COLLABORATIVE KNOWLEDGE CONSTRUCTION

IN PROBLEM-BASED LEARNING:

A CORPUS-BASED STUDY

A thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy in the School of Medicine

OLUKAYODE MATTHEW TOKODE MBChB, MMedSci, MSc,

School of Medicine, Medical Education Unit

July 2016

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

PBL	Problem-based learning
GMC	General Medical Council
KWIC	Key word in context
POS	Parts-of Speech
Semtag	Semantic Tagging
CL	Corpus Linguistics
SCT	script concordance test
ZPD	zone of proximal development
IRE	initiation-response-evaluation
SFL	systemic functional linguistics
DA	discourse analysis
CDA	critical discourse analysis
GT	grounded theory
GEMS	graduate entry medical school
GAMSAT	graduate Australian medical school admissions test
SGT	Small group teaching
F	Female
M	Male
T	Tutor
SDL	self-directed learning
CC	coordinating conjunctions
CS	subordinating conjunctions

DEDICATION

This thesis is dedicated to the memory of my late parents – Pa Asiyanbi Lawrence Tokode and Madam Adedoja Abigail Tokode.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Reg Dennick, for his incredible support during this study. He has been a mentor and a facilitator. He has provided me with rich experiences to nurture my development as an educationist. His indirect statements and pregnant silences always push my thinking and send me back to the drawing board about my project. Similarly, I would like to thank Dr. Paul Rayson of school of computing and communications, Lancaster University, UK, for his valuable support and advice through the course of this project.

I would also like to thank my friends and family who have listened to me and have been a constant source of support throughout this journey. A big thank you to my children for bearing with me during this tedious journey. My wife's love, support, and patience that has made this thesis possible.

ABSTRACT

Background

Effective disease diagnosis and treatment relies on a conceptual knowledge base that is both expansive and well-networked. The problem-based learning (PBL) curriculum is considered as being well-suited to creating this kind of knowledge. The facilitator plays a crucial role in establishing and maintaining the knowledge construction discourse as students interact to resolve case problems. An exploration of tutorial talk could provide opportunities to understand and improve verbal interactions of this nature. Many of the previous studies have only analysed a small amount of tutorial talks owing to methodological constraints, and the existing literature on the subject matter only scarcely touches upon the utility of lexicogrammatical methods for the development of an understanding of knowledge construction in medical PBL tutorials. In this research, a blend of corpus linguistics methodology and a lexicogrammatical approach was employed for the analysis of talk in 8 PBL tutorial groups in order to deepen our understanding of how students jointly construct knowledge and how the facilitator guides the process.

Aims

In this study, a corpus of 2,37,820 comprising eight PBL students' and facilitators' tutorial talk was created to achieve the following aims:

I. To use the students' subcorpus to answer the research question (1) by measuring the frequencies and describing the functions of the frequently occurring (1) referring expression indicators; (2) shared knowledge indicators; (3) knowledge extension indicators; and (4) knowledge enhancement indicators.

II. To use the facilitators' subcorpus to answer the research question (2) by measuring the frequencies and describing the functions of the commonly occurring (1) facilitators' questions; (2) facilitators' directive expression indicators; and (3) facilitators' probability indicators.

III: To make recommendations based on the results of the study.

Methodology

Wmatrix 3 was used to retrieve defined linguistic indicators relating to the research questions. A quantitative analysis of the indicators was performed through word frequency computation and a keyword-in-context analysis. Descriptive statistics with SPSS version 22 was used to computer frequency profile of the indicator functions, and the Log likelihood calculator was used to determine the variation of the functions across the eight PBL groups. Extracts from the dataset were provided to illustrate the indicators' functions.

I. Results of Students' talk analysis

The subcorpus contained 2,10,077 words. The most frequent contents of the students' talk comprised biomedical science and cause-effect vocabularies.

1. Analysis of referring indicators

There were 2,325 referring expression indicators. They were used to mark verbal expressions, amounting to 44.04%; mental expressions, amounting to 42.24%; and learning situation and materials, amounting to 13.72%. The referring expressions were used for providing peer commendation, sharing knowledge, fostering social and cognitive regulation, and for constructing knowledge; the mental referring expressions were used to generate hypotheses, achieve mutual understanding, and define group tasks; and learning referring expressions were used to share learning resources, explain concepts, as well as guide discussions and resolve conflicts.

2. Analysis of shared knowledge indicators

There were 3,437 shared knowledge expression indicators, which are the following: affirmation (73%), negation (17%), and non-lexical content (10%). Affirmative indicators were mostly used for integration-oriented knowledge sharing (42.31%); negation affirmation expressions were mostly used for conflict-oriented knowledge sharing (70%); and non-content indicators were mainly used for idea and information orientation. Shared knowledge was commonly achieved among group members through information addition, repetition and rephrasing, paraphrasing, causal and noncausal elaboration, correction of ideas and information recollection, and by establishing orientation to ideas and information from the group members.

3. Analysis of knowledge extension indicators

There were 6,520 retrieved knowledge extension indicators, which comprised the following: additive 4,227 (63.54%), alternative 1,001 (15.05%), and adversative 1,424 (21.41%). Adversative indicators were more frequently used for knowledge construction compared to additive (33% versus 16%; LL 32.58, $p < 0.01$) and alternative indicators (33% versus 13%; LL 95.74, $p < 0.01$).
The students commonly used additive indicators for simple, temporal, causal-conditional, elaborate, contrastive, and indefinite additions. Alternative indicators were commonly used for offering alternative questions and ideas while adversative indicators were frequently used to link elaborative, contrastive, concessional, and causal-conditional clauses.

4. Analysis of knowledge enhancement indicators

A total of 6,402 indicators were retrieved. The most frequent among the retrieved 6,402 indicators were because, so, as, when, and that. Between 16.94% and 29.24% of the indicators were used for knowledge co-construction. The most frequent indicators' functions were conditional, extension, report, consequence, inference, and feature specification. The reporting functions regularly concerned biomedical theory, previous peer knowledge, research evidence, professional opinion, as well as cognitive tools and criticism; extension function related to biomedical knowledge; and feature specification functions involved biomedical attributes and explanation; the conditional functions were frequently used to state logical conditions for disease presence, manifestation, and treatments; the inferential functions were more consistently used to link biomedical deductions to their premise; and the consequential functions commonly related to the linking of physiological mechanism and organ function to their respective consequences.

II: Results of facilitators' talk analysis

The subcorpus contained 27,743 words. The most frequent content comprised biomedical science and cause-effect vocabularies.

1. Facilitators' questions

There were 35 types of question indicators. The facilitators asked 0.78 lower-order questions per 100 tokens, and 0.25 higher-order questions problem-based per 100 tokens. The questions functioned to stimulate elaboration, elicit information, prompt students, and the offering of suggestions.

2. Analysis of directive expressions

'Should', 'have to', 'need', 'supposed', 'would', and 'can' directive expression indicators were found to be most frequent. They were used to mark expectation, indirect question, and they were commonly used to preface requirement, exhortation, and intention. The indicators functioned frequently to facilitate group process (53.45%) and direct learning (42.00%).

3. Analysis of probability expression indicators

There were of 27 types, out of which 9 were frequently occurring. The probability expression indicators were frequently used to mark possibility, prediction, hedging, and logical deductions. The indicators functioned to preface content information given to the students, to mark process facilitation remarks, and to mark facilitators' questions.

Discussion and conclusion

The study demonstrated the feasibility of using corpus linguistics to study medical students' knowledge construction talk; provided evidence of knowledge construction through prior knowledge mobilisation, knowledge extension, and enhancement; and signified the attainment of shared knowledge. The facilitators frequently asked lower-order questions; the directive expressions indicators were used to mark content-related and learning behaviour expectations and requirements; and the probability expression indicators were frequently used to mark content information given to the students.

This study shows that students construct knowledge in their PBL tutorials. The pedagogic issues that emerged from the study relates to subversion of the PBL facilitation principles. A wholistic understanding of the factors that affect the behaviours of the facilitators in the classroom is important to resolve this problem. This may involve tutor pedagogic education and recalibration of administrative policies and institutional culture to provide an enabling environment for PBL instructional approach.

CHAPTER 1

INTRODUCTION

1.1. Introduction

This chapter describes the background to the thesis. It starts with an exploration of the ancient origin of problem-based learning (PBL), followed by an understanding of the educational context of its modern emergence, its heterogeneity and similarity to other experiential instructional approaches, along with an analysis of its goals and process. The epistemology of knowledge and its theoretical implication in context of the constructivism underpinning PBL are outlined. The chapter concludes with a description of the research problem and questions as well as an overview of the purpose, significance, and scope of the research.

1.2. Background

Classically, PBL is defined as: “the learning that results from the process of working towards the understanding of a resolution of a problem. The problem is encountered first in the learning process” (Barrows and Tamblyn, 1980: p. 1). It is a complete approach to education and not just a teaching technique or tool (Barrett and Moore, 2011a). Savery and Duffy (1996) describe it as the mastery of information in context of the case in which it is being used. PBL uses complex real world problems to introduce concepts and provide a motivating active and cooperative learning environment (Allen et al., 1996).

PBL is considered to be an ancient educational approach used by the likes of Socrates, Aristotle, and Confucius (Kolmos and Graaff, 2007; Lee et al., 2004; Savin-Baden, 2000; Schmidt, 2012); however, its orthodox fashion was introduced at the McMaster University by Barrows and his colleagues more than four decades ago (Azer, 2001; Donner and Bickley, 1993; Neville, 2009; Neville and Norman, 2007). Following its initial adoption at the Maastricht University in the Netherlands, Newcastle in Australia, and New Mexico in the United States (Dolmans et al., 2002; Hung et al., 2008), the popularity of PBL curriculum grew rapidly such that it has been adopted in whole or in part in several disciplines outside medical education in most nations of the world (Azer, 2001; Dolmans et al., 2002; Hung et al., 2008; Kolmos and Graaff, 2007; Neville, 2009).

1.3. Context of Problem-based Learning Curriculum

The popularity and the globalisation of PBL curriculum was based on the need to replace the traditional curriculum that was seen as being inadequate in the knowledge-driven professional world of modern times. The traditional curriculum has been criticised for a number of lapses, including students' information overload, accumulation of unusable knowledge, apathy towards continuous learning, dislocation of basic sciences from clinical practice, excessive focus on scientific research rather than on competences needed in practice, and curriculum overload (Azer, 2001; Neville, 1999; Schmidt, 1983). Aside from the inadequacies traditional curriculum suffered from, a number of other factors influenced the popularity and adoption of PBL curriculum − a need for professionalism; a commitment to lifelong learning in the face of ever-increasing medical information and the dynamic complexities pertaining to disease and treatment; a need for effective inter-professional collaboration and communication; a requirement for critical thinking, problem-solving, and other labour market skills; a means of responding to directed innovation; and a remedial response to the decline observed in the resources for higher education (Azer, 2001; Savin-Baden, 2000). These needs require radical changes in medical education modelled towards knowledge acquisition, application and use, as well as data interpretation and acquisition of lifelong learning skills (Azer, 2001). Thus, PBL seems to present a better prospect for improved medical education when compared to the traditional curriculum. The educational objectives of PBL could help achieve these features (Barrows, 1983). However, there is a challenge PBL faces in terms of its conception and practice.

1.4. PBL and other Experiential Approaches

Significantly, different instructional practices are often categorised under PBL and the resulting confusion can be very frustrating (Margetson, 2003). The situation has significant consequences for research and evaluation, since it would be very difficult to compare the outcomes of PBL in one setting against that of the other (Lloyd-Jones et al., 1998). While some educators considered the orthodox form practiced at the McMaster and Maastricht as the authentic PBL (Harden and Davis, 1998), some others considered other methods, such as case-based learning, project-based learning, discovery learning, and inquiry-based learning as falling under PBL (Azer, 2001; Savery, 2006). All cases may not be problems, and solving these problems may not be the principal purpose of the exercise (Azer, 2001). The term discovery learning connotes the idea that students are supposed to discover knowledge that is already known (Boud and Feletti, 1999).

According to Boud and Feletti (1999), PBL has been described as an approach to curriculum structuring where students are confronted with authentic problems in order to stimulate learning. The curriculum and process for teaching and learning could take several forms and still be compatible with PBL elaborated (Boud and Feletti, 1999). PBL is seen as a way to learn rather than as a learning method or technique. It is considered that its goal is to help students build capacity for modern life and to enable them to contribute productively to their society (Engel, 1999; Neufeld and Barrows, 1974). Barrows describes PBL as an instructional strategy (Boud and Feletti, 1999) that could have many different meanings, depending on the "design of the educational method employed and the skills of the teacher" (Barrows, 1986: p. 481). For Barrows (1986), there is no single version of PBL, and he devised a taxonomy to classify the various forms of PBL. Similarly, Harden and Davis (1998) locate PBL towards the end of a continuum in which traditional lecture lies at one extreme and task-based learning lies at the other one.

1.5. PBL Heterogeneity

PBL's orthodox model has undergone tremendous mutation over time (Gijselaers 1995 cited in Dochy et al., 2014; Gijbels et al., 2013). As PBL curriculum migrated from medical school at the McMaster University in Canada and got adopted and planted in different nations and disciplines with different academic cultures and resources, the orthodox fashion mutated rapidly to appropriate to its new environments. Thus, PBL instructional strategy became a genus with many species (Barrows, 1986; Harden and Davis, 1998), or a syndrome with many forms (Azer, 2001; Walton and Matthews, 1989), or a coat with many colours (Lloyd-Jones et al., 1998), if you please, such that no two PBL curricula are the same in terms of meaning, content, and implementation.

Several models of PBL curriculum have been developed. These models include the orthodox PBL, where education is delivered wholly as PBL as in the McMaster and Maastricht PBL programme models (Kwan, 202; Schmidt et al., 1993); hybrid PBL, where conventional lectures are combined with PBL (Armstrong, 1999; Houlden et al., 2001; Steele et al., 2000), and isolated PBL, where PBL is applied to one subject or an aspect of a subject within traditional curriculum (Lobb et al., 2004). Variation also exists in the number of days spent to complete a PBL cycle. In the McMaster and Maastricht models, students meet twice a week (Barrows, 1988; Schmidt et al., 1993); in Harvard medical school, the students meet three times in a week (Engel, 1999; Silver and Wilkerson, 1991), whereas PBL is completed in a day at Singapore Polytechnic (O'Grady et al., 2012; Yew and Schmidt, 2009, 2011). Other models of PBL vary in term of class size. The traditional and popular variant of PBL in terms of class size comprises a small group. Exley and Dennick (2004) described four categories of small group−teaching with PBL groups having 8–12 students per group (Table 1.1). There is a general agreement that the optimum size for small group−teaching (SGT) is 5−8 students per group, and for tutor-led and tutorial type−SGT, a group of 6 students is the best (Booth 1996 cited in Exley and Dennick, 2004). It is considered that a group membership of below 5 is not optimally suited for the proper utilisation of the diversity and the richness of interpersonal interaction while the input from some group members start declining when the group size is more than 8 per group (Exley and Dennick, 2004). However, due to limited faculty resources, medium to large classes or a whole class−PBL model has also been introduced (Kingsbury and Lymn, 2008; Pastirik, 2006; Roberts et al., 2005). While classical PBL is conducted in face-to-face learning environments, online PBL models have been introduced where the facilitator and the students engage in tutorial discourse via the synchronous online environment (Dennis, 2003; Sendag and Odabasi, 2009; Sterling and Centre, 2004). The classical model of facilitation requires the facilitator to stay with the students throughout the tutorial session; but in floating facilitation, the tutor, as the floating facilitator, moves from one small group of students engaged in a PBL tutorial discussions to the others, asking questions and probing the understanding of the students (Duch, 2001).

Table 1.1: The four categories of small group teaching (Exley and Dennick, 2004)

Category of small group teaching	Example of small group Approaches	Typical student numbers
Tutor-led SGT	Tutorials	4–12
	Seminars	10–25
	PBL groups	8– 2
Student-led SGT	Tutor-less tutorials	4–8
	Learning sets	4–8
	Self-help groups	4–8
Virtual SGT	Virtual tutorials	4–12
	Email discussions	4 upwards
SGT in large groups	Syndicate work	10–100
	Problem classes	10–50
	Group practicals	10–100
	Workshops	10–40

SGT, small group teaching

Despite these differences, there also seems to be a shared understanding pertaining to the educational goals and the underpinning principles and philosophy of PBL. These are highlighted in the following sections.

1.6. Goals of Problem-based Learning Curriculum

The general goals of PBL in terms of outcomes for the students have been well-summarised in the existing literature (Azer, 2001;

Barrows, 1996; Hmelo-Silver, 2004; Neufeld and Barrows, 1974). PBL is designed to help students achieve the following:

- construct an extensive, integrated, and flexible knowledge base;
- develop effective problem-solving and clinical reasoning skills, typical of an expert physician;
- develop self-directed, lifelong learning skills;
- become effective collaborators;
- become intrinsically motivated to learn;
- develop humanistic competences in the students, thus enhancing their sensitivities to the medical and psychosocial needs of patients;
- promote independent critical thinking skills; and
- enhance awareness of and enable work in a variety of health care settings.

1.7. Principles of PBL Practice

The accomplishment of these objectives would require a set of instructional principles that can guide the practice of teaching and the design of learning environments. Learning environment is defined as the classroom- and school-based psychosocial and pedagogic situations that influence students' academic achievements and attitudes (Fraser, 1998). These principles include (Savery and Duffy, 1996) the following:

- All learning activities should be anchored to a problem.

- Learners should be supported to develop ownership of a task or a problem.
- The cognitive demands required to solve the problems presented to the students should be consistent with the cognitive demands in the environment for which the learners are being prepared.
- The task and the learning environment should be designed to reflect the complexity of the environment they are expected to function once they are through with their learning.
- Learners need to be given ownership of the process used to develop a solution to a problem.
- Learning environment should be designed to support and challenge the learner's thinking.
- Students need to be encouraged to test ideas against alternative views and contexts.
- Learning environment needs to provide opportunity for and support reflection on both the content learned as well as the learning process.
- Students need to work together collaboratively to solve the problem presented.

1.8. PBL Tutorial Process

There is a consensus regarding the stages that should be included in the PBL tutorial process (Azer, 2001). These are elaborated below.

Generally, the tutorial process starts with problem presentation in form of a trigger text or a scenario or video. The problem involves a set of phenomena in need of explanation (Dolmans et al., 2001). The students analyse the problem, employ reasoning skills to work on it by generating questions and trying to determine the underlying causal mechanisms, principles, or processes that might help explain the anatomical, physiological, biochemical, or behavioural dysfunction responsible for the patient's problem (Barrows, 1985, 1988; De Grave et al., 1996; Dolmans et al., 2001). To do this, says Barrows (1985), the students need to learn the concepts and mechanisms that are involved in normal basic sciences (anatomical, physiological, biochemical, psychosocial). The students discuss their ideas about the problem; generate hypothesis about basic mechanisms involved in the patient's problem; engage in an inquiry strategy concerning which hypothesis best fits the information collected from the patients' history and clinical examination; and problem synthesis. The students employ their prior knowledge and experience for the clinical reasoning exercise (De Grave et al., 1996; Dolmans et al., 2001). The students constantly question themselves, focusing on what they need to know in order to better understand the basic mechanisms underlying the problem of the patient, and some of these questions are transformed into learning issues. According to Barrows (1985), the hypotheses that the students are asked to develop should be those of the underlying basic mechanisms (not diseases) responsible for the patient's symptoms and signs. The questions on history and examinations are investigations aimed at working out and verifying the correct mechanisms responsible. The learning issues simultaneously developed by the students are those that relate to the preclinical disciplines.

This is followed by a period of self-directed study during which the students distribute the issues at hand among themselves and research upon the learning issues they identified at the initial stage, using a variety of learning resources, including textbooks, monographs, journals, and other library resources. After the initial research, the students can then approach their basic science faculty members for further learning in the form of seminars (Barrows, 1985). The students then return to the group with the gathered information along with other materials that may help augment their learning in the group. According to Barrow (1985), the students also critique their learning sources and the information that the students bring to the tutorial. With the new understanding and knowledge derived from the newly acquired information, the students go over the problem again. They apply their reasoning skills again and refine their hypothesis and problem-solving process in the light of the new knowledge that have just obtained. Through this process, the new information is restructured in a clinically meaningful manner. The students conclude the tutorial process with a reflection on their learning (Barrows, 1985). This involves a review of what has been learnt (in basic sciences, comprising anatomy, physiology, biochemistry, and psychology or behavioural science) along with the learning objectives to ensure that adequate coverage has been achieved. The role of the facilitator, who is considered an expert learner, is to model good strategies for learning and thinking and not dispense knowledge as is expected of them in the traditional curriculum (Hmelo-Silver, 2004). In other words, the facilitator structures learning environments. Table 1.1 shows the seven essential stages of the PBL process.

Table 1.2: Seven stages of the PBL process (Schmidt, 1983)

Step 1:	**Clarify terms and concepts not readily comprehensible.**
Step 2:	Define the problem.
Step 3:	Analyse the problem.
Step 4:	Draw a systematic inventory of the explanations inferred from step 3.
Step 5:	Formulate learning objectives.
Step 6:	Collect additional information outside the group.
Step 7:	Synthesise and test the newly acquired information.

1.9. Epistemology of knowledge

Classroom learning environments shape the educational process as well as human thought and action (Fraser, 1998; Hewitt and Scardamalia, 1998). Practical initiatives might be ineffective without a careful consideration of educational philosophy and theory of knowledge that determine the structuring of the learning environments (Fan, 2014).

It was only in the last century that psychology broke away from contemplative philosophy to establish itself as an empirical science, and it is not surprising that there is a convergence between how the philosophers of science view the nature of knowledge and its acquisition in science and how the educational psychologists view learning and instruction (Schmidt, 1993). Thus, PBL has its root in both philosophy and psychology (Fan, 2014; Schmidt, 1993). According to Schmidt (1993), the psychology-oriented views on how we know our world are based on two philosophical conceptions – empiricism and rationalism.

Empiricism has its root in the idea forwarded by Aristotle, and it postulates that knowledge is acquired entirely from the senses (Dennick, 2015). Rationalism, on the other hand, has its root in the works of Plato. Plato conceived that "true knowledge was located in, and could be generated by, the rational, thinking mind" (Dennick, 2015: p. 39). Dennick (2015) elaborated that Plato conceived that our minds are born with ideal forms of knowledge, and the aim of human development is to understand these ideal and universal forms. For Plato, reasoning is superior to feelings, and according to him, the world of senses was considered erroneous, incomplete, and uncertain; and knowledge originating from them was seen as being unreliable. The mind contains innate knowledge, and true knowledge can be created by reasoning alone.

The influence of these two philosophical ideas on the thinking of later philosophers, and the ensuing philosophical controversy were well elaborated by Dennick (2015): For example, advocates of empiricism, such as Bacon, Locke, and Hume, denied the existence of innate knowledge and argued that knowledge is derived entirely from experience. Rationalists (e.g., Descartes and Spinoza) insisted that the human mind was an intrinsic source of reason, capable of creating knowledge by thought alone, without the requirement of sensory inputs. It was considered that innate knowledge was the law of God, implanted in the human mind. Rational knowledge was considered superior to knowledge derived from the senses because it was knowledge originating from God, and hence, pure and truthful.

The two philosophical positions were synthesised by Immanuel Kant in the 18th century (Dennick, 2015; Dennick, 2016). According to Dennick (2015), Kant contended that knowledge is acquired from the sensory experience, which is then filtered through and structured by the rational processes of the mind. Kant asserted that there is an interaction between reason and sensory experience, leading to the construction of knowledge of the world. Kant maintained that knowledge of the world is mediated by innate rational processes and intuitions along with certain mental frameworks or categories that enable us to see the world through space-, time-, and causality-tinted spectacles; and thus, we can never actually know the real world or things-in-themselves. Psychological conceptions with regard to learning and instruction emerged from this philosophical view of Immanuel Kant, and they served as the basis for the work of constructivists philosophers, such as Piaget, Vygotsky, Dewey, and the like (Dennick, 2015; Dennick, 2016).

1.10. Implications of epistemology of knowledge

The epistemological philosophies, as outlined above, gave rise to behavioural and constructivism theories of education as outline below.

1.10.1. Behaviourism

The theory of behaviourism forwarded by Thorndike and Watson represents instructional and learning theorising that has its root in empiricism (Schmidt, 1993). Behaviourism considers learning as a system of response to external stimuli, which is, in turn, observable as changes in behaviours (Skinner, 1968 cited in Fan, 2014). Learning occurs when a proper response is made to external stimuli through conditioning and reinforcement, or through punishment, practice, and external motivation (Fan, 2014; Fosnot and Perry, 1996). Learners are seen as empty slates upon which the features of the world are written (Schmidt, 1993; von Glasersfeld, 1995).

The educational implications of the conception of the learner and the learning process is profound. Students are viewed as the tabula rasae, upon which educational processes inscribe the features of the world (von Glasersfeld, 1995). It is assumed that students learn by mere observation, by listening to teachers' explanations, or by engaging in practice sessions with constructive feedback from teachers (Fosnot and Perry, 1996). Examinations or drillings are used to assess whether or not the minds of the students are capable of accurately reflecting the gamut of the curricular materials they have accumulated (von Glasersfeld, 1995). Psychological tactics in the form of appropriate motivation, practice, and reinforcements are used to create the desired student behaviours (Fosnot and Perry, 1996). This view of learning may have implications when it comes to inducing behaviour changes; however, it lacks the capacity to explain the cognitive processes of human learning and understanding (Fosnot and Perry, 1996), and thus, it has largely been abandoned (Duit and Treagust, 1998; Fan, 2014; Fosnot and Perry, 1996).

1.10.2. Constructivism

Constructivism has become the metaphor of human learning since the 1970s (Applefield et al., 2001; Liu and Matthews, 2005; Phillips, 1995). It has its root in the philosophy of Immanuel Kant, as previously highlighted (von Glasersfeld, 1995). Constructivism, as a psychological conceptions of learning, largely stems from the work of Piaget (Glasersfeld, 2000; Piaget, 2011) as well as Vygotsky and his followers (Maclellan and Soden, 2004; Phillips, 1995; Rogoff, 2003; Vygotsky, 1930/1978).

For the constructivists, learning is meaning-making, and it involves individuals constructing their own new understandings based on interactions between their prior experience of the world and the knowledge of the world that they keep encountering (Richardson, 2003; Savery, 2006; von Glasersfeld, 1995). Learners play an active role in the learning processes, and thus, knowledge is constructed by learners themselves with their own interpretations of contradictory situations and through integrating new knowledge with prior experiences, suggesting that knowledge is relative as well as subjective. One is seen as knowledgeable if knowledge is constructed from experiences that are richer than others and full of variety (Fan, 2014).

In recent times, constructivism has been criticised for several reasons (Phillips, 1995). Several variants of constructivism exist, such as cognitive constructivism, radical constructivism, social constructivism, critical constructivism, and contextual constructivism (Geelan, 1997; Phillips, 1995). These variants have been compared to religious sects with strong ideological positions (Phillips, 1995). Nevertheless, they generally share a common view of learning as an active process of meaning-making (Fan, 2014).

The two dominant forms of constructivism are the Piagetian and Vygotskian models (Phillips, 1995). There is much debate between the advocates of the two models regarding their similarities and differences, merits and limitations of the accounts of learning (Cobb, 1994; Cobb and Yackel, 1996; Fosnot, 1993; O'Loughlin, 1992; O'Loughlin, 1993; Packer and Goicoechea, 2000). For instance, opponents of Piagetian constructivism contend that Piagetian constructivism, as an extension of Kantian idealism, ignores the subjectivity that each person brings to the reasoning process (Buck-Morss cited in O'Loughlin, 1992; Sampson, 1981). Additionally, it reduces knowledge to individual cognition rather than to the same being products of social, cultural, and historical constitution (Sampson, 1981). Its concept of adaption promotes societal status quo; its commitment to logico-mathematical problem-solving and abstract reasoning sanctions and universalises technical-rational type of knowledge (Fosnot, 1993; Sampson, 1981); its emphasis on individualism and the fact that primacy is given to assimilation over accommodation, indicates the non-dialectical nature of thinking under Piagetian constructivism (O'Loughlin, 1992); its model is hegemonic and structuralist at heart (O'Loughlin, 1993); and his staging of cognitive development is considered inadequate because it was based on invalid conclusions drawn from flawed clinical interviews with young children (Matthews, 1997). Strong supporters rose up to defend Piagetian constructivism; they are strong in their criticism of Vygotskian constructivism. For instance, sociocultural constructivism has been criticised for promoting cultural relativism (Fosnot, 1993). It is seen as a danger to scientific rational thought (Fosnot, 1993; Liu and Matthews, 2005). The epistemological or ontological nature of the two models is another area of intense debate. For Von Glasefeld (1988 cited in Fosnot, 1993; Matthews, 1997), Piagetian constructivism is an epistemology. Packer and Goicoechea (2000) believe that the two constructivist models are not just epistemological but ontological models. Piagetian constructivism is ontological in as much as it is based on dualism (Packer and Goicoechea, 2000); and it is epistemological because its epistemic subject deals with knowledge creation (Matthews, 1997). The Vygotskian constructivist model is ontological and epistemological because it deals with being and the process of becoming a being; learning does not only result in changes in knowing; it also involves changes in being (Packer and Greco-Brooks, 1999; Packer and Goicoechea, 2000). This view was shared by Vygotsky when he said, “To encompass in research the process of a given thing’s development in all its phases and changes – from birth to death – fundamentally means to discover its nature, its essence, for it is only in movement that a body shows what it is.” (Vygotsky, 1978 cited in Wertsch, 1985: p. 17).

In the midst of the heated debates between Piagetian and Vygotskian constructivists, some thinkers suggested that the two models “are not irreconcilable but complementary” (Packer and Goicoechea, 2000: p. 227). They advocated a synthesis of the two perspectives; the work of Piagetian constructivists is analogous to the internalist program, whereas that of sociocultural constructivists is similar to that of externalists (Fosnot, 1993; Greeno, 1997). Piagetian constructivism focuses on structural perspective of understanding, whereas Vygotskian focuses on the functional aspect of understanding (1996 cited in Packer and Goicoechea, 2000), and thus, each perspective tells only half a portion of a good story (Cobb, 1994). Debate also exists regarding whether constructivism is a teaching or a learning theory. Resnick (1989 cited in Richardson, 2003) sees constructivism as a general theory of learning. Thompson (2000 cited in Richardson, 2003) suggests that constructivism is not a learning theory but a model of knowing that could be employed to build a theory of learning. Colliver (2002) sees constructivism as an epistemology and not as a pedagogy. He advocated the teaching of the constructivist nature of knowledge in schools, but he rejected the dichotomous pedagogy of realism and constructivism, arguing that students always learn the same way no matter how knowledge is viewed. According to Richardson (Richardson, 2003), the current interest in the constructivist teaching method was initiated by Fosnot (1989), and he described this teaching model as a facilitation of learning. The next section presents an overview of the three theoretical foundations of PBL, connected by constructivism.

1.11. Theoretical Foundations of PBL: An overview

The three main theoretical perspectives of learning that underpin the problem-based instructional approach are highlighted in this section. Dennick (2016) observed that the constructivist model serves as a bridge between different epistemological and pedagogic theories.

1.11.1. Information Processing theory

The information processing theory concerns the intrapersonal processes that lead to knowledge development – it is an individual that constructs knowledge (Gijselaers, 1996). Three key principles of information processing theory are considered central to learning. First, the effect of prior knowledge on knowledge acquisition. Prior knowledge activation provides a model for understanding new information as well as for constructing new knowledge or restricting existing knowledge (Hmelo-Silver and Eberbach, 2012a; Schmidt, 1983, 1993). Second, the effect of specificity encoding. This principle states that the retrieval and use of knowledge is easier when it is encoded with the cues during learning similar to the ones in the context of its use. The third principle relates to the effect of knowledge elaboration on learning. It states that knowledge is better understood and processed, and it is easier to retrieve if the learners have the opportunity to elaborate upon their existing pool of knowledge. Elaboration of knowledge could be carried out by answering questions, taking notes, discussing subject matter with peers, and by explaining to peers what has already been learnt. The way in which knowledge is structured in one's memory makes it accessible for further use. Knowledge is structured in semantic networks (Hmelo-Silver and Eberbach, 2012a; Schmidt, 1983). Information theory provides insights into how the mind process information, but not where the mind derives the information from.

