on with a small compact car. During the collision:
7. *A large Humvee collides head-on with a small jalopy. During the collision:*
- (A) the Humvee exerts a greater amount of force on the jalopy than the jalopy exerts on the Humvee.
- (B) the Humvee exerts a greater amount of force on the jalopy than the Humvee exerts on the jalopy.
- (C) neither exerts a force on the other, the jalopy gets smashed simply because it gets in the way of the Humvee.
- (D) the Humvee exerts a force on the jalopy but the jalopy does not exert a force on the Humvee.
- (E) the Humvee exerts the same amount of force on the jalopy as the jalopy exerts on the Humvee.

8. A woman exerts a constant horizontal force on a large box. As a result, the box moves across a horizontal floor at a constant speed “ v ”. The constant horizontal force applied by the woman:
8. *A courtesan exerts a constant horizontal force on an ebony casket. As a result, the ebony casket moves across a horizontal floor at a constant speed “ v ”. The constant horizontal force applied by the courtesan:*
- (A) has the same magnitude as the weight of the ebony casket.
- (B) is greater than the weight of the ebony casket.
- (C) has the same magnitude as the total force which resists the motion of the ebony casket.
- (D) is greater than the total force which resists the motion of the ebony casket.
- (E) is greater than either the weight of the ebony casket or the total force which resists its motion.
9. If the woman in the previous question doubles the constant horizontal force that she exerts on the box to push it on the same horizontal floor, the box then moves:
9. *If the courtesan in the previous question doubles the constant horizontal force that she exerts on the ebony casket to push it on the same horizontal floor, the ebony casket then moves:*
- (A) with a constant speed that is double the speed “ v ” in the previous question.
- (B) with a constant speed that is greater than the speed “ v ” in the previous question, but not necessarily twice as great.
- (C) for a while with a speed that is constant and greater than the speed “ v ” in the previous question, then with a speed that increases thereafter.
- (D) for a while with an increasing speed, then with a constant speed thereafter.
- (E) with a continuously increasing speed.
10. If the woman in question 9 suddenly stops applying a horizontal force to the box, then the box will:
10. *If the courtesan in question 9 suddenly stops applying a horizontal force to the ebony casket, then the ebony casket will:*
- (A) immediately come to a stop.

(B) continue moving at a constant speed for a while and then slow to a stop.

(C) immediately start slowing to a stop.

(D) continue at a constant speed.

(E) increase its speed for a while and then start slowing to a stop.

11. An empty office chair is at rest on a floor. Consider the following forces:

11. *A vermillion pouffe is at rest on a floor. Consider the following forces:*

1. A downward force of gravity.

2. An upward force exerted by the floor.

3. A net downward force exerted by the air.

Which of the forces is (are) acting on the vermillion pouffe?

(A) 1 only.

(B) 1 and 2.

(C) 2 and 3.

(D) 1, 2, and 3.

(E) none of the forces. (Since the pouffe is at rest there are no forces acting upon it.)

12. Despite a very strong wind, a tennis player manages to hit a tennis ball with her racquet so that the ball passes over the net and lands in her opponent's court.

12. *Despite a very strong wind, a hackey sack player manages to kick a hackey sack with her foot so that the hackey sack passes over a stile.*

Consider the following forces:

1. A downward force of gravity.

2. A force by the "kick".

3. A force exerted by the air.

Which of the above forces is (are) acting on the hackey sack after it has left contact with her foot and before it touches the ground?

(A) 1 only.

(B) 1 and 2.

(C) 1 and 3.

(D) 2 and 3.

(E) 1, 2, and 3.

13. A boy throws a steel ball straight up. Consider the motion of the ball only after it has left the boy's hand but before it touches the ground, and assume that forces exerted by the air are negligible. For these conditions, the force(s) acting on the ball is (are):

13. *An earl throws a candelabra straight upward. Consider the motion of the candelabra only after it has left the earl's hand but before it touches the ground, and assume that forces exerted by the air are negligible. For these conditions, the force(s) acting on the candelabra is (are):*

(A) a downward force of gravity along with a steadily decreasing upward force.

(B) a steadily decreasing upward force from the moment it leaves the earl's hand until it reaches its highest point; on the way down there is a steadily increasing downward force of gravity as the candelabra gets closer to the earth.

(C) an almost constant downward force of gravity along with an upward force that steadily decreases until the candelabra reaches its highest point; on the way down there is only a constant downward force of gravity.

(D) an almost constant downward force of gravity only.

(E) none of the above. The candelabra falls back to ground because of its natural tendency to rest on the surface of the earth.

4.3.3. Methodology

The Control and Experimental Questionnaires were distributed randomly to a class of freshmen students at the beginning of their physics lecture. About half the class became the Control Group ($n = 92$) and the other half became the Experimental Group ($n = 90$). Students were given 25 min to answer the 13 questions.

4.3.4. Findings

Figure 15 shows the results for all the questions in the control and experimental questionnaires. A more palatable rendering is shown in Figure 16. Here the difference between the control group and

the experimental group's scores are plotted for each question (i.e. control group scores – experimental group scores). Thus, for those cases where the difference is negative the experimental group's score is higher than the control group's score. Highlighted with texture are the questions flagged as potentially generating a significant difference between control and experimental groups; these are questions 1, 3, 4, and 13. (Note: in this representation, an error bar that overlaps with the axis (with zero) would indicate non-significant difference between control and experimental).

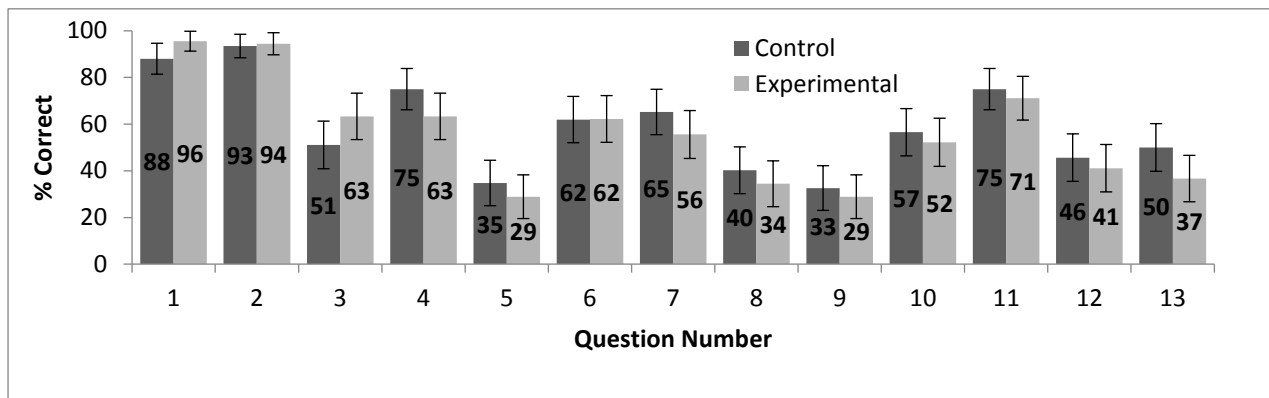


Figure 15: Question by Question Comparison of Control Group (n=92) and Experimental Group (n=90)

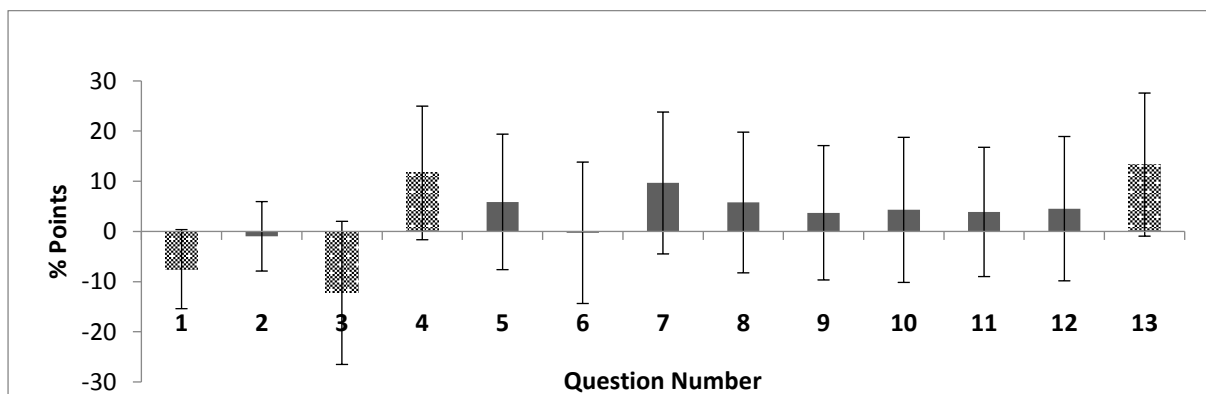


Figure 16: The Difference Between Control and Experimental Groups' Scores. (Control – Experimental).

Figure 17 shows the overall result: the control group performed marginally better (3 percentage points) than the experimental group, but this difference is not statistically significant (χ^2 test yielded a p-value of 0.11). A variety of further tests were conducted in order to discover possible salient factors. Figure 18 shows a similar comparison of the control and experimental groups, but only includes the five most difficult questions; q5, q8, q9, q12, and q13. These were the questions for which students performed poorest (see Figure 15). Here, a χ^2 test yielded a p value of $p = 0.038$, which indicates statistical significance. Note that in this Chapter, **bold** error bars indicate statistical significant difference.

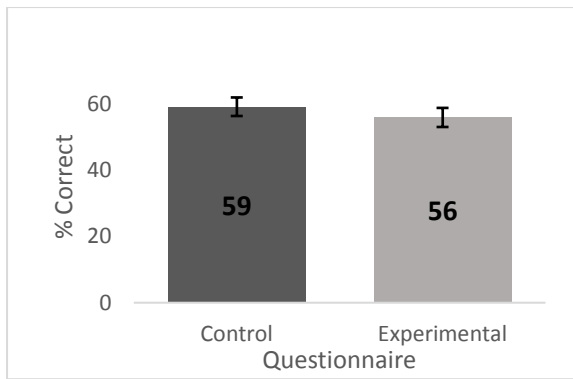


Figure 17: Comparison of Control and Experimental Groups' **Overall Scores**

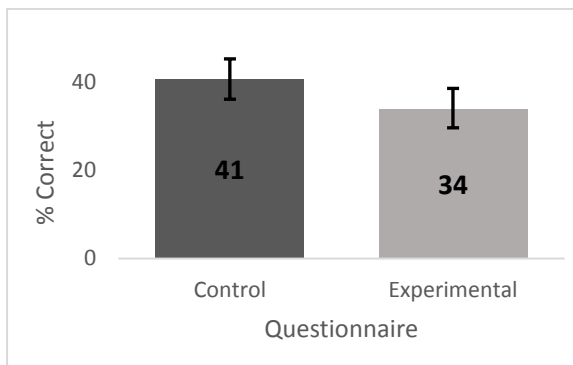


Figure 18: Comparison of Control and Experimental Group's Scores for **Difficult Questions** (5,8,9,12,13)

Figure 19 and Figure 20 compare the first three questions which involved explicit calculation, with the other ten questions which involved only conceptual reasoning. At first sight it seems that an asymmetry exists between these two types of questions: students seem to have performed better in the experimental group on the calculation questions, while the opposite trend can be observed for the conceptual questions. χ^2 tests yield p values of $p = 0.04$, and $p = 0.008$ for the calculation and conceptual questions respectively. The former is not statistically significant (Bonferroni correction, $\alpha = .025$) while the latter is statistically significant.

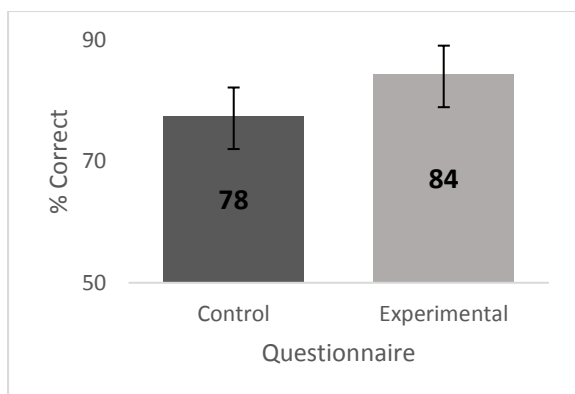


Figure 19: Comparison of Control and Experimental Group's Scores for **Calculation Questions (1,2,3)**

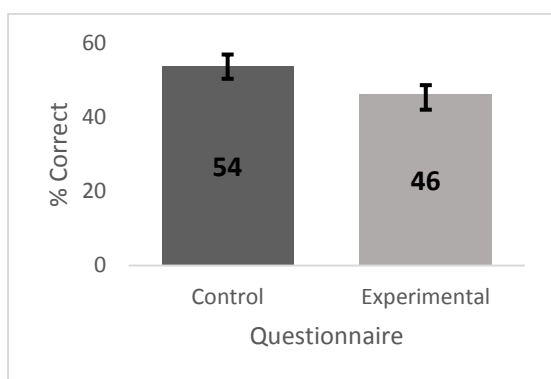


Figure 20: Comparison of Control and Experimental Groups' Scores for **Conceptual Questions**

Other factors such as students' home language and subject stream were explored, but neither yielded significant results. For example, Figure 21 shows a similar "Control – Experimental" plot to Figure 16, but here the results for the Chemical Engineering Students are compared with those of students in the Life Sciences. What were initially flagged as potentially interesting phenomena, such as the asymmetry in observed in questions 5 and 6, did not turn out to be statistically significant.

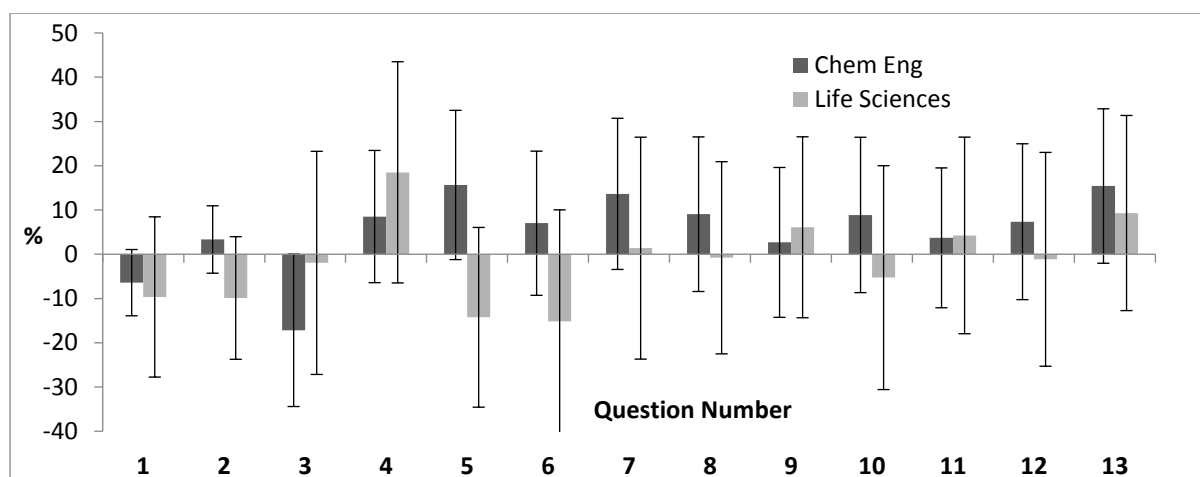


Figure 21: The Difference Between Control and Experimental Scores for Chemical Engineering Students ($n=123$) in Green, and Life Sciences Students ($n=59$) in Purple

4.3.5. Discussion

The hypothesis that students would perform more poorly in the experimental condition was true for only for the conceptual questions (q4 – 13). As mentioned above, this hypothesis was based on two theoretical ideas: (i) Simulation (Lawrence Barsalou, 2003): in order to make sense of a question, students simulate the situation presented in the question. Therefore, the unfamiliar words of the experimental condition hamper this simulation. (ii) Working Memory (Baddeley, 2007): in the effort to make sense of the unfamiliar words, less working memory is available for conceptual reasoning and/or calculation. Our results seem to indicate that while these hypotheses might be true for the conceptual questions, it is not true for the calculation questions. Therefore, a further hypothesis may be required for the asymmetry observed in calculation and conceptual questions. A possible hypothesis is that calculation questions trigger “shallower” cognitive processes related to association and established, algorithmic routine. When presented with a calculation question students may simply focus on the numbers of relevant physical quantities, and manipulate these numbers in relevant equations, without paying much attention to the context in which these numbers are embedded. In other words, calculation questions may prime the “plug and chug” approach (Kuo, 2013). On the other hand, conceptual questions may activate cognitive processes associated with simulation and deeper meaning-making. Supposing this hypothesis is correct (i.e. that calculation questions trigger shallower cognitive processing, and conceptual questions trigger deeper processing), we may speculate on the possible effects of unfamiliar words on student cognition. Such speculation hinges on the effects of familiarity on cognition. This is a large topic and will be explored in detail in Appendix A. For the purposes of this study, we suggest the following idea: a rich, familiar context may distract students from performing efficient algorithmic calculation.

(This is related to a central thesis of Kaminski, Sloutsky, & Heckler (2009) and will be explored in detail in Chapters 5, 6, and 7). If a calculation task can be efficiently completed using shallow cognitive processing, then engaging the slower, deeper cognitive processing of simulation may activate unproductive resources, and distract students from the well-worn algorithmic calculation. Thus, in our study, the calculation questions involving unfamiliar words may have prompted students to ignore the context, and focus on the algorithmic routine. If sufficient information is available to perform the shallow algorithmic processing, the unfamiliar words may therefore aid in focusing students' attention on the calculation rather than the context. Thus, somewhat counterintuitively, the unfamiliar words may enhance student performance in the calculation questions. To test this hypothesis, further calculation questions were added to the next iteration of the instrument, for a more robust comparison between "calculation" and "conceptual".

Figure 15 shows that students performance varied from question to question. Interestingly, no individual question yielded a significant difference in performance. (For example, a Fisher's Exact Test for questions 3 and 13 yielded p values of $p = 0.10$ and $p = 0.07$ respectively). Nevertheless, there were some questions that obviously contributed more (or less) to the overall observed significant difference. We considered possible reasons for why particular questions yielded little or no effect, and how one might alter these questions such that they could yield a significant effect.

Numbers introduced too early (Question 2)

Question 2 only produced a difference of 1 percentage point. This might be because in question 2 the numbers were introduced immediately ("200N" was the 6th word in the sentence, and it was followed almost immediately by a mass). This may have prompted students to immediately engage algorithmic processing, regardless of whether the ensuing context contained familiar or unfamiliar words. In other words, it would not have been necessary for students to make deeper sense of the scenario, and engage their slower simulation system, regardless of whether the scenario description contained familiar or unfamiliar words. This hypothesis will be tested by rewriting question 2 such that a detailed scenario is first presented, before numbers are introduced. This new questions is shown in Study 2 below.

Brevity of Simulation (Question 6)

Question 6 yielded a zero percentage point difference. One possible reason for this, is that the is one sentence long, and therefore does not really stimulate a simulation. The modified question, shown below, contains a more detailed scenario:

A man picks a stone off the ground and puts it in his pocket. He then walks to the first floor of the UCT physics building and finds an open window. Making sure that there is no one below him, he drops the stone to the ground. The stone:

A Sambian picks a stone off the ground and puts it in his sporron. He then walks to the mezzanine floor of the Higgovale Heights and finds an open dormer. Making sure that there is no one below him, he drops the stone to the ground. The stone:

Follow on Question (Question 10)

This question yielded a small difference of 4 percentage points. It referred to a scenario first presented two questions before. It is possible that the bulk of the simulation processing has already taken place during the first (and perhaps second) presentation of the scenario. Thus, this question was discontinued.

All Experimental Questions Occur Together

One concern is that the fact that all questions involving unfamiliar words occurred in the same questionnaire. This may result in student building up a “resistance” to the intended difficulty of simulation. In the next study we mixed control and experimental questions together.

Not sufficient difference in simulation complexity (Q 11)

Question 11 yielded a small difference of 4 percentage points. It was simply one sentence (An empty office chair [a vermillion pouffe] is at rest on the floor). It is therefore discontinued, and replaced with a scenario of more simulation complexity.

4.4. Study 2: October 2013

4.4.1. Sample

The sample consisted of 220 freshmen enrolled in an engineering degree, taking a one year physics service course. All these students would have taken physics at high school.

4.4.2. Instrument

The instrument was similar to that used in the pilot study. One major difference is that instead of putting all the questions with unfamiliar words into one “experimental questionnaire”, these questions were distributed equally between two questionnaires. Thus all students answered some “control” questions and some “experimental” questions. In other words, the “control group” is not a group of students, but the answers to all the “control” questions.

Four questions were discontinued from Study 1, and four new questions were added. One question (Q 6) was modified significantly. This instrument contained five calculation questions (q 1 – 5), and eight conceptual questions (q 6 -13). Two pairs of calculation and conceptual questions (q 2 and q12; and q 4 and q10) contained similar scenarios. Note that the question numbers have changed. For example question 6 of Study 1 is question 8 of Study 2. See Table 6 for details. The new questions are given below:

Study 1 Questions	Action	Study 2 Questions	Type	Scenario
1	retained	1	calculation	car accelerating
2	discontinued	2 (new)	calculation	displacement in park
3	retained	3	calculation	car towing
		4 (new)	calculation	woman pushing box
		5 (new)	calculation	speed of boat
4	retained	6	conceptual	two balls fall from a height
5	retained	7	conceptual	same two balls roll of table
6	modified	8	conceptual	stone dropped from window
7	retained	9	conceptual	truck collision with car
8	retained	10	conceptual	woman pushing box
9	retained	11	conceptual	same woman doubles force
10	discontinued			
11	discontinued			
12	discontinued	12 (new)	conceptual	displacement in park
13	retained	13	conceptual	ball thrown straight upward

Table 6: Relationship Between Questions in Study 1 and Questions in Study 2

2. You are standing at the southern end of an antique fair. You walk 50 m north until you reach a few old chapeaus. You then walk 30 m east and arrive at some phinae. Finally you walk 50 m south to a stand of habiliments. What is the magnitude of your total displacement from your starting position?
 2. You are standing at the southern end of a large park. You walk 50 m north until you reach the swings. You then walk 30 m east and arrive at the sand pit. Finally you walk 50 m south to a hot dog stand. What was the magnitude of your total displacement from your starting position?
- (A) 0 m (B) 30 m (C) 50 m (D) 70 m (E) 130 m

4. A woman pushes a box from the bedroom to the kitchen over a carpeted floor. The force the woman exerts on the box is 200 N. If the mass of the box is 20 kg and frictional force on the box is 50 N, what is the magnitude of the acceleration of the box?

4. *A Dresden pushes a kist from the parritique to the privy over a teppe surface. The force the Dresden exerts on the kist is 200 N. If the mass of the kist is 20 kg and frictional force on the kist is 50 N, what is the magnitude of the acceleration of the kist?*

(A) 10 m.s^{-2} (B) 2.5 m.s^{-2} (C) 7.5 m.s^{-2} (D) 2.0 m.s^{-2} (E) 4.0 m.s^{-2}

5. An old gentleman attends a boat race in Cape Town. The winning boat has an initial acceleration of 6 m.s^{-2} . What is the boat's change in speed in the first 2 s of the race?

5. *A justsmith attends a traxoline derby in Trisos. The winning traxoline has an initial acceleration of 6 m.s^{-2} . What is the traxoline's change in speed in the first 2 s of the race?*

(A) 4 m.s^{-1} (B) 12 m.s^{-1} (C) 24 m.s^{-1} (D) 3 m.s^{-1} (E) 6 m.s^{-1}

12. You are standing at the western end of a large park. You walk east and come to the swings. You continue walking east and arrive at the sand pit. Finally you walk south to a hot dog stand.

Remember: $\text{average speed} = \frac{\text{total distance covered}}{\text{time taken}}$ and $\text{average velocity} = \frac{\text{total displacement}}{\text{time taken}}$.

12. *You are standing at the western end of an antique fair. You walk east and come to a few old chapeaus. You continue walking east and arrive at a collection of phineae. Finally you walk south and stop at a stand of habiliments. Remember: $\text{average speed} = \frac{\text{total distance covered}}{\text{time taken}}$ and $\text{average velocity} = \frac{\text{total displacement}}{\text{time taken}}$.*

Considering your entire trip, which one of the following statements is correct?

- (A) when you reach the *habiliments*, the magnitude of your average velocity and your average speed are equal.
- (B) when you reach the *habiliments* the magnitude of your average velocity is less than your average speed.

- (C) when you reach the *habiliments* the magnitude of your average velocity is greater than your average speed.
- (D) when you reach the *chapeaus*, the magnitude of your average velocity is less than your average speed.
- (E) Both (B) and (D) are true.

4.4.3. Findings

Figure 22 shows the difference between the control scores and the experimental scores for each question. i.e. Control Score – Experimental Score. Questions 3, 8 and 11 were flagged as being potentially statistically significant. χ^2 tests yielded p values of 0.01, 0.002 and 0.1 respectively. Thus, only questions 8 yields a statistically significant difference after Bonferroni correction ($\alpha = 0.004$).

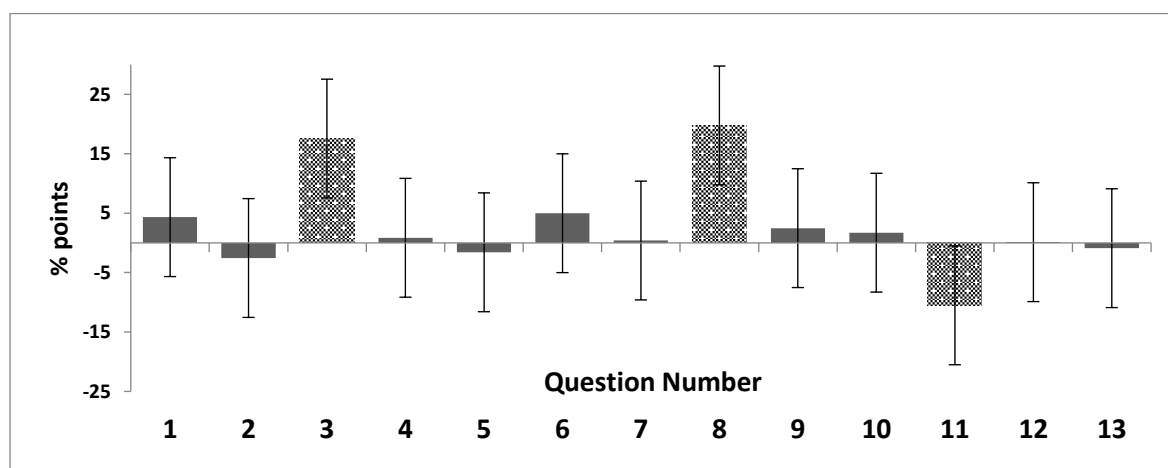


Figure 22: Difference between Control and Experimental $n = 220$

Figure 23 shows the overall comparison between the groups. While the overall score of the control group score was 2 percentage points higher than the experimental group, this was not statistically significant (χ^2 test, p value of 0.051). Figure 24 shows this comparison for the six more difficult questions (i.e. the six questions in which the students performed poorly – questions 3, 6, 8, 10, 11 and 12). A χ^2 test yields a p value of 0.03, which is low, and not statistically significant with a Bonferroni correction ($\alpha = 0.025$). Figure 25 shows this comparison for the calculation questions (questions 1 – 5). Here the students in the control group outperformed those in the experimental group by 4 percentage points. A χ^2 test yields a p value of 0.047, which is not significant with Bonferroni correct, $\alpha = 0.025$). Similarly, there was no significant difference when comparing the conceptual questions ($p = 0.26$).

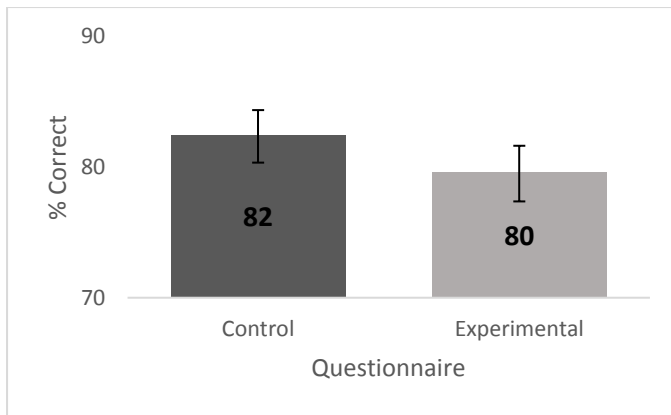


Figure 23: Comparison of Control and Experimental Groups' **Overall Scores**

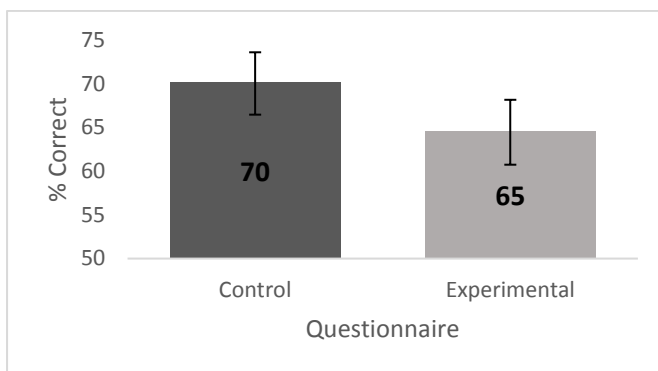


Figure 24: Comparison of Control and Experimental Groups' Scores for **Difficult Questions**

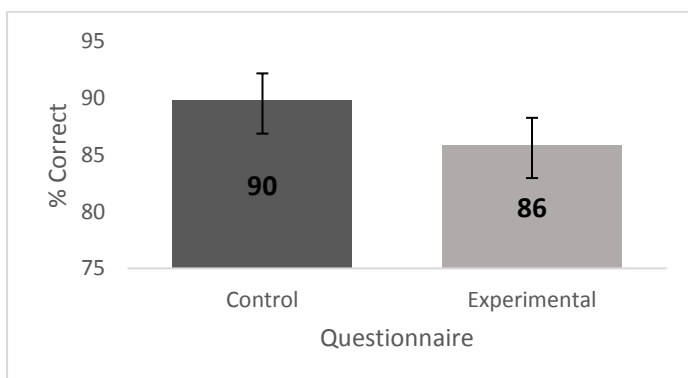


Figure 25: Comparison of Control and Experimental Groups' Scores for **Calculation Questions**

4.4.4. Discussion

As in Study 1, no significant difference was found between performance in the control and experimental questionnaires. However, the difference in performance does seem to be moderately correlated with the difficulty of the question. (The difficulty of the question is simply measured as

the number of students who got the question incorrect.) Figure 26 shows this correlation which has a general upward trend; the more difficult the question, the more likely there is to be a difference in performance between control and experimental groups. Correlation coefficient $R = 0.42$.