1.11.2. Socioconstructivist theory

The Socioconstructivist theory relates to the idea that the mind of an individual is used as a substrate for knowledge construction. Learning occurs as the mind interacts with the world (Glasersfeld, 2000; Huitt and Hummel, 2003). Thus, learners should be actively engaged in their own learning, leading to an active pedagogy, involving exploration, experience, and experimentation (Dennick, 2016). The encounter of a learner with knowledge (or experience) that she/he does not understand, or one that contrasts with what she/he knows, could trigger processes (searching, reading, asking questions, and so forth) that could lead to acquisition of new knowledge or a modification of the existing knowledge to attain newer understandings.

1.11.3. Sociocultural theory

This theory relates to the effect of social and cultural interactions on knowledge construction (Hmelo-Silver and Eberbach, 2012a; Vygotsky, 1930/1978). The theory lays emphasis on the fact that knowledge is constructed in collaboration with other people. This is done through the employment of a scaffolding, which is the help that a learner receives from someone who knows (teacher or peer) how to solve the problem that she/he is not equipped to solve independently (Mercer, 1995; Rojas-Drummond and Mercer, 2003; Wegerif et al., 1999). This theory also proposes that cultural tools mediate learning (Vygotsky, 1930/1978). Halliday (1993) believes that the learning of language is essential to the acquisition of content knowledge. Aside from language as the principal tool that mediates learning, the use of other cultural tools (diagrams, mnemonic, formulae, computer, whiteboard, and so forth) also mediate learning (Hmelo-Silver and Eberbach, 2012a; Wells, 2007).

The integration of these theories provides a foundation for the constructivist model of PBL.

1.12. Statement of the Problem

PBL offers several potential advantages for students' learning, including high motivation, better problem-solving skills, superior self-regulation, basic and clinical knowledge integration, as well as superior knowledge construction and collaboration skills (Hmelo-Silver, 2004; Norman and Schmidt, 1992). These claims find theoretical support from the literature available on the subject of psychology of learning (Norman and Schmidt, 1992). However, the sound theoretical underpinnings of PBL are yet to translate to superior practical benefits. One possible cause of this problem is that outcome studies have significant methodological flaws, including the use of subjective ratings by the tutorial participants and the use of inappropriate students' achievements measures to judge the products that comprise the PBL curriculum (Albanese, 2000; Dolmans et al., 1994; Schmidt et al., 1993; Schmidt and Moust, 1995). Medical education commentators have advocated more research into this subject matter in order to better understand the workings of the PBL tutorials, especially regarding the link between theories and praxis (Albanese, 2000; Morrison, 2004; Norman and Schmidt, 2000). It is argued that this need could be met by process-oriented research that unpacks the “black box” of PBL by describing and qualitatively analysing all the relevant phases of the PBL cycle in order to clarify the learning processes (Hak and Maguire, 2000; Koschmann and MacWhinney, 2001).

All the attributes of PBL are subsumed under four concepts that view learning as being constructive, self-directed, collaborative, and contextual (Dolmans and Wolfhagen, 2005b; Hmelo-Silver, 2004; Norman and Schmidt, 1992; Yew and Schmidt, 2009). Collaborative learning entails that the students interact together for the purpose of achieving a common learning goal (Yew and Schmidt, 2009). One of the principal goals of PBL is to help the students develop effective collaborative skills (Hmelo-Silver, 2004). The potential benefits of working collaboratively in small groups have been widely documented (Albanese, 2000; Hmelo-Silver, 2002, 2004). Only limited evidence exists that can establish that PBL helps students become better collaborators in medical problem-based learning tutorials (Hmelo-Silver, 2004). Given that (1) competencies in collaboration and knowledge-sharing is one of the goals of PBL; (2) tutorial participants spend about

5−6 hours a week talking among themselves; (3) collaboration can promote students' learning and motivation; and (4) the mechanisms of knowledge-sharing in the PBL tutorials are not fully known (Albanese, 2000; Da Silva and Dennick, 2010; Hmelo-Silver, 2002, 2004). And therefore, a study to improve our knowledge on how knowledge is collaboratively constructed in medical education PBL tutorials over a wide spectrum of tutorial facilitation is needed.

1.13. The Purpose of the Study

The purpose of this study is to identify and describe evidence supporting collaborative knowledge construction and scaffolding in graduate-entry PBL tutorial discourse. This is very important because instructional effort in medical education aims to develop a correct understanding among learners, and shared understanding is very important for collaborative problem-solving, as embodied by PBL instructional methods. Students bring to the classroom, errors and misconceptions encountered in their understanding of medical science that need to be rectified. The PBL tutorial is an ideal learning environment where this can be achieved (Moust et al., 1989). Using transcriptions of the talk conducted in the tutorials, the processes of collaborative knowledge construction employed by the graduate-entry basic sciences medical students can be described, and evidence supporting the facilitation brought about through these processes can be highlighted. Thus, the objectives of the study are:

- To create a representative corpus of words that comprise PBL tutorial talk.
- To use the students’ subcorpus to describe how graduate-entry medical students collaboratively construct knowledge.
- To use the facilitators’ subcorpus to describe how facilitators guide the students’ knowledge construction processes.
- To use the study results to make recommendations for practice, policy-making, and further research.

1.14. Significance of the Study

This study offers the potential advantage of helping us in filling a gap in our existing knowledge regarding the processes that are employed to actualise collaborative knowledge construction in medical PBL tutorials, because there is insufficient literature available on this aspect of PBL in medical education. Aside from this intellectual benefit, the study may hold practical value. Administratively, the findings of the study may provide a framework that can be used to examine, reflect, and research upon the quality of collaborative activities practiced in medical PBL tutorials. It could also provide intellectual resource that can be used in facilitator education programmes to sensitise intending facilitators and in-service facilitators to the quality tutorial discourse.

1.15. The Scope of the Study

This study is exploratory in nature for two reasons. First, a reasoned argument has been advanced to establish that there is a need to understand what goes on in the PBL tutorials to ascertain how theory aligns with practice (Albanese, 2000; Hak and Maguire, 2000). Second, there is little data in the literature regarding the types of collaborative interactions that graduate-entry medical students employ in the process of knowledge construction. This may serve as the basis for studying the effects of other contextual factors on these interactions.

This study has not been designed to address the non-verbal components of the collaborative knowledge-building interactions. The results of the study would only be generalizable to the domain of pre-clinical graduate-entry medical education and to the theory of collaborative learning in PBL. The aim of the study is to describe, in some details, what graduate-entry PBL medical students verbalise when they collaborate together to construct knowledge and share an understanding.

1.16. Summary of the chapter

This chapter has highlighted the ancient root of PBL, its modern advent in the field of education, globalisation, and current heterogeneity. The cardinal principles of PBL are described along with its process. This chapter has also highlighted the epistemological trend of knowing and has described the theoretical foundations of PBL. The chapter concludes by describing the purpose, significance, and scope of this study.

The second chapter describes corpus linguistics and corpus analysis, while chapter 3 comprises the literature review. In chapter 4, research methodology is described while the research design is highlighted in chapter 5. The results of the study are presented in chapters 6 and 7, while chapter 8 focuses on discussion, the pedagogic issues that emerge from the study, recommendations, and conclusions.

CHAPTER 2

CORPUS LINGUISTICS AND CORPUS ANALYSIS

2.1 Introduction

This chapter presents the theoretical foundation of corpus linguistics (CL) and sheds light upon its relationship with the linguistic theories along with that it shares with computer software. The purpose of this chapter is to provide a background for the method of analysis that has been used for this research.

2.2. Corpus Linguistics

CL can be defined as a methodology (Gries, 2009; Lindquist, 2009; McEnery and Hardie, 2012; McEnery and Wilson, 2001) designed for analysing a set of machine-readable texts that is determined to be an appropriate basis on which to study a specific set of research questions (McEnery and Hardie, 2012). It is language research methodology that involves the exploration of large quantities of naturally occurring texts that are stored electronically (Meyer, 2002; Thompson and Hunston, 2006). It is a language research methodology comprising a large number of related methods that could be aligned to any theoretical approach to language study (Gries, 2009; Lindquist, 2009; McEnery and Hardie, 2012; Thompson and Hunston, 2006). The two prominent approaches to the study of language that have attracted the most attention and have been frequently addressed and employed in literature pertaining to the field of both linguistics and applied linguistics are generative or intuition-based grammar (Chomskyan generative grammar) and (Hallidayan Systemic Functional) functional- or observation-based grammar (Aarts, 1991; Bourke, 2005; Sadighi and Bavali, 2008; Wells, 1999).

2.3. Corpus-based vs. Corpus-driven Approaches

There are two main traditions that fall under corpus linguistics – corpus-driven and corpus-based approaches. The differences between corpus-driven and corpus-based approaches relate to the use of the data, significance of representativeness, and corpus annotation (McEnery and Hardie, 2012; McEnery et al., 2006). When following the corpus-based approach, data is used to explore a theory or hypothesis in order to validate it, refute it, or refine it, whereas in corpus-driven approach, data is approached without a preconceived categorisation and data is used to generate theories (McEnery and Hardie, 2012). In the corpus-based approach, emphasis is placed on representativeness, balance, and corpus annotation – all of which are considered as being less important in the corpus-driven approach (McEnery et al., 2006)

2.4. Corpus Linguistics and Linguistic Theory

Corpus linguistics, as a field of linguistics, involves the study of language as it is used in the "real world", i.e., it is observation-based

(Aarts, 1991; Leech, 1991; Lindquist, 2009; Meyer, 2002). This is in sharp contrast to intuition-based grammar, which focuses on the internal structure of language, whereas corpus linguistics focuses on external language. The intuition-based approach, especially, as outlined by the theories of Chomsky and other generative grammarians, has wielded more influence on the development of language and language structures (Aarts, 1991; Bourke, 2005; Sampson and McCarthy, 2004). The generative approach to language describes language as an error-laden data source, and it contends that meaningful language study can only be undertaken under experimental laboratory conditions, where errors can be controlled and minimised. The implication of this is that research inevitably involves a detailed and fine-grained exploration of small data set (Meyer, 2002). Corpus linguistics, on the other hand, uses naturally occurring texts, such as data sets, to understand language and its structures. It sees errors as a manifestation of the richness of natural language and considers that they ought to be analysed as opposed to being eliminated; and if the sample of data collected is large enough, language errors will not have a significant effect upon the conclusions drawn thereafter (Leech 2005 cited in Da Silva, 2013).

The other key difference between Chomskyan generative grammar and the functional grammar of Halliday relates to the adequacy level they address. The three levels of adequacy that Chomsky describes (Chomsky, 1965; Sadighi and Bavali, 2008) represents a classification of linguistic theories according to their potency (Da Silva, 2013). The three levels of adequacy are as follows: (1) observational adequacy, (2) Descriptive adequacy, and (3) explanatory adequacy (Meyer, 2002; Sadighi and Bavali, 2008):

An observationally adequate theory is able to describe reality; a descriptively adequate theory describes reality and specifies the underlying abstract principles; and within the purview of explanatory adequacy, a theory does not just describe reality, but explains and predicts it as well. Corpus linguistics focuses on the observational/descriptive levels while generative grammarians classify theories at the explanatory adequacy level (Boeckk and Hornstein 2003 cited in Da Silva, 2013). Leech (1992c) considered that the three levels of adequacy could be collapsed to represent the two distinct paradigms of linguistic research.

CL is not homogenous in terms of its methodology for doing research and answering a set of research questions (McEnery and Hardie, 2012; Thompson and Hunston, 2006). For example, there are divergent views regarding data analysis: Sinclair advocates minimal annotation of corpora so that the text is afforded the opportunity of revealing its true meaning, whereas Quirk and some others advocate extensive annotation of texts to improve upon the rigour and to enhance an understanding of the texts at hand (Da Silva, 2013). Despite the differences in the orientation of the analytical methods, experts in corpus linguistics shares a set of broad principles and core assumptions of corpus linguistics as providing the best approach to the study of language use when compared against the methods adopted by the generative grammarians.

Generative grammar relates to the origin or psychological aspect of language, while corpus linguistics considers language as a tool and focuses on its communicative social functions (Sadighi and Bavali, 2008). Generative grammar focuses on language competence, whereas corpus linguistics concerns itself with language performance (Leech, 1992c; Meyer, 2002). These distinctions make corpus linguistics versatile and applicable to other fields. The advantages of corpus linguistics, as highlighted by Jan Svartvik (1992 cited in Lindquist, 2009: p. 9) include the following:

- Corpus data are more objective than data based on introspection.
- Corpus data can easily be verified by other researchers, and researchers can share the same data.
- Corpus data is needed for studies of variations between dialects, registers, and styles.
- Corpus data provides the frequency of the occurrence of linguistic items.
- Corpus data provides illustrate examples and is a theoretical resource.
- Corpora provide the possibility of total accountability of linguistic features.
- Computerised corpora give researchers all over the world access to the data.

2.5. Corpus Linguistics and Functional Linguistics

Both Corpus linguistics and systemic functional linguistics deal with language description. Systemic functional linguistics is a theory of language, while corpus linguistics is a method for investigating language (Thompson and Hunston, 2006). There is a natural symbiotic and synergistic relationship between functional language theory and corpus linguistics (Halliday, 2006; Thompson and Hunston, 2006). This is because all functional theories of language have the similar objective of demonstrating how speakers and writers use language to achieve communicative goals (Meyer, 2002).

Despite the synergy between systemic functional (SFL) and corpus linguistics (CL), there are some differences between the two. For instance, SFL is theory is heavy in many respects, but CL is not attached to any theoretical framework, but it is rather compatible with any linguistics theoretical framework; and others who subscribe to the corpus-driven approach argue that it should be theory neutral (Thompson and Hunston, 2006). Another area of difference relates to how analyses in SFL have to be inserted into the existing categories (Thompson and Hunston, 2006). This systemic view of language closes and constraints observations, such that findings that do not fit into the three-part existing system are not accounted for at all. In contrast, CL's relative lack of theoretical attachment means that new insights and findings can be easily incorporated into the language descriptions.

Despite the differences, the commonality between SFL and CL far outweigh their differences (Thompson and Hunston, 2006): First, both SFL and CL concern themselves with naturally occurring language and with language as a text. Both focus on language in the “real world”.

Second, the traditions of both commonly share the concern over context, but context is articulated differently in the two traditions. SFL sees language as being contextual, and as a system that interacts with other social semiotic systems. Language influences the choice made and how the available language resources are utilised. On the other hand, CL takes variation between corpora as the beginning of further enquiry (Conrad and Biber, 2000). The register is determined by the context in which the language is being used (Halliday, 1991a). Corpus research can show differences in the relative frequencies between language items, or between co-occurrences, or between corpora (Thompson and Hunston, 2006).

Third, both traditions adopt a probabilistic view of grammar (Halliday, 1991a, 1991b; Thompson and Hunston, 2006). For Halliday (1991a), a linguistic system is inherently probabilistic, and frequency in text is the instantiation of probability in the grammar. He defines register variation “as systematic variation in probabilities; a register is a tendency to select certain combinations of meanings with certain frequencies, and this can be formulated as the probabilities attached to grammatical systems, ….” (Halliday, 1991a: p. 33). CL concerns relative frequencies (Biber et al., 1998; Conrad and Biber, 2000), and Matthiessen argued that relative frequencies indicate systemic probabilities; and by determining the frequencies, probabilities can be established (Matthiessen, 1999). In the study of linguistic change, frequency patterns as revealed in the corpus studies explain historical change because when interpreted as probabilities, they show how text instance maintains or changes the language system (Halliday, 1991a). The notion that SFL and CL share a probabilistic view of grammar was demonstrated by Matthiessen (2006), when he analysed a corpus to reveal not only the probabilities of occurrence of lexicogrammatical features but also the probabilities of the features of co-occurring. Thompson and Hunston (2006) note that the occurrence and frequency of language items within text play a role that is mediated through register in the establishment and evolution of the language system – and hence, the increasing importance of corpus-based studies in SFL (Thompson and Hunston, 2006).

2.6. Corpus Linguistics: Historical Perspectives

Although the availability of electronic, machine-readable corpora of 1960s marked an explosion in language research, corpus-based linguistic analysis of various kinds existed in the fields of biblical and literary studies, lexicography, dialectic studies, language education studies, and grammatical studies before that time (Kennedy, 1998). The post-Bloomfieldian era was a period when some structuralist linguists, for e.g., Harris and Hill in the 1950s, were under the influence of a positivist and behaviourist view of science. The American structuralists considered that the quantitative results from the corpus were sufficient evidence for linguistic investigation, following which intuitive evidence was disregarded.

The American structuralist period was followed by an era of discontinuity in the late 1950s, when Chomsky "put to flight the linguistics of the earlier generation" (Leech, 1991). Generative grammar became influential and very dominant among theoretical linguistics (Lindquist, 2009). Chomsky's opinion that corpora are inadequate at language description, and that intuition is sufficient formed the doctrine to follow for the succeeding generation of theoretical linguistics (Chomsky, 1965; Leech, 1991). He claimed that the findings of corpus linguistics are trivial (Halliday, 1991a; Kennedy, 1998; Lindquist, 2009).

The modern era of corpus linguistics commenced with the works of Randolph Quirk and Francis and Kučera at the beginning of the 1960s (Leech, 1991; Lindquist, 2009). Randolph Quirk and his colleagues surveyed English usage and compiled equal proportions of written and spoken English in 1959. The half-a-million-word spoken corpus contains fully transcribed recordings of private conversations and meetings to produce the London Corpus (Kennedy, 1998; Lindquist, 2009).

In the United States, Nelson Francis and Henry Kučera at the Brown University compiled and completed, in 1964, the first electronic collection of American English to be used for linguistic research (Leech, 1987; Leech, 1992a; Lindquist, 2009). The Brown Corpus consists of one million words of written texts, extracted from various sources, to represent the American English that was prevalent at that time. The text also employed an innovative complex computational analysis based on knowledge from a number of disciplines (Leech, 1992c). These two corpora were compiled in the pre-computer era when research in CL was very burdensome and time-intensive: corpus collection was very difficult, few transcription facilities were available, annotation, tagging and classification were done manually by the researchers (Da Silva, 2013).

The compilation of the London-Lund corpus in 1975 heralded the era of electronic corpora. Jan Svartvik of Lund University, Sweden, upgraded the spoken English aspect of the London corpus from 87 5,000-word texts to a 100 5,000-word texts (Kennedy, 1998). The corpus is transcribed in a detailed prosodic transcription, and with a total of about half a million words. This corpus constituted the largest and most widely used electronic corpus of spoken English available at the time (Kennedy, 1998; Leech, 1991, 1992c).

Advanced computer and word-processing technology brought increased capability for text data collection, transcription, and corpora annotation closer to researchers, leading to increased use of larger corpora that could be distributed among the research community (Alexa, 1997). The 1980s and 1990s witnessed the appearance of megacorpora of natural spoken English (Leech, 2000). Leech's review of some of the electronic corpora of spoken English collected from various sources and diverse contexts and countries at the time was in excess of 43 million words in total. For example, the COBUILD (Bank of English) Corpus, collected in 1987, contains over 20 million words of transcribed spoken English; the British National Corpus (BNC), compiled in the early 1990s, contains more than 10 million words of spoken English; Longman Corpus of Spoken American English, collected in 1997, and orthographically transcribed, contains about 5 million words of spoken language; and Wellington Corpus of Spoken New Zealand English contains about 1 million words of largely informal dialogue.

Advances in the technology for computerising texts pose a new set of challenges. First, there is a requirement for the old methods of corpus annotation and compilation to be tested and adjusted to the new technological situation. Second, new analytic methods and software have been developed in close collaboration with linguists, thus relocating computational linguistics from computer science to the domain of linguistics (Adolphs, 2006; Alexa, 1997). Finally, the application of corpus linguistics has extended to other fields of study, including health care and educational contexts (Da Silva, 2013; Da Silva and Dennick, 2010).

2.7. The Corpora

A corpus refers to a collection of specimens of a language as used in real life, in speech or writing, selected as a sizeable fair sample of the language as a whole or of some linguistic genre, and hence as a useful source of evidence for research on the language. (Sampson and McCarthy, 2004). Nowadays, corpora are collected and stored electronically in machine-readable forms (Kennedy, 1998) as electronic corpora.

2.7.1 Types of Corpora

According to Kennedy (1998), corpora have been compiled for many different purposes, and these, in turn, influence the design, size, and the nature of individual corpora. The types of available important corpora are well-outlined in Lindquist (Kennedy, 1998; 2009):

Spoken corpora: This contains spoken language recorded in different conversational contexts, for e.g., LLC of Spoken English, BNC of Spoken English, and London Teenage Speech.

General Corpora: General corpora is compiled to be used for general purposes. It contains spoken and written language samples and is meant to represent language use in different aspects of the society, for e.g., British National Corpus (BNC), American National Corpus (ANC), International Corpus of English (ICE), and Corpus of Contemporary American English (COCA). The sources of Corpus of Contemporary American English (COCA) include American TV, radio, books, magazines, newspapers, and journals divided into the following five registers of equal size: (1) spoken – majorly comprising transcribed conversations from televisions and radios, (2) fictions, (3) popular magazines, (4) newspapers, (5) academic journals from different fields.

Specialised Corpora: Unlike the general corpora, specialised corpora is used for researching into certain genres or registers and to study certain language use domains, such as academic language or language of sports journalism. However, researchers tend to create specialised corpora to answer the specific research questions that they are interested in. Examples include the Michigan Corpus of Academic

Spoken English (MICASE) and the International Corpus of Learner English (ICLE).

Historical/Diachronic Corpora: This is used for studying language change that takes place over a long period of time by comparing older texts with the modern ones. Examples include the Helsinki corpora and Lampeter Corpus of Early Modern English Tracts (LCEMET).

Parallel and multilingual Corpora: A corpus can also consist of more than one language, wherein, usually, one language is a translation of the other language. The corpora can be used for comparative language studies and for translational studies.

Others: Other types of corpora include dictionaries, text archives, and the World Wide Web. All of these sources can be used for linguistic investigations.

2.8. Computers in Corpus Linguistics

The advent of the computer revolution in linguistic research has changed the way language data is processed analysed. Hitherto, the manual analysis of huge bodies of text was associated with being prone to error, inconsistent, and less exhaustive (Kennedy, 1998). Nowadays, easy access to computers have made access to machine-readable text fairly straightforward and a considerably less complex and laborious process (Alexa, 1997). The immense benefits of computers to linguistic studies, as enumerated by Kennedy (Kennedy, 1998), are summarised below:

Data Storage. Information technology has changed the way corpora are handled and stored. Instead of sorting hundreds of thousands of dictionary slips and index cards into “shoe boxes” (Lindquist, 2009), lexicographers and grammarians can easily use computers to store large amount of texts (Kennedy, 1998).

Text Retrieval. The pre-computer text retrieval involved sorting through large amounts of dictionary slips and index cards, containing examples of sentences and phrases. This time-consuming and exhausting process has been simplified by information technology. Computers are now able to retrieve particular words, phrases, or whole chunks of text in context, quickly and exhaustively.

Data Sorting. Linguistic items can be arranged or categorised in many ways. This operation, which is very cumbersome, when conducted manually, has been simplified by the use of computers that are able to sort linguistic items according to a defined algorithm, taking account of items they collocate with and their typical grammatical behaviour.

Corpus linguistics is thus associated with the computers that can perform diverse operations at an incredible speed, total accountability, accurate replicability, statistical reliability, and the amazing ability to handle large quantities of data.

The availability of modern software implies that electronic corpora are easily accessible and can be shared among researchers, thus drastically reducing the drudgery and the bureaucracy of dealing with the ever-increasing large amounts of data required to compile dictionaries and other sources of information. Computers also offer high reliability when it comes to basic tasks, such as searching and counting; and they can show accurately the probability of occurrence of linguistic items in texts, thus fostering the emergence of mathematical bases for automatic natural language processing, bringing to linguistic studies, a high degree of accuracy necessary in all branches of science.

Moreover, computers have enabled language researchers to work with a large variety of texts, and this has enabled them to seek generalisations about language and language use. Corpus-based quantification of language use has led to interesting scientific generalisations, which has enhanced the links between linguistic description and various applications. This, in turn, has benefited several linguistic domains, including machine translation, text-to-speech synthesis, content analysis, and language teaching.

2.9. Software Tools in Corpus Linguistics

Corpus linguistics is one of the dominant methods in use today for language analysis (Anthony, 2013). A key feature of corpus linguistics is that it relies on the computer software for text analysis (Biber et al., 1998; Hardie, 2012).

The type of research questions that can be answered in corpus linguistics studies is determined by the combination of corpus and search software (Anthony, 2013; Hardie, 2012; McEnery and Hardie, 2012). A corpus of language is considered as being virtually useless without a computer software tool to process it and display results in a way that is easily understandable (Anthony, 2005). Anthony (2013) observed that it is vital to make very clear the distinction between corpus data and the corpus tools used to analyse data. In his astronomical example, he said that an object could be observed with observation tools that range from the human eye and simple binoculars to advanced reflector telescopes positioned in space; however, as is obvious, what can be seen and the results of the observation would depend on the tool selected. Thus, the research questions that can be asked, and the results of analysis of a corpus depend on the sophistication level of the software used.

First generation Tools. Historically, the first software tools for corpus analysis was developed by Roberto Busa (McEnery and Hardie, 2012). Since then, several corpora analytic tools have been developed. In a historical review of corpus analysis tools, McEnery and Hardie (2012) described four generations of corpus analysis tools (Anthony, 2013; Hardie, 2012; McEnery and Hardie, 2012): The first-generation, which emerged in the 1960s and 1970s, ran on mainframe computers. They were limited to analysing English corpora, as they could only process ASCII codes. Most of the tools were designed for a single function, such as counting the number of words in a text or producing KWIC concordance lines. Any other analysis required would require the use of other programmes. Examples of first generation corpus tools include Concordance Generator, Discon Drexel Concordance Program, and CLOC.

Second generation Tools. The second-generation of corpus tools appeared in the 1980s and 1990s. They ran on IBM-compatible PCs, and thus, could run on readily available computers, which enabled researchers to carry out small-scale studies. They were limited to processing ASCII and had restricted functionality. However, they could sort alphabetically the left and right context of the word searched, produce a word list, and calculate some basic descriptive statistics about the corpus. Corpus analysis tools in this category include Oxford Concordance Program, Longman Mini-Concordancer, Kaye Concordancer, and MicroConcord.

Third-generation Tools. The third-generation tools include most of the current tools, such as WordSmith Tools, MonoConc Pro, AntConc, and Xaira. The tools in this category started to emerge in the 1990s, and they are continuously being developed and improved upon till date. These tools are multifunctional, include statistical methods, are capable of dealing with larger corpora, offer multi-language support, and include user-friendly interfaces suitable for those with limited computer experience. Their limitations include a lack of capacity to process corpora in excess of 100 million words, and they require installation on personal computers. The second limitation has become a great problem as publishers are increasingly reluctant to have their dataset used for research purposes, and so, texts can no longer be compiled and distributed for the purpose of analysis on personal computers.

Fourth-generation Tools. The fourth-generation software tools were developed in response to the problems encountered by third-generation tools. Examples of fourth-generation tools are corpus.byu.edu, Wmatrix, CQPweb, and SketchEngine. Their strengths include the provision of protection relating to copyright issues, as the underlying text belonging to the corpora are concealed from the users. These tools decouple corpus searching from the limits of the memory and processing power of the users' computers, can run on all operating systems because they are web-based, require less computer technical skills to use, and require no installation. However, these tools have a number of limitations: They require internet connection for data uploading and for analysis. They require data cleaning, processing, and reformation. They are less appropriate for very sensitive data.

2.10. Corpus Analysis

Corpus analysis refers to the investigation of a corpus to answer definite research questions. A body of texts is called a corpus. Corpus analysis, as a research methodology, deals with the analysis of collections of texts as the source of evidence for linguistic description. An integral part of the analytic procedure is the quantification of the distribution of linguistic terms (Kennedy, 1998). It focuses on linguistic performance, rather than on competence, as well as on observation of language in use (Leech, 1992c). Researchers have always needed evidence for theories about the nature, elements, structure, and functions of language; and this is derived directly from texts in the case of corpus-based research (Kennedy, 1998). To arrive at the evidence, corpus analysis studies texts – spoken and written – by categorising words in a text, according to predefined linguistic categories, and it analyses both grammatical and semantic relationships between them (Adolphs, 2006).

The procedure of corpus analysis aligns with the generic model of scientific methods (falsifiability, objectivity, simplicity, strength, and completeness) set out in the following five core steps (Leech, 1992b; Rayson, 2003, 2008): question, design, annotation, retrieval, and interpretation.

2.10.1 Research Question

This involves developing research questions that a researcher wants answered. Three main types of research questions that can be addressed through observing the corpus methodology fall into the three distinct approaches to corpus analysis (Rayson, 2003, 2008):

Type I Approach. This concerns microscopic text analysis and focuses on linguistic features, such as words/multi-word units or specific linguistic structures. This approach is the traditional method commonly adopted by linguistics (Rayson, 2008).

Type II Approach. This concerns macroscopic text analysis and concerns itself with the whole text or texts. It deals with differences and similarities across different texts, for example, grammatical structures' variations across texts.

Type III Approach. Type II approach was proposed by Rayson in his PhD thesis (Rayson, 2003). This approach combines type I and type II, and it is described as a data-driven approach, in the sense that macroscopic analysis is first conducted, and then the decision is made on which linguistic feature should be further studied microscopically

(Rayson, 2003, 2008). With increasing computer software sophistication, this approach has become more feasible and popular in CL (Rayson, 2003). According to Rayson (2003), type III as an iterative inquiry, is cyclical and spiral in nature until a satisfactory level of analysis is reached. This cyclical and spiral analysis process consists of initial question or plan and is followed by an initial analysis. The results of the analysis are appraised and new aspects to be further explored are identified. This leads to a new plan or question, and a new cycle is initiated, with each cycle adding more details to the analysis (Rayson, 2003, 2008).

The first approach is adopted in this project because the aim of the research is to study the language features of collaborative knowledge building from the linguistic point of view. This study involves the identification of the defined elements of collaborative knowledge building, first quantitatively, and then microscopically, exploring the context of these elements.

2.10.2 Corpus Design and Compilation

A corpus refers to a systematic collection of naturally occurring machine-readable texts of both written and spoken language. The term "systematic" implies that the structure and contents of the corpus are based on certain extra-linguistic principles (Nesselhauf, 2005). Sinclair (1996 cited in McEnery et al., 2006) describes corpus as a collection of pieces of language that are selected and ordered according to specific linguistic criteria in order to be used as a sample of the language. The linguistic criteria or principles relate to the intended use of the corpus, and they influence how the texts are compiled. Once the research questions are formulated, the corpus construction can begin (Reppen, 2012). The texts to be included in the corpus for analysis need to meet explicit criteria, such as representativeness, size, sampling, balance, and comparability (McEnery and Hardie, 2012; McEnery et al., 2006; Meyer, 2002; Reppen, 2012).

Representativeness implies that the corpus needs to be representative of the language being investigated. It is related to the research question and purpose. For instance if the goal of the research is to study newspaper editorials, a collection of personal letters would not be representative of the language of newspaper editorials (Reppen, 2012). There needs to be a match between the language being investigated and the type of material being collected (Biber 1993 cited in Reppen, 2012). Representativeness will be ensured by capturing the correct data and ensuring an appropriate sampling method. This involves ensuring that the compiled corpus covers several discussions of clinical cases across different modules with different facilitators.

Another related criterion to representativeness is the corpus size. Corpus size shapes the research feasibilities, and it is, in turn, itself influenced by theoretical and practical constraints (Da Silva, 2013). The fundamental presumption of CL is that samples of a language used for analysis are large enough to be representative of the language feature being researched upon. According to Leech (1991 cited in McEnery et al., 2006), a corpus is considered “to be representative of the language variety it is supposed to represent if the findings based on its content can be generalised to the said language variety." The representativeness does not mean generalizability to the larger population or group other than the one being studied; on the contrary, it signifies an ability to represent the key elements of the language or the reality being studied. Consequently, a large enough corpus provides an assurance that sufficient recurrences of each language feature are present to enable the extraction of representative patterns (Adolphs, 2006).