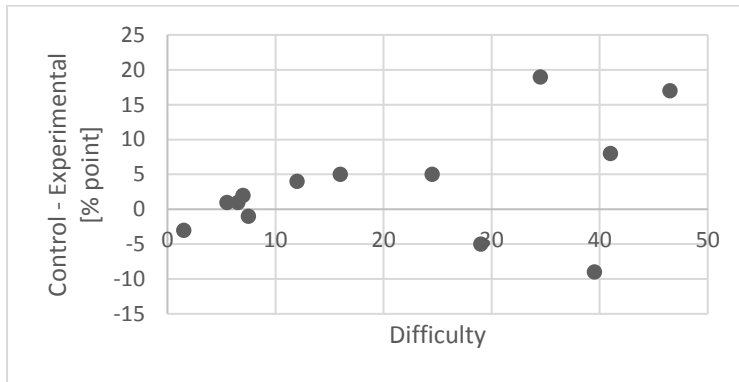


Figure 26: Correlation Between Difficulty of Question and Difference in Performance Between Control and Experimental Groups. $R = 0.42$ ($p < 0.001$)

Importantly, the hypothesis regarding the asymmetry between calculation and conceptual questions was not supported by the results of Study 2. In fact, the difference in performance for calculation questions was the reverse of what was observed in Study 1; the control group outperformed the experimental group.

Finally, all the above χ^2 comparisons assume parametric data. Figure 27 shows a histogram in which the number of students are binned according to the difference in their individual performance on control and experimental questions. The figure clearly shows that student performance is normally distributed around an average of 0.18, where a score of “0” means that a student answered a similar percentage of control and experimental questions correct.¹⁸ In Study 1, it was hypothesised that different types of students (e.g. Chemical Engineers versus Mechanical Engineers) might be affected differently by the presence of unfamiliar words. Figure 27 seems to indicate that all students are effected in a similar way, and on average they tend to perform marginally better on the control questions.

¹⁸ Note that half the students received 7 control questions and 6 experimental questions, and vice versa for the other half. These were normalised such that a score of “0” meant that a student got all questions correct. A positive score entails that the student answered a larger percentage of control questions correct.

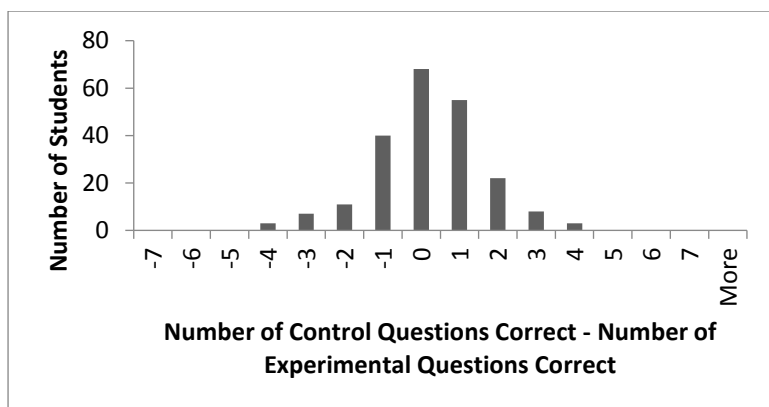


Figure 27: Histogram of Number of Students in Bins Associated with the Difference in Their Performance on Control and Experimental Questions

The only consistent finding to emerge across Studies 1 and 2, is that difficult questions seem to elicit a difference in performance on control and experimental questions. This might be explained by the dual processing theories. The slower processing system is fully engaged only when it is determined that the faster, associative processing system cannot provide a satisfactory solution. A “difficult question” may perhaps be defined as a type of question which is not easily answerable by means of our “associative machine” and engages the slower processing system.¹⁹ Thus, if a difficult question engages the slower processing system, and the unfamiliar words hinder this system (by hindering simulation), this would result in poorer performance. This also seems to be true for difficult calculation questions. (See question 3 in Figure 22).

One further important test performed in this study was to see if there was a relation between the effect of unfamiliar words and “physics ability”. Here “physics ability” is crudely defined as the mark obtained on the final exam. We obtained the exam scores for all students, and divided students into two groups; those who obtained above average exam results, and those students who obtained below average exam results. χ^2 tests of comparison on the control and experimental questions yielded p values of 0.2 and 0.09 for the below average and above average cohorts respectively. This shows that “physics ability” does not seem to be salient factor in determining which students might be more affected by the unfamiliar words.

Note that of the individual questions in Study 1 that were flagged as yielding potentially significant differences (these were questions 1,3,4 and 13), only question 3 showed any potential of yielding

¹⁹ Of course, it is possible for a question to be answered incorrectly through a “misfire” of the associative system. For example, a question might look easy but contain a twist, or a distractor.

such a difference in Study 2. Because many of the findings across Study 1 and Study 2 were not consistent, a final study was performed with a larger number of students.

4.5. Study 3: February 2014

4.5.1. Sample

The sample consisted of 420 freshmen enrolled in an engineering degree, taking a one year physics service course. All these students would have taken physics at high school.

4.5.2. Instrument

As in study 2, the control and experimental questions were equally distributed in two questionnaires. Because the proposed asymmetry between calculation and conceptual questions did not turn out to be significant, two calculation question of Study 2 (question 4 and 5) were discontinued. Table 7 shows the relationship between the questions in Study 2 and Study 3.

Study 2 Questions	Action	Study 3 Questions	Type	Scenario
1	retained	1	calculation	car accelerating
3	retained	2	calculation	car towing
4	discontinued		calculation	woman pushing box
5	discontinued		calculation	speed of boat
6	retained	3	conceptual	two balls fall from a height
7	discontinued		conceptual	same two balls roll off table
8	retained	4	conceptual	stone dropped from window
9	retained	5	conceptual	truck collision with car
10	retained	6	conceptual	woman pushing box
11	retained	7	conceptual	same woman doubles force
12	retained	8	conceptual	
13	retained	9	conceptual	ball thrown straight upward
2	retained	10	calculation	displacement in park

Table 7: Table Showing Relationship Between Questions in Study 2 and Study 3

4.5.3. Findings

Figure 28 shows the difference between the control and experimental scores for each question. None of the questions yielded a significant difference. (For example a χ^2 test for question 10 yielded

a p value of 0.067). Figure 29 shows the overall difference between the control and experimental questions. This three percentage point difference is statistically significant with χ^2 test p value of 0.022. Figure 30 shows a six percentage point difference for the five more difficult questions (questions 3, 8, 9, 10, 11). This is statistically significant with a p value of 0.005 (Bonferroni correction $\alpha = 0.025$). Separate χ^2 tests for calculation and conceptual questions did not yield significant differences between control and experimental results. (p values of 0.25 and 0.046 respectively). Further comparisons between English and non-English students, as well as Chemical Engineers and Electrical Engineers were made. No comparisons yielded significant differences.

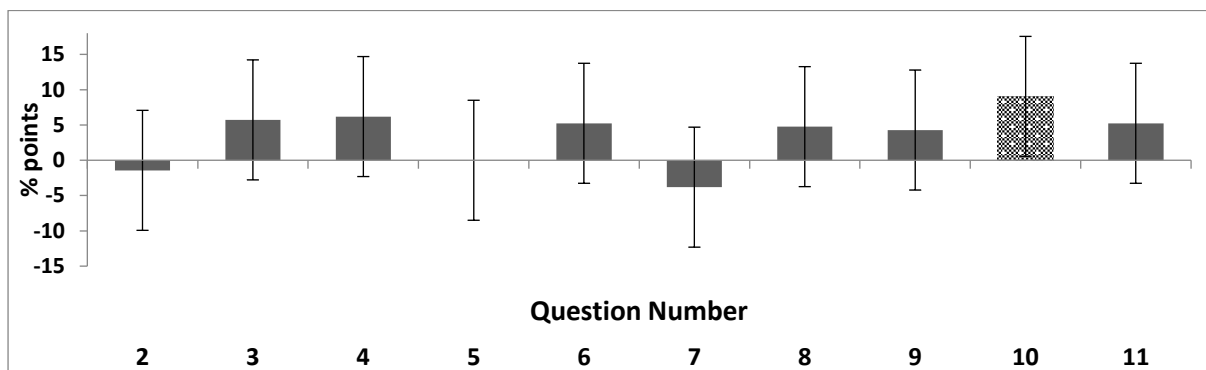


Figure 28: Difference between Control and Experimental $n = 420$

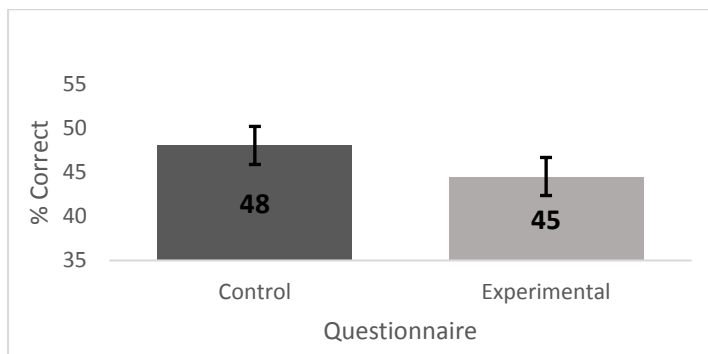


Figure 29: Comparison of Control and Experimental Groups' **Overall Scores**

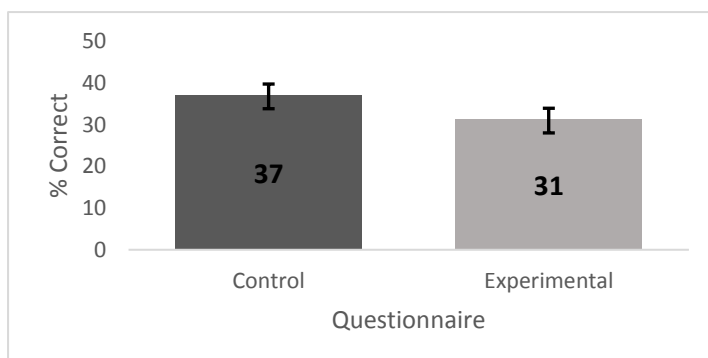


Figure 30: Comparison of Control and Experimental Groups' Scores for **Difficult Questions**

4.5.4. Discussion

The use of the larger sample size ($n = 420$) yielded satisfactory results that seem in keeping with general themes of the previous two studies. There was a significant difference between the overall scores of the control and experimental questions, and this difference seems to be mostly accounted more by the difficult questions. Figure 31 shows that there is a moderate correlation between the difficulty of the question and the difference in performance between the control and experimental questions; $R = 0.45$.

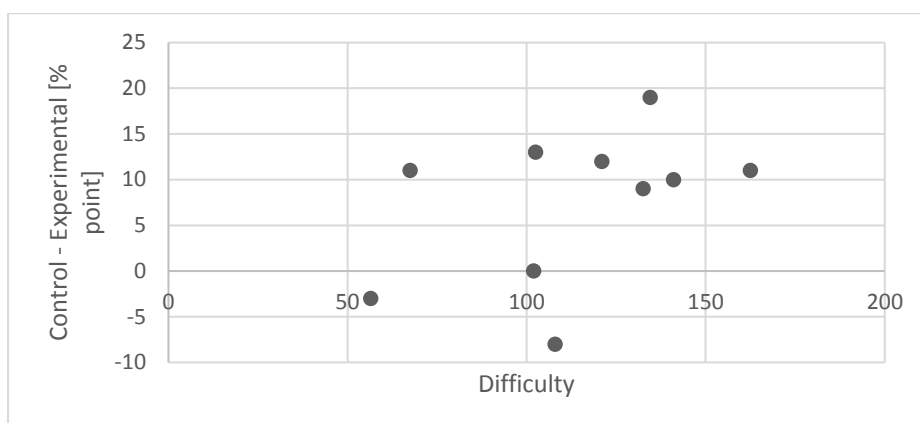


Figure 31: Correlation Between Difficulty of Question and Difference in Performance Between Control and Experimental Groups. $R = 0.45$ ($p < 0.001$)

4.6. Study 4: “Post Script” on Familiarity

A final study was done that tested for students’ familiarity with five questions from Study 2. While this was a different cohort of students to those involve in Studies 1,2 and 3, they are enrolled for the same service physics course as freshmen engineering students; $n = 224$. The questionnaire is shown in Appendix B. Students were instructed not to answer the questions, but simply to rate their familiarity on a Likert scale from 1 to 5; 1 being very unfamiliar, and 5 being very familiar. Two control questions were used; one involved quantum mechanics, and the other involved simple harmonic motion. These topics are not covered in high school physics and should be unfamiliar to the students. The results are shown in Table 8. The questions from Study 2 all receive high familiarity ratings, while the control questions receive very low familiarity ratings.

Study 2 Questions	Type	Scenario	Familiarity Rating
1	calculation	car accelerating	4.00
4	calculation	woman pushing box	5.00
6	conceptual	two balls fall from a height	4.75
9	conceptual	truck collision with car	5.00
10	conceptual	woman pushing box	5.00
control question		qm	1.00
control question		shm	1.50

Table 8: Familiarity Ratings of Five Questions taken from our FCI study ($n = 224$)

4.7. General Discussion and Conclusion

“My students understand basic mechanics”. What does this sentence mean?

Previous research has shown that students’ understanding of basic mechanics is sensitive to context. For example, previous research has shown that students are sensitive to textual contexts, such as substituting the word “table” for the word “spring”. This type of context sensitivity may be modelled in a similar way to the various textual sensitivities observed in the studies on vector addition of Chapter 2.3 – resource activation and simulation. The question asked of the studies in this chapter is: what if a textual prompt is completely unfamiliar and activates very few useful resources/semantic network/simulations?

Section 4.6 is an important addendum to this series of studies. It demonstrates that while the individual words, such as “skurrit”, might have been unfamiliar, the overarching context was very familiar. Consider the following sentence:

teh qukci bowrn fxo jmpued ovre hte layz dgo.

While the individual word forms might be unfamiliar to you, the sentence is probably sufficiently familiar that your “top down” processing allowed you to reconstruct: “the quick brown fox jumped over the lazy dog”.

This series of studies has not tested what it purported to test, as described in the Introduction. i.e. These studies do not test “unfamiliarity” in the sense that a newly introduced concept (such as “Hermitian”) would be unfamiliar to an upper division student. For the upper division student, the word (e.g. Hermitian) *and* the larger context (e.g. vector spaces) would be unfamiliar. Instead, these series of studies test for how students understand *very familiar* basic mechanics problems which contain unfamiliar words. And the main conclusion seems to be: students seem to perform only

slightly worse in familiar problems that contain unfamiliar words. There should perhaps be a further caveat here around *which* words are unfamiliar. For example, if we had substituted the verbs in the various sentences with nonsense words, most problems would become unintelligible, or, at the very least, ambiguous. The substitutions used in the study were of benign nouns.

How should we model students' understanding in these series of studies? The high familiarity rating of the questions means that we should be looking at types of reasoning processes, and framing. High familiarity entails fast, associative, linguistic processing, with minimal simulation (cf. the traxoline example in Chapter 2). The questions were sufficiently familiar that the unfamiliar words did not interfere with this fast thinking, associative processing. In other words, the questions were framed as being similar to many previously encountered questions, and a method of answering was ready to hand. The correlation of the difficulty of the questions with the effect of the unfamiliar words can also be explained using these models. A difficult question forces a student to slow down, to consider the question more closely, and, perhaps, to attempt to simulate the scenario. The unfamiliar word would hinder this simulation, and thus, adversely affect performance.

Note how the model of simulation differs in its contribution to "understanding" in this series of studies, compared with the "vector addition" studies of Chapter 3. When considering the difference between "total force" and "net force" in the previous series of studies, the slower, simulation system lead students to choose the *incorrect* answer. However, in this series of studies, it is assumed that an unhindered simulation system would *aid* students in arriving at the *correct* answer for the more difficult questions.

In conclusion: We hypothesized that the model of "LASS" would be relevant in considering the effect of unfamiliar words on student understanding. This study shows that it is crucial to consider other factors in this particular model, besides the unfamiliarity of individual words. Two salient factors were found to be the familiarity of the larger context, and the difficulty of the question at hand.

INTERMISSION: Prelude to Chapters 5, 6 and 7

Chapters 3 and 4 deal with students' understanding of particular physics concepts in different contexts. Chapters 5, 6 and 7 all deal with the concept of a mathematical group in different contexts. Like chapters 3 and 4, chapters 5, 6 and 7 deal with the context sensitivity of understanding, and the various models of we might use to better incorporate the notion of "context sensitivity" into educational research. An additional focus is on the notion of "an instantiation of a concept", and the accompanying notions of "sameness" and "difference". (i.e. two *different* instantiations of the *same* concept). Figures Figure 32Figure 33Figure 34 represent three of the main findings from chapters 3 and 4 using the notion of "instantiation" of a concept. Note that in each case, an expert would see both contexts (marked in **bold**) as instantiations of the same concept, while learners fail to see this "sameness".

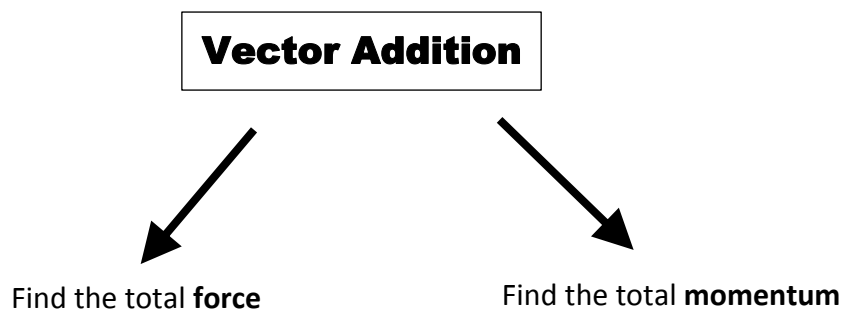


Figure 32: Two Different Instantiations of the Concept of Vector Addition in Two Different Contexts

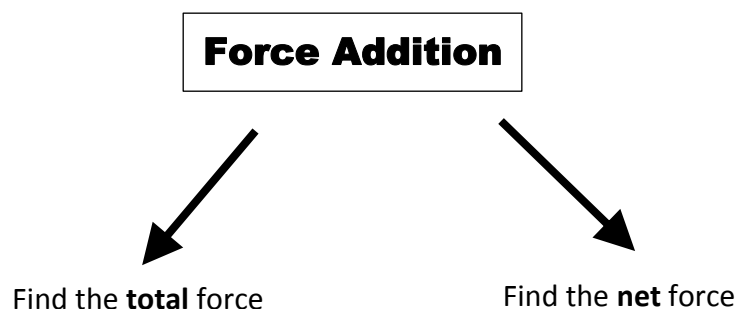


Figure 33: Two Different Instantiations of the Concept of Force Addition in Two Different Contexts

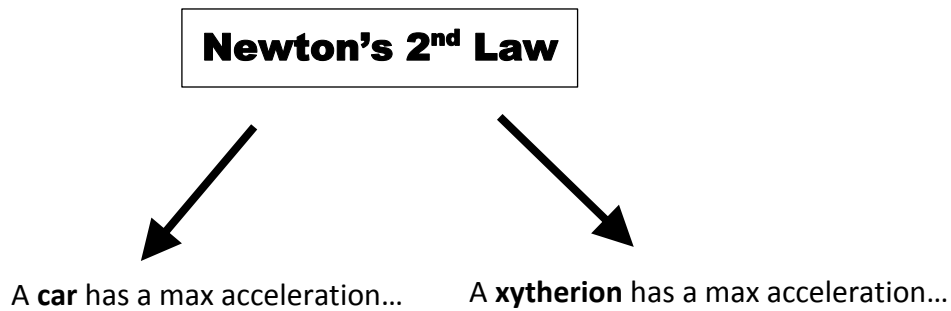


Figure 34: Two Different Instantiations of the Concept of Newton's 2nd Law in Two Different Contexts

Chapters 5, 6 and 7 centre on a series of studies by Kaminski, Sloutsky, & Heckler (2013) which deals with two instantiations of the concept of a mathematical group, one abstract and one concrete. (See Figure 35). Kaminski et al. investigate the efficacy of each instantiation as a learning tool to promote transfer of the concept of a mathematical group.

Chapter 5 presents the results of a replication and extension of their series of studies. Chapter 6 focuses on re-interpreting the results from their study (as well as results from Chapter 5 and De Bock et al. (2011)) from a Resources/Knowledge in Pieces perspective. Chapter 7 focuses specifically on how we might interrogate the notions of “instantiation”, “sameness”, and “difference” in physics and mathematics education research.

As noted in the “Notes to Examiners”, these chapters are intended to be three separate papers.

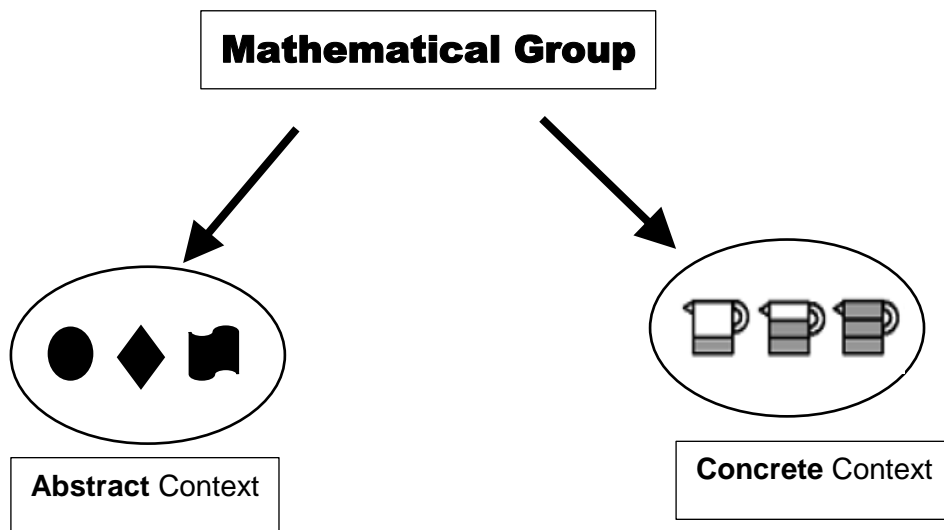


Figure 35: Two Different Instantiations of the Concept of a Mathematical Group in Two Different Contexts

CHAPTER 5: A Qualitative Study of the Transfer Experiments of Kaminski et al. and De Bock et al.

5.1. Abstract

A perennial issue in STEM education is the role of concreteness in the teaching of science concepts. After a series of experiments involving sequestered problem solving, Kaminski et al. concluded that presenting learners with abstract instantiations of mathematical concepts is more likely to result in successful transfer. Their results have been widely cited, and have received both positive and critical commentary, with the most comprehensive critique being De Bock et al.'s extension of the original study. Much of the critical commentary is concerned with the similarities/dissimilarities between what is learned in the learning domains, and what is tested in the transfer domain. This paper seeks to add to this discussion by replicating the study and providing qualitative data, analysed by an approach suggested by grounded theory. Our data provides new insight into possible differences between the learning domains, and raises questions around the appropriateness of using the abstract/concrete distinction in this particular context.

5.2. Introduction

This paper is an extension of the work done by Kaminski, Sloutsky and Heckler (2008, 2009, 2009a, 2013)²⁰ and De Bock et al. (2011). Kaminski et al.'s work focused on the effect of concreteness on transfer, and their key findings were notably published in *Science* in 2008. Contrary to the widely held belief that maths should be taught by means of concrete examples, Kaminski et al. argue that "students might be better able to generalize mathematical concepts to various situations if the concepts have been introduced with the use of generic [i.e. abstract] instantiations" (Kaminski et al. 2008, p.455). Their central experiment focuses on the learning and transfer of the concept of a mathematical group of order 3. Their findings have stirred considerable debate amongst education communities, eliciting both positive and critical commentary. Much of the critical commentary is concerned with the similarities/dissimilarities between what is learned in the learning domains, and what is tested in the transfer domain. (See Mourrat (2008), Cutrona (2008), McCallum (2008), De Bock et al. (2011), Jones (2009a), Jones (2009b)). The only critique which provides further data is a comprehensive study conducted by De Bock et al. (2011) which both replicates and extends Kaminski et al.'s central experiment, and provides results which may be interpreted as contradicting Kaminski et al.'s central claim. De Bock et al. provide both quantitative and qualitative data. Their

²⁰ Also (Sloutsky et al., 2005) and (Jennifer A. Kaminski & Sloutsky, 2012)

qualitative data consisted of written responses to an open-ended question, and probed what students actually learned in the learning domains. Student responses are analysed by a scoring system consisting of pre-determined categories. Our study seeks to add to the discussion by providing further qualitative data analysed using an approach suggested by grounded theory (Strauss & Corbin, 2007). In our approach, student responses are coded according to the key ideas expressed. These codes are later grouped into super-ordinate categories. In other words, in our qualitative analysis the categories *emerge* from the student responses, rather than being imposed thereon. This analysis provides new insights into the differences between the learning domains.

5.3. Key Results from Kaminski et al. and De Bock et al.'s Experiments.

Kaminski et al. and De Bock et al.'s central experiment is represented graphically in Figure 36. Two student cohorts were assigned to two different learning domains, and their performance was measured in a shared transfer domain. Both learning domains are posited to help students learn the mathematical concept of a commutative group of order 3. One learning domain is thought to be an abstract (or generic) instantiation of a mathematical group, while

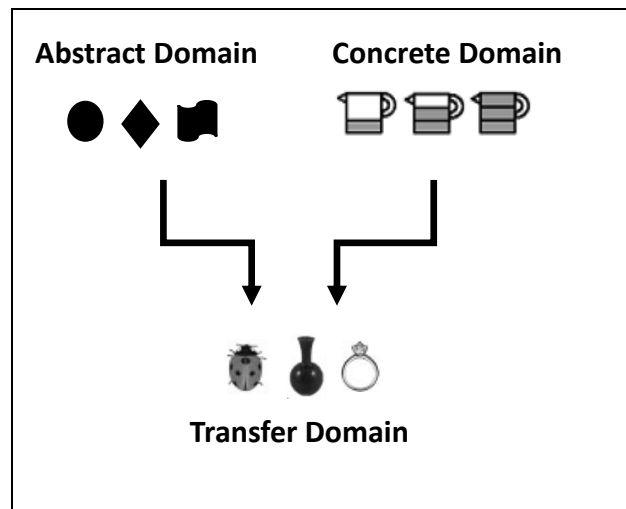


Figure 36: Schematic Representation of Kaminski et al. and De Bock et al.'s Experiment

the other is thought to be a concrete instantiation. The learning domains could vary in two ways: (i) the objects used, and (ii) the storyline used. In their central experiment, the abstract learning domain makes use of geometrical figures, while the concrete learning domain makes use of containers filled with a solution. See section 5.5 below for further details regarding these domains.

Kaminski et al. found that students in the abstract learning domain performed significantly better in the transfer domain than students from the concrete learning domain. (Mean transfer test scores: abstract domain ~79%; concrete domain ~54%). Furthermore, they did not find a significant difference between the performance of those in the concrete learning domain and those in a baseline (without learning) condition. They conclude: “These findings suggest that the concreteness of the learning instantiation hindered participants’ ability to align the common structure of the learning and transfer domains,

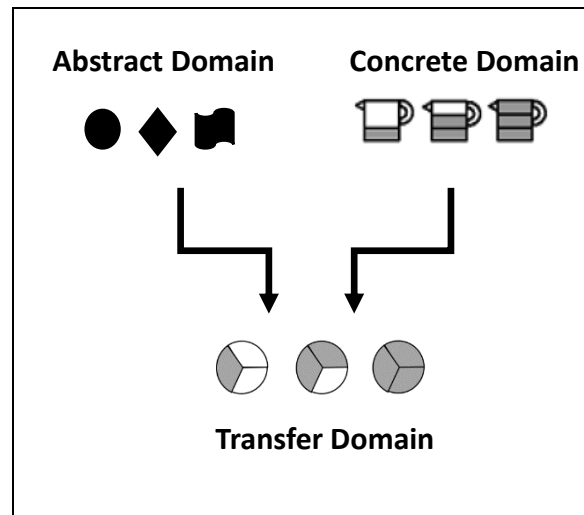


Figure 37: De Bock et al.'s Extension of the Original Experiment

which in turn led to transfer failure” (J. a Kaminski et al., 2013, p.21.) Similarly, de Bock et al. found that students in the abstract cohort significantly outperformed those in the concrete cohort in the transfer domain. (Results for their transfer domain were 75% and 50% respectively).

De Bock et al. then extended Kaminski et al.’s study by using an alternative transfer domain (see Figure 37). Here they found the opposite result to the original experiment: the concrete cohort significantly outperformed the abstract cohort in the transfer test. (The transfer test results being 84% and 73% respectively). In other words, De Bock et al.’s extension of Kaminski et al.’s study shows that student performance in the transfer test is not due to the concreteness of the learning domains *per se*, but rather due to an interaction between the learning domain and the transfer domain.

De Bock et al. also collected qualitative data by probing students’ problem solving strategies in both learning domains. They presented students with a typical problem from the learning domain, and asked them to explain how they arrived at their answer. Student responses were analysed using a pre-determined scoring system consisting of four categories: (i) “Group”, (ii) “Modulo”, (iii) “Rules”, (iv) “No”. Responses were therefore categorized according to whether they contained mention of: (i) one (or more) of the axioms of a commutative group; (ii) properties of modulo 3 arithmetic; (iii) combination rules (almost) literally repeated; (iv) no explanation. They found that 65% of students in the concrete learning domain mentioned some aspect of modulo 3 arithmetic in their explanations, while none of the students in the abstract learning domain did. They also found that 94% of students in the abstract learning domain contained an (almost) literal repetition of the combination rules, compared with only 10% of students in the concrete learning domain.