The optimal corpus size has been a contentious issue (Kennedy, 1998; Koester, 2012; Lindquist, 2009; McEnery et al., 2006; Meyer, 2002). The unwritten assumption that the "biggest is best" characterised the early years of CL, when larger and larger corpora were analysed (Kennedy, 1998).This was attributed to the excitement of the possibility of compiling data set including millions of words and the ability to meet the specific needs of the lexicographers who were the early users of corpus data and needed large data sets to extract sufficient examples of infrequent words to enable them to produce reliable descriptions of their use (Evison, 2012). This position has been abandoned by subsequent applied linguists with different interests. A small corpus is considered appropriate for high-frequency grammatical patterns or discourse features, such as pronouns, prepositions, auxiliary, and modal verbs (Koester, 2012; Meyer, 2002). A large corpora could also mean that too large an amount of data sets may become unmanageable if the analyst is interested in searching for high frequency linguistic items or carrying out a detailed analysis (Evison, 2012). This may force the analyst to conduct random sampling or to work with a smaller subsample that will allow for all occurrences of the items of interest to be examined effectively (Evison, 2012; Koester, 2012).

The small corpora approach, which was successfully used by Biber (1990 cited in Meyer, 2002 & Evison, 2012), showed that just 1,000 words of data is large enough to provide valid and reliable information on the distribution of the frequently occurring linguistic features, such as pronoun and verb forms. Tribble (1977 cited in Evison, 2012) argued convincingly that, for a very specialised register, a smaller corpus is adequate for providing insights into the features of the register. The use of small corpora has been demonstrated in two studies by Koester and O'Keeffe (cited in Evison, 2012). Koester, in 2006, used a corpus of just under 34,000 words to investigate workplace discourse; and O'Keeffe, in 2003, based his study of media discourse on a sample of 55,000 words of phone-in data.

Small corpora are useful when the aim of the study is not to generalise conclusions across a language but rather to understand a phenomenon through the study of the language used. This is the case in a specialised corpus, when CL is applied to other areas of study (Da Silva, 2013). The smaller more specialised corpora, in this instance, allow for a much closer link to be established between the corpora and the contexts in which the texts in the corpora are produced (Koester, 2012). Large corpora describe the general lexico-grammatical patterns, whereas small specialised corpora offer insights into the patterns of language use in particular settings. The compiler is often the analyst of small specialised corpora, and as such, is familiar with the context of the corpus. This enables the analyst to balance and complement the quantitative findings revealed by the corpus analysis with qualitative findings (Flowerdew 2004 & O'Keeffe 2007 cited in Koester, 2012). Flowerdew (2002) and Tribble (2002) (cited in Koester, 2012) argued that smaller, more focused corpora, set up for a particular pedagogic purpose, are a lot more useful and are more likely to yield insights that are directly relevant for teaching and learning for specific purposes.

The use of CL in this research adopts a similar focus of making use of small corpora to understand a phenomenon in a specific setting. The aim is not to generalise the result beyond the research context. A mixed methods approach is employed, and the same is done to allow qualitative insights into the data set in order to enrich qualitative findings. The anticipation is that the results of the research would be used to improve upon the performance of the students and the teachers in the problem-based tutorials.

Sampling is unavoidable because it is impossible to analyse every utterance or sentence of a given language (McEnery et al., 2006). It is, therefore, important to ensure that the sample chosen is representative of the language being investigated. A convenient sampling method – sampling the available individuals – is used in this study, based on practical reasons. Like sampling is the issue of balance in corpus compilation. This implies the range of text categories included in a corpus. This is important if the corpus is made up of texts belonging to different genres. For example, a corpus of political ideology needs to include texts of political views collected from television, newspapers, political debates, party manifestos, and so forth. Similarly, a general corpus needs to include texts of written and spoken language. Balance of corpus is not relevant in this study because the transcribed problem-based learning (PBL) corpora contained only transcriptions of PBL discussions.

The last characteristic of a corpus to be considered is comparability. This simply means comparing likeness against likeness. Corpora can only be compared if the constituent texts belong to the same type of data, for e.g., written vs written and spoken vs spoken. This will ensure that the differences found between the corpora is due to the variables being studied and not due to the differences in the type of the material used within the corpora. For example, differences between written and spoken corpora may be due to the known differences in the grammatical and lexical structures that exist between the two kinds of texts and not due to any variable being investigated. Comparability could be guaranteed by conducting appropriate sampling, data collection, and through preparation (Adolphs, 2006; Adolphs and Knight, 2012).

2.10.3 Corpus Annotation

Corpus annotation is the process of adding interpretative linguistic information to an electronic corpus (Leech 1997 cited in McEnery et al., 2006). It is a prerequisite for corpus analysis and involves the addition of descriptive and or analytical information to the text (Rayson, 2008). Through annotation, word characteristics are made explicit to the software for further analysis. This enables the corpora to be researched and analysed, and comparison between corpora is made possible (Adolphs, 2006; Leech 2005 cited Da Silva, 2013).

Corpus annotation offers a number of advantages for corpus analysis (McEnery et al., 2006): First, it makes information extraction from the corpus easier, faster, and more reliable. Second, an annotated corpus could be reused, thus minimising cost and time of annotation for each analysis. Third, annotation documents linguistic analysis objectively. Fourth, annotation is multifunctional, meaning that corpus annotated for a purpose may be used for several other purposes. Finally, corpus annotation provides a stable base for linguistic analysis, such that successive studies can be compared and contrasted following a common basis.

Corpus annotation has been criticised for several reasons, including production of a cluttered corpora, imposition of linguistic analysis upon a corpus user, and the tendency of the annotation to overvalue the corpus. These criticisms have been "dismissed, with caveats, quite safely" (see McEnery et al., 2006: p. 31 - 32). Annotation adds different information to the textual material based on the research questions and purpose (Adolphs, 2006). Any annotation process commences with grammatical and semantic annotations in CL, and this involves characterising each word on the basis of its syntactic and semantic value (Rayson, 2008). Additional forms of annotation may be added subsequently.

There are three main methods of corpus annotation, which are as follows: automatic, semi-automatic, and manual (McEnery et al., 2006): In automatic annotation, the computer functions independently as an annotator by following the rules and algorithms predetermined by the programmer. Once automatic annotation is completed, large quantities of data can be annotated rapidly and consistently. However, it is time-consuming and costly. Semi-automatic annotation involves the human correction of less reliable or less accurate automatic annotation. This can occur when the human analyst is required to resolve ambiguous cases when the machine is not certain of the same. Semi-automatic annotation may produce more reliable results than fully automatic annotation can. Pure manual annotation occurs when there is no available tool for automatic annotation, or where the accuracy of the available tool is not high enough so as to make the time required for manual correction less than pure manual annotation. Manual annotation is expensive, inconsistent, and time-consuming; and it is only suitable for a small corpora. Corpus annotation could be undertaken at different levels and may take several forms. A list of common annotation types is presented below:

Part-Of-Speech (POS) Tagging. POS tagging (or grammatical tagging or morpho-syntactic annotation) refers to assigning each word or word combination in the corpus to grammatical categories or POS tags, such as noun, adjective, adverb, pronoun, and the like (Rayson, 2003, 2008). It is the commonest, most basic, and the first widely used form of corpus annotation; and it forms the basis for further forms of analysis, such as parsing and semantic annotation (McEnery et al., 2006). According to McEnery et al. (2006), POS tagging is advanced, and it can be performed with high precision suitable for most research questions. The annotation tool which automatically assigns POS tags to words is called a tagger. CLAWS (the Constituent Likelihood Automatic Word-tagging System), developed in Lancaster University, is one of the most popular and reliable taggers used for English language (Garside, 1987). An example of POS tag scheme in Wmatrix 3 is shown in the Table below.

Semantic Tagging (Semtag). Semantic annotation is also referred to as word-sense tagging, and it is particularly useful for content analysis (McEnery et al., 2006). It describes the procedure for adding meaning or semantic value to words. A fundamental part of this process involves distinguishing the lexicographic senses of a word by combining an analysis of the grammatical tags as well as of the context surrounding the word (Rayson, 2008). Semantic tagging is more difficult than POS tagging, because it is knowledge-based requiring resources like dictionaries and thesauri (McEnery et al., 2006). It involves matching each word with its possible dictionary or thesaurus definition. Corpus analysis software is supported by extensive databases, containing collections of many different dictionaries, such that the determination of the semantic value of a word is made possible. More often than not, the meaning of a word depends on the context, and it is important to ensure that the senses of word use are differentiated – a process termed, disambiguation. The sophistication of modern computer software, when it comes to dealing with huge databases and background linguistic knowledge gained from previous extensive research, allow for this process to be carried out with minimal human interference. The capacity of the computer software to determine the semantic value of a word is made possible by the POS tags of the surrounding words. For instance, the word “party”, surrounded by democratic, conservative, and liberal is assigned to a semantic category that is different from the one under which it is categorised when it is surrounded by words like birthday, Christmas, fun, and so forth (Rayson, 2003, 2008). An example of semantic tagging scheme in Wmatrix 3 is shown in the Table 2.1 below.

Table 2.1: Example of POS tag set used by Wmatrix 3

POS	**POS Domain**
BCL	before-clause marker (e.g. in order (that), in order (to)
CC	coordinating conjunction (e.g. and, or)
CCB	adversative coordinating conjunction (but)
CS	subordinating conjunction (e.g. if, because, unless, so)
CSA	as (as conjunction)
CSN	than (as conjunction)
CST	that (as conjunction)
CSW	whether (as conjunction)
VM	Modal auxiliary (can, will, would, etc.)
VMK	Modal catenative (ought, used)

Table 2.2: Example of Semantic tag set used by Wmatrix 3

Semtag	Semantic Domain
A5.1	Evaluation: Good/bad
A5.1+	Evaluation: Good
A5.1-	Evaluation: Bad
A5.2	Evaluation: True/false
A5.2+	Evaluation: True
A5.2-	Evaluation: False
A5.3	Evaluation: Accuracy
A5.3+	Evaluation: Accurate
A5.3-	Evaluation: Inaccurate
A5.4	Evaluation: Authenticity
A5.4+	Evaluation: Authentic
A5.4-	Evaluation: Unauthentic

Grammatical Parsing. POS tagging is a step that often leads to other types of annotation, such as parsing (McEnery et al., 2006). Parsing is a process that concerns text analysis with the aim of understanding its syntactic structure. It assigns grammatical structures to corpus sentences, based on combinations of POS tags of its words (Rayson, 2008). Parsing is about the commonest annotation after POS tagging (McEnery et al., 2006). It is useful for studying clause types, but because of its much lower precision rate compared to POS tagging, automatically parsed corpus needs human corrections invariably (McEnery et al., 2006).

2.11. Retrieval and analysis

This section deals with the techniques and practices for quantitative and qualitative analysis of a corpus. It is common to differentiate these two scientific research methods in context of a corpus analysis (Lindquist, 2009):

Qualitative analysis involves a close analysis of text items or grammatical constructions that could lead to category and/or theory generation. Quantitative methods, on the other hand, count linguistic items and employ frequencies and percentages in order to describe language. According to Lindquist (2009), it is important to bear in mind that the cleavage between quantitative and qualitative is not that sharp for each methods tend to include the elements of the other. For example, to calculate frequencies and compute percentages in quantitative studies, the linguistic items to be computed need to be carefully defined and categorised, which falls under a qualitative analysis. Before we count the frequency and compose the percentages of modal verbs, we need to define what modal verbs are. On the other hand, quantitative analysis hardly stops at computing frequencies and statistical analysis. To make the best use of the figures, it is important to conduct a close observation of the realities behind the figures. Examples of quantitative investigations are presented followed by that of qualitative methods.

2.11.1 Quantitative methods

Basic Information. The first step in quantitative corpus analysis is the computation of basic information about the text or a collection of texts

(Adolphs, 2006). Such information includes total number of words, sentence length, word length, number of paragraphs, number of individual running words or tokens, number of different words or types, as well as the number of lexical and grammatical items in tagged and untagged forms (Adolphs, 2006). Descriptive information is very important in the collection of texts, more so, if the analysis involves a comparison of texts. The basic calculations could be used to establish a provisional picture of the corpus, which could then be elaborated upon and enriched further through quantitative and qualitative analyses (Adolphs, 2006).

The two basic and most common ways of retrieving and interpreting data in corpus analysis are frequency list and concordancing. They can be used exclusively or in combination for data analysis (Da Silva, 2013).

Frequency List. The most useful way to conduct a preliminary survey of the corpus is to compute the frequency list of the contents of the corpus. The frequency list shows the words or other language units that make up the texts in the corpus, together with their absolute or relative frequencies of occurrence (Adolphs, 2006; Barnbrook, 1996; Rayson, 2003, 2008). The frequency list can be produced in several different sequences, as chosen by the researcher (Barnbrook, 1996). The frequency lists include words list of single items or recurrent sequences, or keyword list of single words, or key sequences (Adolphs, 2006; Rayson, 2008).

Single word list. Single word list is produced by identifying each word form found in the text or in different texts, counting identical forms and listing them with their frequencies in a chosen sequence (Barnbrook, 1996). The list can be recovered for all the words in the corpus, or the recovery may be done for only the most frequent words, based on the syntactic or semantic tags used in the text annotation (Adolphs, 2006; Rayson, 2008).

List of recurrence sequences. This is otherwise termed collocations, and it refers to the list of words that co-occur frequently together in clusters (Scott 1997 cited in Da Silva, 2013). The collocations can be either common clusters because they are common expressions in a language (e.g., I don't know, you know), or because they are common clusters found in a particular corpora (e.g., human anatomy, cardiovascular system in medical corpora) (Da Silva, 2013).

List of keywords and key sequences. This refers to a list of frequency that describes not only the words in the text but a particular word, group of words, or sequence, and they are very helpful when the analysis concerns a particular aspect of the corpus (Da Silva, 2013).

Concordancing. The second aspect of basic statistical description in corpus analysis is Concordancing. Concordancing is the process of recovering words and their context from an annotated corpus. A concordancer is a computer software program that rapidly searches a collection of texts to produce a list of incidences of a given linguistic item (e.g., word, phrase, and the like). A concordance, otherwise known as Key Word in Context (KWiC), refers to a piece of text in which a linguistic item is displayed in the centre with its immediate context on both sides (Rayson, 2008; Yoon, 2011). The computer software program makes it possible to change the length of the context that is recovered from the corpus (Rayson, 2008). When read horizontally, concordance outputs information on how particular linguistic items are used in context; and when read vertically, information relating to repeated patterns of the linguistic items are provided (Tognini Bonelli 2010 cited in Yoon, 2011). Rayson (2003) observed that reading a range of 120 words is adequate in most cases to understand the contextual meaning of a linguistic item; however, depending on the research question, occasionally, it might be more rewarding to read entire paragraphs in which the items appear. Concordance provides opportunities for contextual analysis of lexical patterns, provides meaning to words, and facilitates disambiguation and interpretation of analytic findings (Adolphs, 2006; Evison, 2012; Flowerdew, 1993; Rayson, 2003, 2008; Yoon, 2011).

2.11.2 Statistical measures in corpus analysis

One of the outstanding advantages of CL over intuition lies in the provision of reliable quantitative data (McEnery et al., 2006). The use of this quantification in corpus analysis transcends mere counting (McEnery and Hardie, 2012). Statistical measures are used to demonstrate patterns in order to understand relationships within the data and to compare the different corpora being analysed (Rayson, 2008). Statistical measures in corpus analysis could be of the two following types: descriptive (or basic) statistics and inferential statistics (Levon, 2010; McEnery et al., 2006; Oakes, 1998).

Descriptive statistics gives information about the general shape or quality of the dataset (Levon, 2010). The commonly used descriptive statistics in corpus analysis include frequency (raw and normalised), measures of central tendency (e.g., mean, median), and measures of dispersion of a dataset (e.g., range, variance and standard deviation) (Levon, 2010; McEnery et al., 2006).

Inferential statistics, on the other hand, is used to decide whether or not apparent patterns in a dataset really are patterns (Levon, 2010). In other words, whether differences observed between two corpora are genuine differences, or they have come forth by chance due to inherent variability in the dataset (McEnery et al., 2006). Some of the commonly employed inferential statistics in corpus analysis include factor analysis, cluster analysis, Chi square (or Pearson Chi square) test, and log-likelihood (or log-likelihood Chi square or G-square) test (McEnery et al., 2006; Oakes, 1998). These statistical techniques offer information about the relationships of linguistic items (e.g., words, phrases, structures, and the like) within a corpus. They can also be used to investigate the relationships between linguistic items within the corpus and other extra-corporal variables, such as demographic information (Rayson, 2003).

The focus of the present study is not to compare linguistic variables, since it is mainly descriptive in orientation. Therefore, mostly, descriptive statistics will be employed. The descriptive measures will be complemented with qualitative analysis.

2.11.3 Qualitative methods

Qualitative research methods "is concerned with structure and patterns, and how something is" (Rasinger, 2010: p. 52). Qualitative analysis has been undertaken in this study not following any of the discourse analysis approaches but following the functional perspective. The aim is to understand how the linguistic items generated through concordancing are used in context by a close examination of the concordance lines. The aim of this approach is to provide a more comprehensive understanding of the output of quantitative methods.

The KWiC generated by the corpus analysis can be exported into the qualitative analysis software and analysed based on the circumstance of occurrence.

2.12. Summary of the chapter

This chapter discusses the meaning of CL, and it highlights its relationship with systemic functional linguistics theory. It highlights that CL is compatible with fundamental concepts of systemic functional. A historical overview of CL, types of corpora, and the relationship of computer software and CL are also discussed. The chapter then presents the definition of corpus analysis, highlights the principles and types of corpus compilation, and finally, it discusses the different types of corpus analysis.

CHAPTER 3

FACILITATION OF COLLABORATIVE LEARNING IN PROBLEM-BASED TUTORIALS: A QUASI-SYSTEMATIC REVIEW OF THE EXISTING LITERATURE

3.1 Introduction

The aim of the literature review is to elucidate upon the existing theoretical and empirical research conducted on the facilitation of collaborative learning in problem-based learning (PBL) tutorials, thus mapping out the domains of research focus, the analytic techniques employed, and the theoretical perspectives that inform the research.

A systematic review methodology was used to carry out the literature review in this study. A systematic review is a type of secondary study – a study based on the published studies. It is a research methodology designed to gather, evaluate, and analyse all available literature relevant to a research question or a phenomenon (Dickson et al., 2013; Khalid et al., 2011). A systematic review was carried out in this study to improve the quality of the review process. By standardising the review process as compared to the less formal review, it was anticipated that the review would help in identifying the knowledge gaps in the existing literature, and eventually, orientate the current research process.

The literature review is followed by the identification of the gaps in the existing literature, a highlighting of the research questions, and a discussion on the theoretical foundation of the current study.

3.2. General Trend in PBL Research: An Overview

Research on the PBL curriculum has passed through a number of stages (Svinicki, 2007). The initial focus of the research relates to the confirmation of concepts by the curriculum originators and early adopters. Two systematic reviews on the outcome studies (Albanese and Mitchell, 1993; Vernon and Blake, 1993) were undertaken in this period. Albanese and Mitchell's study found mixed results in the review of the studies comparing PBL with traditional curriculum. Vernon and Blake came to similar conclusions but gave better ratings to the PBL curricula. This period of research was followed by a stage of dissension, during which the education research community was polarised into supporters and antagonists of the PBL curriculum (Albanese, 2000; Berkson, 1993; Cobb, 2002; Colliver, 2002; Colliver, 2000; Norman and Schmidt, 2000). The defenders of PBL advanced several reasons as to why the PBL curriculum had not shown superior outcomes in comparison to traditional curricula. These reasons included inferior assessment techniques and previous students' learning under traditional curricula. It was suggested that the PBL curriculum needed to be treated as a curriculum in its own right and should be assessed based on its promises (Albanese, 2000). It was advocated that process-oriented research was needed (Hak and Maguire, 2000; Koschmann and MacWhinney, 2001), and that the supporters of the PBL curriculum needed to keep stick to its philosophy (Dolmans et al., 2001). Since that time, research has focused on understanding the interactions in the PBL tutorials. The research presented here is a process-oriented study, exploring the joint knowledge construction processes of the students inducted in PBL tutorials.

3.3. Definition of Collaborative Learning

The cognitive effects of PBL are linked to peer interactions taking place in the small-group learning environment (Van der Linden et al 2000 Dolmans and Schmidt, 2006). According to Van der Linden (2000 cited in Dolmans and Schmidt, 2006), collaborative learning is said to occur when students share a common goal and responsibilities, are mutually dependent and need to arrive at agreements through open interactions. Collaborative learning diverges from the traditional division of labour because its essential condition is mutual interaction and a shared understanding of a problem (Dillenbourg, 1999). PBL fits this definition of collaborative learning based on its core characteristics (Dolmans and Schmidt, 2006). It involves contributing to discussions and explanations, externalising point of views, bringing prior experiences to bear on the discourse, negotiating with the differences in ideas, effective communication with group members, and knowledge-sharing among group members (Azer and Azer, 2015).

3.4. Methods of Literature Review

While carrying out this literature review, I adopted a scoping of the methodological framework. Scoping is defined as an attempt to map rapidly the key concepts underpinning a research area, the main sources, and the various types of evidence available (Mays et al. 2001 cited in Arksey and O'Malley, 2005). This approach to literature review is considered well-suited for identifying research gaps in the existing literature (Arksey and O'Malley, 2005). According to Arksey and O'Malley (2005), the following are the stages of the review process: (1) identifying the relevant questions, (2) identifying relevant studies, (3) study selection, (4) charting the data, and (5) collating, summarising, and reporting the results.

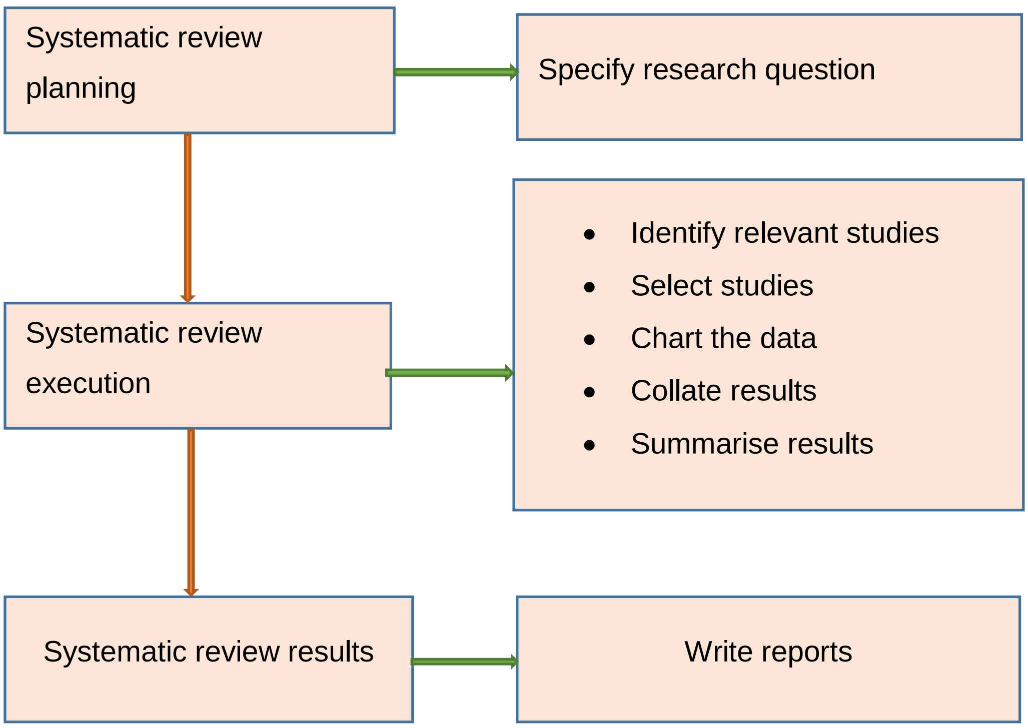


Figure 3.1: Systematic review process

Stage 1: Specify research question

How do the medical students collaboratively construct knowledge, and how do the facilitators guide the process in the PBL tutorials?

Stage 2: Identifying relevant studies

Studies relevant to the review were identified by searching electronic databases and edited books on PBL.

The electronic databases searched were PubMed, Embase, Medline, PsychoInfo, and Web of Science. These databases were searched from the year of their inception until 26 May 2016. To complement the electronic search, two edited books (Evensen and Hmelo, 2000; Hmelo-silver and Eberbach, 2012b) on PBL were searched and relevant articles were identified and retrieved. Relevant references from the retrieved articles and edited books on PBL were also retrieved and reviewed. Searches were limited to search terms relevant to the research questions, and both quantitative and qualitative studies were included.

To locate the studies in line with the focus of the review, search terms were generated and used in combination for the literature review (as described in Table 3.1). The implication of the search strategy is that studies conducted in the health-related field should only be the ones to be identified. The purpose of the review is not to retrieve all the studies that have ever been conducted on tutorial process in the health care field, but I believe, what is important is that the procedure for relevant study identification is to be rigorous enough to identify studies that would enable a general mapping out of the research on the PBL interactive processes. The results of the search are shown in Table 3.1.

Stage 3: Select studies

The criteria used for including studies in the review included the following: (1) studies that describe students' collaboration or interaction in the PBL tutorial, (2) studies that addressed knowledge construction as a process variable, and (3) studies that addressed the role of the facilitator in group functioning. Quantitative and qualitative studies that dealt with issues outside of the review focus were excluded. Such studies include letters, personal opinions, tutor quality and training reports, assessments in PBL, review articles, conference abstracts, curricular comparison, and PBL long-term assessment articles. The review was also limited to articles published in English.

The retrieved articles were exported into the Endnote version 7.3 software. Duplicates articles were then removed. Abstracts of the articles were read and irrelevant studies were removed. Copies of the relevant articles were retrieved and read. Relevant articles from the references of the retrieved studies were also retrieved and read. Information was extracted from the articles for the data charting stage of the review process.

Table 3.1: Search terms and result output

Search Terms	PubMed	Embase	Web of science	PsychoInfo	Medline	Total
PBL AND collaboration	255	448	445	28	32	1208
PBL AND facilitation OR facilitator OR tutor	1964	1063	1412	1634	1842	7915
PBL AND interaction	365	45	337	54	53	854
PBL AND group process	1237	215	1219	204	118	2993
PBL AND knowledge AND construction	66	13	139	38	40	296
PBL AND reasoning	508	280	708	659	686	2841
PBL AND elaboration	33	63	65	66	78	305
Total database articles						16412

Stage 3: Charting the data

An Excel spreadsheet was developed to record the information extracted from the articles. The extracted information was in line with the questions raised in the stage 1 of the review process, including: (1) first author and date, (2) country of publication, (3) source of publication – journal or book, (4) type of study design, (5) study population, (6) whole or partial PBL tutorial evaluation, (7) method of data-capturing, (8) analytic theory, (9) analysis method, (10) type of process variable studied, and (11) the results of the study (Appendix 1). The categories of the extracted information were then prepared and exported into Statistical Package for Social Sciences (SPSS) version 22 to generate descriptive reports. Table 3.2 shows the characteristics of the studies reviewed.

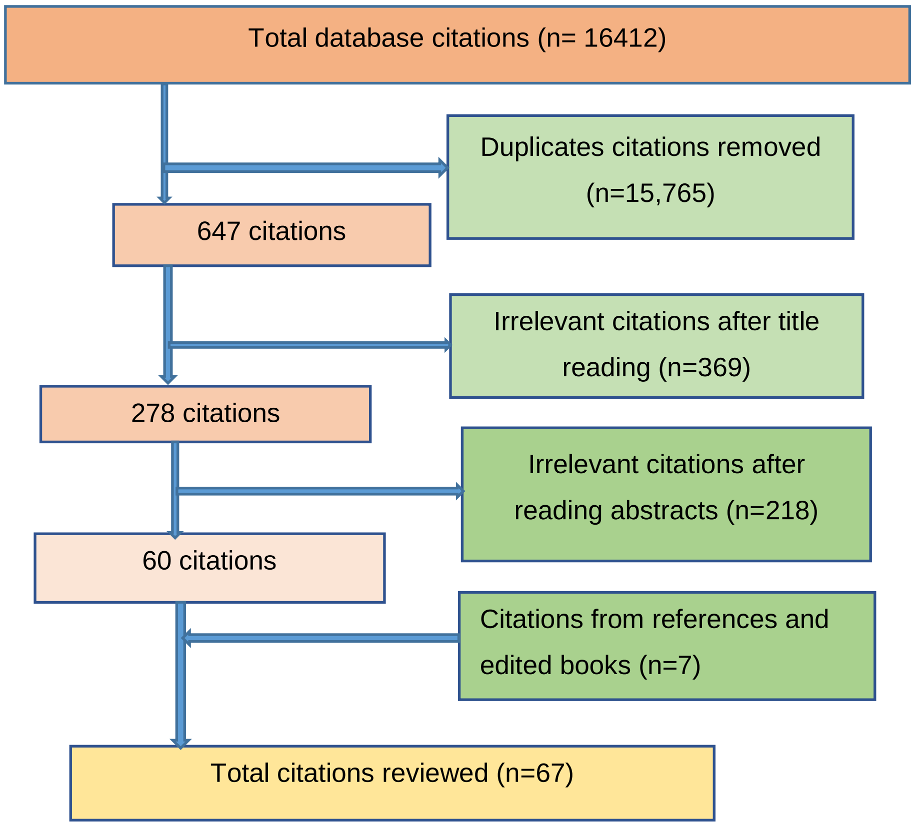


Figure 3.2: Process of citation selection

Stage 5: Collating, summarising, and reporting the results

First, the theoretical perspectives underpinning the interactive learning in PBL is first reviewed. Second, a review and evaluation of the methods that have been used to research upon the facilitation of collaborative learning in medical PBL are conducted. Third, the issue of collaborative knowledge construction is taken up and the theoretical perspective and methodological approach that inform this study are discussed.

Table 3.2: Description of the studies analysed

Category	Subcategory	Frequency
Geographical origin of publication	Asia	13
	Australia	5
	Africa	1
	Europe	29
	America (south and North)	18
	Transcontinental	1
Type of research	Observational	60
	Experimental	7
Research design	Quantitative	39
	Mixed methods	8
	Qualitative	20
Data collection methods	Participant observation	12
	Interviews	14
	Recall/stimulated recall	10
	Written text	5
	Questionnaire	37
	Video-recording	13
	Audio-recording	7
	Written report	7
Methodology of verbal data analysis	Use of coding scheme	5
	Grounded theory	8
	Grounded theory + discourse analysis	2
	Conversation analysis	1
	Discourse analysis	3
	Thematic analysis	1
	Corpus linguistics	1
Analysis technique	Statistical	42
	Software assisted coding (ATLAS.ti/NVivo)	6
	Manual coding	17
	Automated analysis (e.g. Wmatrix)	1

3.5. Results of the Literature Review

This section presents the results of the literature review and are discussed under several themes.

3.5.1 General description

Generally, more than half of the studies were published between 2001 and 2010; about one-fifth were published in the decade before, while about a quarter were published in the following decade (Figure 3.2). Now, 43% of the studies reviewed were published in Europe, while about a quarter were published in America, mainly in North America. Observational studies were the most common types of study, and they were mostly quantitative in nature. Mixed-methods studies were not commonly used (8/67). More than half of the studies used questionnaire for data collection. In nearly 10% of the studies, participant observation, interviews, stimulated recall, and video-recordings were each used. Analysis of group conversations in the studies were mostly done with grounded theory or a designed coding scheme. Only one study had used the corpus linguistics (CL) methodology. Nearly two-thirds of the studies had used quantitative statistical software package for data analysis, about a quarter used manual coding, qualitative software analysis was used in 6 studies, while only one study had used the CL software (Wmatrix).

3.5.2 Theoretical Perspectives on cognitive effects of PBL

This section describes the theoretical perspectives used in the studies for the cognitive effects of PBL.