5.4. Our Study

Our experiment was almost identical to that of Kaminski et al. and De Bock et al. described above: two student cohorts were assigned to two different learning domains. One learning domains is thought to facilitate the learning of a mathematical group via abstract instantiations, while the other learning domain is thought to use concrete instantiations. The main purpose of the study was to elicit what types of strategies students used to solve the problems.

5.5. Instrument

The abstract learning domain included the following text:

“On an archaeological expedition, tablets were found with inscriptions of statements in a symbolic language. The statements involve these three symbols: ● ◆ ◀ and follow specific rules.”

The rules are shown in Figure 38 and include associativity, the existence of an identity, the various rules of combination, and commutativity.

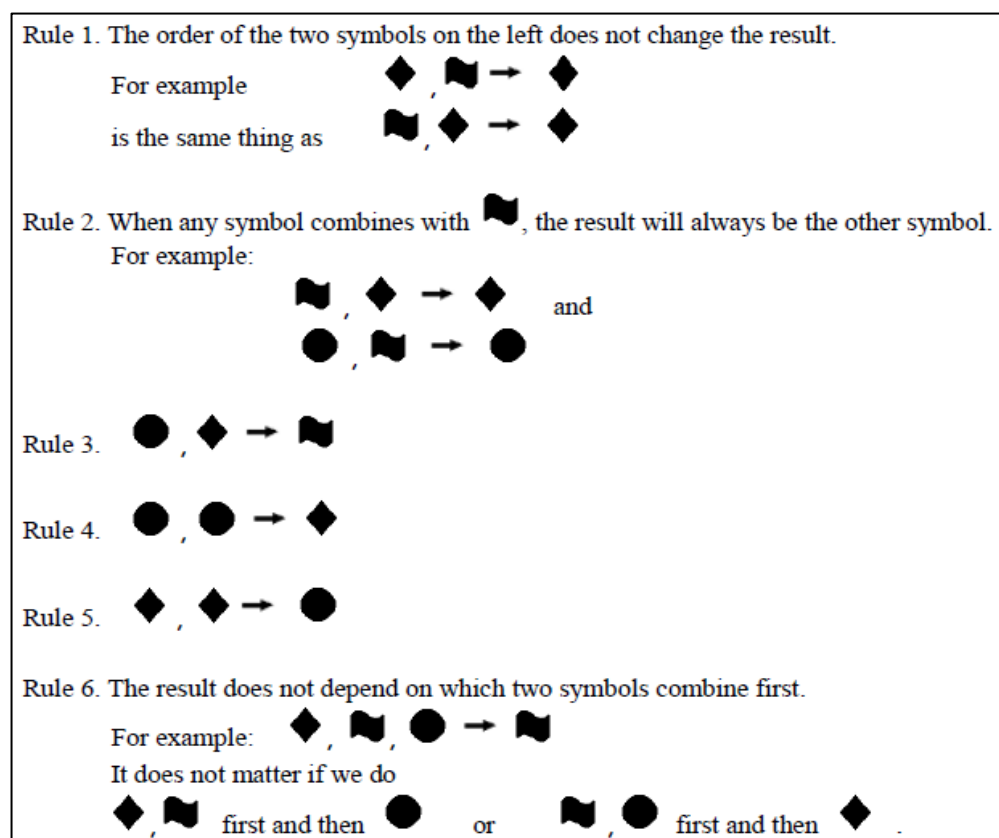
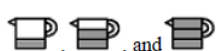


Figure 38: Rules Used in Learning Domain 1. Taken from Supporting Online Material (SOM) (Jennifer A. Kaminski et al., 2009b)

The storyline for the concrete learning domain included the following text:

“A company makes detergents by mixing three different quantities of solutions, represented as:



The company is testing the mixtures and wants to know what amount of solution is left-over in the mixing process.”

The rules are isomorphic to the abstract learning domain. Three of these rules are shown below in Figure 39.













<p>Rule 1. The order by which two cups of solution are combined does not change the left-over result. For example, combining  with  has a left-over quantity of . And combining  with  has the left-over quantity .</p>
<p>Rule 4. A combination of  and  does not fill a container, so the left-over is .</p>
<p>Rule 5. A combination of  and  fills one container and has  left-over.</p>

Figure 39: Three of the six rules used for Learning Domain 2. Taken from SOM (Jennifer A. Kaminski et al., 2009b)

Twelve questions were asked in each learning domain, and these questions were isomorphic (see Figure 40).













<p>2. Find the resulting symbol: , ,  → _____</p> <p>Choose: 1.)  2.)  3.) </p>
<p>2. What is left-over when the following cups of solution are combined?</p> <p>  </p> <p>Choose: 1.)  2.)  3.) </p>

Figure 40: Isomorphic Questions from Learning Domain 1 and Learning Domain 2 Respectively. Taken from SOM

Although our instrument was developed independently of De Bock et al. (2011), it shared an additional feature; an additional transfer domain. The first transfer domain is labelled the “abstract” transfer domain, and is identical to that used by Kaminski et al. The storyline provided was that of a children’s game:

“In another country children play a pointing game that involves these three objects: Children point to objects and the winner points to the correct final object. The rules of the last system you learned are

like the rules of this game. So use what you know about the last system to help you figure out the rules of this game.”

The rules for the game are shown in Figure 41.

If the kids point to these:	 	 	 	 
Then the winner points to this:				

Figure 41: The Rules of Combination for the Transfer Domain. Taken from SOM.

The second transfer domain was labelled the “concrete” transfer domain. It involved the same storyline as the “abstract” transfer domain, but used the symbols shown in Figure 42. Twelve isomorphic questions were asked in each transfer domain.













If the kids point to these:	 	 	 	 
Then the winner points to this:				

Figure 42: Rules of Combination for the Transfer Domain in Cohorts AC and CC

Thus students were assigned to one of four domains as shown in Figure 43. We have followed the labelling convention of De Bock et al. where “A = Abstract” and “C = Concrete”. We have also followed their convention in referring to the original transfer domain (ladybird, vase, ring) as “abstract”, and the new transfer domain as “concrete”. For example, “AC” refers to the cohort of students in the “abstract” learning domain, and the new “concrete” transfer domain. Note that this labelling convention is used here only for convenience. The appropriateness of these labels will be addressed in the Discussion Section 5.9.1 below.

























	AA	CA	CC	AC
Learning Domains	  	  	  	  
Transfer Domains	  	  	  	  

Figure 43: Objects Used in the Four Domains of Our Experiment

Together, the AA and CA cohorts constitute the original experiment performed by Kaminski et al. The two additional cohorts (CC and AC) were given the new transfer domain. The objects used in our

transfer domain are groups of three circles as shown in Figure 43 and Figure 42 . The storyline used was identical to the original transfer domain: “... *Children point to a sequence of symbols, and the winner points to the correct final symbol*”. Note that the rules of symbol combination follow modulo-3 arithmetic. This is similar to the arithmetic of the concrete learning domain (jugs).²¹

The question of primary interest occurred at the end of our instrument, and asked students to comment on any particular strategy they adopted in the learning and transfer domain. The question was deliberately open ended to encourage students to describe in their own terms how they approached solving the problems in the various domains. Finally, students’ beliefs about the similarity of the learning and transfer domains were also probed. This is explained fully in Judgements of Similarity Section 5.8.3 below.

5.6. Methodology

258 freshman engineering students participated in our study which was held during the first 20 minutes of their physics lecture. The questionnaires were randomly distributed, and the questions were answered under exam conditions. There were four sections to each questionnaire: (i) the rules for the training domain; (ii) twelve training domain questions; (iii) twelve testing domain questions; (iv) reflection and feedback. Students were given precise time limits for the various sections of the questionnaire: 2 min reading time for the rules of the learning domain; 6 min for answering the questions in the learning domain; 6 min for the transfer domain; 6 min for reflection. At the end of each time period, students were instructed to turn to the next section of the questionnaire, regardless of whether they had completed the section they were currently working on. They were instructed not to turn back to previous sections. On average, students answered ten questions in each task. Students were only chosen for analysis if they had answered at least seven of the twelve questions in each task, reducing the number of students in the analysis to 232.

Importantly, the aim of the qualitative analysis was not to determine whether students’ responses could be categorised according to pre-determined categories ala De Bock et al. Instead, our aim was to identify patterns in the variety of problem solving strategies used by students. Therefore, students’ comments were analysed using an approach suggested by grounded theory. After transcribing all students’ free responses, the “key points” or “main ideas” were extracted from each response. These main ideas were assigned codes. Over 30 codes were assigned in the original

²¹ Note that the objects and storyline used in our transfer domain were different to those of De Bock et al. The objects used in De Bock et al.’s transfer domain are burnt slices of pizza, and the storyline involves the rather odd scenario of determining the number of burnt slices of pizza from a pizza order. Importantly, the combination rules for the symbols in both our and De Bock et al.’s transfer domains followed the rules of modulo-3 arithmetic.

analysis. In other words, we identified around 30 different “key ideas” regarding students’ strategies involved in solving problems. Patterns of similarity were then sought, and codes were grouped into categories. These categories were, in turn, grouped into super-ordinate categories.

5.7. Quantitative Results

As Figure 44 illustrates, the data were not normally distributed. This was confirmed in Shapiro-Wilk’s tests where $W < W_{critical}$ for all four transfer conditions. This meant that a comparison between the transfer conditions was not meaningful, as shown by the error bars in Figure 45. The error bars represent 1σ above, and 1σ below the mean. Mann -Whitney tests for non-parametric data yielded non-significant differences between both pairs of learning and transfer conditions.

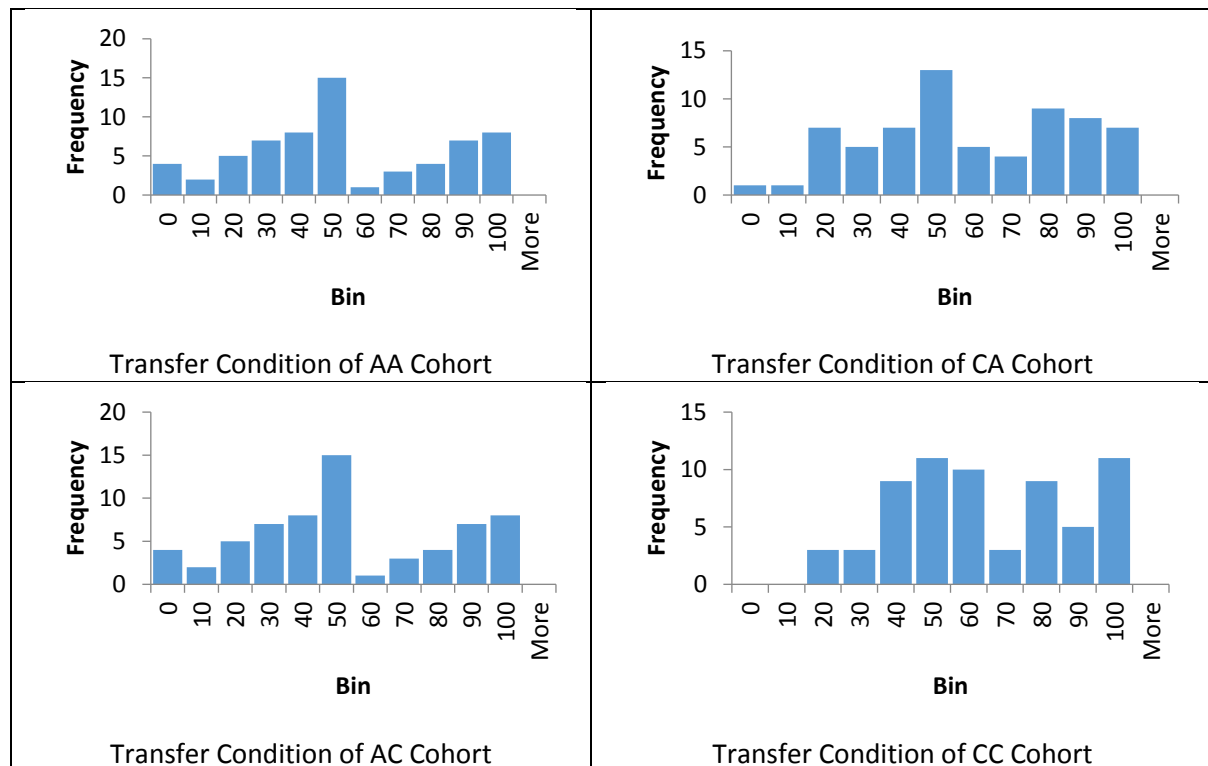


Figure 44: Histogram of the Transfer Conditions of the Various Cohorts

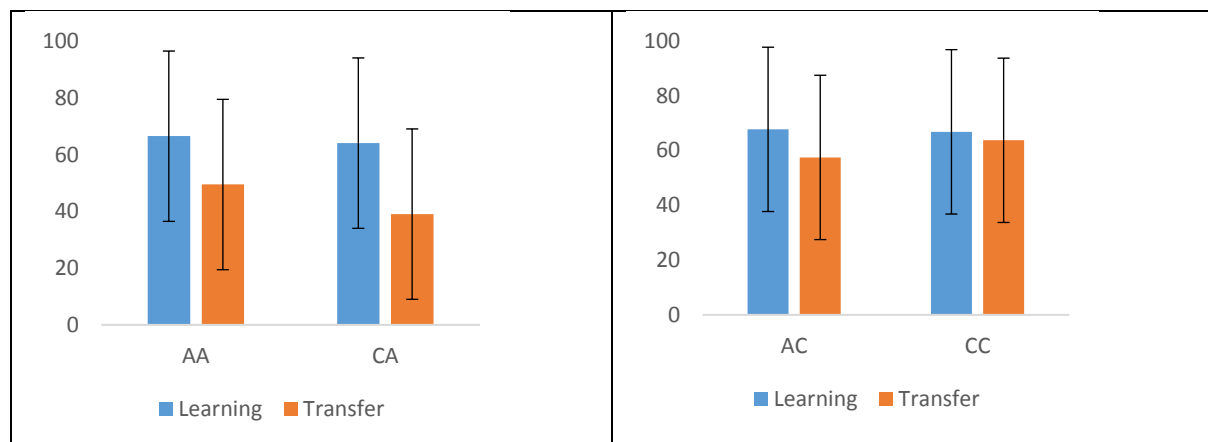


Figure 45: Comparison of Learning and Transfer Conditions for the Various Cohorts

5.8. Qualitative Results

One of the most striking findings from the qualitative results were the number of different strategies used by students in solving both the abstract and concrete tasks. However, after coding for over 30 different ideas present in students' responses, two dominant patterns emerged. The first was the explicit or implicit mention of some type of symbol association. i.e. symbol combination that did *not* involve counting or addition. Examples include: "Both games involved the combination of patterns of symbols" (AC68); "Combine two of the same to get the other" (AA05); "The idea of combining symbols in specific combinations to create certain outcomes" (AA56); "It all had to do with combinations of different visual images to produce a final result" (CA11); "You had been given a set of basic patterns (i.e how different objects combined together) and asked to make judgments on combinations of objects" (AC17). These type of responses were found predominantly in the AA cohort, and seldom found in the CC cohort (see Figure 46).

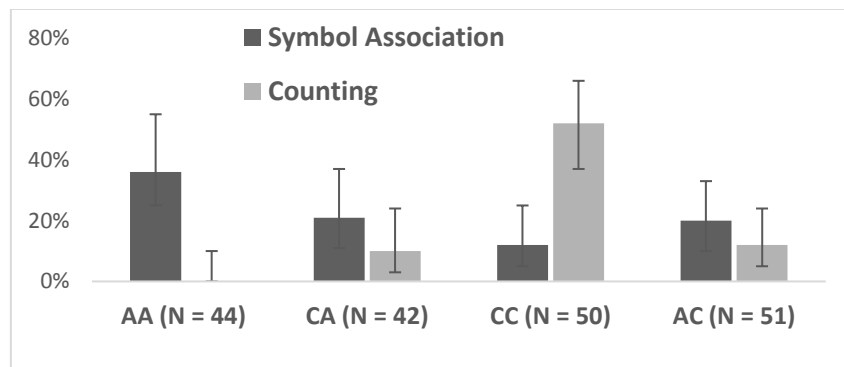


Figure 46: Percentage of students in different cohorts whose free writing responses were categorised as exhibiting characteristics of "symbol association" and "counting".

The other dominant pattern of responses to emerge concerned the strategy of counting or addition: "Count the dark spots and then divide by three" (AC40). "It was all about addition of the circles shaded or the level of the detergent" (CC09); "[the tasks] worked on the same counting principle" (CC44); "It used the same method of 'addition' in each case." (CC45). These type of responses often also included a mention of a leftover or remainder: "Both tasks consist of adding amounts and finding the 'left over'" (CC32); "I thought both games worked with adding the shaded areas and giving the leftover" (CC63); "...add up and find the remainder" (CC40). These responses were overwhelmingly found in the CC cohort, and none were present in the AA cohort (see Figure 46). Certain students explicitly mentioned using both of these strategies: "The first one seemed to be based on addition whereas the second seemed based on a combination system" (CA34). As a point

of clarification, a handful of responses were ambiguous as to whether the student had used counting or symbol association. e.g. “Both games involved combining two or more things” (CC55) These responses were coded as using a “symbol association” strategy unless the student explicitly mentioned some type of counting strategy. For example “Both games involved combining two or more things to get a total number and a leftover” would be coded as using a counting strategy.

5.8.1. The Counting Strategy May Be Predictive of Success in the CC Cohort

As mentioned in the description of our instrument above, the storyline used in both transfer domains framed the activity as a pointing game. In the transfer domain of the AC and CC cohorts the symbol combination of the three circles shared a similar arithmetic (modulo 3 arithmetic) to the jugs. However, because the transfer domain was framed as a pointing game, students may not have detected this underlying arithmetical structure. Instead, they may have framed the transfer domain as a type of “symbol association” game. This is exemplified in the following response: “In the first task, we were to fill cups and check for left overs, while in the 2nd task, we were to correlate symbols” (CC49). Figure 46 shows that only 12% of students in the AC cohort used a counting strategy. This implies that the AC cohort were particularly liable to miss the underlying arithmetical structure of the “three circle” transfer domain. Consider the following response from a student in the AC cohort who uses a symbol association strategy in both the learning and transfer domains: “Both [A and C domains] required knowledge of matching symbols but the second task [C transfer domain] was more complicated as the symbols were similar.” (AC01). In other words, both symbols association and counting strategies are used by students in the “concrete” transfer domain. Of particular interest is whether adopting a particular strategy was predictive of success in a particular domain.

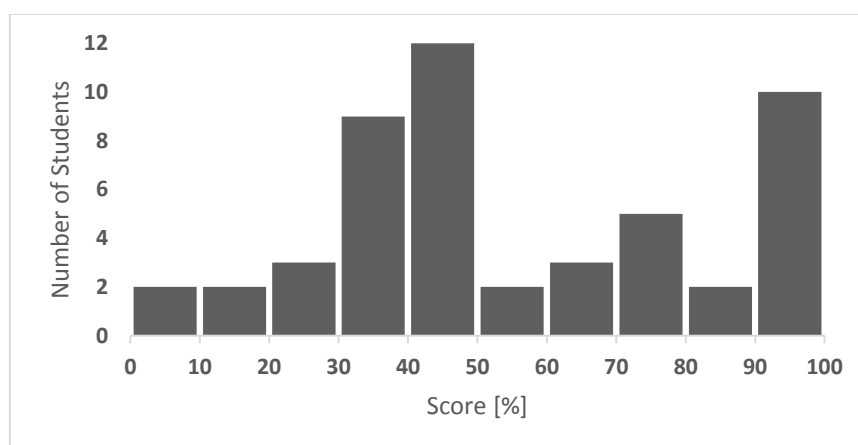


Figure 47: Histogram for “Concrete” Transfer Domain of CC Cohort


Consider the histogram for the transfer domain of the CC cohort in Figure 47 above. Note its striking bimodal character. Of the 20 students who scored above 60%, 17 explicitly mentioned using counting or addition as a strategy for answering the questions (the other three students did not mention a particular strategy). Indeed, in the transfer test, the 26 students (52%) who explicitly mentioned counting/addition obtained an average score of 71%, while the remaining 24 students averaged 40%. The use of a counting strategy thus seems to be a salient factor in determining student success in the concrete transfer task. Other findings agree with this hypothesis: (i) In cohort AC, the six students (12%) who explicitly mentioned using a counting strategy for the transfer test obtained an average score of 75% in their transfer test, significantly above the cohort average of 45%; (iii) In cohort CC, four students explicitly mentioned *not* using a counting strategy in the transfer test. They averaged 21% for the transfer test.

5.8.2. A Variety of Strategies

As mentioned above, there were several strategies used by the students in solving problems in the various domains. For example, one pattern of responses to emerge was the animating of symbols with causal characteristics: “Two ‘strong’ objects that when combined gave you the flag” (AA15); “I tried to determine which of the symbols were most ‘dominant’ in an operation and memorise them” (AA29); “...there was a passive object and two dominant objects” (AA37).

Of particular interest were students that explicitly mentioned a mapping between the objects in the two domains. This mapping was made in several different ways. Some students made explicit connections between objects: “The ring behaved like the [flag] being overridden by any other. The jug and the beetle behaved like the circle and the diamond” (AA06); “The way in which the rules worked were the same the images were just different but they had an exact counterpart”(AC24); “The jugs were filled in thirds, just as the triangles were. Therefore they have the same set of values, just

represented differently” (CC38); “If the volume of the cup was split into 3 smaller cups then the tasks would be identical (CC39)”. Some students simply drew a diagram of this mapping as shown in Figure 48. However, other students assigned symbols. Some assigned letters: “try to break down the

code into letters. e.g. X,y,z, like  = x” (AA29). One student assigned two words to the three objects: “Yes+No = No; Yes + Yes = No ;

No + No = Yes” (AA53). Another included a multiplication strategy: “ $A \times B = C$; $A \times A = B$; $B \times B = C$; $A \times C = A$; $B \times C = A$ ” (AC61). Multiplication strategies were also used with the assignation of numbers: “I set

the  to be like 1. So that multiplying by it gives the same result.” (AC27).

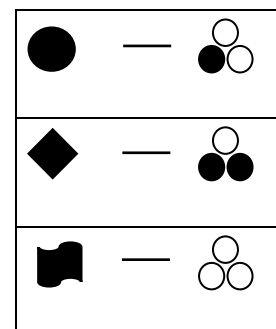


Figure 48: Mapping of symbols used by student AC34.

There were also a variety of ways in which the counting strategy was implemented. Some students used cancellation “I first located symbols that cancel out. The excess is my final amount.” (CC56); “One can simply cross out all the [open circles], then add the rest (CC09)”. “I used elimination.” (AC27). Some proposed adding all the dots and then divided by three: “Finding total, and dividing by three” (CA54). Others noted this strategy, but preferred another: “Maybe I could have added everything and divided by 3, but it would probably have taken more time.” (CC56).

These results demonstrate that it is difficult to deduce what students *learned* from the different domains. Instead, they indicate that students used a variety of creative problem solving strategies.

5.8.3. Judgements of Similarity

After students had completed the mathematical tasks, they were asked to reflect on the similarities (if any) of the learning and transfer domains. The question was posed as follows:

Three Students have just completed the tasks you have completed, and have the following conversation:

Student A: *I thought the principles involved in the first task were identical to the principles involved in the pointing game that followed.*

Student B: *I thought the principles involved in the first task were vaguely similar to the principles involved in the pointing game, but I don't think they were identical.*

Student C: *I thought the principles involved in the first task were very different from the principles involved in the pointing game.*

With whom do you most closely agree? Explain your choice.

The results are shown in Figure 49 below. There are no clear conclusions to be drawn from these results. Compare the results across the different cohorts: there is no significant difference between how students answered in the different cohorts. One might expect students in the AC and CA cohorts to think of their domains as dissimilar, and vice versa for the AA and CC cohort. However, consider that only 6% of the AC cohort claimed that the learning and transfer domains were very different, while 13% of the CC cohort thought the domains were very different. The key issue may be that students are not given a well-defined metric of comparison.

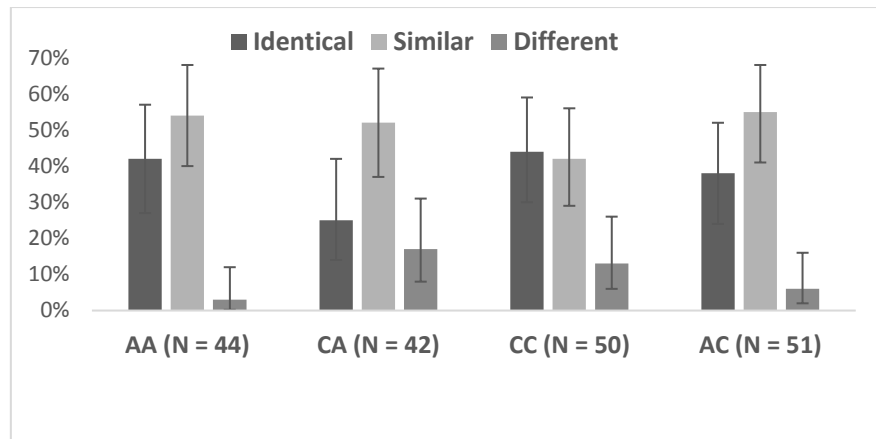


Figure 49: Percentage of Students in Different Cohorts to Rate the Learning and Transfer Domains as Either Identical, Similar or Different.

5.9. Discussion

5.9.1. Metrics of Comparison: Concreteness or Counting.

We will argue that concreteness is not sufficiently well defined, nor sufficiently predictive of performance, to warrant its use as a metric of comparison between domains.

The notion of concreteness is defined differently by different researchers. See Appendix A for a review. Kaminski et al. define the level of concreteness of an instantiation as “the amount of information activated in the mind of the observer”. That is, a concrete instantiation activates more information in the mind of an observer than an abstract instantiation. However, using the same working definition for concreteness, DeBock et al. challenge the construal of the original transfer domain as being concrete:

Kaminski et al. call the children’s game transfer domain a concrete domain, based on the superficial characteristic that “perceptually rich elements” are used. However, at a deeper level it satisfies those researchers’ own definition of generic (abstract): “Instantiations that communicate minimal extraneous details, beyond the defining structural information, are generic instantiations, while those that communicate more extraneous information are concrete.” p.116

Referring to another of Kaminski et al.’s experiments, Jones (2009a) claims that Kaminski et al. unwittingly use a learning domain which is concrete according to their working definition, and yet promotes transfer: “Kaminski et al. have found a concrete instantiation that *does* promote transfer”. And Cutrona (2008) refers to their definition of concrete as “a somewhat confused distinction”.

“Concreteness” does not seem to correlate with performance in this particular study, as evinced by the following three findings: (i) De Bock et al.’s quantitative results demonstrate that student performance is not tied to the concreteness of a domain *per se*, (ii) De Bock et al.’s qualitative results show a clear dissociation between the use of modulo-3 arithmetic between domains, (iii) our qualitative results demonstrate that two dominant strategies (symbol association and counting) are used by students in the two different domains, and there is strong evidence to suggest that it is the use of a particular strategy that is predictive of performance.

Thus, given the possible ambiguity of the notion of “concreteness” and the lack of its predictive power in this particular study, we propose that the “A” and “C” labelling convention used by De Bock et al. (AA, AC, CC, CA) is more appropriately interpreted as referring to domains involving predominantly (Symbol) “Association” or “Counting”, rather than “Abstract” or “Concrete”.

5.9.2. Sameness, Difference and Grain Size

As mentioned in the Introduction Section 5.2, most of the critical commentary of Kaminski et al.’s studies is concerned with the similarities/dissimilarities between the learning and transfer domains. Therefore, the notions of “sameness” and “difference” require close analysis, and will be explored in more detail in Chapter 7. In this section we introduce the inherent ambiguities of “sameness” and “difference” as well as the notion of “grain size”.

Are the numbers “5” and “6” the same or different? They are both real, rational and positive. But 5 is odd and 6 is even, and they have different magnitudes. Judgements of sameness and difference are meaningless without a well-defined metric of comparison. Are the learning domains in our study the same or different? It depends on the metric of comparison. An expert mathematician can show that the mathematical structure of the two domains is isomorphic. However, Figure 49 indicates that this similarity is not readily observed by students. Marton (2006) predicts this opacity of the underlying mathematical structure to novice students:

“In order to grasp a general principle by handling an instance of that principle, the general principle must be discerned by separating it from the specific instance in which it is embedded ... The main point is that in order to discern the general principle to be used in the second problem [in a transfer task], empirically at least two examples are needed. The traditional idea of transfer – learning something in Situation A (discerning a general principle) and using it in Situation B – is logically untenable.” (Marton, 2006, p.513)

We have suggested that, from a pedagogical perspective, the two domains are different, owing to the different strategies adopted by students due to the different affordances of the two sets of

objects. We have further suggested that “concreteness” is not sufficiently well-defined nor predictive metric to be useful in this particular study. Instead, we have suggested that the strategies or activities in which students are engaged is a useful metric of comparison. However, this metric also has its complications. Consider the “counting” strategy described above. It is clear that students adopted a variety of *different* counting strategies (cancellation, adding up, etc.) which were all grouped together under the same heading, “counting”. In other words, we chose a particular grain size at which to define our metric of comparison. The motivation for choosing a particular grain size is explored more fully in Chapter 6, where a Knowledge in Pieces (Hammer et al., 2005) theoretical framework is introduced. Suffice to say that the choice of grain size should result in categories which have explanatory value. This does seem to be the case for “counting” and “symbol association” as evinced by Figure 47.

5.9.3. Novice and Expert

De Bock et al. conducted a qualitative analysis in which student responses were scored according to pre-defined categories. For example, they scored student responses according to whether they mentioned some property of modulo-3 arithmetic. They found that 56% of student responses from the concrete learning cohorts (i.e. CC or CA) mentioned some property of modulo 3 arithmetic, while 0% of student responses in the abstract learning cohorts (i.e. AC and AA) did. This is comparable to our result in which 52% of students in the concrete learning cohorts (CC or CA) mentioned some aspect of “counting” while 0% of students in the abstract learning cohorts (AC and AA) did. (See Figure 46). The difference between our results is that, by using an approach suggested by grounded theory, we did not categorize student responses according to pre-existing categories, but allowed the categories to “emerge” from the data.