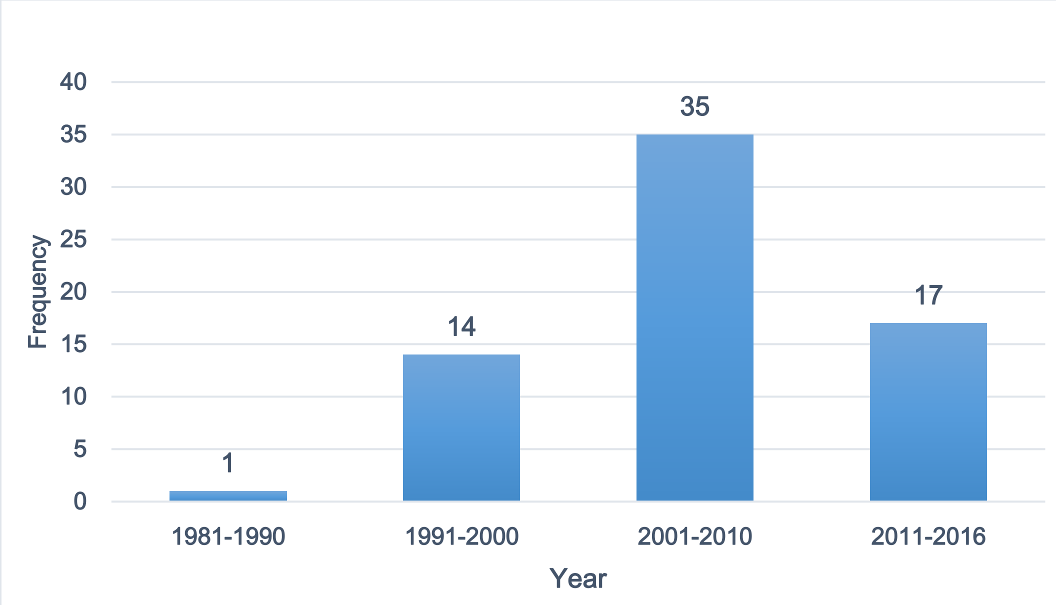


Figure 3.3: Distribution of studies by year of publication

Collaborative learning in small groups can be studied from different theoretical perspectives. For example, Slavin (1996) classified the theoretical perspectives into the following four categories: motivational, social cohesion, developmental, and cognitive elaboration perspectives. Dolmans and Schmidt (2006) combined developmental and cognitive elaboration of Slavin (1996) to discuss the cognitive effects of small group learning in the PBL setting. In this review, I have discussed the empirical evidence for collaborative learning under information processing and developmental perspectives.

Information Processing Perspective

The information processing perspective is a cognitive approach to understanding how knowledge develops in the context of human interactions. It assumes that knowledge develops as a student processes information from the environment. The process of learning and comprehension involves encoding, organising information, forming connections, and creating mental models (Martin et al., 2016).

Empirical evidence of information processing perspective. Research that employed information processing perspective to analyse data on PBL tutorial discussions in health-related educational settings have shown a consistent positive effect to have been exerted by the activation of prior knowledge, encoding specificity, and knowledge elaboration in context of student learning. In this review, pre-intervention and post-intervention tests were commonly used to assess the achievement of learning.

In a study to investigate whether or not PBL leads to conceptual change, De Grave et al. (1996) investigated the conceptual change of a group of second year (n = 5) medical students. The students were trained for the study and were experienced in the PBL process. The students underwent stimulated recall test, following a session of problem analysis. Analysis of problem analysis session data and stimulated recall test transcripts demonstrated the thinking of the students to signify the presence of conceptual change and metacognitive reasoning in the students' talk. De Grave et al. (2001) randomised 48 first-year medical students into two groups to demonstrate the effect of prior knowledge on subsequent recall. The experimental group discussed blood pressure regulation and the control group discussed a problem pertaining to vision. The two groups then read a text on blood pressure regulation. The subsequent recall test showed that the experimental group recalled 25% more than the control group could from the text read on blood pressure regulation. The De Grave et al. (1996) presented extracts from the transcripts to demonstrate the cognitive and meta-cognitive processes of the students, but the sample size was very small – a group of five medical students. On the other hand, the mechanism through which group discussion led to a superior recall for the experimental group in De Grave et al. (2001) study remains unknown.

Another aspect of information processing perspective that has commanded the attention of researchers is knowledge integration. This is because medical experts possess well-integrated biomedical and clinical knowledge that are transferable for diagnosing and resolving clinical problems (Charlin et al., 2000; Schmidt, 1993). How medical students develop this sort of integrated knowledge has been a focus of research (Diemers et al., 2015). Diemers et al. (2015) conducted a study to determine knowledge transfer in third-year medical students (n = 13). After a 10-week course with real patients, the students were asked to think aloud while diagnosing the two cases encountered and discussed during the course as well as the two cases that described the biomedical mechanisms that were taught during the course but were not seen or discussed by the students while explaining the underlying pathophysiological mechanisms of the patients’ features. They determined the diagnostic accuracy and the time taken to think about the cases. They also conducted a pre- and post-intervention test. Qualitative analysis of the transcripts of the think-aloud protocols showed that the students in the course cases used more model concepts and less wrong concepts and links compared to transfer cases. The findings suggested that an integration of the biomedical and clinical knowledge took place during the 10-week course.

In 2009, Collard et al. (2009a) published a paper that investigated maturational increase in biomedical reasoning capacity in comparison with factual knowledge among medical students drawn from the year-three to year-six PBL curricular. The students were tested on script concordance questions and through true/false questions along with knowledge ascertainment tests based on the clinical vignettes and scenarios they have used in their individual curriculum years. The results of their performance were compared with the panel results generated by the tutors. The results showed that the script concordance (SCT) scores were higher in the years 5 and 6 than in the years 3 and 4. Year 3 also showed higher SCT scores on questions in a new context. The scores of year 3 and 4 were significantly higher on the true/false tests than the ones scored by year 4, 5, and 6. For year 3 and 4, there was a positive correlation between scores on the true/false tests and SCT questions. The ascertainment scores for correct answers on true/false questions were higher than what was observed for incorrect questions. The results indicated towards the presence of biomedical reasoning, which had supposedly increased along the period of training following the PBL curriculum. This was accompanied by a decrease in factual knowledge retention. Schmidt and Boshuizen (1992), in an earlier study, found out that experts used less biomedical knowledge to explain clinical problems with increasing level of expertise. These studies are not without limitations. The study of Diemers et al. (2015) was based on a small sample size (n = 13) amounting to 16.3% of the cohort. The study by Collard et al. (2009a) was of a cross-sectional design and had compared the performance of students who differed in terms of curriculum years.

Yew et al. (2011) conducted a study in which they investigated whether or not learning in PBL tutorials is cumulative. The assumption was that the students would be building concepts networks, involving concepts from different phases of the PBL cycle. There were 218 biomedical students. A week before, students had a discussion on the problem – the structure and functions of DNA and RNA – students had an essay pre-test to assess their prior knowledge. They were required to write an essay to describe and explain the structure of DNA and RNA. Each essay was divided into idea units. A score of 2 was awarded to a completely correct idea unit; 1 was given to a partially correct idea unit, and 0 was awarded when an idea unit was completely incorrect. A concept recall procedure test was also carried out. Each concept – keyword or terminology related to DNA and RNA – was given one point. The total scores for each student was summed up after each phase of the PBL. The scores on the essay tests and scores of the concepts mentioned in the pre-test essay, after the problem analysis phase, the self-directed phase, and the reporting phase were compared. T-tests were used to compare pre- and post-test results, and one-way ANOVA was used to analyse the results at the end of each learning phase. The data was also analysed through structural equation modelling. The results of the data analysis suggested a good fit for the hypothesis that learning that occurred in the previous phase of the discussion influenced learning at the subsequent phase of the PBL cycle.

Van Blankenstein et al. (2013) carried out a controlled experimental study to investigate into the effects of elaboration during a problem-based discussion on recall for high and low ability students. The participants were 167 students from health sciences discipline. They were randomly allocated into superficial question group (control group) and elaboration question group (experimental group). All the students had to take an MCQ test to establish their prior knowledge. The students were divided into low ability group (passed the test after more than one attempt) and high ability group (passed the test at the first attempt). Subsequently, students observed a video-recorded PBL tutorial discussion. A tutor in the video asked the experimental group elaborated questions, and they were encouraged to answer as extensively as possible, while the control group students were asked superficial questions. After the tutorial discussions, the two groups read a text relating to the case discussed. Immediately after the text was read, they had to take a recall text. After one week, they returned to have a delayed recall test. The results of the experiment showed that elaborative questions had no significant effect on the recalling of idea units. High ability students outperformed low ability students, but this was independent of the experiment. The result of the experiment contradicted the theoretical benefits of elaboration in learning (Kalyuga, 2009; Oliveira and Sadler, 2008; Schmidt, 1983). This result may be due to the limitation of the study: First, the sample size was not calculated before the experiment, and so, it is possible that the sample was too small to show the effect size. Second, the experiment had low fidelity in that the PBL discussion lasted only 15 minutes, which is not the case in reality. Third, there was no room for the self-directed phase which occurs in normal PBL cycle.

Developmental Perspective

The developmental perspective of collaborative learning assumes that the interaction with the environment leads to the attainment of knowledge. Piaget et al. (1985) held that students learning during interaction with others and peer interactions could lead to cognitive conflict, which may lead to the mental model elaboration, and therefore, a growth of the existing knowledge base. According to sociocultural perspectives, learning is produced during social processes, and it is a function of task, context, and culture in which it occurs. Mastery of concepts or task is attained as leaners are scaffolded through the zone of proximal development (ZPD) by a senior or a more able peer (Hmelo-Silver and Eberbach, 2012a; Oliveira and Sadler, 2008). Vygotsky conceived the ZPD as the distance between the actual level of development, as indicated by independent performance, and the level of potential development, as determined by problem-solving, with the assistance of an adult or more capable peers (Vygotsky, 1930/1978). As the social interaction and cognitive development may not be mutually exclusive, the empirical evidence for development perspectives are discussed together.

There is a great deal of literature on the benefit of students' interactions on cognitive development in the PBL setting. Moust et al. (1987) investigated the effect of active participation of first-year students in problem-based tutorial on their learning. It was anticipated that the students who elaborated more during the group discussion would score more on the subsequent tests. The students were randomly assigned to four groups of about six students each. A problem was discussed by the groups and the discussions were tape-recorded. Thereafter, the students read the problem-relevant text and took a free recall and cued recall test. The results of the two tests did not show any significant difference among the students. The conclusion was that students who spoke less elaborated as much as the students who were verbally more active. Subsequent studies have shown that silent students learn better in PBL tutorials (Jin, 2014; Remedios et al., 2008).

The results of the studies focusing on information processing suggested that group discussion stimulated activation of prior knowledge, knowledge elaboration, information recall, and conceptual change.

Visschers-Pleijers et al. (2004) analysed the verbal interactions of students in three tutorial sessions to determine the presence of elaboration and co-construction in their conversations using a coding system developed by Van Boxtel et al. (2000). The group discussions were videotaped and transcribed verbatim. Elaboration was defined as a stretch of talk involving one student, and co-construction was defined as a stretch of talk involving more than one students. Elaboration and co-construction could occur when students asked questions, reasoned about, or solved knowledge conflicts. The results showed evidence of elaboration and co-construction in the tutorial discussions, indicating that collaborative knowledge construction occurred in the PBL tutorial investigated. However only 2 or 3 students were verbally involved.

Hmelo-Silver and Barrows (2008) investigated how a group of five second-year medical students built knowledge under the guidance of a master facilitator. The tutorial that lasted 5 hours in two sessions was videotaped and transcribed verbatim. The transcripts were analysed at coarse-grained and fine-grained levels. The results of the analysis showed that extracts of the discussions demonstrated how facilitators asked open-ended questions, and how students built on each other's knowledge in the process of joint knowledge construction. The result of the study provided evidence of collaborative knowledge construction in the PBL tutorial group. However, the weakness of this study is that it studied only one PBL group facilitated by an expert facilitator – Professor Barrows.

Visschers-Pleijers et al. (2006a) videotaped four tutorial group sessions of year-two medical students to investigate how much time was spent on each type of interaction, and how the interaction types were distributed over the reporting phase meeting of the PBL cycle. A coding system was used to analyse the transcripts of the verbal interactions. The results showed that learning-oriented discourse accounted for 80% of the interactions: cumulative reasoning accounted for 63%, exploratory questioning for 10%, and conflict-handing for 7%.

Exploratory questioning and cumulative reasoning occurred throughout the meeting, while conflict-handling occurred after the first 20 minutes. The conclusion was that even though the students devoted their time to task-oriented discussions, exploratory questioning and knowledge-related conflicts were less common in the group studied. In this study, the nature and rate of participation of the facilitator are unknown. The study only analysed one session of four groups, consisting of 32–40 (10.7–13.3%) students out of a total of 300 students. It is unclear whether the results could be generalised to other groups.

Gukas et al. (2010) conducted a study to investigate verbal and non-verbal indices of learning in a medical student curriculum in the UK. The study participants were five (38%) out of the 13 tutors and 50 out (38.5%) of the 130 students. The group discussions were observed by the tutors. A coding system consisting of exploratory questioning, cumulative reasoning, and handling of knowledge-related conflicts was used to rate the interactions. The facilitators also rated the threshold for the interventions. The results showed that when the interactions involved exploratory questions and cumulative reasoning, students tend to score high on learning and the threshold for intervention was high. When the interactions involved handling of conflicts, the students tend to score high on learning and the threshold for intervention was low. The conclusion was that the interactions suggest that learning occurred during collaborative interactions. The weakness of the study lied in the small sample size, which limited its generalisability potential and precluded the use of sophisticated statistical techniques for data analysis.

Visschers-Pleijers et al. (2005b) conducted a study to investigate the students' views on the presence and desirability of types of learning-oriented interactions in 22 PBL tutorial groups, using a validated coding system. The results showed that the scores for the occurrences of learning oriented interactions (i.e., exploratory questioning, cumulative reasoning, and handling of conflicts) were reasonable; however, the desirability scores were significantly higher for exploratory questioning and cumulative reasoning, suggesting that students wanted improvement in the group interactions. The limitation of the study was that the data was based on subjective perceptions of the students only and the opinions of the facilitators were not a part of the data.

Romito and Eckert (2011) investigated the relationship between dental students' biomedical knowledge acquisition and the students' level of PBL group interactions. Biomedical knowledge recall and application about the case were assessed with quizzes. Students and facilitators completed questionnaire to assess the students' group interactions. There were 92 (92% of the cohort) year-one students and 99 (97% of the cohort) year-two students. Eleven (68%) year-one, year-six (35%), and year-two facilitators participated. The results of the investigation showed that the year-one participants had significantly higher assessment scores. For both groups of students, the mean score for a recall test was higher than what was obtained in the tests, testing knowledge application. There was a correlation between the recall score and the interaction role for year-one cohort, but no relationship was established between recall score and interaction role for the year-two cohort. There was no correlation between the students' role and score in the application test. The findings suggest that the students' level of group involvement did not significantly affect their assessment performance.

Visschers-Pleijers et al. (2006b) surveyed the opinions of year-one (n = 23) and year-two (n = 25) medical students from 6 focus groups pertaining to the characteristics of an effective discussion during the reporting phase of PBL. The interviews were audiotaped and transcribed verbatim. In the results, the students felt that effective discussions in the PBL tutorials should involve asking for, giving, and receiving of explanations; integrating and applying knowledge; discussing differences with respect to the learning content; and guiding and monitoring content and group process. The learning effects of such discussions mentioned by the students included retention, understanding, as well as knowledge integration and application. The generalisability of the result is, however, limited because of the small sample size.

Another study was designed by Visschers-Pleijers et al. (2005a) to develop and validate a questionnaire to assess the quality of learning-oriented interactions in PBL tutorials. The instrument was rated by 240 year-two medical students. The questionnaire looked at three types of learning-oriented interactions: exploratory questioning, cumulative reasoning, and the handling of learning-related conflicts.

Analysis of the data by regression statistics showed that exploratory questions and cumulative reasoning factors together explained 26% of the variance of the tutorial group's productivity. The data for the study was based only on the students' ratings, thus leaving out the rating of the facilitator – a very crucial factor for the effectiveness of tutorial discussion. However, two studies provided further empirical evidence for the collaborative conception underpinning the PBL instructional approach.

In an exploratory study by Da Silva and Dennick (2010), a full PBL cycle, consisting of three sessions, was analysed with a powerful language analytic online software – Wmatrix. The participants comprised seven medical students and one facilitator. Results of the analysis showed evidence of reasoning, explaining, and questioning in the students' discussion. A small sample size was a clear limitation faced by the study.

Imafuku et al. (2014) studied knowledge construction and interactive experiences of three PBL groups in which students from nursing, medicine, pharmacy, and occupational health departments participated. The data that consisted of transcribed videotapes of tutorial sessions and e-portfolio transcripts were manually analysed, using functional discourse analysis and grounded theory techniques. Results of the data analyses showed evidence of elaboration and co-construction of knowledge in the students' discourse.

3.5.3 Use of Cognitive Tool

De Leng and Gijlers (2015) conducted a study to investigate how a collective drawing of diagrams, using computerised mapping diagram, affects the discussion and knowledge construction during the reporting phase in PBL undertaken by four tutorial groups. The tutorial discussions were videotaped and transcribed, students completed the perception questionnaire, and the perceptions of the tutors were sought in two, more interactive, focus group interviews that involved stimulated recall from video segments. The focus group interviews were audiotaped and transcribed verbatim. The coding system for social mode of knowledge construction by Fischer et al. (2002) was used to analyse the tutorial talk. The results showed that all the tutors unanimously felt that diagram drawing enhanced knowledge elaboration, promoted peer interaction, fostered focus and discussion depth, and helped promote shared understanding. The limitation of the study was the small number of the tutors involved and the self-report nature of the data.

Veronese et al. (2013) conducted a randomised pilot study to investigate the feasibility, acceptability, and effectiveness of concept maps in a PBL tutorial comprising first-year medical students. The students were randomised into a concept map group and a no-concept map group. Subsequently, students were surveyed through questionnaire, following which the tutors were interviewed and the students' examination scores were recorded. A mixed-method approach (qualitative and quantitative) was used to analyse the data. Qualitative data showed that the use of concept diagram was associated with knowledge integration, challenging students' knowledge of the material, and knowledge gap identification. Quantitative analysis showed that students in the concept map group performed better than the ones in the no-concept map group. These two studies showed that cognitive tool, in form of concept maps, fostered medical students' learning. Koschmann et al. (1997) in a descriptive study analysed video transcripts of segments of a PBL group discussion performed by medical students, using the conversation analysis approach. The results of the study showed evidence of questioning, idea negotiation, collaborative knowledge-building, as well as tutor scaffolding in the tutorial transcript. The generalisability of the result was limited because the segments of the analysed discourse were about 2 to 5 minutes in duration and only one group was analysed.

Hurk et al. (2001) conducted a study to investigate the quantity and quality of learning issues generated by the two groups of first-year medical students. The learning issues were collected by one of the students in each group during phase one of the PBL cycle. The learning issues were then scored according to a coding system (Van den Hurk et al., 1998 cited in Hurk et al., 2001) by 12 year-two medical students, who had just completed the cases discussed in the tutorials. The results showed that 21% of all learning issues for one problem and 32% for the other problem scored high on all characteristics, while 5% and 2% of all learning issues scored low on all characteristics. Most of the learning issues were formulated ambiguously and lengthily. The results of the study may be due to the immaturity of year-one students with regard to the PBL process, and it must also be taken into account that there was no evidence that the rating of the assessors were validated by the tutors.

In the study of Yew and Schmidt (2009) to investigate the presence of constructive, self-regulatory, and collaborative processes in the PBL, verbatim transcript of talk in one tutorial group was analysed, using a coding scheme of Van Boxtel et al. (2000). Data analysis showed that learning activities relating to collaborative process amounted to 53.3%, self-directed learning amounted to 27.2%, and knowledge construction amounted to 15.7%. The generalisability of the result is rather limited because only one PBL group was analysed.

3.5.4 Facilitation of the PBL process

The PBL tutor plays a central role in PBL tutorial discussions (Schmidt et al., 2011), as they scaffold learning of the students and model problem-solving, good learning strategies, and thinking, rather than providing content knowledge (Dolmans and Ginns, 2005; Hmelo-Silver and Eberbach, 2012a; Lycke, 2002; Remedios et al., 2008). PBL tutors also scaffold knowledge construction and co-construction (Chng et al., 2011). The role of the facilitator in the PBL environment has attracted rigorous research and a large amount of literature is now available discussing the role of the facilitator on the PBL process variables. The literature review showed that there is no difference in the effects exerted by tutor expertise on the PBL tutorial process. Some studies showed that content experts were superior to non-content experts, some showed that non-content experts were superior to content experts, whereas the effects were mixed in some studies, while in some other studies, the levels of expertise did not influence the PBL tutorial process.

Effects of levels of expertise

Eagle et al. (1992) studied the influence of the PBL tutors' levels of expertise on the number of learning issues generated by the students and on the amount of time devoted to self-study, by analysing students' reports and tutors' self-rating of levels of expertise. They found that expert groups generated more learning issues and spent more time on self-study. However, non-expert tutors with facilitation experience and the ones who studied for the case before the tutorial achieved results similar to those attained by expert tutors. Similarly, Schmidt et al. (1993) found out that students tutored by content experts spent significantly more time on self-study than students tutored by non-content experts. They also found that content experts used a combination of subject matter and process skills to facilitate learning. Couto et al. (2015) investigated the perceptions of medical students on the influence of expertise on their PBL tutorial process. The students felt that content experts were statistically superior to non-content experts in terms of helping students to construct knowledge, guiding the learning processes, generating learning issues, and motivating self-study.

On the other hand, Silver and Wilkerson (1991) investigated the effects on content expertise on the tutorial process in a medical PBL curriculum. Data analysis was based on audiotaped tutorial interaction and tutor self-rating of the expertise level. The results showed that content experts were more directive and dominated tutorial talk. In a prospective randomised trial, Peets et al. (2010) investigated the effects of content expertise and process expertise on the PBL tutorial process, using a student-rated questionnaire on tutor behaviours. The results showed that process experts were perceived as being significantly superior on all items present on the tutor behaviour evaluation instrument. Kaufman and Holmes (1998) investigated whether or not expert tutors differ from non-expert tutors in terms of the extent to which they dispense knowledge in a PBL tutorial, using a self-rated questionnaire. The results showed that less than 50% of the tutors said they almost never dispense knowledge. The tendency to present or explain case knowledge increases with tutor's level of expertise. The tutors who did not dispense knowledge rated PBL more highly than the tutors who dispensed knowledge. However, the students did not rate the two groups of tutors differently.

Other studies did not find any difference in the performance of the content experts and non-content experts. Regehr et al. (1995) conducted a participant observation to investigate the effects of expertise on students' learning, group process, and students' satisfaction in a medical PBL curriculum. They found no significant difference in the tutorial process variables or in the levels of students' satisfaction. Similarly, Davis et al. (1994) did not find any significant differences in the group process performance of the groups facilitated by content experts and non-content experts. The study of Gilkison

(2003) showed that content experts and non-content experts used similar techniques to raise students' awareness, facilitate group process, and direct students' learning. While content experts questioned the students, non-content experts expected the students to question each other.

Effects of background factors

Other studies have compared the effects of staff and peer facilitation on the group learning process. Again, the results have been inconclusive. For example, Schmidt et al. (1994) investigated the effects of peer and staff facilitation on the medical PBL processes. The results showed that the students perceived that staff facilitators were more knowledgeable, made more relevant contributions, and asked more stimulating questions than the peer tutors. However, year-one students felt that their social and cognitive experience aligned with that of the peer tutors, while more senior medical students felt that their social and cognitive image aligned with that of the staff tutors. A contradictory finding was noted in the study of Kassab et al. (2005). The students in their study felt that peers were superior to the faculty when it came to providing feedback and with regard to understanding their difficulties. They perceived that tutorial atmosphere, decision-making, and support for the group were better handled by peer tutors.

Curet and Mennin (2003) investigated the effects of short-term and long-term facilitators on the group learning process in a medical PBL curriculum, using student rating of tutor behaviours with a questionnaire. The students felt that long-term facilitators were superior to short-term facilitators in terms of cognitive, metacognitive, and motivational behaviours. The study of Groves et al. (2005) examined the influence of tutor background on PBL tutorial process, using student-completed questionnaire. The results showed that students perceived that clinically qualified facilitators used content for facilitation, and that they were socially more congruent than the university staff. The non-clinicians (i.e., university staff) emphasised upon assessment and established a more collaborative atmosphere. Chng et al. (2011) researched on the effect of tutor behaviours on group processes in a biomedical PBL curriculum, using a validated questionnaire. Data analysis showed that social congruence had a stronger effect on students' learning while cognitive congruence and use of subject matter expertise did not affect students' learning to the same extent. The literature on the effects of tutor background on group process seems to suggest that social congruence is very important. The inconclusiveness of the studies on the effect of tutor expertise on group process may be due to the methodological flaws in the studies, including sample size, subjective ratings of tutors' behaviours, and inconsistent definitions of expertise (Schmidt and Moust, 1995).

Effects of contextual factors.

To resolve the inconsistency in the research findings on the relative effects of the expertise level on the group processes, research effort has been devoted to exploring the contextual factors that may be playing a compensatory role, modifying the efficiency level of expertise-group process relationship. Davis et al. (1994) conducted a study in which highly focused cases were used to remove the effects of the level of tutor's expertise on students' learning and satisfaction in the medical PBL curriculum. They found that there was no difference in the effects of expertise level on students' learning and satisfaction. The findings in this study agreed with that of Schmidt (1994), which suggested that subject matter expertise could compensate for inadequate students' prior knowledge and unit structure. Similar effects of contextual factors were found by Gijselaers (1997). On the other hand, Dolmans et al. (1996) examined the effects of tutors' expertise on the students' performance in the context of varying curriculum structure and students' prior knowledge. They did not find tutor's expertise as playing a compensatory role. Dolmans and Wolfhagen (2005) investigated how tutor performance, tutorial productivity, and the effectiveness of tutorial unit interact with each other. They found that tutor performance differs across different levels of group productivity; group productivity differs across different levels of tutor performance; and that both group productivity and tutor performance affect the effectiveness of a PBL unit. These studies would suggest that tutor behaviour in PBL is influenced by complex contextual factors. The finding corroborated with the theoretical model of (Schmidt and Moust, 1995; Schmidt and Moust, 2000), which showed that the tutor's social congruence and level of expertise interact with cognitive congruence to influence tutorial group functioning. From this point, the research shifted to connecting the theoretical foundations of PBL facilitation to its practice.

Studies connecting practice to theory

A large number of studies are now available on the interactions going on in the PBL groups with the view to assess how processes in the PBL groups correspond to the intentions and theoretical perspectives of the PBL curriculum. Lycke (2002) analysed videotapes and written reports of three medical PBL tutorial discussions to understand how the processes align with the theoretical foundations of PBL. He found that, in general, the systematic nature of the processes aligns with the ideals of PBL, but that there were significant differences between the groups. In two groups, the interventions of the tutors were facilitative, while it was found to be directive in one group. The group with students who were well-experienced in the PBL processes performed very well – self-directing, negotiating ideas, explaining, and co-constructing knowledge. These learning activities were not well developed in the groups populated by inexperienced students.

Maudsley et al. (2008) carried out a telephonic interview to explore how tutors conceptualised their students' integrated learning agenda. The results showed that almost all the tutors identified a structure-function theme derived from the bio-psycho-social model. However, only half of the participants could differentiate structure-function themes from the other three themes. Only 40% of them managed to articulate on the public health-based theme adequately without confusion, difficulty, or antagonism. The conclusion was that PBL tutors tend to feel insecure when put outside of their comfort zone.

Papinczak et al. (2009) conducted a study to explore the students' perceptions of the tutors' facilitation techniques in a medical PBL curriculum. The students felt that tutors and students were confused about the role of the facilitator, there was variation in the management of sensitive issues, and that the facilitation style was inconsistent – some tutors were directive, while some were not directive. Students reported that they experienced lack of balance in the scaffolding provided with some tutors taking over the group process to dispense knowledge, and some allowing the students to go astray.

Al-Drees et al. (2015) of the college of medicine at the King Saud University, Saudi Arabia, investigated the perceptions of 167 and 108 year-one and year-two medical students respectively, regarding PBL tutorial sessions, using self-completed questionnaire. The students reported that PBL tutorial sessions helped them understand basic science concepts, increased their knowledge of basic science, and encouraged self-directed and collaborative learning. They felt that their decision-making skills improved; however, 54.5% of the respondents said that the students were not well-trained for the PBL curriculum. They reported that the students perceived that only 25.1% of the tutors were well-prepared. Their learning resources included the internet (93.1% of the respondents), lecture notes (76.7% of the respondents), and books (64.4% of the respondents).

Imafuku (2007) of the University of Sydney, Australia conducted an ethnographic case study, involving one group of medical students in their year-three PBL curriculum. Data was collected, using field notes, video-recording of the PBL tutorials, and stimulated recall (follow-up) interviews. Data analysis was done using the functional linguistic theory, popularised by Halliday (Halliday and Matthiessen, 2004). The results showed that the students' perceptions of the tutorial process differed from the real group learning process. The facilitator was directive and predominantly used initiation-response-evaluation (IRE) discourse patterns to facilitate the group interactions. Students felt that the tutor's behaviours in the tutorial session were helpful for their learning. Similar findings were recorded by Faidley et al. (2000). In their study, involving 4 PBL tutorial groups, the utility of the learning team survey (LTS) instrument and a more time-intensive observational checklist for assessing group processing data were compared. The results showed that there were contradictions between the subjective reports of the group processes and more objective observations.

Imafuku et al. (2014) of the Gifu University, Japan, in an ethnographic case study, explored the learning experiences of year-two and year-three students of health sciences. Functional discourse analysis was used to analyse the transcripts of the video recordings, and grounded theory was used to analyse students' reflective journal on learning experiences. They found elaboration, co-construction moves along with knowledge confirmation/clarification requests, and shared-understanding in the students' verbal interactions. There was also evidence of identify formation to be found in the reflection of the students. An earlier study by Imafuku (2012), in which a group of four year-one health sciences (medicine, nursing, pharmacy) students were studied, showed presence of idea negotiation, co-construction, and elaboration in the tutorial transcripts.

The findings in these studies forward a mixed picture in terms of the connection between theory and practice. Although the studies have methodological limitations, such as small sample size and subjective rating of tutors' behaviours, they suggest that there is a gap between theory and practice in context of the PBL curriculum.

Behaviours of effective tutors

This represents another area of research with respect to facilitation and group process. In a cross-sectional survey, conducted through student self-completed questionnaire, on tutor behaviours, AlHaqwi (2014) reported that 75% felt that the tutor figure was essential, and only 58% felt that the role of the tutor was clear and well-defined.

Boelens et al. (2015) studied the perceptions of medical students in the PBL curriculum to examine which tutor tasks were considered most important during tutorial sessions. Data analysis with statistical modelling showed that stimulation of active self-directed learning and case quality affected group-functioning, while stimulation of collaborative learning did not affect the same.

De Grave et al. (1999), in a confirmatory study to define the profile of an effective PBL tutor, found that elaboration, directing learning processes, integration of knowledge, stimulation of interactions, and individual accountability were statistically significant.

Das et al. (2002) conducted a cross-sectional study to investigate how tutors and students perceive tutors' skills in PBL tutorials, using student- and tutor-completed questionnaire. They reported that students rated the tutors as having average to outstanding skills overall; however, students' and tutors' views on tutors' behaviours diverged: students expected more support from tutors; but tutors emphasised upon self-directed learning. There was a suggestion that the opposing differences may be related to the cultural differences that prevailed between the tutors and the students.

Dolmans and Ginns (2005) carried out a study in which a tutor behaviour–rating instrument was designed and validated. The results showed that an effective tutor is one who supports active, contextual, and constructive learning; fosters self-directed learning and promotes collaboration among the students; is aware of his or her own limitation and is very passionate about the tutor role. The review suggests that even though the behaviours of the PBL tutors are known in theory, in reality, this knowledge is not well-shared among the students and the tutors.

Effects of tutor role on students' learning behaviours

The relationship between the behaviours of the tutors and group-functioning has also been a focus of research. Van Berkel and Dolmans (2006) investigated the effects of tutors' competencies on students' learning, group functioning, and achievements at the reporting phase of the medical PBL cycle. Data analysis with statistical modelling showed that stimulation of active and constructive learning, self-directed learning, and collaborative learning by tutors enhanced the quality of the problems and that of group-functioning. The quality of the problems promoted group functioning, which, in turn, had a positive effect on the students' achievements.

Reznich and Werner (2004) investigated the facilitator's encouragement or lack of encouragement of the use of internet on students' use of online information resources. The results of data analysis showed that students who used online information resources rated their facilitators as being more encouraging, while students who did not use online information resources gave their tutor lower ratings.