“There is a huge ontological gulf between, on the one hand, protocol coding categories and, on the other hand, knowledge element types or system configurations. Yet our research techniques have yet to clearly distinguish these.” (Disessa, 2002, p.37)

Ours is clearly a “novice-centred” approach. Kaminski et al. consistently emphasized that the abstract and concrete tasks were “mathematically identical” (Reply to Mourrat, 2008). Thus, from an experts’ point of view, the students were engaged in identical activities involving a mathematical group of order 3. In the words of Sfard (2001): “What we seem to need is a conceptualization that would help to account for the fact that the [novice] is unable to see as *the same* what [the experts] are unable to see as *different*.” (Sfard, 2001, p.13). In developing a metric for measuring difference, we advocate the use of an approach which is “novice-centred”. Even though our qualitative results

are comparable to those of De Bock et al., our question was not whether student responses could be categorised according to “expert concepts” (e.g. modulo-3 arithmetic), but how students describe their particular activities from their novice perspective.

5.10. Conclusion

Kaminski et al. performed a series of experiments to test the role of concreteness in the transfer of mathematical concepts. Many commentators have critically engaged with the results of these experiments, but few have provided additional data. De Bock et al. replicated and extended Kaminski’s experiment. They provide qualitative data in which student responses to the question “How did you arrive at your answer?” are coded according to whether they exhibited features of a mathematical group (associativity, commutativity, inverse and identity), or features of modulo 3 arithmetic. They found a clear dissociation. Student responses in the so called “concrete” learning domain tended to exhibit features of modulo 3 arithmetic, while student responses in the “abstract” learning domain tended to exhibit features akin to the axioms of a mathematical group.

Our study adds to this discussion by investigating what strategies are being used by students in solving the problems in the learning and transfer domains. Our qualitative data is analysed by an approach suggested by grounded theory. Instead of categorising student responses according to pre-defined categories, we allowed the categories to emerge from the qualitative data. Our main finding was that students use a variety of different strategies in solving these problems. After grouping student responses into superordinate categories, we found two productive superordinate categories were “counting” and “symbol association”. Similarly to De Bock et al. we found that these strategies were strongly associated with particular domains; “counting” with “concrete” and “symbol association” with “abstract”. Kaminski et al. claim that both learning domains facilitate the learning of the concept of a mathematical group of order 3 (Kaminski et al., 2013, p.17) However, it is not clear which (if any) of the strategies used by students constitute this understanding. What is clear is that students are engaged in different types of activities in the “concrete” and “abstract” domains. These different activities seem to be determined by the affordances of the different types of objects. (i.e. the jugs have an affordance of counting, while the “circle-diamond-flag” patterns do not).

In conclusion, Kaminski et al claim: “The present findings suggest that in the absence of any *overt similarities* between base [learning domain] and target [transfer domain], generic learning domains can outperform more concrete ones with respect to analogical transfer.” p.26. (Our italics). We suggest that the notion of “similarity” requires interrogation. While both learning domains might be regarded as mathematically isomorphic, there are dissimilarities in the strategies used by students to solve problems in the two domains. It is these differences in strategy that may be the key

explanatory variable in student success in the transfer domain, rather than any measure of concreteness.

CHAPTER 6: A Knowledge in Pieces Perspective of Analogical Transfer in the Studies of Kaminski et al. and De Bock et al.

“Combining trends toward increased contextual dependency, toward multiplicity and toward smaller grain size suggests that an application of a concept is likely to be better viewed as the selected activation of particular concept subcomponents, depending on context.” (DiSessa, 2002, p.33)

6.1. Abstract

A perennial issue in STEM education is whether concepts should be taught using abstract or concrete examples. After a series of transfer experiments involving sequestered problem solving, Kaminski et al. concluded that presenting learners with abstract instantiations of mathematical concepts is more likely to result in successful transfer. Their results have been widely cited, and have received both positive and critical commentary. Their experiment has been replicated and extended by both De Bock et al. and Southey et al. This paper introduces a theoretical lens developed in physics education research known as a “Knowledge in Pieces” or “Resources” perspective which provides an alternative conceptual framework for describing concept transfer. As described by Hammer et al. (2005) resources refer to the cognitive elements (piece of knowledge) students have at their disposal in solving problems and constructing new knowledge. By applying a resources perspective, we provide a productive explanatory framework which does not rely on the notion of concreteness, and adequately accounts for all of Kaminski et al., De Bock et al. and Southey et al.’s results.

6.2. Introduction

This paper responds to a publication by Kaminski, Sloutsky, & Heckler (2013): “The Cost of Concreteness: The Effect of Nonessential Information on Analogical Transfer.” This paper follows a long line of previous publications, which began with results from Kaminski’s dissertation, and had key findings notably published in *Science* in 2008. See for e.g., Kaminski, Sloutsky and Heckler (2008, 2009, 2009a, 2013)²². Contrary to the popularly held belief that math should be taught by means of concrete examples, Kaminski et al. argue that “students might be better able to generalize mathematical concepts to various situations if the concepts have been introduced with the use of generic instantiations” (Kaminski et al. 2008, p.455). Their central experiment focuses on the learning and transfer of the concept of a mathematical group of order 3. Their findings have stirred

²² Also (Sloutsky et al., 2005) and (Jennifer A. Kaminski & Sloutsky, 2012)

considerable debate amongst education communities, eliciting both positive and critical commentary. The most comprehensive critique is provided by De Bock et al. (2011) which both replicated and extended the study, and provided results which may be interpreted as contradicting Kaminski et al.'s central claim. Southey & Allie (see Chapter 5) provided further qualitative data which are broadly aligned with the results of De Bock et al.

Kaminski et al.'s studies are concerned with two major themes in education theory: concreteness and transfer. Critiques of their study might be similarly organised. One might take issue with their definition and application of the notions of *concrete* and *abstract*. Or, one might interrogate their use of the notion of *transfer*. Much of the critical commentary in the literature is concerned with the latter critique, with particular attention being paid to the similarities/dissimilarities between what is learned in the learning domains, and what is tested in the transfer domain. (See Mourrat (2008), Cutrona (2008), McCallum (2008), De Bock et al. (2011), Jones (2009a), Jones (2009b)). Kaminski et al.'s latest paper (Kaminski et al., 2013) responds to many of these criticisms. By conducting a variety of further experiments, they argue that their results are not due to "similarity effects" between learning and transfer domains, but due to "concreteness effects" of the learning domains.

This paper seeks to add to this discussion by introducing a theoretical perspective of conceptual change known as a "Knowledge in Pieces" or "Resources" perspective. This perspective includes a family of sophisticated models of conceptual change which have been developed by various physics education researchers over the past three decades. (See DiSessa (2006) for a review). One of the central insights of these models is an explicit consideration of the grain size of the cognitive elements under consideration. For example, while the concept of "density" might be thought of as a neatly packaged, unitary concept in the mind of an expert, a novice is required to integrate a number of smaller cognitive elements before she can be said to have the larger concept of density. For example, she must be acquainted with the ideas of a fixed mass, a definite volume, the mathematical concept of division understood as "packing into" rather than "dividing up", and the phenomena of floating/sinking. Thus, having the "large" concept of density requires the activation and organisation of various smaller pieces of students' prior knowledge. The concept at issue in the studies of Kaminsk et al. is that of a mathematical group of order 3.²³ We wish to interrogate the grain size of this concept used in these studies and argue that it is too large for a productive description of the experimental results. By conducting a qualitative analysis of students' descriptions of their activities during the transfer task, we hope to illustrate that by using concepts of a smaller

²³ A mathematical group is a set of elements that satisfy four axioms: closure; associativity; invertibility; existence of an identity element. A group of order 3, contains three such elements. An example of such a group are the elements {0;1;2} under modulo-3 arithmetic (i.e. $2+1=0$; $2+2=1$).

grain size one can achieve a more comprehensive and productive description of the results. In sum, we argue for a change in theoretical perspective which introduces more appropriate explanatory variables.

The theory section of this paper provides an introduction to the abovementioned “Knowledge in Pieces” or “Resources” model of conceptual development. It describes how this model deals with notion of transfer (see Hammer et al. (2005)), and contrasts this with Kaminski et al.’s model of transfer. It also introduces notions of context, framing and priming. Kaminski et al.’s experiment is then introduced. The results from their experiment are interpreted from a Resources perspective, and further predictions are made using this model. De Bock et al.’s experiment is then described and their results are interpreted as confirming the aforementioned predictions of the Resources model. Next, we present our qualitative analysis of student responses. We argue that our results also support a resources perspective of transfer and that this would be the most productive lens with which to interpret results of these and future similar studies on transfer.

6.3. Theoretical Background

6.3.1. Resources as Pieces of Knowledge

The basic idea of the Resources (or Knowledge in Pieces) perspective was outlined in Chapter 2. An essential question in the “Knowledge in Pieces” perspective is: “How big is a piece?” We will briefly elucidate on what might count as a “small piece” and what might count as a “large piece”. In his seminal paper, *Toward an Epistemology of Physics*²⁴ diSessa (1993) posited small cognitive elements he dubbed p-prims (phenomenological primitives). These were posited as the basic elements of students’ sense of mechanism. Examples are: *force as mover*, *force as spinner*, *ohm’s p-prim (more effort gives more result, and more resistance gives less result)*, *maintaining agency*, *warming up/dying away*, *balancing*. They are primitive in the sense they constitute the bedrock of students’ reasoning: they typically do not warrant further explanation or justification. For example, as far as the student is concerned, the resource “force as mover” does not require further justification by appealing to electromagnetic theory. P-prims are perhaps the smallest pieces of knowledge education researchers need concern themselves with. Following Hammer et al. (2005) we might think of a “large piece” of knowledge as a “locally coherent set of activations” of these smaller pieces of knowledge. After recurrent, simultaneous activations, a set of resources may become a resource in their own right. For example, Gauss’ Law, which makes use of an array of complex ideas such as “flux”, “symmetry” and “charge” may be thought of as a single resource in the mind of an

²⁴ which might better have been titled: “Toward an epistemology of physics learning”.

expert. This idea bears similarities to the notion of “chunking” in cognitive science, and Sfard's (2008) notion of “reification” in conceptual development. Resources also bear similarities to image schemas (Johnson (1987) and Hampe (2005)) and primary metaphors (Grady (2005)). However, following Hammer et al. (2005), our use of the term “resource” refers to a general element of a general manifold cognitive structure. As mentioned above, the primary purpose for introducing a resources perspective is to interrogate the grain size of the knowledge elements assumed (explicitly or implicitly) in the transfer studies of Kaminski et al. In particular, while the concept, “commutative group of order 3” may be considered as a resource (or a chunk of knowledge) for an expert, we shall argue that attributing such a ‘large’, coherent resource to a novice confounds the interpretation of student performance in their various studies. Rather, we argue that it is more productive to posit resources of a finer grain size, such as “counting” and “symbol association/combination”.

6.3.2. Resources and Transfer

Two of the main differences between a naïve transfer and a resources perspective of transfer are: (i) the types of mental objects in play, and (ii) the dynamics of those objects. (i) The knowledge elements considered by a naïve transfer perspective are coherent conceptions. As outlined above, these are of a larger grain size than resources, are thought to be coherent and stable across contexts, and are aligned with expert knowledge. By contrast, resources are of a smaller grain size, fragmented, context sensitive, and aligned with students’ prior knowledge. (ii) The dynamics of a naïve transfer perspective involves the acquisition of a concept in one context, and the subsequent application (or transfer) of that particular concept in (to) another context. By contrast, a resources perspective is concerned not with acquisition and transfer, but with activation and organisation. A particular context activates a number of cognitive elements (resources) already present as students’ prior knowledge.²⁵ The student will hopefully select those resources which might be useful, and organise them in a productive combination. If a student successfully activates similar resources in a similar combination in a new context, one might say that they have “transferred their knowledge”. However, as argued by Hammer et al. (2005), the notion of “resource activation” is arguably less problematic and more productive than a traditional transfer perspective.

6.3.3. Concreteness and Transfer

Kaminski et al. adopt a “concreteness” model of transfer. Kaminski et al. define the concreteness of a representation to be the amount of information activated in the mind of the observer; the more

²⁵ Notice that a resources perspective is fundamentally constructivist, in the sense that new concepts are not acquired externally, but are built from the organisation and combination of cognitive elements (resources) already possessed by the student.

concrete a representation, the more information is activated in the mind of the observer. Their central argument for why concreteness hinders transfer is that “concreteness obfuscates the analogy between the learning and transfer domains” (Kaminiski, Sloutsky, and Heckler 2006, p.1582). The additional information activated in the mind of the student is thought to either obscure the underlying structure common to both domains, or distract the student from perceiving this underlying structure: “...relational structure common to two situations is less likely to be noticed when the situations are represented in a more concrete, perceptually rich manner than in a more generic form.” (Jennifer A Kaminski et al., 2005). Importantly, their model assumes that the underlying structure (i.e. “commutative group of order three”) is *there in the representation*, waiting to be seen by the student. Their model deals with a unitary, coherent model of concepts, aligned with expert knowledge. They are therefore concerned with a cognitive element of a relatively large grain size, which is thought to be more or less clearly perceived by students depending on the concreteness of the representation of that concept.

6.3.4. Resources and Context

As alluded to above, one of the main differences between the unitary notion of a concept and the manifold notion of resources is the latter’s sensitivity to context. The notion of *context* is a major theme in education research, and will not receive a full investigation here. However, we wish to bring attention to two important aspects of the notion of context. First, if our investigation involves modelling elements of student cognition, the notion of context must be tied to the subject:

“In other words, it does not suffice to consider the instructional environment and the material presented to the student to be the “context”. It is the students’ response to the environment and what is presented that has to be considered to be the context for cognitive ability.” (Redish, 2004)

Second, the notion of “students’ response to the environment” must centre on the ensuing *activity*.

“It is not the material, or externally defined qualities of the situation that are critical, but how the situation is experienced/interpreted by the person as a locus for meaningful actions.” (Van Oers, 1998) paraphrasing (Vygotsky, 1994)

In Van Oers' (1998) account of context as viewed from the perspective of activity theory (Engeström, Miettinen, & Punamäki, 1999), he emphasises that “what counts as context depends on how a situation is interpreted in terms of activity to be carried out” p.481. Thus, we shall argue that in

interpreting Kaminski et al.'s study, it is of crucial importance to consider the various affordances²⁶ of the objects presented, as well as the type of activities carried out by the students.

6.3.5. Resources, Framing and Priming

The concept of framing is widely used in cognitive sciences and linguistics. The foundational theoretical framework is usually attributed to the sociologist Goffman (1974). Following Hammer et al. (2005) we will define a frame to be “a set of expectations an individual has about the situation in which she finds herself that affect what she notices and how she thinks to act.” p.98. Note once again the emphasis not only on passive perception, but on ensuing activity. Framing is crucial not only in determining what resources are activated, but which resources are deemed valuable. For example, some students in our study framed the activity as a type of memory game, while others framed it as a mathematical task. This difference in the framing of the activity affects students' perception (e.g. what they notice and what they deem important). Different resources are likely to be activated, resulting in different actions being performed.

Finally, priming is a “nonconscious form of memory that involves a change in a person's ability to identify, produce or classify an item as a result of a previous encounter with that item or a related item” (Schacter, Dobbins, & Schnyer, 2004). The concept of priming is particularly important in studies involving sequestered problem solving, as the first task may prime the student to frame the second task in a similar manner.

6.3.6. Resources and Grounded Theory/Phenomenography

“There is a huge ontological gulf between, on the one hand, protocol coding categories and, on the other hand, knowledge element types or system configurations. Yet our research techniques have yet to clearly distinguish these.” (diSessa 2002, p.37)

Our qualitative data were analysed using a “bottom-up” approach suggested by Grounded Theory (Strauss & Corbin, 2007) and phenomenography (F Marton, 1981). Traditionally, qualitative data are analysed through the lens of a particular theoretical framework. The conceptual tools available in the framework are applied to the data, and used to interpret or “give meaning” to the data. In contrast, Grounded Theory is a methodology of data analysis which does not make use of a prior theoretical framework. Instead, patterns of ideas are allowed to “emerge” from the data. These

²⁶ affordances may be defined as the *perceivable* opportunities for action in the environment (Gibson, 1979)

patterns of ideas are assigned codes, which can then be grouped together as “concepts”, which can then form clusters of superordinate “categories”.

If qualitative data consist of student explanations, a Grounded Theory approach to the analysis seems to square well with a Knowledge in Pieces model of conceptual development. As discussed above, basic-level cognitive resources (e.g. p-prims) do not require further explanation in the mind of the students.; they constitute the bedrock of student explanations. Thus, the patterns of response to emerge in student explanations may be understood as these basic-level resources or pieces of knowledge activated by the students in that particular task.

6.4. Kaminski et al.’s Central Experiment

Kaminski et al. performed various immediate transfer experiments involving sequestered problem solving. (See Schwartz & Bransford (2005) for alternative means of defining and measuring transfer.) Two student cohorts were assigned to two different learning domains, and their performance was measured in a shared transfer domain. Both learning domains are posited to help students learn the mathematical concept of a commutative group of order 3. One learning domain is thought to be an abstract (or generic) instantiation of a mathematical group, while the other is thought to be a concrete instantiation. The learning domains could vary in two ways: (i) the objects used, and (ii) the storyline used. In their central experiment, the abstract learning domain makes use of





Abstract Learning Domain	
Concrete Learning Domain	
Transfer Domain	

Table 9: The Objects Used for Kaminski et al.’s Central Experiment

geometrical figures, while the concrete learning domain makes use of containers filled with a solution. (See Table 9.) The storyline used for the abstract learning domain included the following text: “On an archaeological expedition, tablets were found with inscriptions of statements in a

symbolic language. The statements involve these three symbols:  and follow specific rules.” The rules are shown in Figure 50, and include associativity, the existence of an identity, the various rules of combination, and commutativity.

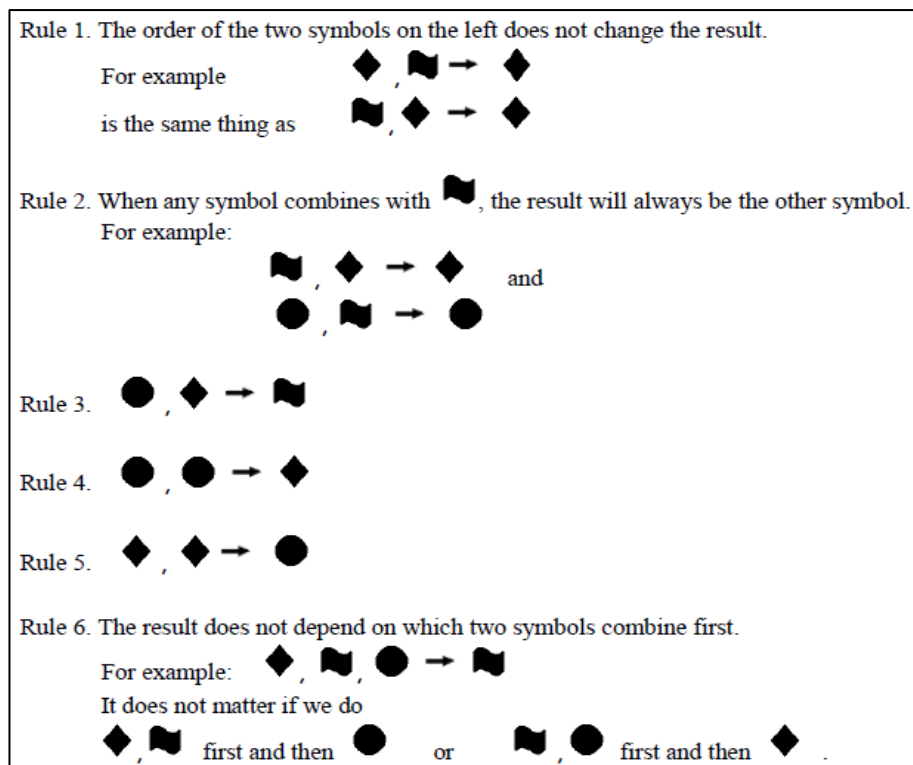





Figure 50: Rules Used in Learning Domain 1. Taken from SOM (Jennifer A. Kaminski, Sloutsky, & Heckler, 2009b)

The storyline for the concrete learning domain included the following text: “A company makes detergents by mixing three different quantities of solution represented as , , and .

The company is testing the mixtures and wants to know what amount of solution is left-over in the mixing process.” The rules are isomorphic to the abstract learning domain. Three of these rules are shown in Figure 51.

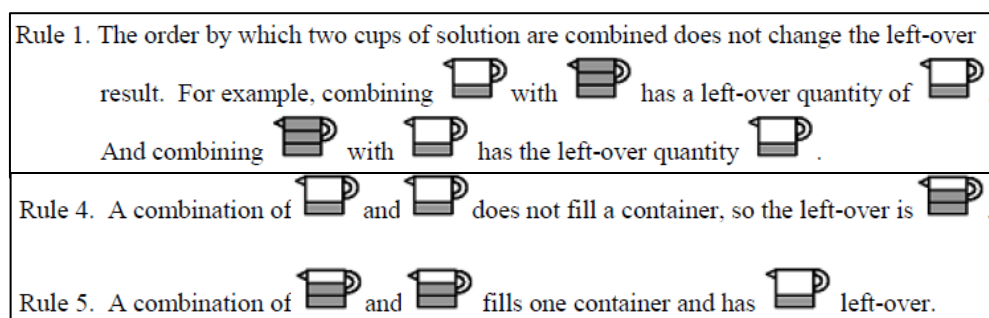


Figure 51: Three of the six rules used for Learning Domain 2. Taken from SOM (Jennifer A. Kaminski et al., 2009b)


For the transfer domain, the storyline provided was that of a children’s game: “In another country children play a pointing game that involves these three objects: Children point to objects and the winner points to the correct final object. The rules of the last system you learned are like the rules of this game. So use what you know about the last system to help you figure out the rules of this game.” The rules for the game are shown in Figure 52.




If the kids point to these:				
Then the winner points to this:				

Figure 52: The Rules of Combination for the Testing Domain. Taken from SOM


24 questions were asked of each learning domain and the training domain. The questions in the learning domains were isomorphic. See Figure 53.

2. Find the resulting symbol:



Choose: 1.)  2.)  3.) 

2. What is left-over when the following cups of solution are combined?






Choose: 1.)  2.)  3.) 

Figure 53: Isomorphic Questions from Learning Domain 1 and Learning Domain 2 Respectively. Taken from SOM

As mentioned above Kaminski et al. define the level of concreteness of an instantiation as “the amount of information activated in the mind of the observer”. A concrete instantiation activates more information in the mind of an observer than an abstract instantiation.

“[The abstract learning domain] is generic in the sense that the symbols are not necessarily meaningful, the storyline is novel, and the rules of the language are arbitrary... [In the concrete learning domain] the elements are familiar with known uses, and the storyline most likely tapped participants’ prior knowledge of containers, quantities, and pouring, which could help to convey the to-be-learned principles of the group structure.” Kaminski, Sloutsky, and Heckler (2013) p.18

They do not explicitly comment on whether the transfer domain is considered a concrete or generic instantiation of the mathematical group. Although, in a separate “concreteness calibration” experiment, students were asked to rate the concreteness (i.e. amount of communicated information) for various symbols used in their experiments, and gave the symbols used in the transfer domain (ladybird, vase, ring) a high concreteness rating, similar to their rating for the jugs used in learning domain 2 (Kaminski, Sloutsky, and Heckler 2013, p.18).

6.4.1. Kaminski et al.’s Results and Conclusion

Kaminski et al. found that students in the generic learning domain performed significantly better in the transfer domain than students from the concrete learning domain. (Mean test scores: generic domain ~79%; concrete domain ~54%). Furthermore, they did not find a significant difference between the performance of those in the concrete learning domain and those in the baseline (without learning) domain. They conclude: “These findings suggest that the concreteness of the learning instantiation hindered participants’ ability to align the common structure of the learning and transfer domains, which in turn led to transfer failure” (J. Kaminski, Sloutsky, and Heckler 2013, p.21)







6.4.2. A Resources Perspective

As mentioned in the Introduction Section 6.2 above, various commentators have pointed to the possible similarity between the generic learning domain and the transfer domain in order to explain the differences in performance of students in the two learning domains. However, Kaminski et al. insist that the difference in performance is not due to “similarity effects” but due to “concreteness effects”. They claim that the generic and concrete learning domains are mathematically identical, and in their response to Mourrat (2008) they state: “... there is every reason to believe that the differences between the examples we used are not mathematical, but cognitive.” As outlined above in Section 6.3.3, Kaminski et al. posit this cognitive difference as the obfuscation of the underlying concept (commutative group of order three) by the concreteness of the representation in the concrete learning domain.













We agree with Kaminski et al. that the two scenarios are mathematically isomorphic. However, in analysing the cognitive differences between the two scenarios, we advocate a resources perspective. Freshman would typically not have “commutative group of order 3” as a coherent, compiled resource and it is therefore an inappropriately large cognitive element to use in analysis. Thus, instead of considering the clarity with which students can perceive (activate) this particular concept (resource), cognitive elements of a smaller grain size must be sought. Furthermore, a resources

perspective would focus less on the representational structure (concrete or abstract) of the domains, and more on the ensuing activities; the tasks give rise to. In other words, in order to determine what resources are being activated, it is necessary to analyse what students are actually doing when confronted with these tasks. Thus, the difference between the domains might be described as being due to their activation of different resources in the minds of the students, which would result in students engaging in different activities (i.e. pursuing different strategies to solve problems). If the transfer domain activates similar resources to the abstract learning domain, this might account for the superior performance of students in that domain. One way of testing this hypothesis is to create two transfer domains; one which is thought to activate similar resources to the abstract learning domain, and one which is thought to activate similar resources to the concrete learning domain. This is the form of the instrument used by De Bock et al.

6.5. De Bock et al.'s Experiment

De Bock et al.'s (2011) study both replicated and extended Kaminski et al.'s study. Like Kaminski et al. their study consisted of two phases: (a) training and testing in a learning domain, and (b) testing for transfer. The learning domains were the same abstract and concrete domains used by Kaminski et al. The extension involved using a further (concrete) transfer domain, in addition to the (abstract) transfer domain used by Kaminski et al. The transfer domain is drawn from one of Kaminski et al.'s prior experiments, and involves the rather odd scenario of determining the number of burnt slices of pizza from a pizza order. The cook is systematically thought to burn a portion of every order. Three possible orders could be placed: 1, 2, or 3 slices represented by ,  and  respectively. The proportion to be burned follows the rules of a mathematical group of order 3. e.g. If an order for 2 + 2 slices is placed ( + ) then one slice will be burnt ().

Students were randomly assigned to one of four cohorts: AA, AC, CA and CC – where the first letter specifies the learning domain and the second letter specifies the transfer domain, and each domain could be either abstract (A) or concrete (C). Finally, students were also given a free response question at the end of the learning phase, in which they were tasked with a typical test question and asked to explain in their own words how they had arrived at their answer.

	AA	CA	CC	AC
Learning Domain	  	  	  	  


Transfer Domain				
-----------------	---	---	--	---

Figure 54: Objects used in the four domains of de Bock et al.'s experiment

6.5.1. De Bock et al.'s Results and Conclusion

Like Kaminski et al., de Bock et al. found that students in the AA cohort of students significantly outperformed the CA cohort in the abstract transfer test. (Results for the transfer test domain were 75% and 50% respectively). However, they also discovered that the CC cohort significantly outperformed the AC cohort in the concrete transfer test. (The transfer test results being 84% and 73% respectively). In other words, the extension of Kaminski et al.'s study showed that when the transfer domain involved the burnt slices of pizza scenario, students in a concrete learning domain outperformed students in an abstract learning domain.

In their qualitative analysis they coded students' responses to the free response question ("How did you arrive at your answer?") according to whether their explanations reflected the usage of modulo-3 arithmetic. They found that 65% of students in concrete learning domain (i.e. CC or CA) mentioned some aspect of modulo 3 arithmetic in their explanations, while none of the students in the abstract learning domain (i.e. AC and AA) did.

6.5.2. A Resources Perspective

De Bock et al.'s result contradicts the idea that the concreteness of the learning domain is always detrimental to performance in a transfer domain. They conclude that performance is optimised when the learning domain is of a similar concreteness to the transfer domain. From a resources perspective, one need not use the abstract/concrete distinction, which has been criticised by some authors as being "a somewhat confused distinction" (Cutrona, 2008). Instead, De Bock's result might better be interpreted as demonstrating that it is the use of a common resource in the learning and transfer domains that leads to better performance. What remains to be shown is what resources might be at play in the different domains. De Bock's qualitative analysis points toward a difference in strategy adopted by students in the concrete learning domain compared with those in the abstract learning domain. Students in the concrete learning domain were likely to use modular arithmetic. From a resources perspective, "modular arithmetic" is arguably at too large a grain-size to be considered as a resource for freshman who have not received specific instruction on this topic. As outlined in the theory Section 6.3.6 above, adopting an approach suggested by Grounded Theory may be advantageous in identifying the resources used by students in these tasks. This type of analysis is performed on our qualitative data below.

Finally, De Bock et al. claim that “the Kaminski et al. study does not provide strong empirical evidence that the participants in the abstract learning domain actually learned the abstract concept of a group instead of just memorizing and mapping symbols and combination rules.” (De Bock et al. 2011, p.115). We would similarly hesitate to describe the learning domain as facilitating the learning (i.e. construction) of a new “piece of knowledge”. A better description might be that the resource activated by the abstract learning domain caused the students to frame the task in a particular way (e.g. the memorizing and mapping of symbols). This may have primed the students to frame the transfer task in a similar way.

6.6. Southey et al.’s Central Experiment

See Chapter 5. In particular the Instrument, Methodology and Results sections 5.5, 5.6, and 5.7 respectively

6.7. Discussion

6.7.1. Two Dominant Strategies, Two Dominant Families of Resources

Our fundamental hypothesis is that the two dominant patterns of response observed in our analysis, may be modelled as the activation of two different cognitive resources, or more accurately, two different “families” of cognitive resources. The first family of resources has to do with symbol association; combinations of input symbols producing resultant output symbols. There is an enormously broad collection of activities which might be

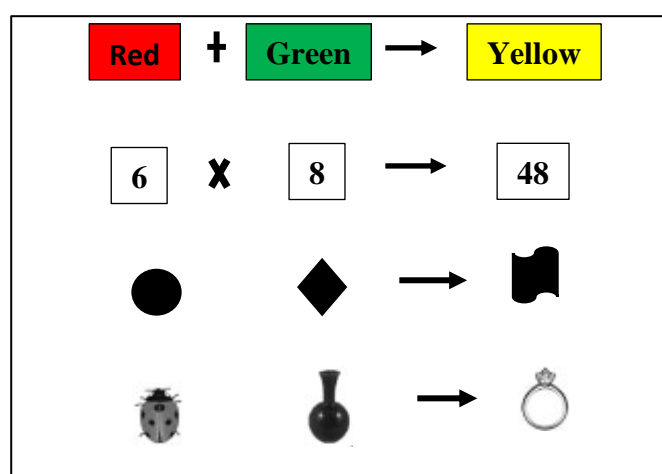


Figure 55: Activities involving the posited family of resources of symbolic combination/association

described as drawing on such resources. Examples may include: combinations of colour or the rote learning of the times-table. (See Figure 55). Of course, there are various other concepts in play which are not shared by these instantiations. For example, colours are not discrete: one could add a large amount of red to a small amount of green to create a reddish-yellow. However, we believe these and other examples bear sufficient resemblance, at this particular grain size, to afford the attribution of a similar family of cognitive resources. (See section 6.7.2 below for a fuller discussion). The other

dominant family of resources has to do with “counting”, as well as “filling up” and “leftover/remainder”. Figure 56 shows three different examples of when these resources may be activated, all of which are taken from various experiments performed by Kaminski et al.

The qualitative results from both De Bock et al. (2011) and Southey et al. (see Chapter 5) suggest that the “symbol association” family of resources is more likely to be activated in the “abstract” domain, while “counting and leftover” is more likely to be activated in the “concrete” domain. They also suggest that either dominant group of resources can be brought to bear on the objects used in the “concrete” domain. i.e. the

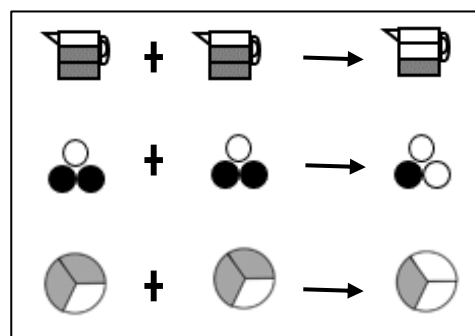


Figure 56: Examples of tasks that make use of a “counting” resource.

objects shown in Figure 56 could be seen either as simple symbols that combine in particular combinations, or as containers that can be filled. In general we might hypothesise that particular symbols do not dictate the activation of a particular set of resources. Rather, symbols have particular affordances, which, together with the context in which they appear, cause particular resources to be activated. Here, context refers to a variety of factors including the storyline used, students’ prior knowledge, the framing of the exercise, prior completed exercises, etc.

Finally, we hypothesise that these two different families of resources may be used as productive explanatory variables in explaining the results of these various studies. If the resources activated in the transfer domain are similar to those activated in the learning domain, students are likely to perform better than either a baseline (students who have not been exposed to a learning domain) or students whose learning domain resource activations differed from the transfer domain. This hypothesis is supported by the results from De Bock et al. who show that students who attempt the “concrete” transfer domain perform better if they had first completed the “concrete” learning domain. And vice versa for the “symbol association” resources (i.e. “abstract” domains). Crucially, none of the transfer domains of Kaminski et al. contained objects with the affordance of the family of “counting and leftover” resources mentioned above. Their transfer domain may therefore be said to be biased toward the “abstract” learning domain, allowing them to conclude that the “abstract” learning domain is most effective for transfer.

6.7.2. Similarity and Grain Size

A crucial issue in any study on transfer is the question of similarity; what makes one thing similar to another thing? Indeed a central critique of Kaminski et al.’s core experiment is that their transfer domain is more similar to the abstract learning domain than the concrete learning domain. This issue

turns on what the notion of similarity consists in. We suggest that any statement of the form “X is similar to Y” is meaningful only when considered in a broader context, including the purpose of making the comparison, and the grain size considered. For example, we agree with Kaminski et al. that both the concrete and abstract domains in their core experiment are instantiations of a mathematical group of order three; they both satisfy the appropriate mathematical criteria. However, we also agree with Cutrona (2008) that only the concrete learning domain involves the richer mathematical concept of modular addition. We would describe Cutrona’s analysis of similarity as occurring at a finer grain size than Kaminski et al. While the two mathematical concepts of “group” and “modular addition” coincide for systems of order 2 and 3, they diverge for higher order systems. McCallum (2008) argues for a similar “fine-grained” distinction in that only the concrete learning domain stresses the cyclic nature of a group ($a+a = b$; $a + a + a = i$). He points out that while groups of order 2 and 3 are cyclic, groups of higher order are cyclic if and only if they are isomorphic to a modular addition group of the same order.

The transfer experiments of Kaminski et al. are notable in that they elicit a difference in student performance. Thus, it seems that the primary purpose of a similarity analysis is to explain a difference in student cognition. Pointing to differences in mathematical structure (cf. Cutrona and McCallum’s comments above) won’t always coincide with differences in student cognition. As suggested above, using a Resources theoretical framework, and conducting qualitative analysis using an approach suggested by Grounded Theory seems better suited for the purpose of identifying similarities or differences in student cognition.

Our qualitative data could also be interpreted at different grain sizes. Both dominant strategies of “symbol association” and “counting and remainder” could be subsumed under a larger grain-size, superordinate category of “inputs producing outputs”. At this grain size they could be said to be similar. Indeed, it is this similarity which allows them to be thought of as instantiations of a mathematical group. They could also be analysed at a finer-grain size; for example “counting and remainder” could be seen as composed of two dissimilar strategies: “adding up and dividing by three” and “cancelling out and finding a remainder”. As has been repeatedly emphasized, the choice of grain size should be determined by the productivity of the analysis it affords. It is this choice of grain size that determines the “level” at which the concept of similarity is applied. Our claim is that describing the differences in cognitive activity (rather than mathematical structure) at the grain size of “symbol association” and “counting and remainder” provides a coherent explanation for Kaminski et al., De Bock et al., and our results, as well as a means of predicting the results of future, similar studies.

A final illustrative example of the importance of grain-size is a further experiment performed by Kaminski et al. which addresses the question of whether the possible numerical nature of the original concrete learning domain (the jugs) might constitute a salient difference in structure from the transfer domain (ladybird, vase, ring). The experiment involves a learning domain which teaches the students modulo-3 arithmetic. i.e. the learning domain used the integers $\{0, 1, 2\}$ instead of geometrical shapes (abstract learning domain) or jugs (concrete learning domain). Training and testing were isomorphic to that of their original experiment. They obtained a significant difference in effect size compared to the original concrete learning domain involving the jugs. Transfer scores for modulo-3 and concrete learning domains were 68% and 54% respectively. Following the tests participants were asked to match analogous elements across the learning and transfer domains. (e.g. the ring would map to the full jug (for the concrete learning domain) as they are both identity elements). Significantly, 55% of the modulo-3 group correctly mapped elements in the two domains, while only 24% of those in the original concrete domain did. The students who made the correct mapping did far better than those who did not: 87% compared with 49%. Kaminski et al. use these result to counter the hypothesise that structural differences exists between their original concrete and abstract instantiations: “transfer difficulty in the transfer domain of [our core experiment] is unlikely to stem from perceived structural differences between a numerical and non-numerical instantiation” (p.23). We would argue that (a) the concept of “numerality” is too large a grain size for a productive analysis, and (b) the emphasis should not be on the structural differences, but on cognitive differences. (See the discussion on Concept Definition and Concept Image (Tall & Vinner, 1981) in section 7.3.2). From a resources perspective, the fact that the jugs and the integers share a numerical structure does not entail that similar resources are activated. Consider the numeric symbol combination: “ $7 + 5 = 12$ ”. For the expert, this may be processed as simple pattern recognition, in the same way “ $6 \times 8 = 48$ ” is processed by one who has drilled their times-table. (“Six-eights-are-forty-eight”). However, the novice might understand the sum using the “making 10” strategy taught in many primary schools: “borrow” three units from the 5, and “fill up” the 7 until it reaches 10. You have two units “leftover”, resulting in $10 + 2 = 12$. Numerality, or the presence of numbers does not necessarily activate a coherent resource or family of similar resources. Our interpretation of the abovementioned results would be that different students employed different strategies when confronted with modulo-3 arithmetic. Some of these may have been categorised as belonging to the family of “counting and leftover”, such as “adding the numbers one at a time, ‘filling up’ to 3, and cycling back to 0”. Others may have used a type of symbol association where a ‘2’ followed by a ‘2’ results in a ‘1’, a ‘2’ followed by a ‘1’ results in a ‘0’, etc. Our hypothesis is that students were more likely to use, or at least more likely to notice the affordance of “symbol

association” in modulo-3 arithmetic than when working with the jugs. Those who activated this resource of symbol association, either by using or noticing this affordance of the numbers, would be more likely to make the mapping with the elements in the transfer task.

6.7.3. The “Reduced Concreteness” Study

As mentioned above, Kaminski et al. (2013) present five further studies in order to argue that the salient factor in their original study was a “concreteness effect” not a “similarity effect”. In offering a Resources Perspective as a more productive theoretical tool than the notion of “concreteness”, it is incumbent on us to illustrate how the results of each of these five studies might be better explained using this perspective. For the sake of brevity, a detailed analysis will be done of two of the five studies, although we believe the key arguments can be extended to the other three studies. One of the five studies, which involves numerality and modular arithmetic, is discussed in Section 6.7.2 above. We will now examine another of the five studies; this particular study involves the notion of “reduced concreteness”.

This experiment altered the concrete learning domain. The jugs were changed to more abstract objects (i.e. rectangular shapes with shaded areas, as shown in Table 10), and the storyline was changed to be identical to that of the abstract domain (i.e. symbolic shapes found on tablets during an archaeological expedition). Participants in this learning domain performed better in the transfer domain than those in the original concrete domain involving jugs (scores of 70% and 54%




Abstract Learning	
Concrete Learning	
Transfer Domain	

Table 10: Objects Used in Experiment 4

respectively). Kaminski et al. concluded that “replacing the familiar measuring cup storyline with the more generic storyline of a symbolic language resulted in more efficient transfer...”. Following the tests, participants were asked to match analogous elements across the learning and transfer domains. Importantly, 75% of students in this reduced concrete domain correctly mapped elements between domains, while only 24% of students in the original concrete domain did. From a resources perspective, the reason more students could correctly map elements between the domains is because the storyline and objects used caused similar resources to be activated in the learning and transfer domains. The rectangular shapes have the affordance of activating either “symbol association” or “counting/filling up” resources. However, the *storyline* involving “the combination of symbolic shapes” is likely to activate resources regarding “symbol association”. Since the transfer domain (“ladybird, vase, ring”) also activates resources involving “symbol association”, students are

more likely to perceive the similarities between the learning domain and the transfer domain. Thus, where Kaminski et al. attribute enhanced student performance to the reduced concreteness of the symbols, we attribute the enhanced performance to an activation of common “symbol association” resources due to the new storyline. A similar effect is arguably demonstrated in our experiment; see Section 5.8.1.

6.8. Conclusion

Kaminski et al. conducted a series of controlled experiments, involving sequestered problem solving, in which students in two different learning domains were compared in a common transfer domain. One learning domain was thought to be a generic instantiation of the mathematical concept of a group of order three, and involved geometrical symbols which combined in particular patterns. The other learning domain was thought to be a concrete instantiation, and involved jugs which could contain $\frac{1}{3}$, $\frac{2}{3}$, or $\frac{3}{3}$ of a particular solution; when combined, the jugs produced a certain amount of leftover solution. The transfer domain involved three objects (ladybird, vase, ring) which combined in particular patterns. Kaminski et al. found that students in the generic learning domain outperformed students in the concrete learning domain and concluded that learning a concrete instantiation of a concept may hinder subsequent transfer. Their studies attracted both positive and critical commentary. The main critique was that the studies did not control for similarities between the learning and transfer domains. Kaminski et al. (2013) responded with a series of new experiments which sought to address these various concerns of similarity between domains. They concluded: “The present findings suggest that in the absence of any overt similarities between base and target, generic learning domains can outperform more concrete ones with respect to analogical transfer.” p.26.

We suggest that the notion of “similarity” requires interrogation. Our study did not focus on the *structure* of the particular instantiation of the concept (abstract or concrete), but rather on what students were *doing* when solving the problems in the various instantiations. As shown in both our qualitative results (see Chapter 5), and De Bock et al. (2011), students are engaged in very different *activities* in the different learning domains. We dubbed two of the dominant activities “symbol association” and “counting/filling up”, which were associated with the “abstract” and “concrete” domains respectively.

From the perspective of a Knowledge in Pieces theoretical framework, we describe the different domains as activating different resources in the minds of the students. The first conclusion to draw from this perspective is that, although both abstract and concrete domains may be described as being mathematically identical in the mind of an expert, they are significantly different in the mind

of a novice. Kaminski et al. argue this difference (in the minds of novices) occurs because a concrete representation renders the underlying mathematical structure opaque, while an abstract representation is more transparent. From a Resources perspective, the notion of an “underlying mathematical structure” is replaced with notions of cognitive resources activated in the minds of students. Indeed, the resources associated with “symbol association” are no more structurally aligned with “commutative group of order 3” than the resources associated with “counting/filling up”. The type of resource activated is dependent on the affordances of the objects used, and the framing of the storyline used. Students were explicitly instructed to use what they knew about the learning domain to help them answer questions in the transfer domain. Thus, if a learning domain activated resources that were productive in the transfer domain, students would likely perform well in that transfer domain. Conversely, if a learning domain activated resources that were not productive in the transfer domain, performance would suffer. There are a variety of explanations for the effect of similar resource activation on performance such as framing, priming, prompting and practise. Crucially, none of the transfer domains of Kaminski et al. contained objects with the affordance of activating the “counting” resource mentioned above. Their transfer domain may therefore be said to be biased toward the “abstract” learning domain, allowing them to conclude that the “abstract” learning domain is most effective for transfer. De Bock et al. show that students in the “concrete” learning domain perform better in the transfer domain, if that transfer domain is likely to activate similar “counting” resources.

An interesting line of further research may be to interrogate the effect of the storyline on resource activation. In particular, the “concrete” instantiations (jugs and tennis balls) have the affordance of both “counting” and “symbol association”; one can see them either as containers being filled in thirds, or as arbitrary symbols occurring in a particular pattern. For example, our “concrete” transfer domain contained symbols with the affordance of modular addition, but was framed with a storyline involving “symbol association”. The same could be said of Kaminski et al.’s “Reduced Concreteness Study” discussed above. This led to the activation of the “counting” resource for some students and the “symbol association” resource for others. If students were given identical symbols, but different storylines designed to activate one or other resource, one could determine the effect of the storyline on resource activation and performance.

We have argued that a Resources or Knowledge in Pieces Perspective provides a richer and more productive analysis than a “traditional transfer” perspective. Importantly, the Knowledge in Pieces perspective bring attention to the grain size of the unit of analysis (i.e. the “piece of knowledge”). The choice of grain size will vary depending on a variety of factors including the aim of the study, the population group sampled, and the audience of the study. Importantly, a choice of grain size may

only be described as being more or less useful; there is no grain size that is more “true” than another. For example, one of our units of analysis, “symbol association” was posited as a cognitive resource. However, it may be understood as consisting of “finer-grained” resources including “dominance”, “neutrality”, and “inversion”. Similarly, our “counting resource” may be thought of consisting of finer-grained resources such as “adding fractions”, “adding whole numbers”, “adding up and dividing by 3”, “cancelling out”, each of which could be broken down into “finer-grained” resources. We have also argued that one should align the grain size of the unit of analysis with those *answering* a question, rather than with those *setting* the question. We conclude that for the purposes of the studies of Kaminski et al., analysis at the the grain size of “symbol association” and “counting” is more useful than analysis conducted at the grain size of the concept of “mathematical group of order 3”.

Finally, it may be argued that a Knowledge in Pieces perspective makes the notions of “abstract” and “concrete” redundant in these particular series of studies. Importantly, the notions of “abstract” and “concrete” are arguably polysemous, and even with the definition provided by Kaminski et al., is disputed by De Bock et al. While we agree with Kaminski et al. that a greater amount of information activated in the mind of an observer may interfere with their ability to notice the underlying structure of a problem, we believe that this phenomenon is better understood through the theoretical lens of Working Memory (Baddeley, 2000) or Cognitive Load Theory (Sweller, Ayres, & Kalyuga, 2011), rather than via the “abstract/concrete” distinction. This is explored further in Appendix A.

CHAPTER 7: Instantiations of Concepts – Sameness and Difference

“Mathematics is the art of giving the same name to different things” – Henri Poincare

“What we seem to need is a conceptualization that would help to account for the fact that the [novice] is unable to see as the same what the [expert] is unable to see as different” (Sfard, 2001, p.13).

7.1. Introduction

To instantiate is to “represent (an abstraction) by a concrete instance”²⁷. The expression “an instantiation of” is related to expressions such as “an instance of”, “an example of”, “a representation of”, “a realization of”. These are common phrases used in many educational contexts. We shall argue that despite its ubiquitous lay-usage, the notion of an instantiation carries with it implicit assumptions which deserve interrogation when being used in the more formal context of education research. The logical form of “instantiation of a concept” implies sameness and difference. A particular instantiation of a concept is *different* to other instantiations of the *same* concept. This is similar to the logical form that underpins the type-token distinction in philosophy. For example, how many letters appear in the following string: “aaa”? There is one letter type, but three letter tokens. One might say the string contains three *different* instantiations of the *same* letter “a”. Similarly we might say that “5-3” and “3-5” are two *different* instantiations of the *same* concept of subtraction. A central question of this paper is: what constitutes “sameness”? In other words, when one speaks of “an instantiation of a concept”, one implicitly assumes criteria of sameness and difference. It is these criteria we wish to interrogate.

This chapter provides three illustrative examples of the notion of “an instantiation of a concept”. It then provides six different perspectives as to how we might model the notions of “sameness” and “difference” as they apply to each example. It is argued that certain perspectives are more productive for physics education research purposes, than others.

²⁷ (Merriam-Webster, 2016)

7.2. Three Illustrative Examples

7.2.1. Subtraction

Imagine you are teaching primary school, and have to explain the expression $5 - 3 = 2$ to your class. What would your explanation look like? Typically, this expression is explained via some notion of “taking away”. For example: “5 sheep *take away* three sheep, leaves you with 2 sheep”. Now suppose you are teaching junior school, and have to explain the expression $3 - 5 = -2$ to your class. It should be obvious that the notion of “taking away” is not applicable to this particular expression. Instead, this is typically explained using a number line. One “begins” at “3” on the number line, and then one moves 5 units to the left, and lands up at “-2”. Thus, in this case, “subtraction” is understood as an instruction to “move left” along a number line. What do “taking away” and “moving left along a number line” have in common? Well, for a novice, nothing. These are completely different types of activity, drawing on completely different types of cognitive resources. However, for an expert, one sees $5 - 3 = 2$ and $3 - 5 = -2$ as two *different* instantiations of the *same* concept of subtraction. For a novice, the difference between these expressions might be understood in terms of the type of ideas which are brought to bear when first coming to understand these expressions (“take away” and “moving left on a number line”). In other words, the difference between these expressions may be understood in the type of cognitive *activities* in which the learners are engaged. And the sameness of these expressions might be accounted for by the fact that they both satisfy the definitional criteria of subtraction²⁸ and/or that a mapping can be found between the ideas used in both cases. E.g. “Take away” might be seen as equivalent to “moving left”, and “negative numbers” might be understood in terms of “owing”. Note that while subtraction might be a coherent, stable concept for an expert, there is no a priori reason why novices should understand these two cognitive processes (“taking away” and “moving left on a number line”) as instantiations of the same concept of subtraction. See Figure 57 below.

²⁸ Subtraction is a non-associative, anti-commutative binary mathematical operation.

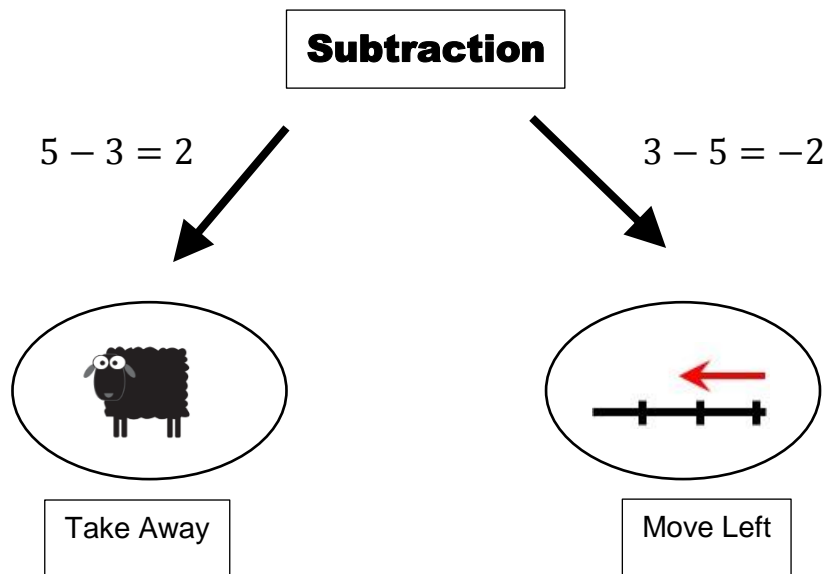


Figure 57: Two Different Instantiations of the Same Concept of Subtraction

7.2.2. Multiplication: The Coconut Seller

Nunes, Schliemann, & Carraher (1993) conducted a study involving Brazilian street vendors, which has now become an exemplar for the situated learning perspective (Lave & Wenger, 1991). One particular study involves a coconut seller who accurately performs multiplication when selling coconuts (e.g. calculating the cost of four coconuts at 35 cruzeiros²⁹ each) but fails to perform abstract multiplication (e.g. “4 × 35”).

Cotext 1: Selling Coconuts

The vendor was asked how much four coconuts would cost. (One coconut cost 35 cruzeiros). He replied: “There will be one hundred five, plus thirty, that’s one thirty-five... one coconut is thirty five... that is... one forty!”

Context 2: Abstract Multiplication

Moments later, the vendor was given paper and a pencil, and asked to multiply 4 by 35. He responded: “Four times five is twenty, carry the two; two plus three is five, times four is twenty... [answer written = 200]”

²⁹ Cruzeiro was the name of Brazilian currency in the early nineties. It refers to the constellation “The Southern Cross”.

Note that both of these contexts may be seen as containing two *different* instantiations of the *same* concept of multiplication. See Figure 58. This type of study has been replicated with Nepalese shopkeepers, American shoppers and Kpelle rice sellers (see Sfard (2001) for a review). The qualitative data often seem so ridiculous, they beggar belief. One exasperated researcher exclaims: “How can people be so dumb and so smart at the same time?!” (Cole, 1996, p.74)

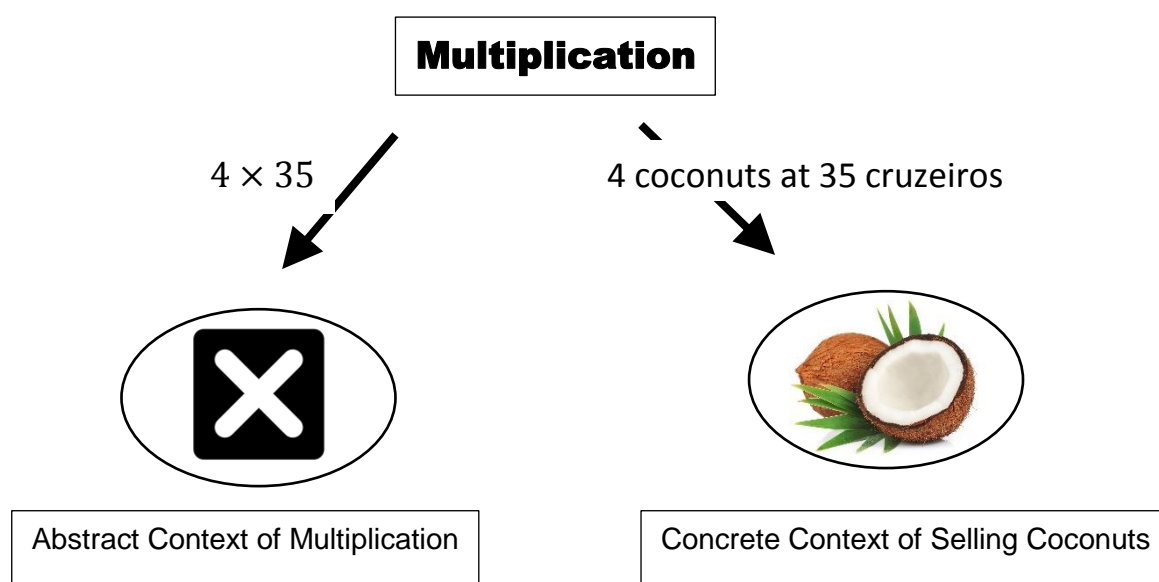


Figure 58: Two Different Instantiations of the Same Concept of Multiplication

The data from the coconut vendor may be read in a number of ways. One might understand them as the vendor’s failure to *transfer* the concept of multiplication from one context to another. Or, one might understand them as illustrating a fundamental difference between *concrete* and *abstract* representations of concepts. I would like to emphasise a possible interpretation that relies on a close reading of the transcripts. In the coconut context, the vendor replies: “There will be one hundred five...”. He seems to have memorised that *three* coconuts gives him a total of 105. He then seems to add the extra coconut (35 cruzeiros) by first adding 30 (to get 135) and then the 5 to get 140. So, in this context, he uses a particular strategy that involves remembering a particularly helpful total (105 cruzeiros for 3 coconuts) and then performing a “two-step” addition calculation. However, in the latter “abstract” context, he uses a completely different strategy; the algorithmic method of multiplication taught in schools. A close reading of his transcript reveals that his mistake is in adding the 2 to the 3 *before* multiplying the 3 by the 4. (By *first* multiplying the 3 by the 4, he would have got 12, which added to the 2 he “carried over”, would have given him the 14 required for the correct answer of 140. Try it!)

So, on this close reading of the transcripts, the difference between the two scenarios is captured by the fact that the vendor is engaged in different cognitive *activities* in each case. And the sameness is captured by the fact that both contexts satisfy relevant objective criteria to be seen as involving the same multiplication problem, " 4×35 ".

7.2.3. Kaminski et al.'s Study

See Chapters 5 and 6 for details regarding Kaminski et al.'s series of studies. For the purposes of this chapter, we wish to re-emphasise the issues of "sameness" and "difference" in their study. The "concrete" and "abstract" conditions were two *different* instantiations of the *same* concept; a mathematical group of order 3. See Figure 59. A variety of authors has criticized the experiment for not controlling for salient differences between the two learning domains. (See for example (See Mourrat (2008), Cutrona (2008), McCallum (2008), De Bock et al. (2011), Jones (2009a), Jones (2009b)). For example, Cutrona (2008) claims that the numerical nature of the concrete task (the jugs) stands in stark contrast to the lack of numerality in the abstract and transfer tasks. However, Kaminski et al. insist that the difference in performance is not due to "similarity effects" but due to "concreteness effects". They argue that the generic and concrete learning domains are mathematically identical, and in their response to Mourrat (2008) they state: "... there is every reason to believe that the differences between the examples we used are not mathematical, but cognitive." It is important to note that the distinction made here between *mathematical* structure and *cognitive* structure, might be associated with locating the criteria for (dis)similarity "outside" or "inside" the student respectively. In other words, one might locate criteria for similarity in mathematical definitions, or in a student's conceptual architecture. Kaminski et al. theorise that the concreteness of an instantiation obfuscates the underlying mathematical structure, resulting in poorer transfer. "These findings suggest that the concreteness of the learning instantiation hindered participants' ability to align the common structure of the learning and transfer domains, which in turn led to transfer failure" p.21. In other words, they posit an underlying mathematical structure which is "noticed" by students, to a greater or lesser extent, depending on the concreteness of the representation/instantiation.

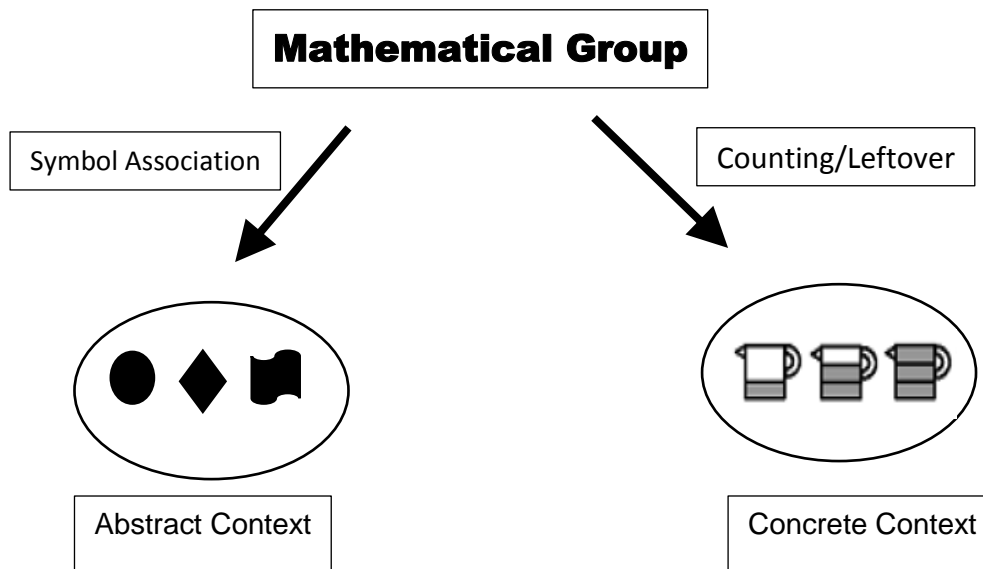


Figure 59: Two Different Instantiations of the Same Concept of a Mathematical Group

As outlined in Chapters 5 and 6, we have argued that a salient difference between the conditions is the type of *activities* (symbol association and counting) that the students were engaged in. Chapter 6 further illustrates how these activities could be related to the activation of different families of cognitive resources.