Yoshioka et al. (2005) carried out a randomised study to investigate the effects of an intervention to facilitate case-based problem finding in a medical PBL curriculum. The intervention group comprised 89 year-one medical students, and the no intervention group comprised 95 year-two medical students. The intervention group received problem-finding lectures, encouragement on problem generation, and nonverbal reinforcement. The two groups were compared on the basis of the total number and categories of problems extracted from the case. The intervention group generated significantly more problems than the no intervention group. The intervention group also generated more questions in a greater number of specified categories.

Chng et al. (2015) carried out a survey-based study to determine the extent to which tutors' behaviours influenced students' learning in a biomedical science curriculum. Data analysis showed that tutors' behaviours affected recall after a problem analysis phase but not after a self-directed learning phase and reporting phase. Tutors' behaviours also affected average students but not stronger and weaker students. Three of these studies showed positive effects as a result of tutors' behaviours on group functioning and students' learning behaviours, and only one showed mixed results. Taken together, tutor behaviours exert positive effects on group functioning and students' learning behaviours.

Characteristics of productive group interactions

In a study carried out at the University of Maastricht medical school in the Netherland, Visschers-Pleijers et al. (2006a) audiotaped and transcribed verbatim four tutorial groups of year-two medical students at the reporting phase of the PBL cycle. Data was analysed using an adaptation of the Van Boxtel et al. (2000) coding system. The tutorial interactions were characterised as being effective, and they were marked by explanatory discussions, knowledge integration and application, different learning content discussion and guiding, and monitoring of the content and process.

The same group of researchers (Visschers-Pleijers et al., 2006b) at the Maastricht medical school audiotaped six focus group interviews, involving 48 year-one and year-two medical students. The audiotapes were transcribed verbatim and the transcripts were analysed with a computer software package (ATLAS.ti version 4.1). The four primary themes that defined effective tutorial interactions were asking for, giving, and receiving explanations; integrating and applying knowledge;

discussing different opinions and perspectives with regard to learning content; and guiding and monitoring the discussion.

Visschers-Pleijers et al. (2005a) carried out a study at the Maastricht Medical School, in the Netherland to develop and validate a questionnaire to identify learning-oriented interactions in the PBL tutorial. This rating saw participation from 242 medical students. Regression analysis showed that exploratory questions and cumulative reasoning factors explained 26% of the variance of the tutorial group's productivity.

Factors affecting learning in PBL group

Students' interactions are a key element of the PBL tutorials, promoting critical thinking processes and spurring the students into deeper learning (Azer and Azer, 2015). Engagement in collaborative discourse provides opportunities for conflict of knowledge that needs to be argued upon and negotiated with in order to achieve a deeper conceptual understanding and an increased capability in terms of scientific reasoning (Osborne, 2010). Participation in collaborative verbal interactions promotes learning in the sense that it imposes on the student, the need to translate their ideas into forms that can be communicated and comprehended by others. This process imposes great mental effort on the students, and it may reveal a problem in the student's knowledge structure that might be in need of correction (Swain, 2000). However, conflicts could arise in the tutorial group that could hinder students' learning (Azer and Azer, 2015). The role of the facilitator includes creating a supportive group atmosphere, encouraging students' participation in interactive discourse, and addressing group problems as and when they arise (Hendry et al., 2003).

Skinner et al. (2012) reported five cross-site ethnographic case studies in Australia and Ireland dental schools, using multimodal data collection techniques and grounded theory to analyse the data. The participants were first-year undergraduate dental students. They reported differences in terms of the social cohesion and functionalities of the groups. In one group, students' functions were almost individualistic; in another group, students could overcome their differences and attained functionality. Additionally, one group functioned more like a cooperative group, eschewing almost all forms of disagreement; one group was very loud and noisy, and they regarded the loudness as being responsible for their functionality; while the fifth group could be considered as an ideal PBL group, with the members getting along well and operating with functionality. In all four groups, some students felt a sense of social isolation which resulted in insufficient participation. One theme that runs through all the groups is that social cohesion is related to group functioning.

Duek (2000) of the University of California, Los Angeles, USA, reported an ethnographic study, involving three groups of medical students in anatomy PBL tutorials. Students' and tutors' interview transcripts and field notes were analysed, using grounded theory approach. They found that group participation was uneven and was subject to dominance, social exclusion, and sometimes, even conflicts.

There were discussion dominators, hyper-contributors, and withdrawing students. But the tutors felt that the groups were doing well generally, only that some students talked more than the others, and they never saw this as a problem.

A study by Imafuku (2012), in which a group of four year-one health sciences (medicine, nursing, pharmacy) students were studied over a year with video-recording of tutorial discussions and post-PBL discussion interviews on learning experiences, showed that a number of factors influenced students' participation and learning in PBL tutorials. These include prior learning experiences and anxiety about peer perception of communication and ideas; identification as a marker of professional motivation; students' perceptions on learning in a PBL tutorial; and social relationship with peers along with positioning within the group.

Jin (2014), in an ethnographic study of one group of year-one dental students in PBL curriculum, analysed questionnaires, interview transcripts, field notes, and transcripts of tutorial spoken discourse, using a combination of grounded theory and critical discourse analysis. He found that silence performed specific roles in group communication and learning. Silence was seen not as a verbal disengagement but as a productive resource, a collaborative practice, a platform for dealing with conflicting understanding and a signal of shifting power relations. In an earlier study, Moust et al. (1986) had observed that contrary to popular notion, silence may be useful in the PBL tutorial, as it may indicate covert learning and knowledge elaboration.

Remedios et al. (2008), in an ethnographic study explored the reasons for silent participation in a Physiotherapy PBL tutorial discussion. They found that silent participation might take place due to several reasons, including learning style, motivation, inadequate preparation, cultural literacy, concern for face and group dynamics. They concluded that silence did not preclude learning.

Lee et al. (2009) of the Catholic University in Taiwan analysed the transcripts of video-stimulated recall and field notes of data collected in ethnographic studies of health sciences PBL tutorials, using grounded theory approach. The objective of the study was to explore when and how a tutor intervenes in the PBL tutorials. There were 366 episodes of tutors' interventions. The themes identified were conflicts relating to tutorial group process and quality of discussion, including quality and quantity of the materials discussed.

(Ahmed, 2014) conducted a cross-sectional survey to explore the perceptions of medical students and their tutors on the frequency and effect of conflicts in PBL tutorials. The respondents reported dominancy, personality clash, quiet students, insufficient commitment, lateness, and absenteeism as the leading problems.

Aarnio et al. (2013) explored types of knowledge conflicts and methods of dealing with them in medical and dental students PBL groups, by coding conflict episodes in tutorial interaction video-recordings. They detected 43 conflict episodes, accounting for 7.6% of tutorial time. Factual conflict was 58%, and conceptual conflict was 42%. They reported that conceptual conflicts tend to last longer and ended up getting elaborated, but elaboration was induced individually rather than collaboratively. In a later study, Aarnio et al. (2014) explored how PBL facilitators helped students resolve conflicts. They found that 43 out of 92 tutor interventions were conflict episodes. Tutors intervened in 24 episodes and were able to resolve 13 conflicts. The methods they used to resolve the conflicts were explanation, questions, and confirmation. They reported that tutors often resolve factual conflicts but rarely resolve conceptual conflicts.

Kindler et al. (2009) interviewed medical PBL tutors to define the categories of difficult incidents and interventions that skilled tutors used in response and to determine the effectiveness of the response. They used thematic method to analyse the interview transcripts. The incidents were divided into individual student incident and group dynamics incident. The responses were categorised as feedback in the tutorial, feedback outside the tutorial, and student or group interventions. Interventions were effective with individual problems, but group dynamics problems were difficult to resolve.

3.5.5 Methodological approach in group process studies

In this section, the methodological approaches that have been used to study PBL process in the reviewed literature are discussed.

Data collection methods

Most of the studies used Questionnaire (e.g. Boelens et al., 2015; Jin, 2014; Van Berkel and Dolmans, 2006; Veronese et al., 2013), interviews (e.g. Duek, 2000; Jin, 2014; Remedios et al., 2008;

Skinner et al., 2012), research participant written test/report (e.g. Eagle et al., 1992; Papinczak et al., 2009; Schmidt et al., 1993; Silver and Wilkerson, 1991), and stimulated/cued recall interviews (Chng et al., 2011; Lee et al., 2009; Remedios et al., 2008) to indirectly collect research data. These indirect methods of data capture may be more practical, quick to carry out, and might provide opportunity to collect large amounts of data. They equally suffer from the limitation of only revealing the subjective opinions of the respondent, which may diver from reality. For example, Imafuku (2007) carried out an ethnographic case study in Australia to investigate students and tutor participation in tutorial discussion. She reported that the tutor was very directive in the tutorial, and that the initiate-response-evaluate discourse pattern dominated the discussion. However, in post-tutorial interviews, the students felt that the tutor role conformed to the PBL ideal and felt that the tutor's behaviour was very helpful for their learning. Another ethnographic study by Duek (2000) showed similar findings. The study was conducted to explore how medical students participated in PBL discussions. The participant observer observed inequality in terms of group participation with evidence of dominance, social exclusion, and sometimes, conflicts. However, in the follow-up interviews with the tutors, they felt that the students were doing well, only that some of them talked too much which, they did not consider as being a serious problem.

Other studies employed direct methods for data collection, such as participant observation (e.g. Davis et al., 1994; Imafuku, 2012;

Skinner et al., 2012), video recording (Aarnio et al., 2013, 2014; Da Silva and Dennick, 2010; Koschmann and Evensen, 2000; Koschmann et al., 1997; Lee et al., 2009), and audio recording (e.g. Diemers et al., 2015; Duek, 2000; Visschers-Pleijers et al., 2006b; Yew and Schmidt, 2009). Direct data collection methods are excellent for data collection in context of complex situations and have high reliability and validity. However, they can be expensive in terms of time and finance. They can generate a large amount of data that may be time consuming to analyse. Thus, most studies tend to have a small sample size.

Methodology of verbal data analysis

PBL is an interactive method, and language is the main tool of carrying out the process of collaborative learning. In PBL curriculum, students talk to each other for 3 or 4 hours over a PBL cycle (Da Silva and Dennick, 2010). This could generate large amounts of verbal data to provide opportunities for understanding the processes going on in the tutorial. In the move to comprehend the full impact of the PBL discussions and the extent to which learning is taking place in the group interactions, research effort has been focused on opening the “black box” to ascertain the interactions that take place during tutorial discussions (Da Silva and Dennick, 2010; Yew and Schmidt, 2009). Several methodological approaches have been used to analyse tutorial talk.

Coding scheme. De Leng and Gijlers (2015) used a coding scheme to investigate the impact of collaborative diagramming on the basic science knowledge construction of medical students in a PBL curriculum. The coding scheme was developed by Weinberger and Fischer (2006) to analyse argumentative knowledge construction in computer-supported collaborative learning.

In 2004, Visschers-Pleijers et al. (2004) published the result of a pilot study in which a coding scheme, developed by Van Boxtel et al. (2000), was employed to analyse collaborative knowledge construction in the students' verbal interactions in a PBL setting. They concluded that the coding system was well-suited for analysing verbal interaction in PBL tutorials. The coding was done manually, which is laborious and time consuming.

Grounded Theory. Grounded theory is a qualitative research methodology, and it assumes that people have a patterned experience, and that they order and make sense of their world. The order is derived from the shared social and symbolic interactions (Charmaz, 2006; Hutchinson, 1986). A few studies have used this methodology to study PBL group conversation either alone or in combination with other methodologies. Remedios et al. (2008) analysed interview transcripts of physiotherapy students and used constructive grounded theory to explore the experience of silent participants in problem-based tutorials. Duek (2000) used grounded theory methodology to explore equity of participation in group discourse in medical PBL tutorials. Skinner et al. (2012) explored the experiences of dental students in PBL tutorials, by using grounded theory to analyse the interview transcripts.

Discourse analysis. According to Eggins and Slade (1997), discourse analysis describes and relates conversation structures to other units, levels, and structures of language. Discourse analysis (DA), developed from the work of J.R. Firth, aimed at offering functional interpretations of discourse structure, as expressions of sociocultural context. In discourse analysis, discourse is described as a level of language that is different from grammar, and it is divided into units or moves. Discourse analysis aims to describe and relate these units to grammatical units such as clause. Hmelo-Silver and Barrows (2008) analysed the discourse moves of medical students in PBL tutorials to investigate how they construct medical knowledge under the guidance of an expert facilitator. Visschers-Pleijers et al. (2004) analysed discourse moves of medical students in PBL tutorial discussions, using a coding system.

Critical discourse analysis. In critical discourse analysis, language is analysed from the position of how social structures influence language use, and how this use affects social structures (Fairclough, 2003). In other words, critical discourse analysis focuses on the relationship between language, ideology, and power, and the link between discourse and sociocultural change (Fairclough 1992 cited in Eggins and Slade, 1997; Fairclough, 2003). According to Eggins and Slade (1997), it is in the conversational interactions that ideologies are operationalised and transmitted, and the goal of critical discourse analysis is to denaturalise discourse in order to expose the hidden ideology (1997).

Jin (2014), in a study to explore the experience of silent participants in PBL tutorials, used critical discourse analysis to analyse transcripts of tutorial discussions and participants' interviews. He concluded that silence performs specific roles in group communication and learning. He reported that silence was seen as a signal of shifting power relations.

Conversation analysis. Conversation analysis refers to the methods that speakers in conversation relationship use to organise their conversation in the form of a turn-taking system (Eggins and Slade, 1997). Koschmann et al. (1997) used the conversation analysis method to examine the emergence of learning issues in medical problem-based tutorial discussions.

Thematic analysis. Thematic analysis is a method used to identify, analyse, and report patterns called themes within textual data (Braun and Clarke, 2006). In a study to investigate the categories of difficult incidents as well as the interventions that facilitators used to resolve them, Kindler et al. (2009) used thematic analysis semi-structured interviews with PBL facilitators in a medical curriculum.

Multi-methodology. Other studies have analysed PBL-related verbal data, by using more than one methods (Imafuku, 2012; Imafuku et al., 2014; Jin, 2014).

Corpus linguistics. This is a methodology for analysing machine-readable texts, which is considered an appropriate basis on which to study a specific set of research questions (McEnery and Hardie, 2012; McEnery and Wilson, 2001). Da Silva and Dennick (2010) in an exploratory study used CL methodology to analyse transcripts from three sessions of medical students' tutorial discussions. Da Silva (2013), in an educational and transcultural comparative study, used CL methodology to investigate the development of clinical reasoning among medical students.

3.5.6 Techniques of data analysis

The most common technique for verbal data analysis is manual coding. This is error-prone, labour–intensive, and time-consuming, and thus, the studies that used them tend to have a small sample size (e.g. Duek, 2000; Gilkison, 2003). Other studies have used software-assisted coding techniques. For example, Jin (2014) used NVivo to analyse the experience of silent participants in PPBL tutorials, while Visschers-Pleijers et al. (2006a) used ATLAS.ti for their data analysis. Kamin et al. (2001) measured critical thinking in PBL discourse by coding and analysing conversation transcripts, using QSR NUD*IST software. These software tools, though useful for their purposes, are not suitable for linguistic analysis.

Wmatrix 3 is an online-based software. It has the capacity to automatically tag and process large amounts of text data and output results that could make allowance for a powerful linguistic analysis of the transcribed PBL tutorial data. The use of this software has been limited to a group of researchers in PBL (Da Silva, 2013; Da Silva and Dennick, 2010).

3.6. Identifying Gap in the Literature

Azer and Azer (2015) in a systematic review of group interactions in PBL observed that most studies did not explore concepts and theories behind group interactions. He suggested that there is a need for studies that would develop models of group interactions for studying group interactions and establish a link between theory and practice. Howley et al. (2013) observed that studies on the language of collaboration in science are not based on theory of language, but that they apply theory of science to analyse language interaction. They defined linguistic analysis as the type of exploration that employs constructs from the field of linguistics for the purpose of studying language. Although systemic functional linguistics (SFL) has been employed in the study of PBL group discussion in the form of discourse analysis and critical discourse analysis, the potential of the lexicogrammatical dimension of functional theory has not been fully explored as of yet. The studies of Da Silva (2013) and Da Silva and Dennick (2010) examined lexico-grammatical elements in the talk of the PBL students, but these studies investigated clinical reasoning development over time and not how the students construct knowledge together. Therefore, in the study described in this thesis, the medical students' collaborative knowledge construction in the PBL tutorials will be investigated following the language-based conceptual framework, using corpus linguistics methodology and the lexico-grammatical approach for data analysis.

3.7. Language and Learning: A theoretical consideration

In this section, a language-based theory of learning is presented. This is followed by the role of dialogue in collaborative knowledge construction. An overview of the integration of lexico-grammatical method and corpus linguistics methodology is outlined.

Language development is genetic and social in origin (Halliday, 1978; Halliday, 1993; Wells, 1994), and it can be used in two modes for educational development, which are: (1) monologic mode and (2) dialogic mode. The monologic mode focuses on enculturation, while dialogic mode concerns co-construction of meaning (Wells, 2007).

The human genetic propensity to enculturation and building of cumulative cultural resources is shown early in life as children orient themselves to other humans, engage in dyadic interactions, share emotions, show interest in the environment, and engage in inter-subjective sharing (Tomasello, 1999; Tomasello et al., 2005; Wells, 2007). The inter-subjectivity may be primary, referring to the interactions between the child and the caregiver, or it may be secondary, wherein the attentions of the child and the caregiver are directed towards an object (Trevarthen 1979 cited in Wells, 2007). The object of attention has material and symbolic functions, and through the use of these signs, enculturation into the sense-making practices of the community is facilitated.

The construction of linguistic meaning potential, organised in terms of the interrelation of semantics and lexico-grammar, provides the foundation for communicating information (Halliday, 1975). Wells (2007) considers that both Vygotsky and Halliday emphasised that in learning language, the child simultaneously encounters and appropriates the culture's way of making meaning of human experience, since this is encoded in the utterances that accompany a joint activity. This is the culture's way of making meaning, and it becomes the resource for intrapersonal and interpersonal thinking.

By thinking, those types of mental activity are being referred to that are made possible through the mediation of sign, outwardly in interaction with others, or through the medium of the inner sign. The signs are used to interpret events, both external and internal, along the norms of the society, simultaneously as communication is made with others through the same signs (Wells, 2007). Learning the meanings that correspond to the words and grammatical structures of a language also involves learning the concepts that are thereby encoded. Thinking is shaped by the increasing range of signs that become available to an individual and through the appropriation of the sign systems of a community to which an individual belongs (Vygotsky, 1930/1978). The grammatical functions and the word meanings develop during interaction with more mature speakers about their shared situations. The linguistic meanings facilitate reference to particular objects as tokens of more general classes. The connection of the word to thought is vital because without thought, the word is dead.

The meanings that mediate individual's thinking are those that are appropriated from the sign functions of artefacts that mediate the wide range of activities in which people engage together (Wells, 2007;

Wells and Arauz, 2006). According to Vygotsky, signs are transformed as they become part of an individual's resources, in the context of the activity in which they are encountered as well as in relation to the individual's past experiences. The meaning of words or signs do not remain constant for individual persons, but they develop as they are encountered in new contexts of activity, and as connections of various kinds are established with other meanings (Vygotsky, 1934/1986). Word meanings also differ between individuals because of the specific situations in which they are encountered and also due to the affective loading they take on, as a result. Thus, Vygotsky differentiates between the two attributes of a word: meaning and sense. The former refers to the dictionary meaning, while the latter refers to the significance a word holds for the user (Vygotsky 1987 cited in Wells, 2007). Pulhan (cited by Vygotsky 1987 in Wells, 2007) defines a word's sense as the aggregate of all the psychological facts that arise in our consciousness as a result of the word. A word's sense changes in different contexts (Wells, 2007). Vygotsky (1987: p. 249 cited in Wells, 2007) maintained that:

> "……word meaning is not constant; it changes during human development and with different modes of functioning of thought. The fact that the internal nature of word meaning changes implies that the relationship of thought to word changes as well. The meaning potential of words become the resources for engaging in thinking together with others in dialogic setting."

Wells (2007) observed that a dialogue between people makes negotiation inevitable, if the verbal interaction is to be sustained. This is achieved by the participants continually aiming for sufficient inter-subjectivity to allow for the conversation to proceed. The engagement of two or more individuals in conversation assumes the privilege of the speaker to make a speech, and the commitment of the listener to make sense of the speech. Speaker's intention may be obvious from the words spoken, or the intention may be recovered from the immediate and remote discourse context, from nonverbal cues, or from deliberate further clarification (Voloshinov, 1973). According to Voloshinov (1973), besides determining speaker's intention, the participants also need to determine the stance of the speaker. This information informs the position to be taken in response, and this situation is described by Voloshinov (1973) as "a bridge thrown between myself and another. If one end of the bridge depends on me, then the other end depends on my addressee" (1973: p. 86). The words function in a similar manner at the early period of a child's language development. When interacting with a child, the adult provides a bridge that the child is invited to cross in order to enter into the system of shared meanings that enables a group of people to function as a society (Wells, 2007). In appropriating these signs, the child is able to construct "meaning potential" (Halliday, 1975). The context of interaction that accompanies joint actions provide the infants with the opportunity to encounter and learn the sign system of the community's language and, in the process, the child takes over the community's ways of being in, and making sense of, everyday actions and events, in terms of the culture's knowledge and values, as these are represented in the linguistic signs that mediate their interactions.

According to Hassan (2002), the nature of the interactions between the child and the caregiver and the different ways of joint meaning-making inculcate different mental dispositions in the child's assumptions, regarding the focus of attention and material; and linguistic actions considered important in this regard. Additionally, individual life experiences also result in differences that particular words and signs have for the children. Through the "ratchet effect" mechanism, all individuals get conditioned into the knowledge and the practices of the society in which they grow up in (Tomasello, 1999). According to Tomasello (1999), the knowledge and the practices of the society refers to the ways of acting, thinking, valuing, and communicating thoughts and feelings about experiences. These features constitute the cultural resources that the children need in order to master meaningful functioning in their community. Similarly, in an educational setting, monologic direct instruction enables the students to take over knowledge and skills from the previous generations of professionals (Wells, 1998, 1999). However, there is a problem: monologic direct instruction is limited in the sense that it does not permit dialogue to clarify misconceptions of the learner, prevent the learner from offering alternative perspectives on the discourse topic; and further, it does not allow the learners to question the existing norms and propose new ideas with the aim of extending and improving the received knowledge and practices in order to adapt them to the present and future use. To achieve this objective, a dialogic mode of discourse is needed. Thus, education needs a balance of both monologic and dialogic interaction between and among the members of successive generations (Wells, 1999, 2007; Wells and Arauz, 2006).

Although the notion that knowledge is constructed through dialogue has been known since the time of Socrates, Vygotsky and Bakhtin provided the rationale for dialogue as the principal ingredient for knowledge advancement within a society in their cultural historical activity theory. Bakhtin (1986) asserted that our philosophical, scientific, and artistic thought is born in the course of interaction and struggle with others' thoughts. This situation assumes that knowledge is intrinsically bound up in knowing together, which is undertaken by individuals in specific situations with particular ends in view. Thus, books are considered not as a reservoir of knowledge but as mediators of establishing joint knowledge acquisition (Wells, 2007). Knowledge defined as what is known is authoritative and is communicated monologically, whereas knowledge characterised as a dialogic construction is more positively achieved through dialogue between people who engage in collaborative problem-solving, explanation-generation, or by defining a course of action. Such collaborative knowing leads to improved group knowledge and enhancement of an individual's understanding, because it provides opportunities for comparison of perspectives and an individuals' idea reformulation enhances the group discussion (Wells, 1999). This is in line with

Vygotsky's conception that speech does not merely serve to express developed ideas, but thought is reconstructed as it is transformed into speech (Vygotsky, 1934/1986). Dialogic interaction has far-reaching effects: the teachers need to be democratic and less directive; and the students need to uphold equitable participation and need to be coherent in their contributions. The dialogic classroom behaviours may also offer feedback into the wider society in order to improve the democratic behaviours of the future citizens (Wells, 2007).

3.8. Dialogue and collaborative knowledge construction

Dialogue plays a central mediating role in knowledge construction, because it is the major means of arriving at a shared understanding (Wells, 2000). The premise for this idea suggests that knowledge is invented and reinvented between people, as they bring personal experiences and information from other sources to use when solving some particular problem, residing in the mind of experts, waiting to be transmitted to the learners (John-Steiner and Meehan, 2000; von Glasersfeld, 1989, 1995; Wells, 2007). Knowledge is, therefore, an enhanced understanding of the problem situation achieved by the participants, and it can also be defined as the representation of the understanding that is produced in the process (Wells, 2000). To understand how knowledge is built and appropriated during and through collaborative endeavours, said Wells (2000), attention needs to be paid to discourse.

Collaborative knowledge construction has certain distinctive features (Wells, 2000): First, knowledge is created and recreated in a setting of a specific activity that involves human participants and materials as well as intellectual tools employed for the mediation of the activity. Second, knowledge is created between people. This implies that knowledge construction is situated in a discourse, wherein each contribution responds to what preceded it and a further response is foreseen. This mode of discourse fulfils the principle of responsivity by which a structure of meaning is built up collaboratively over successive turns. Bakhtin (1986) asserts that all discourse is dialogic, and that the meanings of words and expressions are borrowed from the speech of others. Additionally, each word is linked to a complex chain of utterances. According to Bereiter (1994), knowledge construction dialogue is a progressive discourse, because it involves the process by which the sharing, questioning, and opinion revision leads to a new understanding that participants reckon as being superior to their own prior understanding. Such discourse is based on the following factors concerning the participants: (1) mutual understanding commitment, (2) the empirical testability commitment, (3) the expansion commitment, and (4) the openness commitment (Bereiter, 1994). The implication of the Bereiter's description is that collaborative knowledge construction involves not just knowledge-sharing and empirical verification of propositions, but it is also a commitment to expansion and openness of propositions, contributed by the participants. By commitment to expansion and openness, it is meant that classroom participants entertain alternative positions and voices. This is in contrast to the authoritative contractive positions that act to challenge or fend off the scope of dialogue by preventing alternative views or by questioning propositions.

This disposition to knowledge is based on the idea that knowledge can be improved upon. The fusion of progressive discourse and improvable object provides a partial explanation of how knowledge is collaboratively constructed in dialogue. There are other two ways by which speaking can enhance an individual's understanding, which are: (1) through producing meanings for others and (2) through producing what is said (Wells, 2000).

To produce meaning for others, according to Wells (2000), speakers have to interpret preceding contribution in terms of the information it introduces, their own stance towards the information, compare the interpretation with their own current understanding of the matter being discussed on the basis of their prior experience, and then, they must formulate a contribution that will add to the common understanding achieved in the discourse, so far, by extending, questioning, or qualifying what someone else has said. In uttering what is said, the speakers interrogate the meaning of what is said, evaluate its coherence and relevance, and begin to formulate further response.

In contributing to the knowledge construction dialogue, a speaker simultaneously adds to the structure of meaning collaboratively created and advances his understanding through the constructive and creative effort involved in saying and responding to what is said. Listening, as well, advances the understanding, because a similar constructive effort is needed to listen responsively and critically to the contributions of others. The discourse involves internalisation of the meanings created during the collaborative discourse and for making contributions to a discourse in response to the contributions of other discourse participants.

The means that collaborative knowledge construction is not limited to speech (Wells, 2000). Knowledge construction could also be carried out across time and space through a dialogue that uses writing and other visuospatial modes of representation (Wells, 2000). Intertextuality is variously defined as the juxtaposition of texts (Bloome and Egan-Robertson, 2004; Varelas and Pappas, 2006), mixing of discourses (Varelas and Pappas, 2006), making sense of texts from other contexts (Pappas et al., 2004; Varelas and Pappas, 2006) and the explicit and implicit relations that a text or utterance has to prior, contemporary and potential future texts (Bazerman, 2004), to mention a few. A text is seen as the product of textualization (Bloome and Egan-Robertson, 2004). Text is viewed in an expansive way that includes written, signed, electronic, and pictoral, equations, scientific formulae, diagrams, and charts (Bloome and Egan-Robertson, 2004; Pappas et al., 2004). It also includes oral texts, including speakers' recounting of previous events or experiences (Wells, 1990). It can also refer to a string of words, conversation, written genre, as well as the genre of social activities or events (Bloome and Egan-Robertson, 2004). These texts have been called "cultural tools" (Wells, 2000; Wertsch, 1991;

Wertsch and Toma, 1995) or “thinking devices” (Lotman 1988 cited in Pappas et al., 2004). According to Vygotsky (1930/1978), thinking and speech are intimately related – one influencing the development of the other, and these cultural tools play a significant role in this relationship (Vygotsky, 1930/1978; Wells, 1994, 2000).

The use of thinking tools is consistent with the scientific inquiry that scientists conduct, for they employ cultural tools in form of written texts as mediators, as they deal with ideas, thoughts, and reasoning of others (Goldman and Bisanz 2002 cited in Pappas et al., 2004). As students engage in knowledge construction dialogue, they make sense of the discourse by using cultural tools to make sense of their experiences. This situation offers the participants opportunities to explore and articulate their sense-making, which is often presented through intertextual links (Pappas et al., 2003).

Although graduate-entry medical students bring a wealth of prior experience to the school, and PBL tutorials are well-furnished with “thinking devices”, how these cultural tools are employed to construct knowledge and make sense of the PBL task at hand is not fully understood. Since most classroom interactions take place through linguistic transactions, the quality of students’ discourse appears to be a reasonable place to examine how collaborative learning is conducted and how cultural tools are used for knowledge construction.

3.9. Applying Lexico-grammatical and Corpus Linguistics

This section discusses how lexico-grammar and CL can be brought together for textual analysis of tutorial discussions. First, the lexico-grammar is discussed.

Lexicogrammar is one of the three dimensions of language analysis under the functional theory of language, the other two being discourse and genre analyses (Morley, 2000). In functional linguistics, register of a language is conceived as having the three following layers: register, tenor, and textual layers. The layer of ideation refers to the subject matter of the discourse such as technical vocabulary – noun, noun phrases, nominal groups, and verbs. The layer of tenor relates to grammatical elements that are used to construct relations of status, power, social contact, and solidarity among group participants (Martin and White, 2007). Examples include mood resources, such as questions, statements, demands; modal resources, such as modal verbs and adverbs; and appraisal devices. The layer of mode refers to the resources in the language for structuring knowledge or text. These devices include references, lexical cohesion, conjunctions, and ellipsis (Halliday and Hassan, 1976).

These devices are used to link language elements together to form a complex whole (Halliday and Matthiessen, 2004).

Halliday described a clause (or phrase, multi-word unit) as a construing experience – abstract or concrete. This, Halliday described, as the representation perspective of clause (Halliday and Matthiessen, 2004). The experience or ideas are woven together by using textual devices (mentioned above) to form a complex body of representation. Halliday described three ways by which conjunctions could be used to expand experience/ideas – elaboration, extension, and enhancement. Elaboration implies that one clause restates the content of the clause to which it is linked, without adding new information. In extension, the clause adds new information to the clause to which it is linked. This could be done through contrast, addition, or through an alternative connective grammatical device. In enhancement, the new clause enhances the previous clause by providing such information as reason, purpose, time, similarity, and the like.

Corpus linguistics is a methodology for analysing machine-readable texts. It involves a group of methods for studying language. The text could be spoken or written words, which are transcribed and converted to an electronically readable format. The machine-readable text, thus created, can be processed with text analysis software to produce linguistics forms that can be analysed. Further information on corpus linguistics is presented in chapter 4. The research questions are highlighted in the next section.

3.10. Research Questions

The following questions are addressed in this research:

Main Research Question: How do the graduate-entry medical students in the Derby medical school collaboratively construct knowledge, and how do the facilitators guide the process? The sub-questions that arise from the main question are, as follows:

1. What are the frequencies and functions of the commonly occurring referring expression indicators in the students' talk?

2. What are the frequencies and functions of the commonly occurring shared knowledge indicators in the students' talk?

3. What are the frequencies and functions of the commonly occurring knowledge extension indicators in the students' talk?

4. What are the frequencies and functions of the commonly occurring knowledge enhancement indicators in the students' talk?

5. What are the frequencies and functions of the questions' indicators in the facilitators' contributions?

6. What are the frequencies and functions of the commonly occurring facilitators' directive expression indicators?

7. What are the frequencies and functions of the commonly occurring probability expression indicators, evident in the facilitator's talk?