7.3. Perspectives on “Sameness” and “Difference”

The following sections shall investigate different ways of understanding how one can recognise two *different* things as being instantiations of the *same* concept.

7.3.1. Conceptual Essence

A Platonic Form is perhaps the most unambiguous example of some-*thing* which provides a criterion of “sameness”. Plato posits a perfect realm, populated by eternal, unchanging Forms. The world we inhabit contains imperfect instantiations of these perfect Forms. For example, all tables in the world are imperfect instantiations of the one Table; the Form “Table”. “We customarily posit a single Form in connection with each of the many things to which we apply the same name.” (Plato's Republic, Edited by Cooper et al., 1997, X 596a). Another way of saying this is that all tables share a common essence:

“Plato... believed that certain kinds of things have an essence (ousia) and that this essence can be known. ‘Essence’ was identified with ‘what something (really) is’ or with a thing’s true nature. Plato had assumed that these essences exist

independently from their sensible instantiations... These essences (ousiai) are also called Forms (eide, ideai).” p.19 (Helmig,2013)

To re-iterate, all tables are instantiations of the same Form; they all share the same essence (Plato, 1998, 389). Compare: all mathematical groups are instantiations of the same concept; they all share the same mathematical structure. While Platonism is no longer taken seriously as a metaphysical theory, it provides a clear illustration of how one might posit a concept³⁰ as a “something” from which all instantiations of the concept “flow”; a “conceptual essence” which all instantiations of the concept share. In the case of Kaminski et al., this “something” is the mathematical structure, or concept definition of the mathematical group. As discussed in Section 7.2.3 above, Kaminski et al. posit the underlying mathematical structure as being *there in the instantiation*, observed more or less readily depending on the concreteness of the instantiation. It seems analogous to the idea that these instantiations contain a type of “conceptual essence”; an essence which is more easily observed in abstract instantiations, and less easily observed in concrete instantiations. For Kaminski et al. it is this “essential” mathematical structure that accounts for the “sameness” of the instantiations.

A critical point is that this account of “sameness” is *independent of a thinking subject*. By focusing on the mathematical structure (or conceptual essence) of instantiations, the criteria for sameness are *out there* in the objective definition. One advantage of locating the criteria “outside” the subject, is that the “sameness” of different instantiations is easily accounted for. Two contexts are instantiations of the same concept just in case they both share the same conceptual essence, or conceptual structure. The disadvantage of locating criteria “outside” the subject is that it does not focus on how *we* (or *our students*) might see two instantiations of a concept as the same, or different. And it is this subjective perception of various contexts that are perhaps of more interest to education researchers. This distinction between the “subjective” and “objective” is well captured by the notions of “concept image” and “concept definition”, explored in the next section.

7.3.2. Concept Image and Concept Definition

“Concept Image and Concept Definition in Mathematics” (Tall & Vinner, 1981) introduces two useful ideas which separate out the notion of a concept *as it is currently understood by the learner*, and a concept *as it is generally defined by community of experts*. In other words, the ideas of “concept

³⁰ The precise relationship between Forms and concepts is discussed at length in Helmig (2013) and will not concern us in this paper. There is sufficient similarity between Forms and concepts to draw an analogy between Form and concept, essence and structure.

definition” and “concept image” captures “the distinction between the mathematical concepts as formally defined and the cognitive processes by which they are conceived” (Tall & Vinner, 1981, p.1). These ideas have been productively used in Mathematics Education for the past 3 decades. (See for example Przenioslo (2004) and Giraldo & Calvalho (2006)). This distinction appears under different guises in other areas of the literature. For example Entwistle & Peterson (2004) make a similar distinction between a concept and a conception: “[a conception is] an individual's personal and therefore variable response to a concept...” (Entwistle & Peterson, 2004, p.407). A concept definition is the form of words used to specify that concept, while the concept image is “the total cognitive structure which colours the meaning of the concept” (Tall & Vinner, 1981, p.1). The concept definition is stable, while the concept image is dynamic, and changes as the student associates more cognitive processes with the concept in question. Consider the example of subtraction from Section 7.2.1. The concept definition of subtraction is “an anticommutative, non-associative, binary operator”. As illustrated above, this definition is a far cry from how the concept is taught in primary school, where it is usually introduced via some notion of “taking away”. Here, the cognitive processes associated with “taking away” or “making less” would be considered the concept image. As illustrated above, subtraction can be understood in different ways. e.g. moving left along a number line. The concept image may thus be composed of many different cognitive processes. Tall & Vinner (1981) dub those particular cognitive processes activated by a student in a particular context, the “evoked concept image”.

Kaminski et al. take the notion of “concept” to be more aligned with “concept definition” than “concept image”. They continually emphasise the isomorphic mathematical structure of the instantiations, and tie this notion of mathematical structure to mathematical definition or mathematical principles. (See, for example, Kaminski et al. (2013) p.18, and their response to Mourrat (2008) above). As with the notion of a “conceptual essence”, an advantage of aligning the notion of a concept with *concept definition*, is that “sameness” is easily explained: Instantiation A is the same as instantiation B just in case they both satisfy the same concept definition. In other words, an instantiation of a concept is easily recognised as that which satisfies the criteria of the definition. For example, the contexts depicted in Figure 57, Figure 58, and Figure 59 all satisfy the criteria (the objective definition) of subtraction, multiplication and a mathematical group, respectively. QED. If one takes the concept definition to be the criterion of similarity, then Kaminski et al. are correct in insisting that both instantiations are saliently similar, since both accurately satisfy the four criteria for a mathematical group. However, we must question whether this is a well-chosen criterion of sameness. Their experiment tested for the *transfer of learned material*. Therefore, the criteria for sameness ought to be “that which was learned”, or, more particularly, “that which was transferred”.

(See Marton, 2006). The results from De Bock et al. (2011) and those outlined in Chapter 5, demonstrate that in the context of the experiment of Kaminski et al., there is little evidence to suggest that students learned the mathematical structure of a group (De Bock et al., 2011, p.123). It seems more likely that they were engaging in activities such as “symbol association” and “counting”.

7.3.3. Knowledge in Pieces

One of the more productive models of conceptual development to emerge in the past 20 years, is from physics education research (PER). This is broadly referred to as a “Knowledge in Pieces” (or “Resources”) model of conceptual change and was developed in opposition to a Misconceptions model. (See for example Hammer (2000), DiSessa (2006) and Scherr (2007)). Instead of modelling student knowledge as occurring in coherent, unitary conceptions (or misconceptions), a Resources model advocates that knowing should be thought of as the activation of multiple fine-grained cognitive elements in a particular context. These fine-grained elements have been dubbed “cognitive resources”, or simply “resources”. The word “resource” derives from computer science, and is a piece of code (or a sub-routine) that may be incorporated into a larger program. The shift is therefore from viewing a concept, such as “subtraction”, as a unitary entity, to viewing it as being composed of a number of fragmented elements such as “take away”, or “moving left on a number line”. These elements may be loosely associated, but have limited coherence in the mind of the learner. They are also highly context sensitive. A “low level” or “basic” resource is thought to be grounded in repeated patterns of experience. This is similar to the notion of image schemas (Johnson (1987) and Hampe (2005)), phenomenological primitives (diSessa, 1993) and primary metaphors (Grady, 2005). Repeated activation of a common set of resources might lead to the formation of a superordinate resource. For example, a full understanding of the concept of a mathematical group, involves activating and coordinating resources associated with “identity”, “inverse”, “closure” and “associativity”.

As argued in Chapter 6, the qualitative results of De Bock et al. and Southey et al. suggest that different resources (or pieces of knowledge) were activated in the two different conditions of Kaminski et al.’s experiment. This lends credence to the idea, that *from the perspective of the learner*, the two conditions are substantively different. Importantly, what counts as a “substantive difference”, involves *what is being learned*. i.e. if what is learned in one condition can be shown to be different to what is learned in the other condition, then the conditions would count as being substantively different. From the perspective of Kaminski et al., both conditions facilitate the learning of the concept of a mathematical group. Note that a resources perspective problematises the notion of “what is learned”. From a resources perspective there is no “thing” that is learnt in

either condition; one does not *acquire* something in one condition, and then apply this recent acquisition in a transfer condition. Rather, one speaks of resources, or pieces of knowledge, being activated in a particular context. Thus, an occurrence of “successful transfer” would be re-interpreted from a Resources perspective as: “the same resources that were activated in the learning condition were *also* activated in the transfer condition”.

This shift in perspective from a large grain size of “unitary conception”, to a smaller grain size of “pieces of knowledge” has substantial explanatory power. For example, when we teach primary school children how to solve the problem “5 – 3” in terms of “5 sheep take away 3 sheep”, we are activating resources associated with “taking away”, “making less”, “discrete objects” etc. With practise, these resources are being coordinated and compiled into a superordinate resource which we might call “subtraction as taking away”. Has the child learned subtraction? Perhaps, like the blind men and the elephant (see Figure 60), we might say that the child has learnt *a piece* of the concept of subtraction. Indeed, the piece of knowledge she possesses is, at first “touch”, very different from other pieces of knowledge she will develop regarding subtraction (e.g. moving left on a number line). So do we have a model for how we eventually manage to coordinate these various pieces of knowledge into a more coherent whole? In other words, how do we account for the fact that experts see the instantiations in Figure 57, Figure 58 and Figure 59 as being instantiations of the *same* respective concepts? The next two sections attend to this issue.

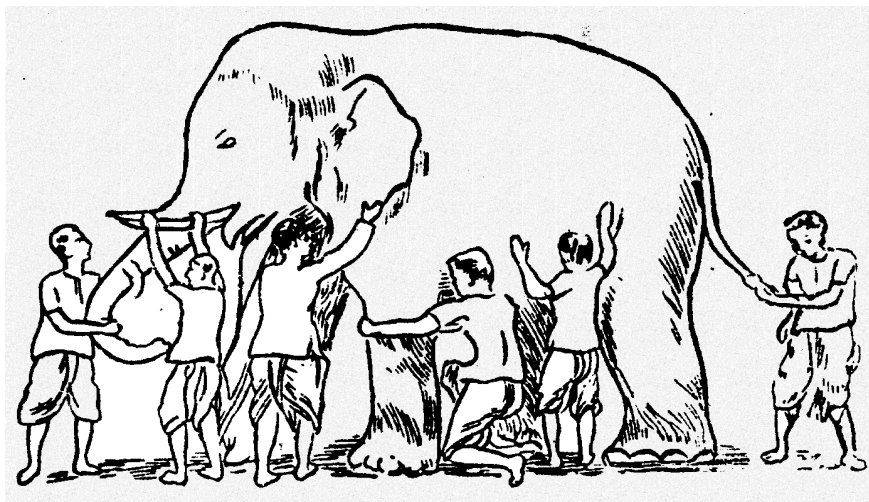


Figure 60: A Picture Depicting the Parable of the Blind Men and the Elephant

7.3.4. Coordination Classes – Seeing Sameness

In “Why Conceptual Ecology is a Good Idea” di Sessa (2002) introduces, in addition to the notion of a p-prim, a further type of mental entity he calls a coordination class. (See also diSessa & Sherin (1998), Levrini & diSessa (2008) and diSessa & Wagner (2005)). The p-prim is more suited for

modelling novice knowledge, while a coordination class is more aligned with expert knowledge: “...coordination classes may well not exist in naïve thinking.” (diSessa, 2002, p.43). A coordination class is thus posited to be something of a larger grain size than the pieces of knowledge activated in novice cognition: “The name ‘coordination class’, in fact, is meant in part to convey precisely that particular relations among knowledge pieces – coordinations – are what define this class.” (diSessa & Wagner, 2005, p126). Crucially, coordination classes are linked to the notion of *perception*: “Succinctly, having a concept, according to coordination class theory, is in essence being able to ‘see’ the information that defines the concept in an appropriate range of relevant situations”. Commenting on the study of the coconut sellers described in Section 7.2.2, Anna Sfard writes: “What we seem to need is a conceptualisation that would help to account for the fact that the Brazilian vendor is unable to see as *the same* what the researchers are unable to see as *different*” (Sfard, 2001, p.13). Coordination class is a theoretical tool that can help in this regard. As a theoretical tool, it contains further structure such as “causal net” and “readout strategies” which shall not be addressed here. It is a tool that is being productively used in education research. (See Levrini & diSessa (2008) and Wittmann (2002)). It’s contribution to this chapter’s discussion on “sameness” is it’s emphasis on “sameness” as something we *see* in different instantiations of a concept, by coordinating various pieces of knowledge. Furthermore, coordination classes emphasise the contrast between novice and expert cognition/perception. They highlight the idea that experts’ ability to see “sameness” consists in their able to coordinate various relevant pieces of knowledge. While novices might be said to “possess” relevant pieces of knowledge, it may be their inability to effectively coordinate these pieces, that results in their inability to see “sameness”. This notion of “seeing sameness” is well illustrated in the now classic study of Chi, Feltovich, & Glaser (1981). When presented with a variety of physics problems, novices focused on superficial similarities, while experts could identify (coordinate) similar physics principles that might be applied. Even though novices might be said to “possess” the relevant pieces of knowledge to solve these problems, it was their inability to coordinate these pieces of knowledge appropriately, that resulted in their inability to “see” the problems as instantiations of deeper, underlying physics principles.

7.3.5. Commognition – Thinking/Talking Sameness

“I show him samples of different shades of blue, and say: ‘The colour that is common to all these is what I call “blue”’.” (Wittgenstein, 1958, §72)

In her paper, “On the Gains and Dilemmas of Calling Different Things the Same Name”, Sfard (2001) discusses the study of coconut seller (Nunes et al., 1993). As mentioned above, she says: “What we seem to need is a conceptualisation that would help to account for the fact that the Brazilian vendor

is unable to see as *the same* what the researchers are unable to see as *different*" (Sfard, 2001, p.13). She references two papers (Dreyfus, Hershkowitz, & Schwarz, 2001) and (Ohlsson & Regan, 2001) which account for this notion of "sameness" by positing the theoretical entities of "structure" and "template" respectively. Sfard notes: "Probably contrary to the intentions of their authors, the words *structure* and *template* may be read as referring to entities that are 'sitting there' in things or heads..." p.10. Her own approach in explaining "sameness" follows her influential theoretical framework of "commognition" or "thinking as communicating" (A Sfard, 2008). According to Sfard, the notion of "sameness" refers to the creation of a new discourse which "...subsumes several former, independently existing discourses, turning them into discourses 'about the same thing'". p.19 For example, a parabola belongs to a type of graphical discourse, while the expression " x^2 " belongs to an algebraic discourse. However, through the process of objectification (by viewing a parabola and " x^2 " as objects in their own right) we can subsume both of these objects within the discourse of "function". i.e. Both a parabola and " x^2 " may be spoken of as instantiations of the same function. Thus, according to Sfard, "sameness" is best understood as residing in our *talk* of objects; in our discourse. Importantly, she guards against thinking of "sameness" as referring to some hidden entity, or some distilled essence. (See also van Oers (2001)). In the qualitative data presented in Chapter 5, it is clear that students did not subsume their descriptions of the learning and transfer conditions under an overarching discourse of mathematical groups³¹. However, a significant number of students did identify a mapping between the conditions, implying that they could relate the two discourses, even if they hadn't yet turned them into discourses about the same thing; a mathematical group. In the example of the coconut seller it is clear that no mapping between the two discourses of selling coconuts and doing arithmetic is made.

7.3.6. Situated Activity – Doing Sameness

In his paper "Sameness and Difference in Transfer" Marton (2006) outlines Lave's (1988) critique of traditional transfer experiments which is captured by four questions: (i) What is learned? (ii) Who defines the relations between situations? (iii) How many situations are involved? (iv) Where does transfer happen? We wish to focus on the first two of these four questions. Recognised as the pioneer of "situated learning", Lave provides an alternative account of "knowledge". Instead of "a set of tools stored in memory for use in different situations" (Marton, 2006, p.502) Lave describes knowing as people's ability to participate in activities. The notion of activity is central to Marton's account of "sameness" in transfer. He cites Greeno, Moore, & Smith (1993): "For an activity learned

³¹ There was in fact one student of the 258 that mentioned the fact that both learning conditions were examples of a mathematical group.

in one situation to transfer to another situation... the second situation has to afford that activity and the agent has to perceive the affordance" (Greeno et al., 1993, p.102).

This final perspective on "sameness" and "difference" has been included to provide an alternative to the predominantly cognitivist perspectives mentioned above, and as a reminder that our criteria for judging student understanding is ultimately their behaviour. Citing a final mark for a test is not as accurate an indicator of student understanding as observing the type of activities she engages in while answering questions. Indeed, in order to better model conceptual change, DiSessa (2002) advocates a "process-data" approach: observing student activities in the process of learning, rather than simply collecting pre/post scores.

In the three examples provided in Section 7.2, it is demonstrated that students are engaged in *different activities* in each of the different scenarios. Lave's critique of transfer serves as an underpinning for viewing these different *activities* as a justification for flagging the scenarios as being substantively *different*. This is an important heuristic for determining salient differences. For example: Suppose we were to conduct an investigation into the effects of concreteness on primary school learners' understanding of subtraction, and used the following two instantiations: (i) The concrete instantiation of "5 coconuts – 3 coconuts", and (ii) the abstract instantiation "3 – 5". Following the "activity heuristic" we could argue that the difference in concreteness is irrelevant because the learners would be engaged in two different *activities*: taking away, and moving left along a number line, respectively. The salience of the difference in activities preclude any comparison of concreteness.³²

Finally, Marton also places emphasis on the idea that criteria for "sameness" should centre on a *learner's relationship* to the situations under comparison, rather than the perspective of an expert. (See also (Lobato, 2003) and (Lobato & Siebert, 2002) for an exposition of *actor-orientated* transfer).

7.4. Conclusion

This chapter has sought to interrogate the notion of an instantiation of a concept. In particular, it has sought to interrogate the notions of "sameness" and "difference" when speaking of two *different* instantiations of the *same* concept. Three illustrative examples were provided, and different perspectives were given as to how one might account for notions of "sameness" and "difference" in those examples. Two key issues emerged: (i) the issue of coherence/fragmentation. In other words, the issue of grain size. And (ii) the issue of expert/novice. In other words, the issue of the *subjective*

³² This is a possible difference between the studies of Kaminski et al. (2013) and Goldstone & Sakamoto (2003). While both test for the role of concreteness in transfer, the former arguably do not control for the difference in student activity.

perception of sameness. Regarding grain size, we have argued that “smaller is better” when considering the process of novice learning. Instead of conceiving of learning as a coming to know/see the essential nature of a concept buried within an instantiation, we have argued that it is more productive to think our knowledge as being composed of pieces, with limited coherence. We then presented three different ways one might link these pieces of knowledge with the notion of “sameness”. Coordination classes provide an explanatory framework for how we coordinate pieces of knowledge in order to *see* sameness. Commognition provides an explanatory framework for how pieces of knowledge (which in this view might be thought of as being linked to different discourses) come to be subsumed under an overarching discourse, allowing us to *talk* (or think) about sameness. And the perspective of situated learning emphasises the *activities* one is able to participate in when one has assimilated a piece of knowledge.³³ Thus if a situation affords an activity in which we are unable to participate, one might say that one cannot yet marshal one’s current abilities (cf. coordinate one’s pieces of knowledge) to participate in the activity.

Crucially, the ability to productively relate pieces of knowledge means that one can see, talk or act as an *expert*. It is a hallmark of expertise that one can see “sameness” where others see difference. Thus, when speaking of “different instantiations of the same concept”, one must be cognisant of the particular perspective that might recognise this “sameness”. It has been argued that in the three illustrative examples outlined above, it is only a relative expert (in subtraction, multiplication and group theory) that can see the pairs of instances as being instantiations of the same concept.

A parting quote. The anthropological work of Schmandt-Besserat (1992) looked at how various peoples (from Fijian Islanders to nomadic Mongolian tribes) conceived of the concept of “number”. They found that several languages have different names for numbers depending on what is being counted. E.g. the “three” of “three coconuts” is a *different* word to the “three” of “three canoes”. As experts, we need constant reminding that what we see as “obviously the same” cannot be taken for granted: “There may be cultures in which it is by no means obvious what three canoes have in common with three coconuts.” (Harris, 2005,p.115)

³³ Note: the warning bells from Chapter 1 are ringing. It is not the “assimilated piece of knowledge” that causes the activity; nor is the “assimilated piece of knowledge” a label for the activity. This sentence simply tries to draw a tentative mapping between the discourse of “pieces of knowledge” and the discourse of “activity” (i.e. the discourses of cognitivism and socio-cultural theory). However, I believe these discourses are sufficiently different that we cannot hope for a subsuming discourse.

CHAPTER 8: General Discussion and Conclusion

This thesis has established a variety of hitherto unpublished context sensitivities exhibited by students in learning physics. In the topic of vector addition, this thesis has demonstrated that students are highly sensitive to the type of physical quantity being added, and exhibit a sensitivity to the textual prompts “total” and “net”. They are also sensitive to the type of object on which two forces are acting (a person or a hard ball). The latter sensitivities were shown to be further modulated by details of the question methodology (e.g. the order of the questions). In answering a modified FCI, students showed a moderate context sensitivity to the familiarity of the words used in the questions.

This thesis then sought to explain *why* these particular context sensitivities might occur. In offering explanations we have modelled student understanding of these particular topics. In particular, we have shown how Resources/Knowledge in Pieces, and embodied simulation, might be productive models of student understanding of these topics. And a further aspect of this modelling enterprise is student reasoning. While one might classify the more interesting or intriguing data as that which involves deep conceptual reasoning (which might be modelled with embodied simulation), it is important to emphasise that much of the data might be classified as involving shallow reasoning. Finally, we have shown how this modelling of student cognition takes place within broader models of students’ framing of the activity (which includes students’ epistemologies).

The notion of “context sensitivity” evokes a particular notion of “context” and “concept”. A concept seems to be an unchanging “something” that might appear in changing contexts. And possession of a concept (i.e. understanding) seems to be the source of a subject’s ability to successfully/accurately perform certain activities with regularity. Indeed, the noting of a context sensitivity is the noting of a disruption of this regularity in a particular context. This thesis has challenged these ideas in at least two ways. Firstly, we have suggested “Copernican turn” in terms of the arrow of “understanding”. We do not possess a “thing” called “understanding” which allows us to successfully perform activities with regularity. Rather, it is this successful regularity which is labelled (or reified) as “understanding”. This cures us of looking for the source of our abilities, and focuses our attention on the type of abilities we deem desirable. Secondly, the notion of “concept” has been interrogated, and the notion of “grain size” has been brought to the fore. This might also be seen as problematising the line drawn between concept and context. For example: Is **vector addition** a concept that occurs in the context of adding **momenta**? Or might **vector addition of momenta** be thought of as a type of concept in its own right? There are perhaps good reasons for defining

concepts at the grain size that we do; “vector addition” satisfies a particular definition (a particular set of axioms) which are applicable in a range of circumstances. However, we have argued that this type of delineation of a concept is only helpful when considering expert understanding. Indeed, we have argued that when considering novice understanding, this delineating concepts through definitions may be unhelpful, and even misleading. We believe this is particularly apparent in our analysis of the concept of a mathematical group as it was used in the studies of Kaminski et al.

Nevertheless, we have not suggested that we completely disregard the notion of a concept being surrounded by a context. The crucial issue is that we recognize the limitation of these terms, and guard against recursively defining context as “that which surrounds a concept”, and a concept as “that which is surrounded by a context”. We must constantly question where it might be most productive to draw the line between “context” and “context”. And we must guard against thinking that there is a “natural” or “true” way of delineating either

This thesis has emphasized the importance and productivity of a Resources perspective. It is this perspective that underlies the choice of an appropriate grain size. A resource, in the mind of a student, represents a compiled, consolidated “piece of knowledge”. In the language of Sfard it represents an objectified discourse. In the language of Lave, it represents an activity in which a student might legitimately participate. These are the “things” that we should be the objects of our research. In other words, in physics education research, it is perhaps student resources that should be delineated as “concept”.

Finally – the substrate to this thesis is, perhaps ironically, the non-cognitivist philosophical outlook introduced in the introductory first chapter. The grounding of the thesis in Wittgensteinian ordinary language philosophy has guided the intention of this research, not toward “truth”, but toward “clarity”. If we accept that concepts such as “understanding”, “concept”, and “context” gain traction, not by referring to specific cognitive or behavioural processes, but by being embedded in our very messy form of life, then we are cured of a desire to find the true nature of “actual” understanding. In the act of modelling “understanding”, we do not more accurately describe some transcendental entity. Rather, the act of modelling is itself the augmentation of our understanding of “understanding”.

I hope this thesis has contributed to an awareness of how we might choose appropriate models of understanding in physics education research.

Appendix A:

Three Uses of “Concreteness” in Education Research

1. Introduction

Although no generally accepted definition exists, the concept of concreteness is widely used as both a descriptive and an explanatory tool in education research; education materials may be described as being more or less concrete, and this concreteness is said to interact either positively or negatively with student learning. This appendix seeks to demonstrate that there are at least three different concepts that underlie the use of concreteness as a descriptor: (i) Tangibility; (ii) Familiarity and (iii) Perceptual Richness³⁴. It shall be argued that these concepts do not align to produce a homogenous concept of concreteness, but that they constitute different axes, each with its own criterion of measurement. Furthermore, it shall be demonstrated that a variety of theories and heuristics provide the framework for the use of concreteness as an explanatory tool, such as working memory, embodied cognition and dual representation theory.

Authors are usually aware of the polysemous nature of the concept of concreteness, and commendably define their own particular use of the term. However, in the wider literature, conclusions such as “concrete instantiations may be more engaging for the learner...but do not necessarily promote transfer”(Kaminski, Sloutsky, & Heckler, 2008, p.455) are uncritically cited without proper attention being brought to the particular definition of concreteness under consideration, and the particular theoretical underpinning of the conclusion being drawn. While it is likely that the abstract/concrete distinction will remain as a *descriptor* in education research, we shall argue that it is of too coarse a grain size to be useful as an *explanatory tool*.

The appendix will centre around two sets of education research: (i) manipulatives used in math education at preschool level. (See Petersen & Mcneil (2013); Brown, McNeil, & Glenberg (2009); Clements (2000); and Sarama & Clements (2009)). And (ii) representations of complex higher order concepts. (See the series of studies by Kaminski et al. and Goldstone & Sakamoto (2003)). It shall

³⁴ Note a further general definition is general vs. particular. This distinction is not explicitly discussed in this paper. It commonly occurs in the education literature in the related, but sufficiently distinct, conversation around “abstraction”.

also pay particular attention to the use of concreteness in the studies of concreteness fading (Fyfe, McNeil, Son, & Goldstone, 2014).

1.1. Origins and Current Usage

The *abstract/concrete* distinction is perhaps as old as philosophy. Philosophers from Plato and Aristotle, to Leibniz, Locke and Reichenbach have all held differing views on the distinction (See Brook 1997). Much of the original philosophical discourse focuses on notions of *completeness*, *containment* and *predication*. *Abstracta* is a Latin word which was originally used to translate Aristotle's term, χωριστά ("separately"), which referred to those entities which were 'incomplete in themselves', or 'separated' from matter (Rovira, 2000). For example colour would be considered abstract, as it requires matter in order to exist, and is therefore incomplete in itself. It should be clear that our use of *concrete* and *abstract* in education research bears little resemblance to its original philosophical use. Indeed, even in contemporary philosophy Hale (1988) identifies a dozen possible ways of construing the *abstract/concrete* distinction. Furthermore, these concepts are now used in disciplines as diverse as philosophy, linguistics, cognitive psychology and education research, each with their own particular discourse and use of these terms. A final contributor to the opacity of these concepts is their widespread lay-usage. Indeed, the sheer number of contexts in which these terms are used should flag the importance of conceptual clarity in academic discourse.

1.2. Concrete Objects and Concrete Concepts

What is the difference between a concrete *object* and a concrete *concept*? While a detailed definition will depend on the notion of concreteness under consideration, we offer the following definition: a concrete **concept** may generally be defined as our subjective concept of a concrete **object**. Simply put, if a football is a concrete **object**, then our concept of a football is a concrete **concept**. As education researchers attempting to model student cognition, our focus is predominantly on concepts. However, in the education literature, the value of *concreteness* has been ascribed to a host of entities: to objects (such as footballs); representations of objects (such as a photograph or a drawing of a football); concepts (such as subtraction); instantiations of concepts (such as the equation " $5 - 3 = 2$ "); and our knowledge of concepts (such as whether our knowledge of a football comes from playing with one or merely having had one described to us). The muddles are manifold: A concrete object such as a football, may be said to have an abstract representation in the form of a line drawing. An abstract concept such as subtraction³⁵, may be thought to have a

³⁵ If the concept of subtraction does not seem sufficiently abstract (even with a definition that begins: "a binary operator..."), consider the concept of a mathematical group of order 3, and the concrete instantiation thereof used by J. a Kaminski et al. (2013).

concrete representation in the form: “five apples take away three apples”. And our knowledge of a concrete object, such as football, might be described as abstract if we have only ever had it described to us with words (“round, hollow, pumped full of air, etc.”) but never encountered the object itself.

One of the main questions education research wants to answer in relation to concreteness, is the advantages and disadvantages of using abstract or concrete *materials* in instruction. A source of potential confusion is that those materials can be **objects**, or **representations of objects** (e.g. the counters used by Petersen & Mcneil (2013) and the physical models used by Fyfe et al. (2014)), or representations of higher order **concepts** (e.g. *competitive specialisation* in Goldstone & Sakamoto (2003) and *mathematical group of order three* in J. a Kaminski et al. (2013)). These studies cross reference each other without noting whether they are similarly concerned with objects or concepts..

2. Tangibility

Tangibility is perhaps the most readily recognisable measure of concreteness, and most commonly used by the layperson. Often introduced in primary school via concrete and abstract nouns, concreteness refers to “something tangible, solid; you can touch it, smell it, kick it; it is real” (Wilensky, 1991). “By ‘concrete’ most practitioners mean physical objects that students can grasp with their hands.” (Sarama & Clements, 2009) Importantly, this notion of concreteness tends to be discrete rather than continuous; something is either tangible or it is not, its measure of concreteness does not lie on a scale. Hale (1988) cites a variety of more definitive ways of stating this particular version of the distinction:

- i. Concrete entities are accessible via our 5 senses whereas abstract entities are not.
- ii. Concrete entities participate in causal networks, whereas abstract entities do not.
- iii. Concrete entities exist independently of human minds or language, whereas abstract entities are human constructions or creations
- iv. Concrete entities are in spacetime whereas abstract entities are not.

Note that concepts, by definition, are intangible (Margolis & Laurence, 2007). Thus a concrete concept is certainly not a type of concept that you can literally grab hold of. However, the notion of tangibility squares well with our general definition of concrete concepts given above: a concrete concept is simply a concept of a tangible object. Clements’ definition of *sensory-concrete* knowledge is also captured by this notion of tangibility: “We have *sensory-concrete* knowledge when we need to *use* sensory material to make sense of an idea” (Clements, 2000, p.47).

Perhaps a more compelling definition for a concrete concept is one provided by proponents of concreteness fading: “concrete materials... are grounded in perceptual and/or motor experiences” (Fyfe et al., 2014, p.9). i.e. grounded in experiences in which we have seen or touched something tangible.

Importantly, the latter definition brings to the fore one of the reasons why describing concepts as concrete (qua tangibility) is considered to have explanatory value. Some theorists argue that concepts grounded in perceptual-motor experiences are thought to be more easily grasped by students, and provide a concrete foundation on which to build more abstract knowledge. Indeed, this is a fundamental thesis of a large network of theoretical frameworks including embodied cognition, embedded cognition and conceptual metaphor theory. (See Pouw, van Gog, and Paas (2014)). It is important to note that in order for concrete materials to have these proposed advantages, it is not enough that they *could be* grounded in perceptual and motor experiences. Rather, a student must actually have interacted with these materials. This highlights one of the important differences between using concreteness as a descriptive and an explanatory tool. One can objectively *describe* a representation of a football as being tangible (because it is potentially grounded in perceptual motor experiences), but in order to *explain* why this representation of a tangible object is beneficial, it is essential to connect it to a subject who has had some sensory familiarity with the football. The importance of this *relationship* one has with the object or concept is stressed both by Clements (2000) and Wilensky (1991): “concreteness is not a property of an object but rather a *property of a person's relationship to an object*” (Wilensky, 1991, p.4) Thus, the notion of familiarity modulates the explanatory power of the notion of tangibility in education research.

3. Familiarity

The notion of familiarity will not be found in serious philosophical discussion regarding the *abstract/concrete* distinction, yet it is foundational in lay and educational discourse. It is also a key factor underlying the conceptual confusion around concreteness. An excerpt from a tutorial at www.projectlearnnet.org provides a clear example of lay usage:

*Discussing similarities and differences between that which is unfamiliar and distant (i.e., abstract) and that which is familiar and close to home (i.e., concrete) is a valuable way to help students grasp the abstract concept.*³⁶

³⁶ http://www.projectlearnnet.org/tutorials/concrete_vs_abstract_thinking.html

Examples from educational research include Fyfe et al. (2014): “Concreteness fading assumes that learners can easily comprehend the concrete materials, [and] are unfamiliar with the abstract materials...” (p.16). Fyfe et al. also define concrete materials as “connecting with learners’ prior knowledge” (p.9). Podolefsky et al. (2008) describe the conservation of energy approach to solving a problem as an abstract approach, and yet they write: “... in practise experts may treat the conservation of energy approach in a concrete way. That is, this more abstract approach is quickly recognized and readily accessible to the expert.” Here it seems that an abstract concept, the conservation of energy, may be treated in a concrete manner, as its familiarity allows the expert to use it as readily accessible tool. Clements (2000) defines a particular type of concreteness, “integrated-concrete”, as referring to “concepts that are ‘concrete’ at a higher level because they are connected to other knowledge”, and this connectedness makes these concepts as “accessible and useful... as a wrench to a plumber” p.48. Conversely, a rugby ball may be described as a concrete concept, yet for a subject with no familiarity with rugby balls, it might be described as an unfamiliar, or abstract, idea. Perhaps familiarity is a figurative extension of tangibility: a concrete concept is one which is not only tangible with our hands, but also one which our minds can easily grasp (quickly recognise and readily access), and which can serve as a concrete foundation upon which we can build further knowledge.

3.1. Familiarity and Tangibility

We have shown that in the educational literature, “concreteness” can be associated with both “tangibility” and “familiarity”, which have different criteria of measurement. We might describe a concept as being concrete because it is grounded in perceptual motor experiences, or we might describe a concept as being concrete because it is connected with prior knowledge. In this section we shall illustrate that “tangibility” and “familiarity” might be thought to constitute two separate axes of a “concreteness space”. Importantly, this is a purely descriptive space.

If we map the literal notion of tangibility to the term “objectively concrete”, and the notion of familiarity to the notion “subjectively concrete”, we might construct an “concreteness space” as shown in Figure 61 **Error! Reference source not found.** with tangibility (“objective concreteness”) on the x axis, and familiarity (“subjective concreteness”) on the y axis.³⁷ Note that in our diagram, the origin is the place of maximum concreteness, and concreteness *decreases* as one moves along the axes. Hence our labelling of the x and y axes with the terms “intangibility” and “unfamiliarity” respectively. To illustrate the use of this diagram, consider a Brazilian high school physics student who is very familiar with footballs and electric fields, but has never encountered a rugby ball, or the

³⁷ Acknowledgement to Christine Lindstrom who helped develop this particular diagram

idea of a gravitational field. A football would be both objectively and subjectively concrete; it is a tangible object which the student has played with on numerous occasions. A rugby ball would be objectively concrete, but subjectively abstract. While an electric field is commonly thought to be an abstract object (or concept), it would be subjectively concrete for this student as she would have encountered this concept at various stages in high school. The gravitational field is both intangible (objectively abstract) and unfamiliar (subjectively abstract).

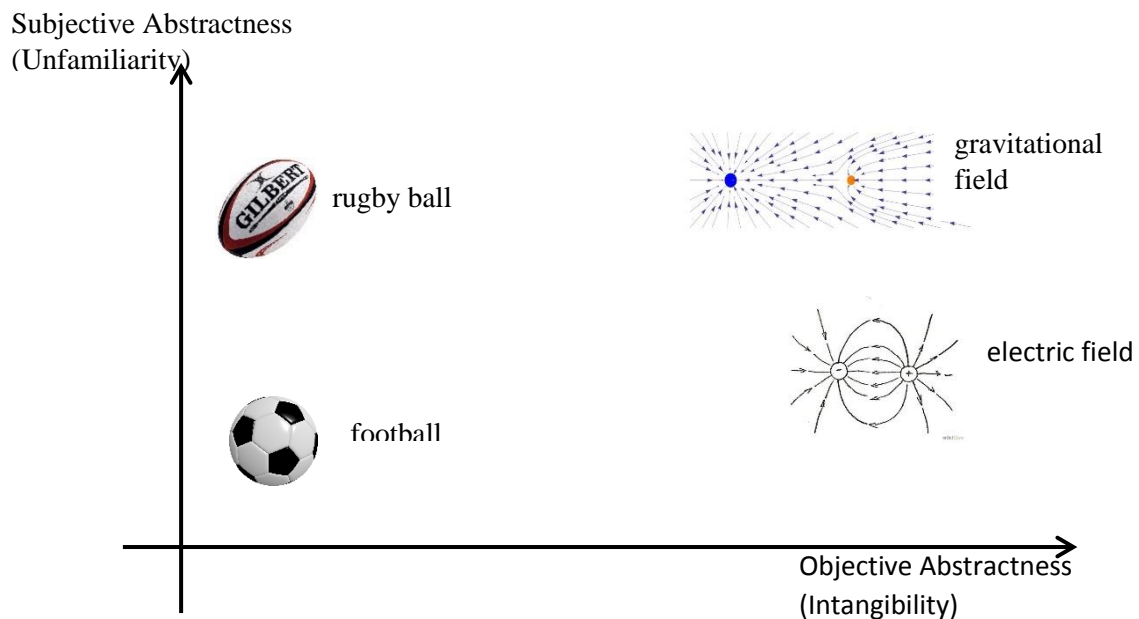


Figure 61: The “concreteness space” for a Brazilian high school physics student, who is unfamiliar with rugby balls and the idea of a gravitational field.

Note that in the diagram a subject’s knowledge is built “upward”; we understand less familiar things by means of more familiar things. Indeed, the student might be taught the meaning of a rugby ball by being told: “Well, it’s like a football, but...”. Thus, her understanding of “rugby ball” would be aided by means of her more readily accessible and familiar idea of a football. Similarly, if the student had never encountered the notion of a gravitational field (making it subjectively abstract), her knowledge could be built on her “subjectively concrete” knowledge of electric fields.³⁸ Note that as one becomes more familiar with an object, that object will *move* along the vertical axis. An object’s “subjective concreteness” is therefore a variable whose value lies on a continuum. However

³⁸ Importantly: there is no *reduction* here: his understanding of magnetic fields is not reducible to his understanding of electric fields. Rather he will recruit knowledge structures that underlie his understanding of electric fields, in his understanding of magnetic fields. Note further that this notion of “knowledge structure” is also not a reducible notion, but a helpful heuristic, in noticing similarities?!

“objective concreteness” is a constant, and according to some definitions, might be thought of as a dichotomous variable; an object is either tangible or it is not.

The distinction between an “objectively concrete object” and a “subjectively concrete object” maps fairly well onto the distinction proposed by Clements between sensory-concrete knowledge and integrated-concrete knowledge: “*Sensory-Concrete* refers to knowledge that demands the support of concrete objects and children’s knowledge of manipulating these objects. *Integrated-concrete* refers to concepts that are ‘concrete’ at a higher level because they are connected to other knowledge...” (Clements, 2000, p.49) As the notion of “subjectively concrete” is related to the notion of knowledge construction, one might argue that it is better suited for the purposes of educational research. Indeed, in his call for a re-evaluation of the notion of “concreteness”, Wilensky argues: “Concreteness, then, is that property which measures the degree of our relatedness to the object, (the richness of our representations, interactions, connections with the object)...” (Wilensky, 1991, p.4).

Finally, one might associate “objective concreteness” with the activity of description and “subjective concreteness” with the activity explanation. As argued above, an object may be described as being tangible or not, but this description is explanatorily irrelevant without knowledge of a subject’s familiarity with the object or concept. In other words, tangibility (“objective concreteness”) of educational materials may be of potential benefit for students, but this benefit can only be explicated by a theoretical framework such as embodied cognition. In order to use such an explanatory theoretical framework, the educational materials must be both tangible *and* familiar (“subjectively concrete”). To re-iterate: “familiarity” modulates the explanatory power of “tangibility”. We now turn to the explanatory power of “familiarity”.

3.2. Two Models of Familiarity

Familiarity is not a technical term, and related concepts such as knowledge, memory and recognition/perception all represent large, complex areas of research in cognitive science. It is nevertheless used as a type of explanatory tool in the education literature. Importantly, familiarity may be seen as being both a help and a hindrance to learning. The most obvious, common sense idea is that familiarity aids learning. However, familiarity can also be seen as a hindrance to learning. Our approach will be to highlight two ideas connected with the notion of familiarity as they occur in education literature. The first follows from Sfard’s (2008) notion of objectification, as well as the notion of “chunking” in cognitive science. These perspectives view familiarity as an aid to learning. We shall dub this idea “**familiarity means less**”. The second follows from a “network activation”

model of knowledge. Our familiarity with a concept may thus be modelled as being proportional to both the number and strength of connections it has with other concepts in a network. We shall dub this idea “**familiarity means more**”.

3.2.1. Familiarity as a Help (Familiarity Means Less)

Clements (2000) and Wilensky (1991) view familiarity as the foundation of knowledge construction. The degree of connectedness of a concept in a knowledge network, is the degree to which it can be used as a building block on which to build further knowledge.

Familiarity with a particular routine or use of a concept can also be seen as fundamental to knowledge construction. Sfard (2008) provides a compelling account of how our activities become reified into concepts. For example, our concept of the number “five” is, according to Sfard, based on the reification of the counting routine “one, two, three, four five”. The pattern of activity becomes sufficiently familiar to earn a label, “five”, and become a unitary mental object.

A further, relevant theoretical framework is that of working memory (Baddeley, 2007). Our working memory capacity is roughly the amount of “stuff” we can hold in mind at one time. This capacity is limited, but can be augmented via the process of “chunking”. Consider the following two strings of digits: 49720361 and 19391945. The former might be considered as consisting of eight chunks of information, while the latter only two chunks – the dates for the commencement and cessation of the second world war. It is the connection of those strings of four digits (1939 and 1945) to other concepts in a knowledge network that allows us to chunk them and free up space in working memory for other cognitive processes. Sfard’s notion of reification may also be thought of a type of “chunking” of an activity. For example, the activity of performing arithmetic on our fingers is “chunked” to become a less burdensome mental activity. This is also the train of thought behind embedded embodied cognition, where we become sufficiently familiar with procedures that we can mentally simulate them with little effort (Pouw et al., 2014).

Thus, familiarity with a concept, or with processes or activities underlying concepts, may be said to reduce the cognitive load on working memory, helping the learner to assimilate new information and build their knowledge network.

3.2.2. Familiarity as a Hindrance (Familiarity Means More)

While gaining familiarity with a concept is arguably synonymous with “learning”, familiarity can also be understood as a hindrance to learning. Consider Figure 62a below. The black square represents a complex concept which is connected to other concepts in a knowledge network. The triangle represents that which is to be learned. The essential idea is that our familiarity with the other

elements in the knowledge network can hinder or distract us from learning the element represented by the triangle.

For example: consider the square to be the complex notion of force, and the triangle to be Newton's first law. Our concept of force is connected to thousands of experiences of forces. One particular pattern of experience which is reified into a type of concept, is the notion of "*force as mover*": when we see an object moving, we assume that there must be a force causing this movement. The notion of *force as mover* is a hindrance to learning Newton's 1st Law which states that if an object is moving at a *constant* velocity then there is *no* (resultant) force acting. This has been well documented in physics education research, and *force as mover* has been labelled as a common misconception of the concept of force. In other words, it is our familiarity with a vast network of experiences which hinders us from learning this particular aspect of the concept of force. Similar effects have been shown for trying to learn circuits (John & Allie, 2013).

These examples can also be understood from the perspective of working memory. If a particular image or concept is familiar (i.e. has many connections) it is likely to invoke a lot of information in the mind of the student (Kaminski et al., 2013). This "weight" of information might overload working memory and thus hinder other cognitive processes. Sanchez and Wiley (2006) refer to this as the "seductive details effect".

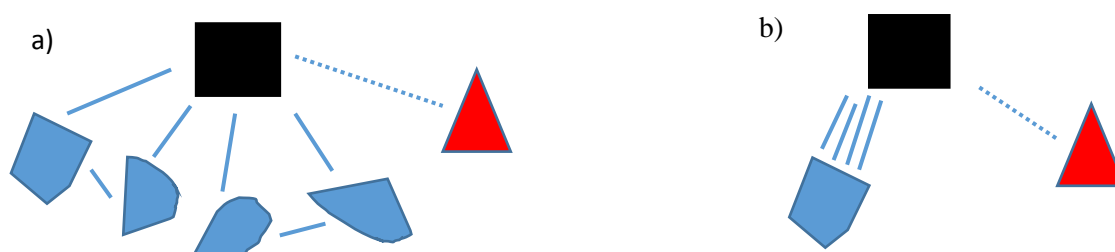


Figure 62: Two models of how familiarity might hinder learning. The black square represents a concept connected to other familiar concepts in a knowledge network. The triangle represents something to be added to the knowledge network.

Figure 62b models a mechanism Petersen & McNeil (2013) call "functional fixedness". A large body of research suggests that children (and adults) often resist changing their well-established knowledge of objects, concepts, or procedures (e.g., Duncker, 1945; Mack, 1995; McNeil & Alibali, 2005). Children seem to have a particularly difficult time using an object symbolically once they have already gained experience with that object in a play setting (DeLoache, 2000). Another way of phrasing this particular hindrance is that familiarity leads to context boundedness.

In summary: (i) familiarity qua chunking, can be thought to help learning by freeing up working memory, and decreasing cognitive load. However, (ii) familiarity qua number and strength of connections, can be thought to hinder learning either by (a) preventing the learner from forming new network connections (see Figure 62a) or (b) by “functional fixedness” (see Figure 62b).

4. Perceptual Richness

This is one of the more widely used notions of concreteness in the education literature. In their paper *Relevant Concreteness and its Effects on Learning and Transfer*, Kaminski et al. (2005) operationalize the term concreteness as “the amount of information communicated through a specific instantiation of a concept” (p.15). By communicated information they mean “information activated in the mind of the observer”. They claim that “perceptual information can be increased by adding visual information or increasing the perceptual richness of the symbols” (p.16). In the literature on concreteness fading the authors provide a related definition: “concrete materials... have identifiable correspondences between their forms and referents” (Fyfe et al., 2014, p.9). For example, in a computer simulation designed to teach competitive specialisation, Goldstone & Sakamoto (2003) use two representations of ants eating food: the first involves iconic representations of ants and fruit, and the second involves dots (to represent the ants) and blobs (to represent the food). The former, perceptually rich scenario, obviously has a greater identifiable correspondence between the forms (graphical representations of ants and fruit) and referents (ants and food).

The perceptual richness of an instantiation of a concept is generally thought to hinder students’ ability to transfer for the reasons illustrated in Figure 62. A perceptually rich representation may be thought to communicate more information than a more generic or sparse representation. This excess of perceptual information occupies space in working memory, increasing cognitive load, and hindering other cognitive processes.

Another avenue of explanation is via the notion of familiarity. A perceptually rich representation of an object or a concept is likely to activate knowledge elements that we familiarly associate with such an object/concept. We may therefore struggle to build or connect new knowledge elements. Petersen and Mcneil (2013) refer to this phenomenon as “reduced salience of the set”, and it is related to the notions of “functional fixedness” and “dual representation” mentioned above.

Although both of these avenues of explanation are used in relation to the notion of perceptual richness, it seems that the conflation of perceptual richness with familiarity is unfounded.

4.1. Perceptual Richness and Familiarity

Kaminski et al. (2013) identify two components of concreteness in their definition: perceptual concreteness and conceptual concreteness. The former is increased by adding perceptual detail, and is a measure of perceptual richness. However, conceptual concreteness is increased “by describing familiar situations or adding detail” (p.16). Conceptual concreteness may thus be thought to be a measure of familiarity, and as “information activated in the mind of the observer” it may be thought of as the number of networks that are activated. As Kaminski et al. state: “Much more information would be activated if the image were of someone familiar” (ibid). According to our taxonomy, Kaminski et al. have confounded two aspects of concreteness: Perceptual Richness and Familiarity. We will argue that familiarity (conceptual concreteness) and perceptual richness (perceptual concreteness) do not form a unified measure of concreteness, but may be thought of as independent measures resulting in different effects.

The most comprehensive demonstration of the parsing of familiarity and perceptual richness is a study by Petersen & Mcneil (2013) (hereafter referred to as the Petersen study). In seeking to clarify whether concrete materials are helpful in teaching children counting, Petersen divided the counting objects into those that were perceptually rich and those that were familiar. Children were randomly assigned to object type in a 2 (familiarity: high or low) X 2 (perceptual richness: high or low) factorial design. Note that Peterson and McNeil used the label “established knowledge” rather than “familiarity”. In one experiment the objects used were a plastic strawberry and pear (familiar and perceptually rich), two pencils (familiar and perceptually sparse), two glittery fuzz balls (unfamiliar and perceptually rich) and two plastic discs (unfamiliar and perceptually sparse). Students were tested in two widely used counting tasks. Their performances in the various conditions are shown in Figure 63. (Note once again that the axes measure “unfamiliarity” and “sparseness”; high familiarity and high perceptual richness occur at the origin). Neither familiarity nor perceptual richness predict performance. Rather, perceptually rich objects result in low performance if the objects are familiar to the children (strawberry and pear), and in high performance if the objects are unfamiliar to children (glittery fuzz balls). And vice versa for perceptually sparse objects.

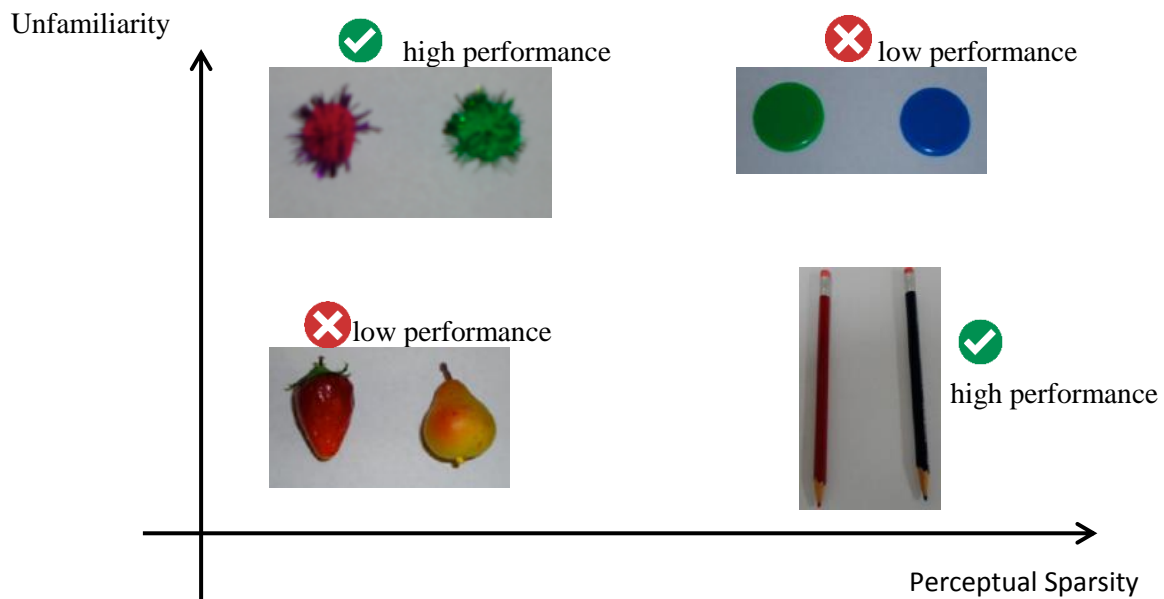


Figure 63: Representation of the results from Petersen and McNeil's study in which student performance on a counting task was found to depend both on familiarity and perceptual richness

One possible salient difference between the studies of Kaminski et al. and Peterson is that the latter deals with concrete objects, while the materials used in Kaminski's study were graphical representations. However, the study by Goldstone & Sakamoto (2003) (hereafter referred to as the Goldstone study) used graphical representations and might be interpreted as having a similar result to Petersen. Students learnt the concept of competitive specialisation through a computer simulation involving ants and food. As described above, and shown in Figure 64 below, half of the students were given a perceptually rich simulation in which the ants and food were represented by graphical representations of ants and fruit. The other half of the students learnt the concept using a perceptually sparse simulation involving dots and blobs. The students were tested both on the initial simulation they had learnt, and on a transfer task. In analysing the transfer scores, Goldstone observed an interaction between the perceptual richness of the initial simulation, and performance on the initial simulation. Consider the high performers on the initial simulation: these students were more likely to obtain a better transfer score if they had learnt on a perceptually rich simulation. The opposite pattern was observed for poor performers, who obtained better transfer scores if they had learnt on a perceptually sparse simulation.

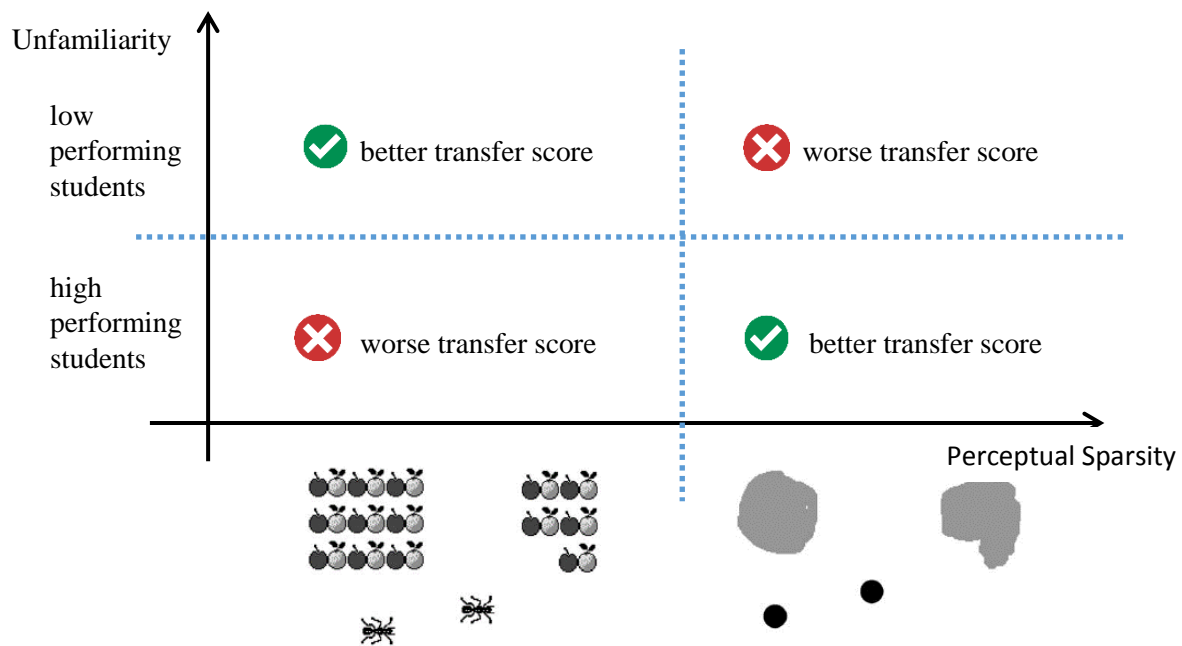


Figure 64: After learning the concept of competitive specialisation through either a perceptually rich or sparse simulation, students were tested in a transfer test. Their results for the transfer test depended both on the concreteness of the initial simulation (x axis), and on their performance in the initial simulation (y axis).

The validity of the comparison between the Petersen study and the Goldstone study turn on our understanding of “higher prior knowledge”. In Petersen’s study “higher prior knowledge” referred to a familiarity with an object; children had played a game with the object and could be said to have a particular use for that object. In Goldstone’s study, “higher prior knowledge” is measured by performance on an initial test. In other words, “higher prior knowledge” may refer to students’ prior knowledge that facilitated their understanding of the concept of competitive specialisation. One might say that their “higher prior knowledge” constituted a type of familiarity with this concept. If “prior knowledge” is judged as sufficiently similar in both studies, then they both demonstrate that “perceptual richness” interacts with prior knowledge, where “prior knowledge” might be thought of as a type of familiarity with a particular object or concept.

In any event, going back to the original definition of Kaminski et al., it should be clear that perceptual concreteness (qua the amount of perceptual detail) and conceptual concreteness (qua familiarity or the extent of a semantic network) are separate measures of concreteness, and might be categorized as perceptual richness and familiarity respectively. This is made further apparent when one considers the underlying theoretical frameworks

4.2. Perceptual Richness as an Explanatory Tool

The abovementioned authors have a variety of ways of explaining why perceptual richness may have an effect on learning. The general idea is that perceptual richness distracts or hinders students from learning that which we wish them to learn. This distraction is usually understood through Baddeley's (2007) model of working memory: We have a limited capacity for the amount of information we can hold in mind; if some of that capacity is taken up by perceptual detail, less is available for learning that which ought to be learnt. One can also make use of Cowan's (2005) model of working memory which focuses on the notion of "attention". For example, in Petersen's experiment one might hypothesize that the sparkling appearance of the fuzz balls will capture the children's attention and distract them from the counting task.

Although the notions of "perceptual richness" and "familiarity" both use underlying theoretical frameworks of working memory, the mechanisms are different. "Perceptual richness" deals with a type of "perceptual information overload", while "familiarity" deals with a surplus of network connections. Furthermore, as argued in section Familiarity as a Help (Familiarity Means Less) 3.2.1 familiarity can also be modelled as *preventing* information overload, through the process of chunking, or streamlining information.

5. Conclusion

This Appendix draws attention to the fact that "concreteness" has numerous meanings and uses within the education literature, and that these are not aligned. We have drawn attention to three possible meanings: tangibility, familiarity and perceptual richness. We have argued that these terms are different from each other, not only as descriptive tools, but also as explanatory tools. As descriptive tools, we have argued that these terms describe saliently different features of a scenario, and as explanatory tools, we have argued that they use different underlying theoretical frameworks.

We have argued that familiarity and tangibility might constitute two different axes of concreteness; a subjective and objective measure, respectively. In education research, we are interested in the effect concreteness might have on students' performance. Therefore, we have argued that it is the students' subjective perception of concreteness that is a more salient measure than an objective measure. We have also argued that familiarity is distinct from perceptual richness. We have pointed to empirical evidence that familiarity modulates the effect of perceptual richness on student performance. Finally we have argued that familiarity itself can be interpreted as having two different types of effect on student performance.

In sum, we have argued for a more nuanced language of description for why students might benefit from or be hindered by particular educational materials. Instead of using the notion of “concreteness” we can use notions of tangibility, familiarity and perceptual richness. And perhaps more importantly, it is incumbent upon education researchers to explain how a particular theoretical framework underpins the term they use.

Appendix B:

Supplementary Material for Vector Addition

First Coding Key 2014 Qualitative Data

Mentionings

DIR	Direction	Some mention of the importance of direction
PYTH	Pythagoras	Some mention of pythagoras
EFF	Effect	Some mention of the phenomenon of effect
	1	Just mention of effect or result
	2	Effect or Result as pertains to the NET force
CMP	Components	Some mention of components
ADD	Addition	Do they make some reference to the type of addition. e.g. 'simply add'
EXPE	Experience	

Types of Explanation

EXPL	Explanatory Value	1 - 10. Does it just restate the distinction, or actually give some reason WHY...
COH	Coherence	Does the explanation make sense? If A: Does it give a reasonable explanation for why they are the same. B: Does it pick up any type of distinction between net and total. C: etc.
INC	Inconsistency	Some inconsistency in reasoning
	1	General Inconsistency
	2	Given their reasoning, should have changed their answer (5 or 7) but didn't
PRO	Procedure	Does their explanation for the difference subsist in giving a different procedure of calculation
	1	Mention of procedure and something else
	2	Mention of only procedure
ONT	Ontologies	Do they have different ontologies for the two separate forces and the combined force
EMP	Emphasis	Is the emphasis on why the total is 7 or why the net is 5
	1	Why is the total 7
	2	Why is the net 5
SIM	Simultaneity	Do they argue that simultaneity of the forces is a factor
AGA	AGA	Do the forces work against each other?