8. How can the result of the research be used to improve the PBL process?

3.11. Summary of the chapter

The review has highlighted several challenges regarding research on the PBL process, including inadequate theorisation, use of subjective data-capturing techniques, and use of error-prone and labour-intensive analytic techniques. The literature review indicates that there is little research that has been done on collaborative knowledge construction in PBL in the UK. The present study aims to fill the gap by using a combination of lexico-grammatical and corpus linguistics methodology to provide a window into seeing the knowledge construction interactions in the PBL tutorial discussions. The study design and methodology are discussed in the next chapter.

CHAPTER 4

RESEARCH METHODOLOGY

4.1. Introduction

Research methodology, as a systematic way to solve a problem, refers to the science of studying how research is to be carried out. It concerns the identification of which method, in terms of accuracy and efficiency, is most suitable for a research problem (Rajasekar et al., 2013). This chapter outlines the research-underpinning philosophy in relation to other worldviews, elaborates the research strategies, including the methodologies espoused, presents an overview of the lexicogrammatical analysis, highlights the methodological options in content as well as the electronic textual analysis and electronic analysis software. This concludes by discussing the strategies for assessing research rigor and trustworthiness.

4.2. Philosophical Worldviews

Several methodological and practical decisions are required to achieve research objectives and to answer research questions. This section presents the proposed decisions along with their philosophical and theoretical underpinnings.

The decision to adopt a research design needs to be based on a careful consideration of the questions the study aims at answering (Creswell, 2009). Several research strategies are available, including qualitative, quantitative, and mixed method strategies. These research strategies are underpinned by distinctive philosophical worldviews (Creswell, 2009).

4.2.1 The Nature of Worldview

The philosophical worldview carries several names, including paradigms, epistemologies, and ontologies, or broadly, research methodologies (Creswell, 2009). Creswell (2009: p. 6) defined worldview as "a basic set of beliefs that guide action". Guba and Lincoln (1994) defined worldview as a set of basic beliefs about the nature of the world, and how it can be perceived. According to Guba and Lincoln (1994), the worldview defines for researchers what is being researched and what falls within and outside the scope of their legitimate inquiry. The basic beliefs that define worldviews can be encapsulated by the responses given by advocates of any given inquiry paradigm to the three fundamental questions (Creswell, 2003; Guba and Lincoln, 1994). These questions are interconnected, such that the answer given to any one question constrains how others may be answered. These questions are ontological (what is the nature and form of reality?), epistemological (what is the relationship between the researcher and reality?), and methodological (how can the researcher set about discovering what can be known?). Ontology refers to assumptions made about the nature of reality; epistemology describes how the knowledge may be gained; and methodology refers to the practical approaches and their rationale behind the consequent apprehension of reality (Lincoln et al., 2011;

Rajasekar et al., 2013). The answers given to these three questions gives rise to four inquiry paradigms, outlined below.

The Positivist Worldview. This is the standard view of natural science research (Robson, 2011). According to Phillips & Burbules (2000 cited in Creswell, 2009), positivism believes in acquiring an absolute knowledge of reality. It holds that there is an objective reality external to the researcher, and that it is possible for every scientist looking at the same bit of reality to see the same thing (Robson, 2011). Ontologically, positivism assumes the existence of an objective reality that is independent of the researcher's belief (Creswell, 2009; Lincoln et al., 2011); that this reality can be known, and that symbols can accurately describe and explain this objective reality (Cohen and Crabtree, 2006). Experimentation, theory, and hypothesis testing are used to acquire the knowledge of reality. This paradigm has been superseded by post-positivism (Creswell, 2009).

The Post-Positivist Worldview. This paradigm has been variously labelled – scientific methods, doing science research, empirical science, and postpositivism (Creswell, 2009). A post-positivist thinker believes in the independency of the researcher and the researched but accepts that the hypotheses, theories, background knowledge, and the value of the researcher can influence what is observed (Reichardt and Rallis, 1994) − "we cannot be positive in our claim to knowledge when researching human actions and behaviours " (Creswell, 2009: p. 7).

Ontologically, post-positivism holds that reality does exist but reckons that it can only be known imperfectly and probabilistically in part because of the researcher’s limitations (Reichardt and Rallis, 1994). Epistemologically, positivists believe (1) in the existence of general patterns of cause-and-effect that can inform prediction and can establish control over natural phenomena, and the goal is to discover these patterns; (2) that accurate data of the world can be obtained through observations and measurement; and (3) that a research can be free of objective and subjective biases by following a strict methodological protocol (Cohen and Crabtree, 2006). Methodologically, positivist research relies heavily on experimental and manipulative methods. This generally involves hypothesis generation and testing, identifying cause-effect relationships, and defining confounders (Cohen and Crabtree, 2006; Creswell, 2009, 2013; Robson, 2011).

The Social Constructivist Worldview. This is also called the social construction of reality, naturalistic inquiry or interpretivism (Creswell, 2009; Robson, 2011). It is an ontological belief that rejects the existence of any objective reality (Creswell, 2009, 2013; Lincoln et al., 2011). It considers that individuals construct subjective meanings of reality (Creswell, 2009). These meanings are multiple, intangible mental constructions, and they are socially, culturally, and experientially based and depend for their form and content on the individual persons or groups holding the constructions (Creswell, 2009; Guba and Lincoln, 1994). These constructions are alterable, as are their associated realities (Guba and Lincoln, 1994).

Epistemologically, the researcher and the researched are assumed to be interactively linked such that reality is subjectively created through the interactions of both (Guba and Lincoln, 1994). Methodologically, the research participants are assumed to help in the construction of reality through their diverse personal, cultural, and historical backgrounds (Robson, 2011). The varying constructions are interpreted through hermeneutical and phenomenological approaches and are compared and contrasted through a dialectical exchange (Guba and Lincoln, 1994; Robson, 2011). Similarly, the background of the researchers shape the research findings and the researchers acknowledge that their interpretations flow from their personal, cultural, and historical experiences (Creswell, 2009).

The Advocacy and Participatory Worldview. Some researchers who felt that interpretivism is not radical enough to cater to the concerns of the marginalised groups in the population espouse the philosophical assumptions of advocacy worldview. Advocacy worldview is a research tradition that grew out of the need for a reform that holds the capacity to change the life of the individuals or institutions in which they work in matters relating to empowerment, inequality, oppression, domination, suppression, and alienation (Creswell, 2009). According to Kemmis and Wilkinson (1998 cited in Creswell, 2009), the four features of advocacy worldview are:

- Advocacy paradigm assumes that participatory action is dialectical, and it concerns bringing about change in practices.
- It focuses on assisting individuals to free themselves from social hindrances found in the media, in language, in work places, and in power relationships within educational settings.
- It is emancipatory in that it helps to free people from unjust and irrational structures that limit self-development and self-determination.
- It is collaborative because inquiry is done in conjunction with others rather than on them.

4.2.2 Paradigmatic Controversy

Paradigm controversy dates to the time of ancient western philosophy. Socrates and Plato assumed universal truths or approaches to viewing the world; the Sophists, such as Protagoras and Gorgias, believed in multiple or relative knowledge of truths; and others believed that truth lies in the balance between or in the mixture of the two extremes (e.g., Aristotle principle of balance and moderate scepticism of Cicero and Sextus Empiricus) (Johnson et al., 2007). This debate continues even till this date.

The ardent debates between the supporters of quantitative and qualitative research paradigms led to the emergence of purists on both sides (Dornyei, 2007; Johnson and Onwuegbuzie, 2004; Lincoln and Guba, 1985; Lincoln et al., 2011; Robson, 2011). The purists argue that the two paradigms are mutually exclusive and incompatible (Dornyei, 2007; Johnson and Onwuegbuzie, 2004). Other researchers (e.g. Dornyei, 2007; Rossman and Wilson, 1991) believed that research questions and topics vary in the ways they lend themselves to micro-level (qualitative) and macro-level (quantitative) analyses. They maintain that both approaches are compatible and useful if applied in the appropriate research context. This led to the emergence of the situationalist approach to research methodology (Dornyei, 2007). According to Dornyei (2007), there were still some other researchers who felt that even though some research questions or topics are more suited to either qualitative or quantitative approaches, questions can be viewed from another perspective using the other approach, thus revealing new aspects of the issue. These researchers contend that some sort of integration of the two research approaches can be beneficial to provide a convergence in findings, provide richness and detail, or offer new interpretations (Rossman and Wilson, 1991). This view has led to the pragmatic position that underlies the mixed method approach (Dornyei, 2007).

4.2.3 The Pragmatism Worldview

The Pragmatism worldview is a philosophical position advanced by the pragmatists to underpin the mixed-method research (Hanson et al., 2005). It owes its origin to the work of pragmatists, such as Dewey, James, Pierce, Murphy, Rorty, and West (Creswell, 2003; Creswell, 2009; Hanson et al., 2005; Johnson and Onwuegbuzie, 2004; Johnson et al., 2007). There are several forms of the worldview, but for many pragmatists, there is no commitment to a particular philosophical position. The pragmatists seek a middle ground of the continuum between the two extremes of post-positivism and interpretivism worldviews (Creswell, 2009; Hanson et al., 2005). Ontologically, pragmatism worldview recognises the existence of a natural world and the emergence of social and psychological world, and it respects the existence of, and influence of the inner world of human experience in action (Robson, 2011). Epistemologically, pragmatism worldview rejects the notion of an absolute truth; considers truth as what works at a point in time; and under this worldview, the current truth, knowledge, and meaning is considered tentative and as something that changes over time (Robson, 2011). Methodologically, the pragmatists use whatever works to grapple with research topics and questions, using qualitative and quantitative research methods that can be combined to understand the research problem (Creswell, 2009; Dornyei, 2007; Hanson et al., 2005).

4.2.4 Research philosophical stance

I accept the view that certain research topics and questions are best researched using either qualitative or quantitative methods, but I have come to learn that flexibility is very important when conducting a PhD research, because things may not work out as planned. Therefore, I have followed a pragmatic worldview in this research, situating my philosophical position in the middle of the continuum between the opposing extremes of quantitative and qualitative paradigms.

Several factors influence this decision. First, the research was modified from the planned quantitative approach owing to the difficulties to secure enough participants. Second, there were difficulties and lack of funding to secure the service of a computer expert to upload a customised dictionary to facilitate a quantitative analysis of all the research questions. Third, the time limit for the programme and lack of funding rendered impractical the procurement of sophisticated software and training to use the software for thematic analysis.

4.3. Research Strategies

The worldviews detailed above are aligned with quantitative, qualitative, and mixed-method research methods (Creswell, 2003; Creswell, 2009; Dornyei, 2007).

4.3.1 Quantitative strategies

Quantitative research strategies are associated with the post-positivist worldview. These strategies include quasi-experiments, correlation studies, and various types of surveys (Creswell, 2009). According to Dornyei (2007), quantitative strategies are highly reputed, rapid to conduct, systematic, rigorous, focused, tightly controlled, and enjoy precise measurements of data that can be generalised. On the other hand, it is simplistic, averages out outliers, requires a large sample size, is reductionist in nature, and fails to capture the views of the research participants.

4.3.2 Qualitative strategies

Qualitative research strategies are related to the interpretivism worldview. The strategies include ethnography, grounded theory, case studies, phenomenology, and narrative research (Creswell, 2003; Creswell, 2009, 2013). Qualitative strategies use a wide range of data (recorded interviews, texts, and images), are conducted in a natural setting; uncover insiders' meanings, use small sample size, and adopt an interpretive analysis (Dornyei, 2007). According to Dornyei (2007), qualitative research strategies are exploratory in nature, make sense of complexity, broaden the interpretation of human experience, create a deep understanding, and offer flexibility when things go wrong, while providing rich materials for the research report. The alleged weaknesses of qualitative strategies include lack of methodological rigour, interference due to the researcher's bias, loss of generalisability because of a small sample size; the fact that it is time- and labour-intensive, and that it could yield too complex or too narrow theoretical models.

4.3.3 Mixed-method strategies

Mixed-method research strategies involve a combination of qualitative and quantitative methods. Several labels have been applied to the combination, including multi-strategy design (Robson, 2011), methodological triangulation, multi-methodological research, and mixed-methods research (Creswell, 2003). The combination of the two research methods rests on several considerations, including increased strengths, while minimising weaknesses; a better understanding of complex issues through convergence of the two research methods; improved validity; and generalizability of the research findings to wide audiences (Dornyei, 2007; Johnson and Onwuegbuzie, 2004). The drawbacks of the mixed-methods research is that the researchers need to learn multiple research approaches, risk loss of methodological rigour; and this can be overwhelming for a single researcher and can be more expensive and time-consuming, (Dornyei, 2007; Johnson and Onwuegbuzie, 2004).

Types of mixed-methods strategies. There are several variations of the mixed-methods research (Hanson et al., 2005; Johnson and Onwuegbuzie, 2004; Johnson et al., 2007; Rossman and Wilson, 1991); but all of them can be classified into two general categories.

Timing classification. This classification is based on the time of the research process, when one research approach is added to the other. Three variants have been defined in this regard, which are (Creswell, 2009):

Sequential. One research method is used to elaborate on the result of the other research method. Either qualitative or quantitative approaches might be the first method or the second method.

Concurrent. There is a simultaneous collection of qualitative and quantitative data. The results of the data analysis are integrated in the overall results interpretation.

Transformative. In context of transformative mixed-methods, the researcher uses a theoretical lens in a study that combines quantitative and qualitative approaches. The theoretical lens provides a framework that guides methods for data collection, analysis, and interpretation of results.

Purpose classification. This classification is based on the purposes and the rationale for mixed-methods designs, as described by Greene et al. (1989):

Triangulation. This purpose seeks convergence, corroboration, correspondence of results from different methods to increase the validity of constructs, and an inquiry into the results.

Complementarity. The purpose is to elaborate, enhance, illustrate, and clarify the results of one method with the results from the other method, thus increasing the interpretability, meaningfulness, and validity of constructs and inquiry results.

Developmental. Here, the purpose of mixing is to use the results from one method to help develop or inform the other method.

Initiation. The purpose of initiation mixed-methods design is to determine paradox and contradiction, new perspectives of frameworks, the recasting of questions or results from one method with questions or results from the other method for the purpose of increasing the breadth and depth of the inquiry.

Expansion. The purpose of expansion in a mixed-methods design is to extend the breadth and the range of inquiry, by using different methods for different inquiry components. This is aimed at increasing the scope of inquiry, by selecting the methods that are most appropriate for multiple inquiry components.

4.3. 4 Research Strategy Stance

A complementary mixed-methods approach is adopted in this study. This provides opportunities to quantify variables, which are then categorised along with concrete examples. This is in line with the philosophical worldview adopted for this research and the view that this provides more comprehensive results.

4.4 Research Design

Research design is the logical structure of inquiry. It concerns how research questions align with the evidence provided as an answer to the questions. The research design, in this study, is influenced by the context of the research. Problem-based learning (PBL) is a heterogeneous instructional approach due to its design and implementation, and no two PBL curricula are the same. Moreover, educational research is besieged with many confounding variables that are difficult to control (Albanese, 2000). The purpose of the study is not to seek explanation or to investigate into the development of variables, but to describe how knowledge is constructed by the students and is scaffolded by the facilitator. Therefore, experimental and longitudinal design is considered inappropriate for this study. A case study of the cross-sectional variety is deemed appropriate because of the context-bound nature of the research.

4.4.1 Types of Case Study

The basic types of case study designs were described by Yin (2009). The general characteristics of these designs serve as a background for determining the specific case design for this study.

Single case holistic. There is a single case, with a single unit of analysis. For example, research is carried out on educational curriculum in an institution, and data is collected from only one group of actors, for e.g., from students only.

Single case embedded. This refers to a situation, where there is a single case but multiple embedded units of analysis. For example, a curriculum is investigated and data is collected from more than one participant, for e.g., students and teachers.

Multi-case holistic. There are multiple cases, but each case has a single unit of analysis. For example, curriculum research carried out in two institutions but collected from one group of actors, for e.g., students.

Multi-case embedded. There are multiple cases, and each case has multiple embedded units of analysis. For example, research is carried in more than one institution, and data is collected from more than one group of actors, for e.g., students, teachers, and administrators. This study is a single case embedded design.

4.5. Data Analysis Approach

The approach to data analysis in this study arose from the consideration of the weaknesses encountered by the previous research works on PBL facilitation. For example, previous comparative outcome studies have disadvantaged the PBL instructional approach, because the assessment tools designed for traditional curriculum were used (Albanese, 2000). Furthermore, previous process research works on tutorial discourse are considered as being under-theorised (Azer and Azer, 2015; Howley et al., 2013), and they have used unsystematic analytic laborious techniques that are prone to error and researcher bias, and they are less feasible for application to a large students' interaction data set (Cockburn and Dale, 1997; Koschmann et al., 2000). Therefore, Howley et al. (2013) suggested an integration of learning and linguistic theories to develop a robust framework for understanding collaborative knowledge building; and Da Silva and Dennick (2010) advocated a computer-based text processing approach for the analysis of students' talk in the PBL sessions. The next section highlights a linguistic theoretical approach to the analysis of PBL tutorial conversations.

4.6. Lexicogrammatical Analysis: An overview

The basic grammatical analytic unit is the clause (Fontaine, 2013). Lexicogrammatical analysis involves an analysis of the clause moods, the lexical units of the clauses (Eggins, 2005; Halliday and Matthiessen, 2004; Thompson, 2004), and the lexical cohesives, binding the clauses together to form a text (Eggins and Slade, 1997; Halliday and Hassan, 1976). At the lexicogrammatical level of analysis within systemic functional linguistics, a clause is analysed at the three register axes of field, tenor, and mode (Eggins, 2005; Fontaine, 2013). A manual analysis of large data when done this way is very time-consuming, expensive, error-prone, and difficult (Barnbrook, 1996; McEnery and Hardie, 2012; McEnery and Wilson, 2001; McEnery et al., 2006). Recent advances in computer-based text processing could assist researchers to perform large amount of data analysis as well as provide a solid theoretical foundation for data interpretation (McEnery and Gabrielatos, 2006). This point takes me to the next issue: text analysis. Two electronic text analysis methodologies – corpus linguistics and electronic content analysis – offer a possible solution (Da Silva, 2013; Da Silva and Dennick, 2010). Text analysis is the subject of the next section.

4.7. Text Analysis: A definition

Text analysis refers to the method of communication researchers that employ to describe and interpret the characteristics of a recorded or a visual message (Frey et al., 1990). According to Frey et al. (1990), its purpose is to describe the content, structure, and functions of the messages contained in texts. Text analysis consists of a class of techniques for the social scientific study of communication (Popping, 2000). Lindkvist (1981: 26 cited in Popping, 2000) distinguishes between three definitions of 'text', which are:

1. Every semiotic structure of meaning (language, music, architect, picture, event, audio-visual language, and social actions);

2. Every linguistic means of expression (music is excluded from the text concept);
3. Written language (audio-visual language e.g., transcripts of broadcasting are excluded).

Text, in this study, refers to verbal linguistic means of expression, including the audio-visual expression.

There are four major approaches to text analysis – rhetorical criticism, content analysis, interactional analysis, and performance studies (Frey et al., 1990). This study concerns content analysis. Content analysis functions to identify, enumerate, and analyse occurrences of specific messages and message features contained in texts (Frey et al., 1990). A brief description of content analysis is presented in the next section.

4.8. Content Analysis

Languages are systems for signifying content (Teubert and Čermáková, 2007). Each utterance has a content, but the content is not the utterance, says Teubert and Čermáková (2007). “The utterance is a sequence of signs which represent the content or stand in place of content” (Teubert and Čermáková, 2007: p. 2). What is analysed in content analysis is the content that the language signifies.

Some researchers use content analysis and text analysis interchangeably, for e.g., Popping (2000). Content analysis (CA) can be defined as “the process of making inferences from symbolic medium such as text” (Weber, 1984: p. 126). It is a research method for constructing replicable and valid inferences from data to their context (Krippendorff, 1989). The conception of content analysis transcends what is said and to whom it is said in order to encompass other communicative circumstances and contexts, such as social and cultural conditions that explain what is said (Krippendorff, 1989). Meanings reside in the context of the language use (Frey et al., 1990; Halliday, 1978). The rules of the inferential process is influenced by the theoretical and substantive focus of the researcher (Krippendorff, 1989; Weber, 1990).

Content analysis seeks to analyse data within a specific context, based on the meaning a group or culture attributes to them (Krippendorff, 1989). According to Krippendorff (1989), communications, messages, and symbols differ from observable events or things, because they convey meanings other than themselves, reveal some properties of their producers, and have cognitive consequences for all concerned. Unlike other social research techniques that focus on describing the manifest behaviours and characteristics, the concern of content analysis transcends observable physical vehicles of communication. and it relies on their symbolic qualities to trace the antecedents, correlates, or consequences of communication to render analysable the unobservable context of data (Krippendorff, 1989).

Content analysis has wide-ranging usefulness, including generation of culture indicators, the possibility to audit communication against set objectives, the identification of intentions and other characteristics of communicator, the reflection of cultural patterns of groups, institutions, or societies, and the capacity to reveal the focus of individual, group, institutional, or societal intentions (Weber, 1990). The central issue in content analysis is the classification of many text words into fewer content categories; and elements of text classified in the same category are presumed to have similar meanings, based on the precise meaning of words or based on words sharing similar connotations (Weber, 1990).

Historically, content analysis has its intellectual roots in the ancient analysis of symbols and texts, and empirical inquiries into the meaning of communication date back to the theological studies of the 1600s, when the printing of nonreligious material was found to be a threat to the established authority (Krippendorff, 2013). The first documented case of a quantitative analysis of printed material, which involved the counting of religious symbols in songs, occurred in Sweden, in the 18th century (Krippendorff, 2013). Analysis of newspaper content, which started at the beginning of the 20th century, is considered to have evolved through the following five methodological stages:

1. Frequency analysis until the 1950s;
2. Emotion type analysis in the middle of the 1950s;
3. Intensity analysis in the 1950s and 1960s;
4. Contingency analysis starting from 1960s;
5. Computer analysis starting from the end of the 1960s.

Researchers in many fields, such as library science, political science, psychology, sociology, for a few to name, have applied content analysis to their work, and in the process, content analysis has been adapted to meet the unique needs of their research questions and strategies, with the result that a cluster of techniques and approaches for text analysis, grouped under the broad term of textual analysis, has emerged (Frey et al., 1990; Mergenthaler, 1996; White and Marsh, 2006). The next section discusses text as the material for content analysis.

4.9. Methodological options in content analysis

The methodological considerations in content analysis fall into two main axes – intent and technology (Kondracki et al., 2002).

The intent axis involves determining whether an inductive or a deductive approach to the research questions is more appropriate (Elo and Kyngäs, 2008; Shepherd and Achterberg 1992 cited in Kondracki et al., 2002). The inductive content analysis approach is used when there is inadequate prior knowledge about the phenomenon being researched upon, whereas deductive analytical approach is used when the analytic structure is operationalised on the basis of previous knowledge (Creswell, 2013; Thomas, 2006; Trochim and Donnelly, 2008). Inductive analysis moves from specifics to broader generalisations and theories, whereas deductive approach aims at theory- or model–testing, and it moves from more general to the specific (Thomas, 2006; Trochim and Donnelly, 2008). The process of inductive content analysis involves open coding and categorisation by key words, themes, and so on. (Kondracki et al., 2002). Following the deductive content analysis approach, the researcher tests existing categories, concepts, models, or hypothesis (Elo and Kyngäs, 2008).

The other issue in the intent axis relates to whether a study will examine manifest (visible at the surface level) or latent (deeper meaning implied in the text) content of the text or a combination of both (Shepherd and Achterberg 1992 cited in Kondracki et al., 2002). Manifest content is determined by using coding and key word searches, and it can be recorded in frequencies, for e.g., word counts while latent content is richer and more complex, requiring the development of constructs and the drawing of conclusions to add broader meaning to the text (Shepherd and Achterberg 1992 cited in Kondracki et al., 2002; Mergenthaler, 1996; White and Marsh, 2006).

Researchers in qualitative tradition focus on latent content and analyse data inductively, while researchers in quantitative tradition concern themselves with manifest content and tend to analyse data deductively (Kondracki et al., 2002). There have been debates regarding the compatibility of qualitative and quantitative approaches (Greene and Curucelli, 1994; Howe, 1988; Mergenthaler, 1996), but this has been resolved largely by taking a pragmatic approach that considers that the two methods are compatible and a more plausible answer to a research question can be provided by mixing the two analytic approaches (Greene and Curucelli, 1994; Johnson and Onwuegbuzie, 2004; Mergenthaler, 1996; Morse, 1991). Inductive and deductive analysis and qualitative and quantitative approaches are not mutually exclusive, and it is often useful to apply both (Kondracki et al., 2002).

The second axis relates to the technology employed for data analysis – manual method, or computerised method, or both. The manual approach to content analysis is tedious, time-consuming, expensive, inconsistent, and error-prone (Adolphs, 2006; Krippendorff, 2013; Popping, 2000; Weber, 1990). The computerised approach to analysis is considered capable of off-setting the limitations of manual analysis (Smith et al., 1996 cited in Kondracki et al., 2002; White and Marsh, 2006). The concept of computer-assisted content analysis is derived from research on artificial intelligence (AI) (Macnamara, 2005). The next section discusses artificial intelligence, followed by computer-assisted content analysis.

4.10. Artificial Intelligence (AI)

AI refers to the theory and development of computer systems capable of performing tasks that normally requires human intelligence. For a computer system to be artificially intelligent, it must be able to think and act like humans as well as think and act rationally (Kok et al., 2009; Russell and Norvig, 2010).

The dream of a machine that could perform human functions date back to the ancient times. Many artificial intelligences or automata appear in Greek mythology to perform the work that the Gods find burdensome (McCorduck et al., 1977). For example, around 850 BC, Homer tells us that Hephaestus, the God of fire and divine smith, who, because he was crippled, fashioned out attendants and endowed them with intelligence, speech, and strength, and then he learned how to do things from the immortal Gods, so that they could help him walk and assist him in his forge (McCorduck et al., 1977). McCorduck et al. (1977) cited several examples of AI systems: Ramon Lull, a 13th century Spanish mystic, went to the Muslims, where he was introduced to Arabic thinking machine, called a zairja. He himself came back to fabricate his own thinking machine called, Ars Magna. The machine is supposed to bring reason to dawn on all subjects to arrive at truths without the trouble of human thinking. In 1843, Ada Lovelace published a description of Babbage's analytical Engine (Morrison and Morrison, 1961 cited in McCorduck et al., 1977). In 1915, Leonardo Torres constructed two chess machines that played the chess endgame (Randell, 1973 cited in McCorduck et al., 1977). Other milestones in AI include the description of language for logic reasoning by George Boole in 1847; description of the Turing-machine by Alan M. Turing in 1936; the creation of artificial neurons in 1943 by Warren McCulloch and Walter Pitts; the determination of the theory of decision in 1944 by Neumann and Morgenstern; the creation of first neural computer in 1951 by Marvin Minsky and Dean Edmonds; the term AI was defined by John McCarthy in 1956; and in 1965, Herbert Simon said machines would be capable of doing any work that a man can do (Benko and Lanyi, 2009). AI systems are now taken to signify the creation of intelligent agents to help us do our work faster and easier (Russell and Norvig, 2010).

Artificial Intelligence was born out of a common effort by computer scientists, psychologists, logicians, and others, whose common objective is to teach computers to think and behave like human beings (Leech, 1987). The skills that such computer system possess include visual perception, knowledge representation, speech recognition, decision-making, motion and manipulation, and natural language-processing (Russell and Norvig, 2010). The convergence of computer science, linguistics, and the interdisciplinary field of computational linguistics for the purpose of teaching computers how to process natural languages led to the emergence of electronic text analysis (Adolphs, 2006).

4.11. Electronic Text Analysis

Electronic text analysis or computer-assisted content analysis (Alexandra, 2013) refers to the use of computer for the analysis of machine-readable text data, thus making replicable and valid inferences from the text to their context (Adolphs, 2006; Popping, 2000). Computer-aided content analysis unifies the traditional content analysis performed by humans with expending research from computer science on computational linguistics, natural language processing, and machine learning (Adolphs, 2006; Monroe and Schrodt, 2008).

Given the tedious, error–prone, and time consuming nature of manual coding and analysis, researchers realised the suitability of systematic analysis of texts by computers (Monroe and Schrodt, 2008). Harvard 's General Inquirer emerged as the first widely used computer software program for automated content analysis (Stone et al. 1966 cited in Monroe and Schrodt, 2008). Since that time, there has been a rapid growth in the computer-based content analysis due to various factors, including advances in information technology, text analysis software development, improved computer accessibility, availability of cost-effective devices for making text machine-readable, proliferation of volumes of electronic and digital texts, developments in other fields, and the growing interest in using electronic resources to complement more traditional approaches to the analysis of language and literature (Adolphs, 2006; Krippendorff, 2013; Oostdijk, 1991; Weber, 1984).

Computer-assisted text analysis has found applications in diverse disciplines of social sciences and humanities (Krippendorff, 2013; Popping, 2000). Electronic text analysis applications to language research is described under the term electronic content analysis (Adolphs, 2006), the term which refers to “all forms of research on language in which computers are used” by Popping (2000)p. 169. The traditions and methodologies of computer-aided language research are described under various terminologies that include corpus linguistics, natural language processing, humanities-computing, textual data–mining, to name a few (Adolphs, 2006). These areas of electronic content analysis have different orientations, based on their research goals. Two electronic text analysis methodologies – corpus analysis and electronic content analysis – are described in this study.

4.12. Electronic content analysis

The development of electronic content analysis software programs, emerged from the research carried out into natural language processing, textual data mining, artificial intelligence, and linguistics (Adolphs, 2006; Leech, 1987). The approaches to electronic text analysis vary in their mechanisms based on the task at hand.

Computer-based content analysis normally proceed from simple word count to modelling operations (Alexandra, 2013): the more simple computer-based content analysis operation consists of word count that relates to a particular outcome of interest and the construction of dictionaries that estimate the desired categories. Modelling approaches – language modelling and statistical modelling – do not rely so heavily on word count and the dictionary construction of categories (Monroe and Schrodt, 2008). Language modelling concerns the syntax and attempts to identify parts of speech in a given document, and it enables researchers to see the who, what, when, where, and how of a message. Statistical modelling, also referred to as machine learning, on the other hand, is based on a “bag-of-words” approach, and it concerns the likelihood of co-occurrence of words and phrases. It has two variants – supervised and unsupervised (Alexandra, 2013; Gentleman et al., 2008).

Supervised machine learning involves a procedure, whereby a human coder trains the machine to infer meaning from a certain pattern of words or categories by manually coding a subset of documents that is then applied to the rest of the corpus (Alexandra, 2013; Kotsiantis, 2007; Sebastiani, 2002). This was the approach used by Sebastiani (2002) for text categorisation, and Shami and Verhelst (2007) for classification of emotions in speech. Unsupervised machine-learning is similar to supervised machine-learning, but without involving human coding of documents. “The machine learns from statistical co-occurrence of words, grouping documents that belong together” (Alexandra, 2013: p. 371). This approach was used by Coates et al. (2011) for text detection and identification of characters in scene images, and by Niebles et al. (2008) for human action categorisation using spatio-temporal words. These approaches can be applied to a large corpora, without the need for prior coding; however, the downside is that less information is obtained from documents as information regarding who, what, when, where, and how could not be obtained (Alexandra, 2013).

4.13. Electronic content analysis software

Many types of software for electronic content analysis are available that support text analysis tasks within a variety of disciplinary contexts in significantly different ways (Alexa and Zuell, 2000; Alexa and Züll, 1999).

The software programmes facilitate data analysis by providing strategies for managing texts and their coding, examining word frequency and usage in context, creating and maintaining categories and categorisation schemes, assigning categories or codes to word strings, words, phrases, sentences, and the like ., note keeping, coding exporting, and providing team support for text analysis project (Alexa and Züll, 1999).

There are comparative similarities, and differences between corpus analysis and electronic content analysis software. Similar to corpus analysis software, electronic content analysis program is capable of automatic text coding, based on previously defined categorisation methods, derived from measures of statistical association and co-occurrence of words in large bodies of texts (Sowa, 2000 cited in Da Silva, 2013; Smith and Humphreys, 2006). Additionally, the software is able to group codes into themes identified from statistical measures of frequent associations in the text, and from previously defined hierarchical network of codes (Smith and Humphreys, 2006). Although both software programs code words, the tags associated with the words and the methods of extracting meaning from texts differ. The difference lies in the techniques employed for text exploration and the information sources used to code the words as well as to identify higher-order categories (Alexa, 1997): For instance, in corpus analysis categories are predefined, but in electronic content analysis, the contents and themes are derived and named based on the co-occurrence in the texts. Electronic content analysis focuses on the semantic network of words, whereas corpus analysis applies grammatical and semantic dictionaries to match the words in the POS and semantic categories (Da Silva, 2013).

Following a review of the literature, Alexa and Züll (1999) categorised text analysis software packages, based on their ability to support different types of research (e.g., qualitative, quantitative, or both) and a combination of functionality (e.g., database managers, archiving programmes, text searching, text retrievers, taggers, code-and-retrieve programmes).

Examples of qualitative software packages include AQUAD, ATLAS.ti, HyperRESEARCH, NUD*IST, QED; WinMAXpro, DICTION, DIMAP-MCCA, KEDS, TEXTPACK, TextSmart, and Wordstat, which are categorised as quantitative ones; and Code-A-Text and TATOE support qualitative and quantitative analysis. The software packages provide a combination of functionality by providing a combination of operations for text analysis. The basic operations that the text analysis software packages support can be grouped into four operations to achieve the following: (1) text import and management; (2) text and coding exploration; (3) export operation; and (4) creation of dictionaries, categorisation schemes, and coding. The differences in the available text analysis software programs relate to whether or not they support these operations, and how they support certain operations (Alexa and Zuell, 2000; Alexa and Züll, 1999).

4.13.1 Strengths of Electronic Text Analysis Software

Some of the advantages of corpus software programs have been well-discussed (Adolphs, 2006; Alexa, 1997; Alexa and Zuell, 2000; Alexa and Züll, 1999; Kondracki et al., 2002), and they include the following, as summarised by Da Silva (2013):

Bias reduction. Corpus analysis software provides opportunities for systematic and automatic tagging and coding, based on explicitly defined rules, thus enhancing analytical reliability and replicability.

Feasibility of large corpus analysis. Increasing sophistication of corpus analysis programs makes possible the analysis of the ever-increasing sizes of the corpora. The analysis of the latest 450 million-word version of Birmingham Bank of English Corpus would have been impossible without corpus analysis software.

Pattern identification. Corpus analysis facilitates identification of patterns in texts. Semantic and syntactic patterns are often too intimately entrenched in the text for unaided human detection; however, text analysis software can easily identify these patterns. The software enables identification of repetitive words as markers of political ideology by Rayson (2008).

Easy manipulation and analytical flexibility. Corpus analysis programs provide enormous opportunities for easy manipulation of text details and provide flexibilities for data analysis. With the software, it is possible to choose between type I, II, and III analytic approaches. It is also very possible to move to and fro between frequency listings and concordancing. It is possible to view KWiC lists from descriptive statistics and vice versa.

Foster research collaboration. Corpus analysis software makes online storage, sharing, and exporting of data in different formats possible. This provides opportunities for researchers and research departments to share the same data, verify previous findings, and explore aspects of data, not yet investigated by previous researchers (Alexa, 1997).

4.13.2 Weaknesses of electronic text Analysis software

Corpus analysis software programs are not without their own limitations. As noted by Rayson (2003), language and communication are dynamic and three-dimensional with other elements, such as speech, pace, tone, and other nonverbal elements, adding to the complexity, richness, and vitality of communication. These other elements are beyond the reach of simple techniques of analysing textual material. Inspite of being well-suited for analysing explicit content and investigating explicit meaning in texts, the corpus analysis software does not seem to be the only method to analyse latent variables, because the software is not sophisticated enough to capture all the richness of human communication (Alexa and Züll, 1999). For this reason, Kondracki et al. (2002) cautioned that the methods need to be used within the limits of their potentialities while recognising their limitations.

Collaborative knowledge-building, in problem-based learning tutorials, creates opportunities for verbalisation, information sharing, and for the building of a collaborative knowledge of the case (Hmelo-Silver, 2003a, 2003b; Hmelo-Silver and Barrows, 2008), and this requires the use of methodologies such as corpus linguistics and electronic content analysis. This study uses Wmatrix3 for corpus analysis, and the QDA Miner as an electronic text analysis software.

4.14. Research rigor and trustworthiness

Research design represents a logical set of statements and a given design can be judged by a set of tests. This section presents the strategies for establishing the quality of the research.

4.14.1 Construct validity

"Validity is the degree to which an instrument measures what it is supposed to measure" (Polit and Beck, 2012: p. 336). A study is conducted with specific exemplars of outcomes, settings, and people that represent broad constructs. Construct validity involves deductions from study particulars to the higher-order constructs that they are intended to represent. Construct validity is very important because constructs link the operations used in the study to a relevant conceptualisation and to mechanisms for translating the resulting evidence to practice (Polit and Beck, 2012). Several strategies are available to enhance construct validity of a research (Polit and Beck, 2012; Yin, 2009). The first step in fostering construct validity involves careful definition of the research topic in terms of specific concepts, and then, they are related to the objectives of the study. Second, the operational measures that match the concepts need to be identified. Shadish and colleagues (2002 cited in Polit and Beck, 2012) extended construct validity to cover persons, settings, and study outcomes.

4.14.2 Reliability

Reliability is the consistency with which an instrument measures a target attribute (Polit and Beck, 2012). It aims to minimise errors and biases in a study (Yin, 2009). Reliability could be assessed using test-rest and inter-rater measures.

Test-retest reliability test measures the stability or reproducibility of the measure: the same assessment instrument is rated by the same individual at two different times and Cohen's kappa, weighted kappa, or correlation coefficients are calculated. A kappa coefficient of 0 indicates agreement due to chance; negative value indicates that it is worse than chance; and a value of 1 indicates perfect agreement (Bowling, 2009; Viera and Garrett, 2005).

Inter-rater (or interobserver) test reliability refers to the extent to which the results obtained by two or more raters agree on the same phenomenon rated. The results are then used to calculate an index of agreement between the raters. Cohen's kappa value of 0.60 is considered as being minimally acceptable, and that of 0.75 and above is considered as being very good (Polit and Beck, 2012).

4.15. Summary of the chapter

This chapter has highlighted the philosophical perspectives and has presented the rationale for the philosophical views and methods adopted in this study. An overview of content analysis and corpus linguistics methodology are discussed.

CHAPTER 5

RESEARCH DESIGN

5.1. Introduction

This chapter describes the steps employed for the conduction of the study. The study is situated within the pragmatic paradigm. This provides the epistemological foundation upon which all parts of the study are based. The research procedure is informed by the research questions, which also authenticate the methodological choices made, the analysis carried out, and the result interpretation and discussions. The chapter concerns the design of the study, participant recruitment and sampling, inclusion and exclusion criteria, and it also highlights ethical issues, data collection, processing, and analysis. This chapter concludes with ay highlighting of the criteria for assessing research rigor and trustworthiness of the study.

5.2. Study Design

The research strategy used for the present study is an embedded case-study design, as described by Yin (2009:p. 46. An embedded case study is a situation where a single case (such as, an organization, institution, or the like) consists of the units of analysis (Yin, 2009). Case study involves exploring a programme, event, activity, process, or one or more individuals in depth (Creswell, 2009). The result of the research is generalisable to the context of the study as well as the theory underpinning the research (Yin, 2009). Several factors influence this decision.

First, the research focus aligns with the description of the case study as an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the context intimately influences the phenomenon (Yin, 2009). In this study, the aim is to observe knowledge construction in the PBL tutorial in its real-life context, with the researcher not exerting any control over the classroom events.

Second, the context-bound nature of the problem-based learning (PBL) curriculum influenced the choice of case study as the nest-suited research approach for this study. Since its introduction in the 1960s, (Schmidt, 2012), the orthodox format of PBL has mutated to heterogeneous variants (Barrows, 1986; Karen et al., 1998), such that no two PBL curricular are now the same across groups and across implementations. The study is less-suited for comparative design, because of the many confounding factors that would be very difficult to control as well as because of the limited sample size (Albanese, 2000). A mixed-methods approach of the type III procedure to corpus analysis, described by (Rayson, 2003, 2008), was used. This involves a combination of macro-analysis and a further micro-analytic exploration of utterances in order to grasp a thorough understanding of the data in its context. This aligns with the definition of mixed methods, as defined by Tashakkori and Creswell (2007), and has characteristics that are compatible with "a truly mixed approach methodology", described by

Tashakkori and Teddlie (2003 cited in Mertens, 2010: p. 294). This approach involves multiple approaches in all stages of the study, and iterative and complementary analyses from different perspectives and viewpoints (Tashakkori & Teddlie, 2003 cited in Mertens, 2010). Data search and retrieval make extensive and frequent use of concordancing for validating the classifications generated automatically by the software as well as to establish an in-depth contextual understanding of the frequencies and co-occurrences.

5.3. Study Site and Setting

The selection of the site of the study was based on the objectives of the study, which involves an exploration of the collaborative knowledge construction of graduate-entry medical students, and how the PBL facilitators guide the process.

Educational context of the Study. Following the recommendation of the General Medical Council (GMC, 1993) for the incorporation of the theories of adult and problem-focused education into medical education, medical schools in the UK began the adoption of the PBL curriculum in the mid-1990s (Maudsley, 1999). The British medical schools that adopt the PBL curricula were located in Manchester, Liverpool, and Glasgow Universities (Maudsley, 1999). Since that time, several other medical schools, including the University of Nottingham, have adopted the PBL curriculum.

Medical Education Programmes at University of Nottingham. The University of Nottingham has three pathways for the conduction of undergraduate medical education – a 5-year BMBS medicine (A100) programme, a 4-year BMBS graduate entry medicine (A101) programme, and a 6-year BMBS medicine (A108) programme. The graduate entry pathway is a four-year programme, consisting of an initial one and a half years of predominantly 4–5 hours per week of PBL sessions in integrated basic medical sciences, along with lectures, clinical skills training, and some sessions of shadowing in community-based sites (Figure 5.1). The following two and a half years are spent in clinical placements, along with their colleagues from the 5-year programme.

Table 5.1: GEM year-one students' weekly timetable

<table>
<tr><th>Time</th><th>Monday</th><th>Tuesday</th><th>Wednesday</th><th>Thursday</th><th>Friday</th></tr>
<tr><td>9-10am</td><td>Lecture</td><td>PBL</td><td rowspan="2">Workshop</td><td rowspan="3">GP visit</td><td rowspan="2">Lecture</td></tr>
<tr><td>10-11am</td><td>Lecture</td><td>PBL</td></tr>
<tr><td>11-12pm</td><td>Lecture</td><td rowspan="2">FP/PS</td><td rowspan="3">FP/PS</td><td rowspan="2">PBL</td></tr>
<tr><td>12-1pm</td><td>Lecture</td><td>FP/PS</td></tr>
<tr><td>1-2pm</td><td>FP/PS</td><td rowspan="2">Clinical</td><td>FP/PS</td><td>FP/PS</td></tr>
<tr><td>2-3pm</td><td>Lecture</td><td>Lecture</td><td rowspan="2">Lecture</td><td>Clinical</td></tr>
<tr><td>3-4pm</td><td>Lecture</td><td rowspan="2">Clinical</td><td>Lecture</td><td>Clinical</td></tr>
<tr><td>4-5pm</td><td>FP/PS</td><td>FP/PS</td><td>Workshop</td><td></td></tr>
</table>

(Student guide 2017. Online:www.nottingham.ac.uk/medicine/teaching)

Graduate Entry Medical Programme in Derby. The graduate entry programme is held at the Graduate Entry Medical School (GEMS), situated in Derby. The programme is based on the pedagogic principles pertaining to adult learning, as specified by the GMC (GMC, 1993), but it adopts a hybrid PBL approach with the inclusion of lectures, clinical skills, and basic/clinical sciences workshops to provide students with diverse learning opportunities, outside of the PBL sessions (Table 5.1). The school was opened in 2003, and it admits ninety students annually from different professional backgrounds. A passing score in the GAMSAT examination and a structured interview is required for admission. It is a highly competitive admission process, and only about 7.5% of the applicants get admitted annually (GAMSAT, 2011 cited in Da Silva, 2013). A principal eligibility criterion for admission is the willingness to participate in a PBL curriculum.

The four-year medical degree programme in the GEM school uses PBL as the vehicle for learning through the initial years. The first eighteen months is dominated by PBL sessions, and this is followed by five semesters of clinical rotations through the hospitals and GP practices in the region. The PBL sessions are organised around eleven main curricular themes, which include a combination of lectures, clinical skills practice, basic and clinical workshops, professional development sessions, and clinical case discussions, even though theme length may vary, depending on the theme specification. The medical school defines the curriculum outcomes, but the students are responsible for defining the learning outcomes for the PBL sessions.

Implementation and Timeline of the PBL Process. Every week, the students spend about 4.5 hours in the PBL tutorials, discussing a clinical case that is selected, based on the themes of the curriculum. The 4.5 hours are divided into three sessions.

Session 1. At the beginning of session one, a computer audio presentation of the case vignette is done. The students clarify the terms and concepts that are unknown or the ones that they do not understand. After this, the students give possible explanations or hypotheses for the case disorder, based on their existing knowledge. This is done by brainstorming. They use whiteboard to organise their learning by listing what they

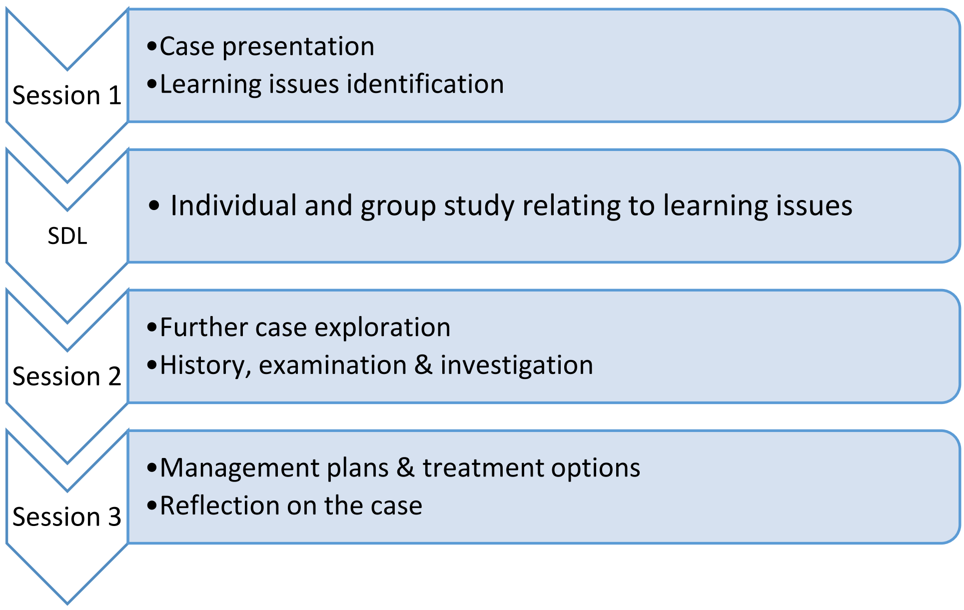


Figure 5.1: Problem-based learning process

know and what they do not know. They provide as many different hypotheses/explanations as possible, using their prior knowledge, practical experience, or their ideas. They then discuss the possible explanations, and they try to establish the connections between them and identify what knowledge is lacking. They then interview the patient, who is usually one of the students or a tutor, who acts as the patient. The students try as much as possible to get the missing knowledge from the patient through an elicitation of clinical symptoms and signs. The remaining knowledge that is lacking is crystallised into learning issues. Thereafter, the students divide the learning issues amongst themselves and proceed to the self-directed learning (SDL) period. They search the literature and other sources of information to gain knowledge and to attain an understanding of the topics formulated as the learning issues. They study theoretical concepts and explanatory model, which are then applied to the problem. Each student reads all the learning topics generally but tries to be an expert on the assigned topic through an in-depth research into the same.

Session 2. In the second session of the PBL cycle, the students converge again with the results they have obtained by researching upon the learning issues. The students take turns to present their results. Based on the results of the self-study, the students can reduce the number of hypotheses to one or two disorders. They discuss the explanatory models (disease mechanisms) that were found for the problem. This stage involves a lot of discussions, and during the same, the students can challenge each other. They can also raise further questions that relate to the presentation of their peers or sources of information. The learning resources and information obtained are normally shared with peers during this session.

Session 3. In session 3, the students explore the investigations that are relevant to the case. They formulate the treatment options for the disorder and summarise what they have learned. This session also involves a reflection on the PBL process for the case: what was done well, what was inadequate, and the quality of the facilitation.

5.4. Recruitment of Research Participants

Having identified the participants, the next step was to gain access to the site and the participants. As the researchers do not work in the graduate-entry medical school in Derby, participants were approached through the head of the tutors in the school. This lack of familiarity with the research participants was considered beneficial, as it reduced the threat posed to the validity of the data collected (Morse, 1994). The supervisor of the project initially discussed the research with the head of the tutors and written information regarding the study was given. The head of the tutors informed the students and the facilitators and arranged a mutually convenient time for the researchers to meet the students and the facilitators.

5.4.1 Recruitment of the student participants

The research team (supervisor and two PhD students) was introduced to the class of first-year medical students after a lecture by the head of the facilitators. A presentation was given by a member of the research team about the project after which participant information sheet and consent forms were distributed. The students and the facilitators were given the opportunity to ask questions about the project. A member of the research team later visited each tutorial group to further elaborate on the project and answer further questions that the participants may have. The consent forms were also collected from the participants.

All the students and the facilitators of the 2012/2013 and 2008/2009 cohorts of University of Nottingham Graduate Medical Entry (GEM) School were individually invited to take part in the study. Recruitment of the participants was done twice, as the data that constituted the corpus was collected on two occasions; 2009 and 2013.

Of the 12 year-one problem-based learning (PBL) tutorial groups approached for participation for data collection in 2013, and 4 groups (3 students and 1 facilitator) declined participation via consent form. Two further students, one from each group, changed their minds and declined participation on the day of data collection. Thus, 6 groups participated in the study. The data was considered insufficient. As it was judged impractical to restart the process of data collection again because of time limitation, the corpus design was modified in view of this logistical problem. Therefore, the data collected was supplemented with samples taken from the video and audio recordings made in the same school in 2009 academic year.

In the 2009 data collection, all the twelve year-one tutorial groups' students were individually approached for participation. Five groups declined participation in the study via the consent form. Seven

PBL groups participated in the study. From the two periods, data was collected from 13 groups. I do not think the time difference in the data collection affected the quality of the data. The project was undertaken in the same medical school operating using the same curriculum and under the same academic supervisor. The same data collection procedure was observed during the two data collection periods and transcription was done by the same outside professional transcriber. The participants in the 2009 data collection have had four months' experience with PBL curriculum, while the participants in 2013 data collection were 3 months into their PBL curriculum experience.

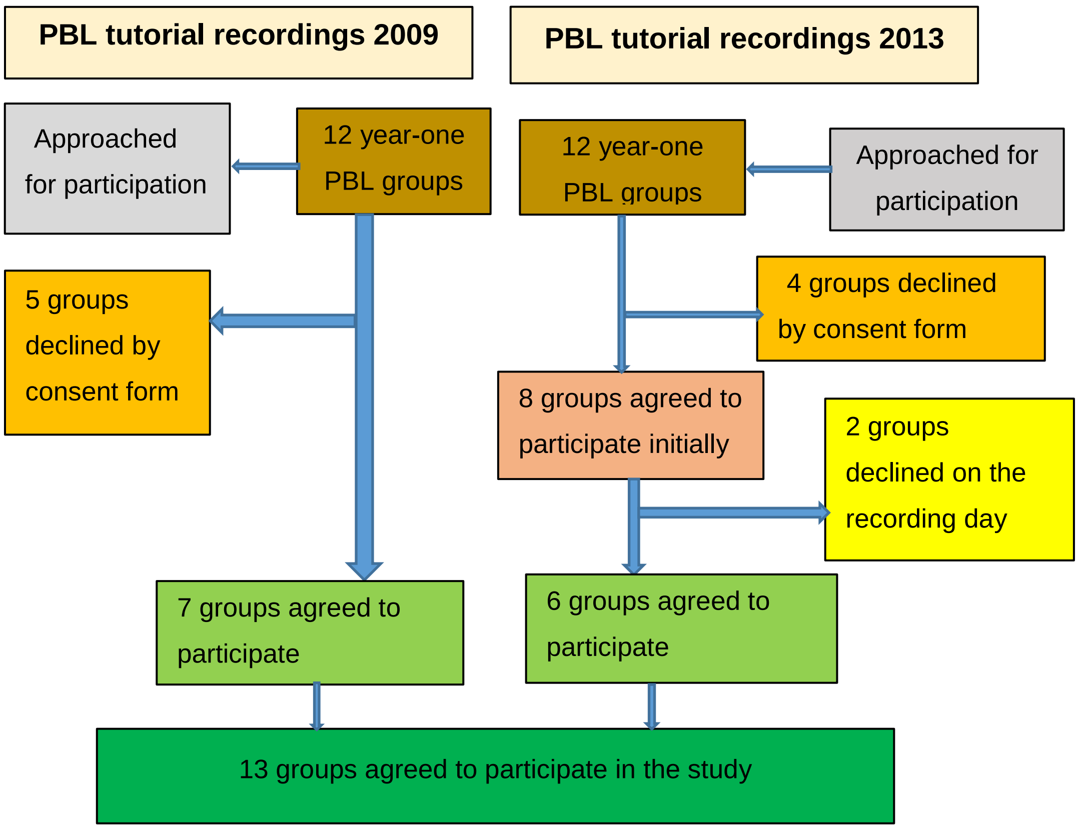


Figure 5.2: Participant recruitment process

5.4.2 Recruitment of PBL facilitators

Recruitment of the facilitators were done simultaneously with that of the medical students. The head of the facilitators facilitated the meeting between the research team and the facilitators. The facilitators were also given the information sheet about the study and were visited in the tutorial groups with the students when any question they had were answered.

5.4.3 Sample selection and sample size

The aim of this study was to gain insight into the process by which graduate entry medical students collaboratively construct knowledge of the case scenario presented to them during PBL tutorial. Therefore, a non-probabilistic purposeful sampling or concept/theory sampling was used (Creswell, 2012). According to Creswell (2012), theory or concept sampling involves sampling individuals or sites, because they would help the researcher understand a concept or a theory better. This aligns with the idea that sampling approach must reflect the nature and goal of the study (Creswell and Clark, 2011; Teddlie and Tashakkori, 2009).

5.4.4 Inclusion criteria

This refers to the attribute of the participant and the tutorial group for inclusion in the study. The medical students included in the study were first year PBL graduate entry medical students who expressed willingness to participate in the study by signing and returning a consent form. The facilitators included in the study were first-year PBL facilitators who expressed willingness to take part in the study by completing and returning a consent form. The PBL groups included in the study were those in which all the medical students and the facilitator consented to participation in the research.

5.4.5 Exclusion criteria

Facilitators and medical students who were unwilling to participate in the study were excluded. Substitute facilitators were excluded from the study. PBL groups in which a member of the tutorial group (a facilitator or a student) opted out of the study were excluded from the study.

5.5 Procedure for Data Capture

This is a three-part process, consisting of data collection, computerisation, and annotation (Meyer, 2002). Several methodological considerations were addressed in the recording process.

The central concern in collecting any kind of speech is the recording of a natural speech. This requires dealing with “reactive (Hawthorne) effect” (Bowling, 2009: p. 174) or “observer’s paradox” (Meyer, 2002: p. 57). This is a situation where the research participants change their natural behaviour because they are being observed.

Although it is impossible to completely make individuals forget that their speech is being monitored and recorded, this effect can be minimised (Meyer, 2002). We followed the recommendations of Meyer (2002) to preserve the naturalness of the PBL discourse during the recordings: the recording of the sessions was done in the natural PBL discussion environment, devoid of the presence of the researcher. The researcher was only in charge of setting up the recording equipment at the beginning of the session and its collection at the end of the session. The equipment collection visit provided the researcher with opportunities to talk to the group members and understand if any problem had been encountered during the recording. The added advantage of this approach to data recording is that the approach aligns with the principle of respect for the learning environment and culture of the group studied, and it upholds the ethics of respect for voluntary participation in research. By giving the group members the control of recording, the intrusion of the researcher into the PBL learning session is eliminated. This recording method is also in line with the problem based learning culture of student autonomy in which students are expected to take control of group activities with minimal interference from the tutor/facilitator. Additionally, the recording method supports the ethics of voluntary participation in that it gives the students the opportunities to decide on which aspect of the discussion they would like to record and which aspect they would not like to include in the reordering.

A test recording of a session was conducted where the students familiarised themselves with the equipment and its operational mode. The test recording session provided dual benefits: It habituates the research participants to the process of being recorded, and thus, further help to enhance the naturalness of the participants' verbal behaviour. It also provided the opportunities to correct problems relating to equipment use and detect malfunction that needed correction. It was at this session that the groups agreed to record the case-based discussions and leave out other social personal and non-academic discussions. The test recordings were not included in the data for analysis.

The quality of recordings for transcription is very important for collecting natural speech data especially of the multi-party dialogue type. This inevitably requires a consideration of the tape recorder to use. Recording quality was guaranteed by using a high-quality digital recorder, an Olympus DS-2500 professional dictation machine with 4G memory, and a Sony full HD camcorder. Digital recordings are considered superior to analogue recording for several reasons. The recording quality did not degrade quickly. Digital recorders could record up to four hours of continuous speech, without the need to turn over the cassette which may interfere with the naturalness of the discourse. Besides, digital recordings can easily be uploaded unto a computer, processed and transferred to the transcriber with a password protection without loss of recording quality. Digital recording also enabled us to use software to improve the quality of the sound files and eliminate background noise.

The quality of the recording was also ensured by correctly placing the recording devices in the correct positions in the tutorial rooms. The simultaneous use of audio and video recording devices placed at two different locations in the rooms enhanced the capture of the participants' speech and enabled the researcher to allocate utterances to the participants accurately and correct the transcripts. Video recording was used alongside audio recording because the videos were used to allocate talk to each participant.

Due to the drop outs, only the PBL sessions (full cycle) of the 10 consenting groups were video and audio recorded to generate a total of 30 sessions of recording giving a combined total of approximately 40-45 hours of recordings.

5.6 Recording File Format

The files were stored in MPG and MP3 formats. They combine file compression, recording quality, and flexible file-handling. The file compression provides opportunities for file uploading and transfer (Thorsten et al., 2012). File conversion software (iSkysoft video converter) was also installed on the computer to facilitate file conversion from one format into other.

5.7 Transcription of Audio Recordings

Transcription is derived from the Latin word trans-scriber, and it refers to the transformation of an audio or video recording into the format suitable for analysis (Taylor and Gibbs, 2010; Thorsten et al., 2012). This is accomplished by transcribing the recordings into texts called transcripts. The decision regarding the details of the transcription and the format of the transcript is influenced by the research question and the method of analysis. Thorsten et al. (2012) and Taylor and Gibbs

(2010) describe three transcription conventions: complex transcription, simple transcription and colloquial transcription.

Complex transcription. This involves adherence to complex transcription rules that involve the inclusion of prosodic elements (e.g., intonation, primary and secondary emphasis, voice volume, speed, and pitch of speech), phonetic elements (e.g., in research on dialect) or non-verbal elements (e.g., gestures and deictic expressions). Analysis requiring complex transcription tend to focus on deeper semantic content of a conversation.

Example:

S1: =<<dim> or whether they'll get divorced ↑'after all.>

S2: ˇhm,

(--)

S1: <<pp> this is still-> ((breathes out for 2.1 sec)) <<p> t'is a ↑' they are a good example for this

Simple convention. In simple transcription convention, para-verbal and non-verbal elements of conversation are omitted. The dialect and colloquial language elements are corrected to standard language. Only the more relevant parts of the speech are transcribed as spoken exactly (verbatim) and information is reviewed into clearer or shorter sentences. The focus of this transcription convention is on readability and it is easier and quicker to produce than complex transcription. It is useful when the focus of the analysis involves grammatical and surface semantic content of a conversation.

Example:

S1: ….. or whether they'll get divorced after all.

S2: Hm ……

S1: This is still …… It is a transition.

Colloquial transcription. This refers to the transcription convention in which colloquial expressions, semi-words and sounds are included in the transcription.

Example:

> "I would'ne leave the kids wi' som'd'y when they were young. I wanted to go oot wi' my friends but I wis waryied aboot what meeght happen while I wis oot. My Granmither offered to look after them but I felt it wis too much for her as her health was nae good".

Complex and colloquial transcriptions are not suitable for corpus analysis. These transcription conventions would make automatic software annotation extremely difficult, and thus, produce unnecessary errors. Simple transcription convention was used for the transcription in this project because it is devoid of the above-mentioned problems relating to complex and colloquial transcription conventions. It is also less expensive, faster to carry out, and is adequate to answer my research questions.

5.8 Transcript Editing Rules

To ensure that the transcripts are consistent, the simple transcription convention guidelines (Thorsten et al., 2012) were followed during transcript editing and cleaning. The guidelines include:

- Transcribe literally without summarising and phonetic transcription and dialectic and colloquial language are reviewed into standard language.
- Sentence discontinuation, pauses, abrupt stops are indicated by ellipses. An ellipsis was also used to mark the point of re-entry into an abandoned statement after a moment of interruption. Punctuations like full stop, semicolon, colon, commas, question mark, and hyphen are kept as in Standard English. Vocal interjections and capitalisations were kept as in Standard English. Emotional and non-verbal utterances were not included and overlapping between speeches were not indicated. Unclear words were not guessed. Inaudible or incomprehensible utterances were marked with inaudible in square brackets; however, the duration of the disturbances that caused the inaudibility was not included.
- Each speaker’s turn starts on a new line. Each participant is marked with a gender initial letter, followed by the number given to the participant (e.g. M1 for male number 1 and F2 for female number 2, etc.).
- Contractions, symbols, abbreviations were kept exactly the way they were said because the software is capable of their tagging

and analysis. Numbers, decimals and other numerals were also kept exactly the way they were spoken.

5.9 Transcript Cleaning

The transcript cleaning stage of data preparation involves the removal of irrelevant materials from the transcripts so that an appropriate transcript quality could be attained in readiness for data analysis (Taylor and Gibbs, 2010). The procedure was carried out manually by the researcher with the assistance of the video files.

The process starts with a careful review of each transcript against the video recording. The simple transcription convention rules were used as guide during this process. It involves correction of mistakes and errors and deletion of irrelevant material in the transcripts in line with the convention rules. PBL discourse is a multi-party dialogue, involving 7–10 students and a facilitator, and thus, vocalised pauses that allow a speaker to pause and plan what next to say, chorus responses, and sub-group utterances are common. There is no universally agreed spelling for vocalised pauses (e.g., uh, uhm, erm, hem, etc) but to maintain consistency, these expressions were corrected to the spellings supported by the software dictionary. Chorus responses were marked with chorus enclosed in square bracket. Moments of sub-group talk were treated as normal contribution. However, side talks (discussion about football etc.) not related to case discussion were considered irrelevant and edited out. However, jokes and humours were included in the transcription because they were considered relevant to the social dimension of collaboration.