- 1 Against each other
- 2 Do not effect each other

Second Coding Key 2014 Qualitative Data

- Direction is the reason for distinguishing [33]
- Direction is the reason they cannot be added [10]
- Direction General (Opposing Forces) [5]
- Force is Lost [4]
- Effect, Result, Movement and Experience? [30]
 - Movement
 - Effect
 - Result
 - Experience
- Ontology of total force and net force. Combination and Splitting [16]
- Explicit Mention of “vector” and “scalar” [14]
- “Just adding” versus “something else” such as “the resultant” [63]
- Net Force is Resultant Force [5]
- Explicit Procedural [31]

Supplementary Material for FCI

Familiarity Rating Questionnaire

This questionnaire contains 7 questions.

Do NOT answer these questions!

ONLY provide a rating of familiarity of the questions.

The familiarity rating scale is from 0 – 5.

A rating of 5 means the question is *very familiar* - you have seen many questions of that type before.

A rating of 3 means that the question is *somewhat familiar* – you have seen some questions of that type before.

A rating of 0 means that the question is *completely unfamiliar* – you don't recognize any of the concepts in the question.

1. A car is on a highway and slows down from a speed of 9 m.s^{-1} to a speed of 5 m.s^{-1} in half a second. What is the average acceleration of the car during this time?

Familiarity Rating (0 – 5):

2. A wavefunction $\psi(x)$ has been expressed as a sum of energy eigenfunctions:

$$|\psi\rangle = \sum_n c_n |u_n\rangle$$

Viewing $|\psi\rangle$ as a vector in a Hilbert space, explain what the $|u_n\rangle$ are.

Familiarity Rating (0 – 5):

3. You exert a force of 200 N on a box of mass 20 kg as you push it over a rough surface. The frictional force on the box is 50 N. What is the acceleration of the block?

Familiarity Rating (0 – 5):

4. Two metal balls are the same size but one weighs twice as much as the other. The balls are dropped from the roof of a single story building at the same instant of time. Will the balls land on the ground at the same time?

Familiarity Rating (0 – 5):

5. For a simple harmonic oscillator, such as a mass on a spring, what happens to the period if the spring constant is doubled?

Familiarity Rating (0 – 5):

6. A large truck collides head-on with a small compact car. During the collision, does the truck exert a larger force on the car than the car exerts on the truck?

Familiarity Rating (0 – 5):

7. A woman exerts a constant horizontal force on a large box. As a result, the box moves across a horizontal floor at a constant speed “ v ”. Is the horizontal force applied by the woman greater than, or equal to the frictional force on the box?

Familiarity Rating (0 – 5):

References

- Abbott, A. . (1969). *Ordinary Level Physics* (2nd Ed.). London: Heinemann Educational Books Ltd.
- Adair, R. K. (1969). *Concepts in Physics*. New York: Academic Press.
- Allie, S., & Buffler, A. (1998). A course in tools and procedures for Physics I. *American Journal of Physics*, 66, 613.
- Baddeley, A. (2000). The episodic buffer : a new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417–423. Retrieved from <http://eprints.whiterose.ac.uk/66016/>
- Baddeley, A. (2007). *Working Memory, Thought, and Action*. Oxford: Oxford University Press.
- Baker, G. P., & Hacker, P. M. S. (2008). *Wittgenstein: Understanding and meaning: Volume 1 of an analytical commentary on the philosophical investigations, part I: Essays* (Vol. 1). John Wiley & Sons.
- Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Retrieved from <https://books.google.co.za/books?hl=en&lr=&id=H41WUfTU2CMC&oi=fnd&pg=PT7&dq=karen+barad+universe&ots=MH-86zzPFW&sig=dwahvviDVhaSrT5Jsv7JFqbt2-Q>
- Barniol, P., & Zavala, G. (2010). Vector addition: Effect of the context and position of the vectors. In *AIP Conference Proceedings* (Vol. 1289, pp. 73–76).
- Barniol, P., & Zavala, G. (2014a). Force, velocity, and work: The effects of different contexts on students' understanding of vector concepts using isomorphic problems. *Physical Review Special Topics-Physics Education*. Retrieved from <https://journals.aps.org/prper/abstract/10.1103/PhysRevSTPER.10.020115>
- Barniol, P., & Zavala, G. (2014b). Test of understanding of vectors: A reliable multiple-choice vector concept test. *Physical Review Special Topics-Physics Education*. Retrieved from <https://journals.aps.org/prper/abstract/10.1103/PhysRevSTPER.10.010121>
- Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(4), 577–660. Retrieved from http://journals.cambridge.org/abstract_S0140525X99002149
- Barsalou, L. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*. <http://doi.org/10.1080/769813547>
- Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1521), 1281–1289.
- Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. *Symbols, Embodiment, and Meaning*, 245–283.
- Beiser. (1973). *Physics*. Menlo Park: Cummings Publishing Company Inc.
- Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Blackwell Publishing.
- Bergen, B. (2012). Louder than words. *The New Science of How the Mind Makes Meaning*.
- Brook, A. (1997). Approaches to abstraction: A commentary. *International Journal of Educational*

- Research*, 27(1), 77–88. [http://doi.org/10.1016/S0883-0355\(97\)88445-5](http://doi.org/10.1016/S0883-0355(97)88445-5)
- Brown, D. E., & Clement, J. (1989). Overcoming misconceptions via analogical reasoning: abstract transfer versus explanatory model construction. *Instructional Science*, 18(4), 237–261. <http://doi.org/10.1007/BF00118013>
- Brown, M. C., McNeil, N. M., & Glenberg, A. M. (2009). Using concreteness in education: Real problems, potential solutions. *Child Development Perspectives*, 3(3), 160–164.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121–152. http://doi.org/10.1207/s15516709cog0502_2
- Clements, D. H. (2000). Concrete manipulatives, concrete ideas. *Contemporary Issues in Early Childhood*, 1(1), 45–60.
- Cole, M. (1996). Cultural psychology. *A Once And future Discipline*. Retrieved from <http://www.jstor.org/stable/pdf/681304.pdf>
- Cooper, J. M., Hutchinson, D. S., & others. (1997). Plato: complete works.
- Cowan, N. (2005). *Working memory capacity*. (H. L. Roediger, J. S. Nairne, J. Neath, & A. M. Surprenant, Eds.) *Cognition & Emotion* (Vol. 2). Psychology Press. <http://doi.org/10.1080/02699938808408071>
- Crummett, W., & Western, A. (1994). *University Physics*. Dubuque: Brown Communications Inc.
- Cutnell, J. D., & Johnson, K. W. (1995). *Physics* (3rd Editio). New York: John Wiley & Sons.
- Cutrona, L. J. (2008). “Concrete” examples a fraction too abstract [Letter]. *Science*, 322, 1632. Retrieved from http://www.sci.sdsu.edu/CRMSE/personal_pages/sreed/ScienceLetter.pdf
- De Bock, D., Deprez, J., Van Dooren, W., Roelens, M., & Verschaffel, L. (2011). Abstract or Concrete Examples in Learning Mathematics ? A Replication and Elaboration of Kaminski , Sloutsky , and Heckler ’ s Study. *Journal for Research in Mathematics Education*, 42(2), 109–126.
- Dervin, B. (1983). An overview of sense-making research: Concepts, methods, and results to date. *International Communication Association Annual Meeting*. <http://doi.org/10.1038/cddis.2011.1>
- Deventer, J. Van, Wittmann, M., & Hsu, L. (2007). Comparing student use of mathematical and physical vector representations. *AIP Conference*. Retrieved from <http://aip.scitation.org/doi/abs/10.1063/1.2820935>
- DiSessa, A. (2002). Why “conceptual ecology” is a good idea. In *Reconsidering conceptual change: Issues in theory and practice* (pp. 28–60). Springer.
- DiSessa, A. (2006). A History of Conceptual Change Research: Threads and Fault Lines. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 265–281). Cambridge University Press. Retrieved from <http://psycnet.apa.org/psycinfo/2006-07157-016>
- Disessa, A. a. (2002). Why “Conceptual Ecology” is a Good Idea. *Reconsidering Conceptual Change: Issues in Theory and Practice*, 28–60. http://doi.org/10.1007/0-306-47637-1_2
- diSessa, A. A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction*, 10(2–3), 105–225. <http://doi.org/10.1080/07370008.1985.9649008>
- diSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education*.

- diSessa, A., & Wagner, J. (2005). What Coordination Has to Say about Transfer in J. Mestre (Ed.), *Transfer of learning: Research and perspectives* (pp. 121-154). Greenwich.
- Dreyfus, T., Herschkowitz, R., & Schwarz, B. (2001). Abstraction in context: The case of peer interaction. *Cognitive Science Quarterly*, 1(3), 307–368.
- Elby, A. (1999). Another reason that physics students learn by rote. *American Journal of Physics*. Retrieved from <http://aapt.scitation.org/doi/abs/10.1119/1.19081>
- Engeström, Y., Miettinen, R., & Punamäki, R. (1999). *Perspectives on activity theory*. Retrieved from https://books.google.co.za/books?hl=en&lr=&id=GCVCY2xHD4C&oi=fnd&pg=PR10&dq=activity+theory&ots=l_EPDD6mV&sig=Y3AnU5ntI8hAEEWZUghD5VapRok
- Entwistle, N. J., & Peterson, E. R. (2004). Conceptions of learning and knowledge in higher education: Relationships with study behaviour and influences of learning environments. *International Journal of Educational Research*, 41(6), 407–428. <http://doi.org/10.1016/j.ijer.2005.08.009>
- Fauconnier, G., & Turner, M. (2008). *The way we think: Conceptual blending and the mind's hidden complexities*. Basic Books.
- Finkelstein, N. (2005). Learning physics in context: A study of student learning about electricity and magnetism. *International Journal of Science Education*. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/09500690500069491>
- Flores, S., Kanim, S., & Kautz, C. (2004). Student use of vectors in introductory mechanics. *American Journal of Physics*. Retrieved from <http://aapt.scitation.org/doi/abs/10.1119/1.1648686>
- Fodor, J. (1975). *The language of thought*. Retrieved from https://books.google.co.za/books?hl=en&lr=&id=XZwGLBYLbg4C&oi=fnd&pg=PR7&dq=fodor+1975&ots=_mmiKKXmHZ&sig=i_JaSW8mkl2k1aL-rt8qPsrMqPU
- Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness Fading in Mathematics and Science Instruction: A Systematic Review. *Educational Psychology Review*.
- Gibson, J. J. (1979). *The ecological approach to visual perception. of Experimental Psychology Human Perception and (Vol. 39)*.
- Gilbert, J. K., & Watts, D. M. (1983). Concepts, Misconceptions and Alternative Conceptions: Changing Perspectives in Science Education. *Studies in Science Education*. <http://doi.org/10.1080/03057268308559905>
- Giraldo, V., & Calvalho, L. (2006). Generic organizer for the enrichment of the concept image of derivative. *Proceedings of the 30th PME International*. Retrieved from <http://files.eric.ed.gov/fulltext/ED496933.pdf#page=193>
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Retrieved from <http://doi.apa.org/psycinfo/1975-09476-000>
- Goldstone, R. L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology*, 46(4), 414–466.
- Grady, J. (2005). Primary metaphors as inputs to conceptual integration. *Journal of Pragmatics*, 37(10 SPEC. ISS.), 1595–1614.
- Greeno, J. G., Moore, J. L., & Smith, D. R. (1993). Transfer of situated learning.
- Hacker, P. (1993). Wittgenstein: Meaning and Mind, Volume 3 of an Analytical Commentary on the Philosophical Investigations, Part II: Exegesis 243-247. Retrieved from

<https://philpapers.org/rec/HACWMA-5>

- Hake, R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*. Retrieved from http://pdfserv.aip.org/AJPIAS/vol_66/iss_1/64_1.pdf
- Hale, S. C. (1988). Spacetime and the abstract/concrete distinction. *Philosophical Studies*, 53(1), 85–102.
- Halloun, I. A., & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics*, 53(11), 1043. <http://doi.org/10.1119/1.14030>
- Hammer, D. (2000). Student resources for learning introductory physics. *American Journal of Physics*, 68(S1), S52. <http://doi.org/10.1119/1.19520>
- Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing*, (January 2002), 169–190. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.114.7178>
- Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. *Transfer of Learning from a Modern Multidisciplinary Perspective*, (Rec 0087519), 89–120. <http://doi.org/10.1002/car.1158>
- Hampe, B. (2005). *From Perception to Meaning: Image Schemas in Cognitive Linguistics*. *Cognitive Linguistics* (Vol. 8). Retrieved from <http://books.google.com/books?hl=en&lr=&id=W7rfP-eliyOC&pgis=1>
- Harris, R. (1990). *Language, Saussure and Wittgenstein: How to play games with words*. Retrieved from https://books.google.co.za/books?hl=en&lr=&id=IU0hLaYBa-MC&oi=fnd&pg=PR9&dq=harris+saussure+wittgenstein&ots=kAaD1_rgvT&sig=VP_4SF04LSw5J_9o9E0Tjfw98
- Harris, R. (2005). *The semantics of science*. A&C Black.
- Heckler, A. F., & Scaife, T. M. (2015). Adding and subtracting vectors: The problem with the arrow representation. *Physical Review Special Topics - Physics Education Research*, 11(1), 10101. <http://doi.org/10.1103/PhysRevSTPER.11.010101>
- Helmig, C. (2013). *Forms and concepts: concept formation in the Platonic tradition* (Vol. 5). Walter de Gruyter.
- Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics*. <http://doi.org/10.1119/1.15129>
- Hestenes, D., & Wells, M. (1992). A mechanics baseline test. *The Physics Teacher*. Retrieved from <http://aapt.scitation.org/doi/pdf/10.1119/1.2343498>
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30(3), 141–158. <http://doi.org/10.1119/1.2343497>
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*. Retrieved from <http://aapt.scitation.org/doi/pdf/10.1119/1.2343497>
- Hintzman, D., Curran, T., & Oppy, B. (1992). Effects of similarity and repetition on memory: Registration without learning? *Learning, Memory, and Cognition*. Retrieved from <http://psycnet.apa.org/journals/xlm/18/4/667/>
- Hintzman, D. L., & Curran, T. (1995). When encoding fails: Instructions, feedback, and registration

- without learning. *Memory & Cognition*, 23(2), 213–226. <http://doi.org/10.3758/BF03197223>
- John, I., & Allie, S. (2013). DC circuits: Context dependence of student responses. In *2012 PHYSICS EDUCATION RESEARCH CONFERENCE* (Vol. 1513, pp. 202–205).
- Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. University of Chicago Press.
- Johnston, W., & Hawley, K. (1994). Perceptual inhibition of expected inputs: The key that opens closed minds. *Psychonomic Bulletin & Review*. Retrieved from <http://link.springer.com/article/10.3758/BF03200761>
- Jones, M. G. (2009a). Examining Surface Features in Context. *Journal for Research in Mathematics Education*, 40(2), 94–96.
- Jones, M. G. (2009b). Transfer, Abstraction, and Context. *Journal for Research in Mathematics Education*, 40(2), 80–89.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux. Retrieved from <http://books.google.co.za/books?id=ZuKTvERuPG8C>
- Kaminiski, J. A., Sloutsky, V. M., & Heckler, A. F. (2006). Effects of concreteness on representation: An explanation for differential transfer. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 1581–1586).
- Kaminski, J. A., & Sloutsky, V. M. (2012). Representation and transfer of abstract mathematical concepts in adolescence and young adulthood. *The Adolescent Brain: Learning, Reasoning, and Decision Making*, 67–93.
- Kaminski, J. A., Sloutsky, V. M., & Heckler, A. (2009a). Transfer of mathematical knowledge: The portability of generic instantiations. *Child Development Perspectives*, 3(3), 151–155. <http://doi.org/10.1111/j.1750-8606.2009.00096.x>
- Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2005). Relevant concreteness and its effects on learning and transfer. In *Proceedings of the ...* (pp. 1090–1095). <http://doi.org/XXVII> Annual Conference of the Cognitive Science Society
- Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The Advantage of Abstract Examples in Learning Math. *Science*, 320(5875), 454–455. <http://doi.org/10.1126/science.1154659>
- Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2009). Concrete Instantiations of Mathematics: A Double-Edged Sword. *Journal for Research in Mathematics Education*, 40, 90–93.
- Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2009b). Supporting Online Materials for “The Advantage of Abstract Examples in Learning Math.” Retrieved from <http://science.sciencemag.org/content/suppl/2008/04/24/320.5875.454.DC1>
- Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2013). The cost of concreteness: the effect of nonessential information on analogical transfer. *Journal of Experimental Psychology. Applied*, 19(1), 14–29. <http://doi.org/10.1037/a0031931>
- Knight, R. D. (1995). The vector knowledge of beginning physics students. *The Physics Teacher*. <http://doi.org/10.1119/1.2344143>
- Kohn, A. (1993). Preschoolers’ reasoning about density: Will it float? *Child Development*, 64(6), 1637–1650. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8624.1993.tb04204.x/abstract>

- Kripke, S. (1982). *Wittgenstein on rules and private language: An elementary exposition*. Retrieved from https://books.google.co.za/books?hl=en&lr=&id=N5I3hNIS_yIC&oi=fnd&pg=PR7&dq=kripke+rules+and+rule+following&ots=euUH258L5e&sig=hBfUwdEFw7UhuQUern-R1jYwo9c
- Kryjevskaja, M., Stetzer, M., & Grosz, N. (2014). Answer first: Applying the heuristic-analytic theory of reasoning to examine student intuitive thinking in the context of physics. *Physical Review Special Topics-Physics*. Retrieved from <https://journals.aps.org/prper/abstract/10.1103/PhysRevSTPER.10.020109>
- Kuo, E. (2013). More than just “plug-and-chug”: Exploring how physics students make sense with equations. Retrieved from <http://drum.lib.umd.edu/handle/1903/14545>
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind. Language* (Vol. 64). <http://doi.org/10.2307/415440>
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by* (Vol. 111). Chicago London.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic books.
- Langacker, R. (2008). *Cognitive grammar: A basic introduction*. Retrieved from <https://books.google.co.za/books?hl=en&lr=&id=UKVVKz0ZRqWC&oi=fnd&pg=PR9&dq=langacker+cognitive+grammar&ots=7Clcsxrs46&sig=2tnykcEYtN-yEFp-5lKtD3qGbNc>
- Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge University Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. (R. Pea & J. S. Brown, Eds.) *Learning in doing* (Vol. 95). Cambridge University Press. <http://doi.org/10.2307/2804509>
- Levrini, O., & diSessa, A. A. (2008). How students learn from multiple contexts and definitions: Proper time as a coordination class. *Physical Review Special Topics - Physics Education Research*, 4(1).
- Lobato, J. (2003). How Design Experiments Can Inform a Rethinking of Transfer and Vice Versa. *Educational Researcher*, 32(1), 17–20. <http://doi.org/10.3102/0013189X032001017>
- Lobato, J., & Siebert, D. (2002). Quantitative reasoning in a reconceived view of transfer. *Journal of Mathematical Behavior*.
- Margolis, E., & Laurence, S. (2007). The ontology of concepts—abstract objects or mental representations? 1. *Noûs*, 41(4), 561–593. <http://doi.org/10.1111/j.1468-0068.2007.00663.x>
- Marton, F. (1981). Phenomenography—describing conceptions of the world around us. *Instructional Science*. Retrieved from <http://www.springerlink.com/index/NX088903180K41UW.pdf>
- Marton, F. (2006). Sameness and Difference in Transfer. *The Journal of the Learning Sciences*, 15(4), 499–535. http://doi.org/10.1207/s15327809jls1504_3
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—Outcome and process. *British Journal of Educational Psychology*. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.2044-8279.1976.tb02980.x/abstract>
- McCallum, W. (2008). Commentary on Kaminski et al., “The advantages of abstract examples in learning math.” Retrieved from http://math.arizona.edu/~wmc/Research/Kaminsky_et_al_commentary.pdf

- McDermott, R. (2001). The acquisition of a child by a learning disability. *Understanding Learning: Influences and Outcomes*. Retrieved from <https://books.google.co.za/books?hl=en&lr=&id=kyN4zGMHsToC&oi=fnd&pg=PA60&dq=mcdermott+the+acquisition+of+a+child+learning+disability&ots=cmQTS9Qin-&sig=ajbLEaBb6Gmo2O9SwYXyXMthpfl>
- Merriam-Webster. (2016). Merriam-Webster Online Dictionary. Retrieved November, 6.
- Mourrat, J. C. (2008). Testing the limits of “concrete” and “generic” [Letter]. *Science*, 322, 1632. Retrieved from http://www.sci.sdsu.edu/CRMSE/personal_pages/sreed/ScienceLetter.pdf
- Nguyen, N.-L., & Meltzer, D. E. (2003). Initial understanding of vector concepts among students in introductory physics courses. *American Journal of Physics*. <http://doi.org/10.1119/1.1571831>
- Nunes, T., Schliemann, A. D., & Carraher, D. W. (1993). *Street mathematics and school mathematics*. Cambridge University Press.
- Ohlsson, S., & Regan, S. (2001). A Function for Abstract Ideas in Conceptual Discovery and Learning. *Cognitive Science Quarterly*, 1(3), 243–277.
- Paivio, A. (1990). *Mental representations: A dual coding approach*. Retrieved from https://books.google.co.za/books?hl=en&lr=&id=hLGmKkh_4K8C&oi=fnd&pg=PA3&dq=pavio+dual+coding+theory&ots=B4zVdFkqpn&sig=tODHjNFXJMqvANY-FQoXloeToll
- Palmer, D. H. (2001). Investigating the Relationship Between Students’ Multiple Conceptions of Action and Reaction in Cases of Static Equilibrium. *Research in Science & Technological Education*, 19(2), 193–204. <http://doi.org/10.1080/02635140120087722>
- Petersen, L. a., & Mcneil, N. M. (2013). Effects of Perceptually Rich Manipulatives on Preschoolers’ Counting Performance: Established Knowledge Counts. *Child Development*, 84(3), 1020–1033. <http://doi.org/10.1111/cdev.12028>
- Podolefsky, N. S., Finkelstein, N. D., Henderson, C., Sabella, M., & Hsu, L. (2008). How Abstract is Abstract? Layering meaning in physics. In *Aip conference proceedings* (Vol. 1064, p. 167).
- Pouw, W. T. J. L., van Gog, T., & Paas, F. (2014). An Embedded and Embodied Cognition Review of Instructional Manipulatives. *Educational Psychology Review*.
- Przenioslo, M. (2004). Images of the limit of function formed in the course of mathematical studies at the university. *Educational Studies in Mathematics*, 55(1), 103–132. <http://doi.org/10.1023/B:EDUC.0000017667.70982.05>
- Quinn, J. R. (2009). *Teacher sense-making and policy implementation: a qualitative case study of a school system’s reading initiative in science*. Maryland.
- Redish, E. F. (2004). A theoretical framework for physics education research: Modeling student thinking. *arXiv Preprint physics/0411149*.
- Reese, R. L. (2000). *University physics*. Brooks/Cole Pub. Co. Retrieved from <http://books.google.co.za/books?id=MNRUAAAAYAAJ>
- Roche, J. (1997). Introducing Vectors. *Physics Education*, 32(5), 339–345.
- Rogoff, B. (1992). Three ways to relate person and culture: Thoughts sparked by Valsiner’s review of Apprenticeship in thinking. *Human Development*. Retrieved from <http://www.karger.com/Article/PDF/277225>
- Rosch, E. (1999). Principles of categorization. *Concepts: Core Readings*, 189–206.

- Rovira, R. (2000). Leibniz's twofold distinction between concrete and abstract and Aristotle's fourfold division of entities. *Studia Leibnitiana*, 68–85.
- Ryle, G. (1945). Knowing how and knowing that: The presidential address. *Proceedings of the Aristotelian Society*. Retrieved from <http://www.jstor.org/stable/4544405>
- Sanchez, C. a, & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition*, 34(2), 344–55. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16752598>
- Sarama, J., & Clements, D. H. (2009). "Concrete" computer manipulatives in mathematics education. *Child Development Perspectives*, 3(3), 145–150.
- Schacter, D. L., Dobbins, I. G., & Schnyer, D. M. (2004). Specificity of priming: a cognitive neuroscience perspective. *Nature Reviews. Neuroscience*, 5(11), 853–62. Retrieved from <http://dx.doi.org/10.1038/nrn1534>
- Scherr, R. E. (2007). Modeling student thinking: An example from special relativity. *American Journal of Physics*. <http://doi.org/10.1119/1.2410013>
- Schmandt-Besserat, D. (1992). *Before Writing: From Counting to Cuneiform, Volume 1*. Austin: University of Texas.
- Schroeder, S. (2006). *Wittgenstein: the way out of the fly-bottle*. Retrieved from https://books.google.co.za/books?hl=en&lr=&id=md-KV6HUueUC&oi=fnd&pg=PR5&dq=schroeder+wittgenstein&ots=LY9V7QBMI5&sig=soFt_1PQP_EltMuCnWJrXLBu9FOk
- Schwartz, D., & Bransford, J. D. (2005). Efficiency and Innovation in Transfer. *Transfer of Learning from a Modern Multidisciplinary Perspective*, 1–51.
- Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*. Retrieved from http://journals.cambridge.org/article_S0140525X00005756
- Serway, R., & Jewett, J. (2013). Physics for scientists and engineers with modern physics.
- Sfard, A. (2001). On the gains and dilemmas of calling different things the same name. *Quarterly of Cognitive Science*, 1(3/4), 359–388. Retrieved from csq.e-reviews.com
- Sfard, A. (2008). *Thinking as Communicating*. New York: Cambridge Univ Press.
- Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing*. Retrieved from <http://books.google.co.za/books?hl=en&lr=&id=cqwx8aDfUuQC&oi=fnd&pg=PA3&dq=sfard+thinking+as+communicating&ots=cxr8rTAVni&sig=GtVSvZWR7DMq89Y7TRywLOm49IE>
- Shaffer, P. S., & McDermott, L. C. (2005). A research-based approach to improving student understanding of the vector nature of kinematical concepts. *American Journal of Physics*, 73(10), 921–931.
- Sherin, B. (2001). How students understand physics equations. *Cognition and Instruction*. Retrieved from http://www.tandfonline.com/doi/abs/10.1207/S1532690XCI1904_3
- Sloutsky, V. M., Kaminski, J. A., & Heckler, A. F. (2005). The advantage of simple symbols for learning and transfer. *Psychonomic Bulletin & Review*, 12(3), 508–513. <http://doi.org/10.3758/BF03193796>
- Southey, P. (2012). *Students' Understanding of Density: A Cognitive Linguistics Perspective*.

University of Cape Town.

- Southey, P., & Allie, S. (2014). Vector Addition in Different Contexts. *2014 Physics Education Research Conference Proceedings*, 243–246.
- Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy Implementation and Cognition: Reframing and Refocusing Implementation Research. *Review of Educational Research*, 72(3), 387–431.
- Stewart, J., Griffin, H., & Stewart, G. (2007). Context sensitivity in the force concept inventory. *Physical Review Special Topics Physics Education Research*, 3(1), 1–6.
<http://doi.org/10.1103/PhysRevSTPER.3.010102>
- Strauss, A., & Corbin, J. (2007). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications, Incorporated.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive Load Theory. *Learning*, 1(1), 57–69.
<http://doi.org/10.1007/978-1-4419-8126-4>
- Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics*, 12(2), 151–169.
- Turner, M. (2000). Backstage cognition in reason and choice. Retrieved from
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1249963
- Turpen, C., & Finkelstein, N. (2010). The construction of different classroom norms during Peer Instruction: Students perceive differences. *Special Topics-Physics Education Research*. Retrieved from <https://journals.aps.org/prper/abstract/10.1103/PhysRevSTPER.6.020123>
- Ungerer, F., & Schmid, H. (2013). *An introduction to cognitive linguistics*. Retrieved from
https://books.google.co.za/books?hl=en&lr=&id=AVIVAgAAQBAJ&oi=fnd&pg=PP1&dq=an+introduction+to+cognitive+linguistics+ungerer&ots=TVfMSbyX1r&sig=RUumlCHQ9Y3ey-Wsj_2w_Jebh_M
- Van Deventer, J. (2008). *Comparing student performance on isomorphic math and physics vector representations*. University of Maine.
- van Oers, B. (2001). Contextualisation for abstraction. *Cognitive Science Quarterly*, 1(3), 279–305.
- Van Oers, B. (1998). The fallacy of decontextualization. *Mind, Culture & Activity*, 5(2), 135–142.
<http://doi.org/10.1207/s15327884mca0502>
- Vinner, S. (2002). The role of definitions in the teaching and learning of mathematics. In *Advanced mathematical thinking* (pp. 65–81). Springer.
- Vygotsky, L. (1994). The Problem of the Environment. In R. Van der Veer & J. Valsiner (Eds.), *The Vygotsky Reader* (pp. 338–354). Blackwell.
- White, A. (1982). The nature of knowledge. Retrieved from <https://philpapers.org/rec/WHITNO-5>
- Whitney, P. (2002). Schemas, frames, and scripts in cognitive psychology. Retrieved from
<https://www.mysciencework.com/publication/show/8d2dfb0ee073b4f9d6fbe91977712892>
- Wilensky, U. (1991). *Abstract meditations on the concrete and concrete implications for mathematics education*. Epistemology and Learning Group, MIT Media Laboratory.
- Wittgenstein, L. (1958). Philosophical Investigations I. In G. E. M. Anscomb (Ed.), *Philosophical investigations* (pp. 2–13).
- Wittmann, M. (2002). The object coordination class applied to wave pulses: Analysing student

reasoning in wave physics. *International Journal of Science Education*. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/09500690110066944>

Wright, C. (1998). Self-knowledge: the Wittgensteinian legacy. *Royal Institute of Philosophy Supplement*. Retrieved from http://journals.cambridge.org/article_S135824610000432X