5.10 Formation of Study Corpus

Of the 7 groups that participated in the 2009 data collection, transcription of one tutorial group discussions was incomplete because of poor recording. Thus, transcription of six tutorial groups was complete, and the transcript texts from these 6 groups were included in the study corpus. Of the remaining 6 groups that participated in 2013

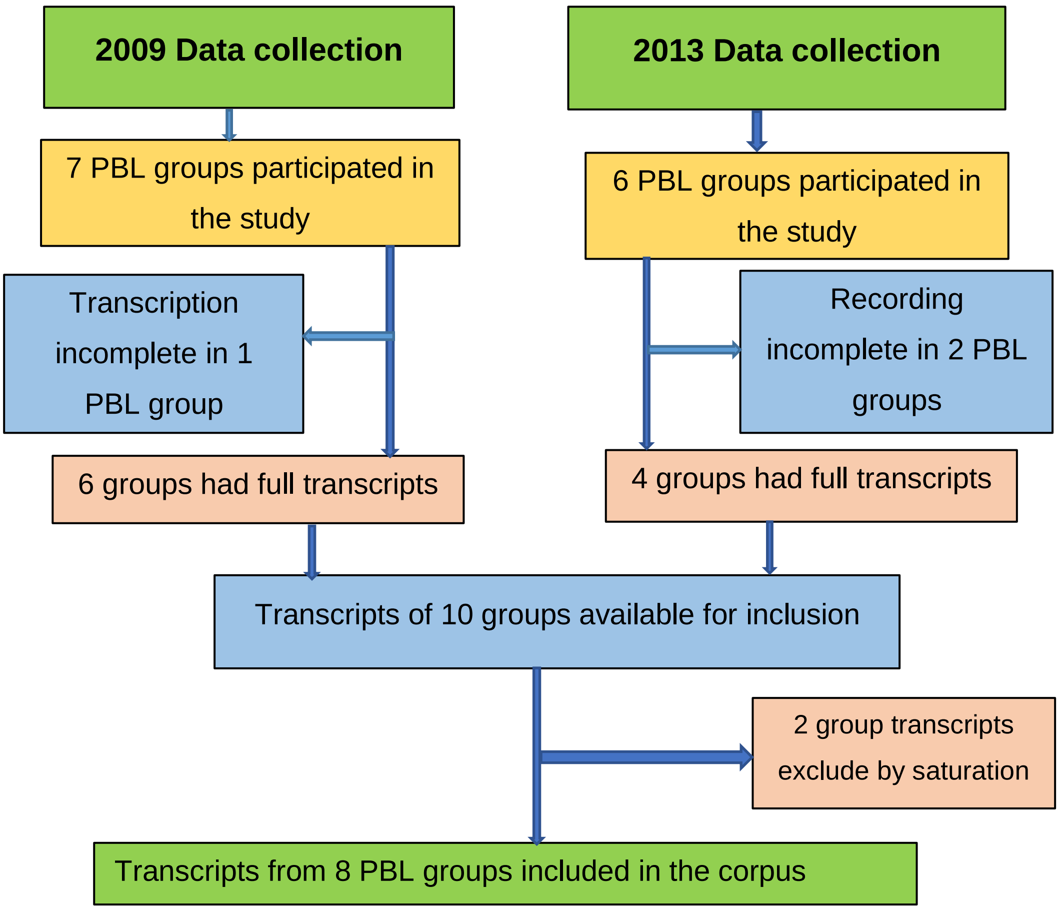


Figure 5.3: The process of study corpus formation

data collection, recordings were incomplete in two groups. In the first group, the facilitator was unwell and the recording suspended by substitute facilitator. In the second group, the recording was fragmentary due to multiple recording machine mechanical failures. Thus, only four groups have full recordings of all the three sessions and the transcript texts from these 4 groups were available for inclusion in the study corpus (Figure 5.3).

5.11 Study Corpus Description

Of the 10 tutorial groups with full transcripts for inclusion in the corpus, transcripts from eight tutorial groups were included and analysed. The determination of the number of groups analysed was based on saturation (Charmaz, 2006; Schreier, 2012). Data analysis was discontinued when no new emerging themes were found. Therefore, the study corpus consists of transcripts texts from 8 PBL tutorial groups (Figure 5.3).

The words that make up the corpus were contributed by the eight tutors and 63 medical students. The corpus was derived from the texts of the transcripts that were generated from audio recordings of 24 tutorial sessions. The recordings represent full PBL cycle of eight year-one medical students' PBL tutorial groups. The study corpus consists of 2,37,820 words divided into two subcorpora – students' subcorpus and facilitators' subcorpus. The students' subcorpus contains 210, 077 words (88.33%) while facilitators' subcorpus contains 27,743 (11.67%)

words. Table 5.2 describes the statistics of the sub-corpora making up the study corpus.

The size of the corpus for this study is 237, 820 words. The corpus size is considered adequate based on the results of previous studies. The use of small corpora has been demonstrated in two studies by Koester and O'Keeffe (cited in Evison, 2012). Koester (2006 cited in Evison, 2012) investigated workplace discourse using a corpus of just under 34,000 words. Similarly, O'Keeffe (2003 cited in Evison, 2012) based his study of media discourse on a sample of 55,000 words of phone-in data. Furthermore, Jennings (2013) in an ethnographic study of the practice and facilitation of problem-based learning studied only two tutorial groups.

5.12 Text Preparation for Analysis

Each group transcript was separated into students' transcript and facilitator's transcript. The text of the facilitator for each group was given a number (e.g. T1 for group 1 facilitator) and the facilitator's transcript was saved in word format and as plain text. Similar procedure was followed for the creation of students' transcript, and this was given a number (e.g. S1, student group 1). The students' transcript for each PBL group was saved in word format and as plain text. The student and the facilitator plain texts were uploaded into the Wmatrix 3 software and saved in folders (Appendix 5).

Table 5.2: Descriptive statistics of the study corpus

Group	Students' word count	Number of students/ group	Facilitators' word count	Number of facilitators/ group	Total word count/group
PBL 1	30, 068	7	5, 850	1	35, 918
PBL 2	30, 195	8	1, 225	1	31, 420
PBL 3	36, 783	7	1, 364	1	38, 147
PBL 4	17, 423	8	2, 103	1	19, 526
PBL 5	21, 279	9	3, 087	1	24, 366
PBL 6	15, 466	8	3, 294	1	18, 760
PBL 7	20, 075	7	7,303	1	27, 378
PBL 8	38, 788	7	3, 517	1	42, 305
Total	210, 077	61	27, 743	8	237, 820

5.13 Corpus Annotation

Corpus annotation is very important if the corpus is to yield maximum benefit to the user. Two types of annotation were carried out before data analysis: the pre-software annotation and software annotation.

5.11.1 Pre-software annotation

The pre-software annotation (Taylor and Gibbs, 2010) or corpus mark-up (Meyer, 2002) involves the addition of structural information to the text before automatic annotation with a software. Mark-up is purely descriptive, implying that it does not instruct the computer to perform any operations on a text but consists of mark-up codes that simply provide names to categorise parts of a document (Meyer, 2002). Mark-up codes are not considered as words by the software, and hence, are not included in the word lists or in any other analysis (Taylor and Gibbs, 2010). The advantage of structural mark-up is that it greatly facilitates the automatic analysis of a corpus (Meyer, 2002). Students were tagged with numbers to identify each student and the gender was identified with either M for male or F for female. Facilitators were designated with T and the group number was added (e.g. T1, T2, etc.).

5.11.2 Wmatrix software annotation

There are two different aspects to software annotation: the Tagset and the tagger. The Tagset refers to a group of symbols that represent various parts of speech. A tagger, on the other hand, refers to the software program that inserts particular tags that make up a Tagset (Meyer, 2002). The software used for corpus annotation in this research is Wmatrix 3. Wmatrix 3 is a software program that provides a web interface to the UCREL Semantic Analysis System (USAS) and Constituent Likelihood Automatic Word-tagging system (CLAWS). USAS is a semantic tagging tool, while CLAWS is a part-of speech (POS) corpus annotation tool. It also provides a web interface for standard corpus linguistic methodologies, for e.g., frequency lists, concordances, grammatical categories, and semantic domains (Rayson, 2003, 2009b). Corpus annotation was made automatically by the Wmatrix 3 software, using pre-existing dictionary.

CLAWS part-of-speech Tagger. Claws or part-of-speech tagger for English (or a grammatical tagging tool) was the first annotation tool to be developed by University Centre for Computer Research on Language (UCREL) at Lancaster. The POS tagging software has been in continuous development since early 1980s (http://ucrel.lancs.ac.uk/claws/). The consistent accuracy of CLAWS lies between 96–97% and the precise of degree of accuracy varies, depending on the type of text (Garside, 1987). When major categories are considered, the tagger has an error rate of only 1.5% with about 3.3% ambiguities unresolved within the BNC.

The CLAWS annotation system assigns part-of-speech designation to each word in a text. For example, in I'm doing the work: 'I' is assigned to the class of pronoun, 'm and doing are assigned to verbs, 'the' to the class of articles and work to the class of nouns. Wmatrix 3 used the latest version of the tagger (CLAWS) to POS tag about 100 million words in the British National Corpus (BNC). The process of POS tagging of the words in a text follows the Word Class Tagging Guidelines (available online: http://ucrel.lancs.ac.uk/bnc2sampler/guide_c7.htm).

UCREL Semantic Analysis System. This is a classification system in which each word or multi-word unit in a corpus is assigned a semantic field tag. The semantic fields categorise together word senses that are related because they are connected at some level of generality with the same mental concept (Figure 5.4). The words included in each category could be synonyms, antonyms, hypernyms, and hyponyms. At the present time, the lexicon contains about 37,000 words, while the template list holds over 16,000 multi-word units (Archer et al., 2002). The large size of lexicons of words and multi-word units is arranged into hierarchical multi-tier structure, containing 232 category labels that represent fine grained sub-divisions of 21 high-level categories. When compared with manual tagging, the classification system has a reported accuracy of 91%. This comprehensive categorisation is considered sufficient for the current research.

A general and abstract terms	**B** the body and the individual	**C** arts and crafts	**E** emotion
F food and farming	**G** government and public	**H** architecture, housing and the home	**I** money and commerce in industry
K entertainment, sports and games	**L** life and living things	**M** movement, location, travel and transport	**N** numbers and measurement
O substances, materials, objects and equipment	**P** education	**Q** language and communication	**S** social actions, states and processes
T Time	**W** world and environment	**X** psychological actions, states and processes	**Y** science and technology
Z names and grammar			

Figure 5.4: USA's high-level categories http://ucrel.lancs.ac.uk/usas/)

5.14 Corpus-based Analysis Process

The process of corpus-based data analysis consists of text transcription, corpus compilation, corpus annotation, development of analytic model or hypothesis to be tested and data analysis (Figure 5.5). Text transcription and corpus compilation and annotation have been discussed above. Next section discusses the analytic model, followed by a discussion of data analysis.

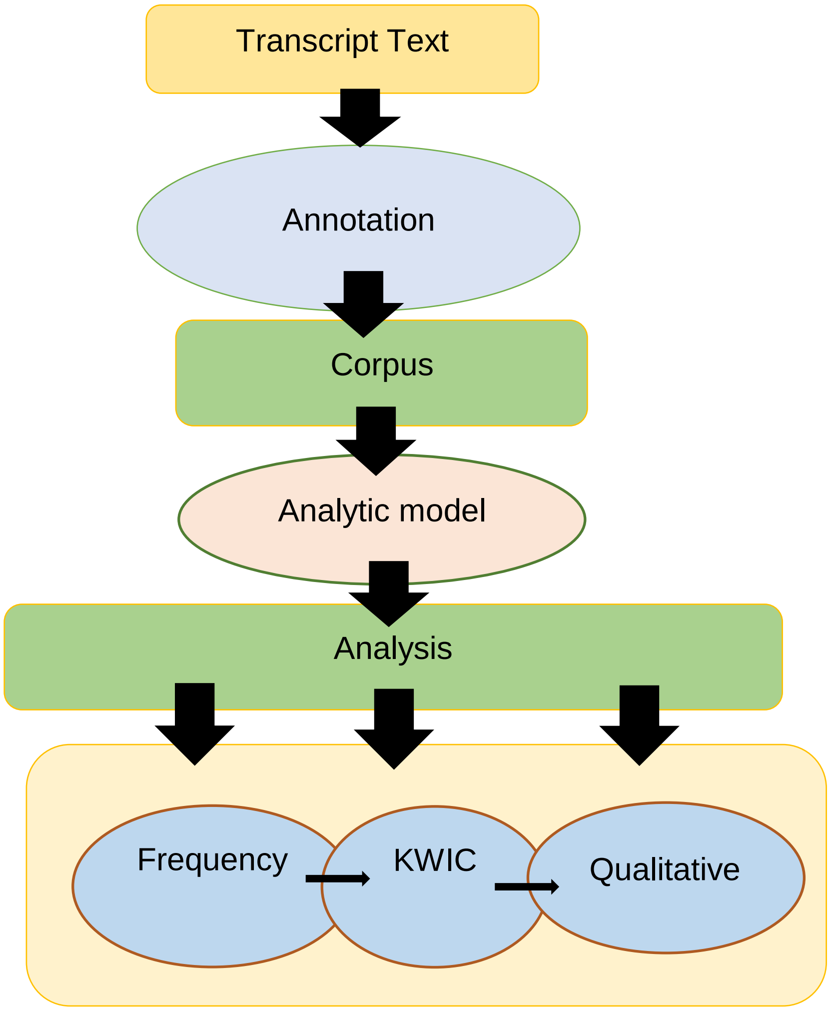


Figure 5.5: Corpus-based analysis procedure (adapted from Wallis and Nelson, 2001: p.312)

5.15 Analytic Conceptual Framework

Educational experts working from the sociocultural perspective believe that knowledge has historical and cultural origin and is negotiated, co-constructed, and scaffolded (Mercer and Littleton, 2007; Rogoff, 2003; Stahl, 2000, 2003a, 2003b, 2006; Vygotsky, 1930/1978). The model of collaborative knowledge construction here presented was developed by Stahl (2000). It allows for the identification of cognitive and social processes involved in collaborative knowledge construction. According to Stahl (2000), collaborative knowledge construction consists of mutually constituting cycles of personal knowledge construction and social knowledge construction. This implies that social processes constitute personal understanding and personal knowledge constitute collaborative understanding.

According Martin Heidegger (1927/1996) and Donald Schon (1993), cited by Stahl (2000) to learning starts from tacit pre-understanding. An individual experiences cognitive conflict as he encounters unfamiliar experiences that render some elements of the tacit pre-understanding problematic. The cognitive conflict can be resolved by reinterpreting meaning structures to arrive at new understanding that forms the tacit understanding for new learning. The resolution of the cognitive conflict is based on some feedback from the world such as experience with artefacts, such as our tools and symbolic representations that are elaborated Stahl (2000).

According to Stahl (2000), the process of personal meaning-making of an individual has its origin in prior social processes derived from interpersonal language and culture. The internal thought processes and structures are derived from prior social interactions (Vygotsky, 1934/1986; Vygotsky, 1930/1978). Our personal interpretive perspective is a consolidation of many perspectives and voices of others we have known (Bakhtin, 1986; Bazerman, 2004; Martin and White, 2007; Scollon et al., 2004; Staarman et al., 2003). At other times, the cognitive conflict may need to be resolved through social process, especially if it is provoked by others. For this to be accomplished, a process of explicit social process is required to create new meanings collaboratively by articulating beliefs in words and express ourselves in public statements (Stahl, 2000).

During this cycle, the processes of knowledge construction of several individuals are advanced and supplemented. An individual's personal belief is articulated in words. The public statement is taken up in social setting and discussed from multiple perspectives by the discourse participants (Stahl, 2000). According to Stahl (2000), the original statements are articulated, refined, elaborated, and subjected to conflicting interpretations, during extensive discussion that involve argumentation, clarification, and negotiation with the discourse gradually converging to shared understanding. The accepted common result that emanates from the negotiation of different perspectives is referred to as ***collaborated knowledge***. Thus "collaboration and undistorted communication mediate between personal belief and accepted knowledge (Stahl, 2000: p. 72). The accepted knowledge forms the tacit pre-understanding of an individual for future discourse.

Conditions for knowledge construction. Collaborative knowledge construction can be defined as a situation in which particular forms of interactions among students are expected to occur, which would trigger knowledge construction mechanisms, but there is no guarantee that the expected interactions will actually occur (Dillenbourg, 1999). The general tendency is to increase the probability of the collaborative interactions that would trigger knowledge construction mechanisms. Several strategies that could promote co-construction of knowledge in

PBL tutorial have been identified by experts in the field (Barrett and Moore, 2011b; Barrows, 1992; Halela and Fagerholm, 2008; Hmelo-Silver and Barrows, 2006, 2008), and they fall into two main categories:

Students' collaborative knowledge construction. To construct knowledge together, it is important that students: benefit from one another's prior learning; brainstorm ideas without censoring; engage in high-quality independent study; evaluate information sources critically; justify arguments with evidence; confirm one another; build on one another's ideas; review their learning; move from their current position based on evidence and ideas; make links between different concepts; work in a range of media; and compare the depth and breadth of their own knowledge with that of other students.

Facilitation of knowledge construction. To promote knowledge construction among the students, the facilitator: must help create a stimulating, encouraging, enjoyable, and warm atmosphere in the tutorial team; must be nondirective; ask open ended questions, and provide hints and prompts rather than providing explanations; help develop critical thinking in the students; help student reflect on their experiences; and help students monitor their thinking processes.

5.16 Language model of knowledge construction

To operationalize the study and help answer the research questions, a language model of collaborative knowledge construction is required. Mercer (1995) suggests that a sociocultural model of guided knowledge construction of knowledge requires an explanation of how language is used to collaboratively construct knowledge and share understanding. Language analysis of conversational data is a vast and problematic field (Schiffrin, 1987). This is because language analysis could be done from many theoretical perspectives (Schiffrin, 1987), many lexicogrammatical tokens could realise a set of indicators and it is impossible to analyse all the linguistic features that can realise a set of indicators (Holtz, 2011). Therefore, it is important to define the assumptions underpinning this study and provide a model of language analysis. This will provide orientation for the analysis and define the framework upon which result interpretation and general conclusions would be based.

5.17 Assumptions underlying Language Model

1. Language encodes knowledge (Gee, 1994; Halliday, 1978), information (Halliday and Matthiessen, 2004), and ideas (Schiffrin, 1987).

2. Language is used to create knowledge**.** As Halliday theorised, there is a close relationship between language and knowledge development (Halliday, 1975, 1978): “Language is not a domain of human knowledge ….; language is the essential condition of knowing, the process by which experience becomes knowledge.” (Halliday, 1993: p. 94).

3. Knowledge construction processes have linguistic markers that are recoverable from the language resources employed by the collaboration participants. Information and ideas in a talk may be marked with linguistic elements (structural units) or they may be unmarked (non-

structural) (Halliday and Matthiessen, 2004). The structural markers may be recovered for analysis (Fontaine, 2013).

4. Language has ideational and interactional contents. That is, conversation fulfils content and interactional purposes simultaneously (Fontaine, 2013; Schiffrin, 1987).

5. Language is context sensitive. This implies that language is used in a context and the meaning of a structural form depends on the other adjoining structural elements (Mercer, 2004; Mercer et al., 2004).

5.18 Dialogue and Knowledge Construction

The dialogic notion of knowledge involves more than treating knowledge as an exchange of ideas and viewpoints – it is an epistemological position that views knowledge as something that is constructed together socially (Barrett and Moore, 2011b). Dialogic knowledge is created, reinforced, elaborated, and developed during problem based learning tutorial discourse. Dialogue enables the tutorial participants to work through the problem and thus proceed to a conclusion by reason and argument. The success of this process depends on the nature of the students' talk in the tutorial.

According to Halliday and Matthiessen (2004), language serves as the tool for representing experience and externalize it, and it serves as the means to structure ideas and information together in complex ways. Experience are represented in language as clauses, groups, phrases, and words (Halliday and Matthiessen, 2004: p. 88). According to Halliday (2002) and Halliday and Matthiessen (2004), the process of knowledge construction entails more than individuals adding ideas together. The process involves (1) coordination through which assertions are clarified and restated; (2) expansion through the process of addition, alternation and contradiction; and (3) enhancement by which assertions are expanded through connection with conditional, concessional, exceptional, antecedential, consequential, analogical, simultaneous, ordered, durative, and spatial linguistic elements. Furthermore, language is used to mobilise prior experiences and other cultural artefacts through naming and referents and serves the purpose of attaining mutual understanding.

The language-based facilitation model used in the study was also inspired by the work of Halliday (Halliday, 1975). He wrote extensively on the enculturation process of a child through scaffolding, idea negotiation, and modelling by an adult.

5.19 Operationalising the Research Questions

The philosophy of PBL proposes that the knowledge construction process is expansive and informed by prior knowledge. The facilitator is required to function in a nondirective manner, to uphold expansive knowledge ideals, and guide students through questioning. Knowledge is constructed as students make contributions to the tutorial discourse. This study focuses on retrieving the linguistic evidence of these propositions from the PBL tutorial discourse.

The first step is the development of research questions, followed by the transformation of the questions into a set of specific behaviours that can be measured (Cohen et al., 2011). This is followed by the enabling of corpus analysis by clearly developed working definition of the grammatical constructions to be studied in such a way that they can easily be extracted from the data set. This will prevent inconsistent analytic process and thus avoid bias results (Meyer, 2002).

5.19.1 Operationalised Questions

Each sub-research question derived from the elements of the conceptual framework is explained and outlined below:

Prior knowledge. The first key concept in collaborative knowledge construction is prior knowledge. The reference to previous experience and knowledge can be made through naming (Clark and Wilkes-Gibbs, 1986a) or reporting of what was said or thought (Halliday and Matthiessen, 2004). Thus, referring markers (nouns and reporting tokens) constitute a set of indicators for analysis.

Operationalised sub-question 1.

To measure the frequencies and describe the functions of the frequently occurring referring expression indicators (e.g., think, say, talk).

Shared knowledge. One of the goals of group learning is for the individual participants to share the knowledge that results from joint efforts of discourse participants. Collaboration is a process, wherein conversational interactions could lead to shared knowledge. Shared knowledge is the process through which discourse participants share mutual understanding through social interactions (Clark and Schaefer, 1989; Schegloff, 1982; Teasley et al., 2008). After an iterative process of refinement of ambiguity and partial meanings, participants in interactional situation come to share meanings (Oliveira and Sadler, 2008). Shared meaning does not imply that the knowledge of the participants is identical, but it suggests overlap of conceptions (Fischer and Mandl, 2005; Roschelle, 1992).

Shared knowledge could be process convergence or outcome convergence. This study analysis process convergence that is studied in the context of grounding (Clark and Brenna, 1991 cited in Teasley et al., 2008). In a conversational interaction, the speaker and the listener need to explicitly demonstrate that they understand each other. This could be done in several ways, such as concept completion, mutual elaboration, and simple affirmative acknowledgement, and repetition, expression of reactive tokens or smooth regulatory move to the next topic. On the other hand, the participants could seek evidence of understanding by asking questions or requesting repairs (Oliveira and Sadler, 2008; Teasley et al., 2008). Thus, evaluation of shared knowledge markers represents another set of indicators for collaborative knowledge construction.

Operationalised sub-question 2.

<u>To measure the frequencies and describe the functions</u> of the commonly occurring shared knowledge indicators (e.g. yes, yeah, no).

Constructing knowledge. According to Fontaine (2013) and Halliday and Matthiessen (2004), the language structural units (word, group, phrases and clauses) are used to represent knowledge during conversation. The structuring of these elements in complex ways to produce knowledge have been compared to the weaving of clothes from yarns (Halliday and Hassan, 1976). The language structural units are structured together with connectives (Fontaine, 2013; Halliday and Hassan, 1976; Halliday and Matthiessen, 2004; Van Dijk, 1977). Van Dijk (1977) categorised these connectives into five units – conjunctions (e.g., and, because, or), sentential adverbs (e.g., yet, nevertheless), prepositions (e.g., due to, in spite of), interjections, and particles (e.g., you know) and predicates (e.g., to conclude, it follows that). However, in this study attention was focused on connectives from conjunction category. According to Van Dijk (1977), these natural language connectives are classified as conjunction, disjunction (or alternation), contrast, concession, condition, causality, and reason and circumstantial (time, place, manner). Halliday and Matthiessen (2004) described the structuring together of language structural units as expansion. This can be done, says Halliday and Matthiessen (2004), through elaboration that does not involve adding new information, extension that involves adding new information or varying the existing information, and enhancement in which one clause (or other structural unit) enhances the meaning of another by qualifying the preceding clause in some ways such as by reference to time, place, manner, cause or condition. This study concerned the extension and enhancement categories. In Halliday and Matthiessen (2004) scheme, extension category encompasses the conjunction, alternation, and contrast categories of (Van Dijk, 1977), while enhancement subsumed concession, condition, causal and circumstantial categories. In discourse situation, knowledge construction could be autonomous in which main clause, and the contingent clause are contributed by the same speaker or it could be co-constructed in which case the main clause is generated by one speaker while the contingent clause is contributed by another speaker (Lustigman and Berman, 2016). Thus, the explicitly co-constructed knowledge, in this sense, is considered collaborative knowledge construction. Based on this idea of collaborative knowledge construction, the markers of knowledge extension and knowledge enhancement could form a set of indicators for investigation.

Operationalised sub-question 3

To measure the frequencies and describe the functions of the commonly occurring knowledge extension indicators (e.g., and, or, but).

Operationalised sub-question 4

To measure the frequencies and describe the functions of the commonly occurring knowledge enhancement indicators (e.g., because, so, if, since).

5.19.2 Problem-based Learning Facilitation

The role of the facilitator is central to the success of the PBL process (Barrows, 1988; Hmelo-Silver and Barrows, 2006, 2008). The facilitator is in the position to model the key features of problem based learning philosophy. Thus, the following themes were explored to investigate the role of the facilitator:

Guided questioning. This involves how the facilitator scaffolds knowledge construction by guiding students through questioning. The types of questions asked by the facilitator could serve as empirical evidence of how the facilitator guides collaborative knowledge construction activities.

Operationalised sub-question 5

To measure the frequencies and describe the functions of commonly occurring facilitators' questions indicators (e.g., what, why, how, do, does).

Stance of the Facilitator. The constructivist philosophy of PBL requires the facilitator to be nondirective in their attitude to problem-based learning tutorial facilitation (Hmelo-Silver, 2003b; Hmelo-Silver and Barrows, 2008). Markers of necessity and obligation in conversation include should, need to, must, and the like. Thus, the analysis of linguistic devices of necessity and obligation could provide a lens to view the directive, otherwise a facilitator guides knowledge construction.

Operationalised sub-question 6

To measure the frequencies and describe the functions of the commonly occurring facilitators' directive expression indicators (e.g. 'have to', 'need to', should).

The next concept in collaborative knowledge construction is that knowledge is improvable. This is the process by which knowledge is created, restructured, and improved. Improvable and expandable knowledge are marked with linguistic tokens of possibility and opinions, such as probably, possibly, I think, and the like, while nonexpendable knowledge are marked with absolute and certainty linguistic tokens (Martin and White, 2007). The knowledge markers constitute another set of indicators for analysis.

Operationalised sub-question 7

To measure the frequencies and describe the functions of the commonly occurring facilitators' probability indicators (e.g. may, probably, possibly).

5.20 Selection of Indicators for Analysis

Many lexicogrammatical tokens may realise each set of the indicators but as it is impossible to investigate all these linguistic features (Holtz, 2011), a decision was made regarding which set of indicators should be investigated. The decision to select the linguistics features for investigation was based on relevant literature (Biber, 2006; Eggins, 2005; Halliday and Matthiessen, 2004; Thompson and Hunston, 2000; Trebits, 2009).

Furthermore, two issues need to be noted regarding the focus of this study, which are: (1) it is a structural analysis as against non-structural analysis. (2) It is corpus-based (concept-driven) as against corpus-driven (data-driven). In structural analysis, language structure is identified and its function is retrieved based on its context. For example, there may be an elaborative link between two clauses, but if there is no conjunction of elaboration between the two clauses to indicate this, the relationship between the two clauses cannot be retrieved and analysed. Corpus-based research test theory/hypothesis with data unlike corpus-driven research that involves generating theory or hypothesis from data. As this research employs a corpus-based approach, the collaborative knowledge construction indicators investigated in this project were based on previous categories found in the literature. The Wmatrix 3 tagging domains are described in Table 5.2.

Table 5.3: Semantic and part-of-speech categories investigated

Indicator	Tag	Domain	Examples
Referring expression	Q1.2	Paper and documents	Notes/diagram
	Q2.1	**Linguistic actions**	Say, talk
	Q2.2	Speech acts	Lecture (s)
	Q3	Speech and grammar	Read/reading
	Q4.1	The Media: Books	Book, textbooks
	Q4.2	The Media: Papers etc.	Journal article
	X2.1	Thought, belief	Think, thought
Shared knowledge	UH	Interjections	Yes, No, oh
Knowledge construction	CC	Additive conjunctions	And
	CCB	Adversative conjunctions	Or
	CS	Subordinating conjunctions	Because, if, so
	CSA	As (as conjunction)	As well as
	CSN	Than (as conjunction)	More than
	CST	That (as conjunction)	that
	CSW	Whether (as conjunction)	Whether or not
Facilitator's question	?	Simple interface search	What, why, how
Facilitator's stance	A7	Probability	Perhaps, maybe
	A7+	Likely	Could, may
	A7-	Unlikely	impossible
	S6+	Obligation or necessity	Must, have to
	S6-	No obligation or necessity	Not compulsory

5.21 Unit of Analysis

This is defined as "the basic unit or focus of a researcher's analysis – typically individual study participants (Polit and Beck, 2012: p. 745). Corpus-based studies concern two types of research goals – describing a linguistic structure and its variants or describing some group of texts. Thus, corpus-based studies typically have two kinds of unit of analysis – occurrence of a linguistic feature or a text (Biber et al., 1998). According to Biber et al. (1998), these units of analysis are called observations for the study. Each observation is an occurrence of the structure in a study that characterises a linguistic structure. In a study that seeks to describe a group of texts, each observation is a text and a unit of observation. The latter unit of analysis type – a text – is followed in this study. It allows normalised count, or rates, of different linguistic features to be analysed as characteristic of each text. The linguistic features analysed are linked to the research questions.

This unit of analysis has been chosen for several reasons. By definition, a corpus is a collection of texts. Reppen (2012) illustrates that in an in-class writing, a text could be defined as all the essays written in the class on a particular day, or a text could be each student's essay. Krippendorff (2004)p. 97 describes units of analysis as "wholes that analysts distinguish and treat as independent elements. The wholeness of a unit of analysis suggests that it is not further divided during an analysis or at a particular stage of an analysis." In this study, the text for each tutorial group talk was divided into two units– students' subcorpus and facilitator's subcorpus. Thus, the text produced by the students and the facilitators were treated as separate units of analysis. It is considered that analysing the students' contributions together will provide better understanding for their interactive verbal actions.

5.22 Data Analysis with Wmatrix 3 Software

The uploaded plain files were automatically annotated by Wmatrix 3 software. The software annotation assigns semantic and grammatical values to each word or multi-word unit and classifies them in hierarchical multi-layered structures of language use. The annotation performed by the software is a generic, fine grained data characterisation, and it is independent of the research focus or the researcher's opinion. This situation makes it mandatory to apply a framework that reflects the research focus and questions to the data. The focus of the analysis in this research relates to the meaning and uses of the words in the text. Therefore, a few tag categories were selected in line with the research questions (Table 5.2). Data combines macro- and micro- levels of analysis in iterative manner – type III approach (Rayson, 2003, 2008).

5.23 Methods for Extracting the Indicators

Following corpus compilation, it is then necessary to plan out exactly what kinds of grammatical information to be extracted from the corpus and to determine how the information will be coded and recorded in a way that can most efficiently help to find the linguistic features being investigated in the corpus (Meyer, 2002).

To extract the indicators relating to the research questions from the corpus, this study used Rayson's (Rayson, 2003, 2009b) corpus analysis tool, Wmatrix 3. For the analysis of the facilitator's questions, the question mark was searched in the simple interface of the software (Appendix 6). This generated the concordance lines of the questions marks (Appendix 7). For other indicators, the retrieval was done at the part-of-speech and semantic tag level. The relevant part-of-speech (Appendix 8) and semantic tag (Semtag) domains (Appendix 9) were opened. This shows the frequency lists of the POS or Semtag categories (Appendix 10). The frequency list of the relevant POS or Semtag domain was then opened to generate the concordance lines (Appendix 11). A mixture of searches on Wmatrix 3 simple interface and advanced interface searches were used. The choice of one over the other was based on practical expediency.

Simple interface search. This was used to retrieve the questions asked by the facilitators. The whole group transcripts were carefully read and all the questions were marked with question marks. The questions were retrieved by searching for the question marks in the simple interface of Wmatrix 3.

Part-Of-Speech Tag (POS) domain. The coordinating and subordinating conjunctions and interjections indicators were retrieved from the POS domain on the advanced interface of Wmatrix 3. They are UH, CC, CCB, CS, CSA, CSN, CST and CSW (Table 5.2)

Semantic Tagging domain. The indicators for referring and stance expressions (Table 5.2) were retrieved from the semantic tagging domain of the advanced interface of Wmatrix 3.

5.24 Data Preparation for Analysis

After the frequency list for either POS (Appendix 8) or Semtag (Appendix 9) has been generated, the relevant item on the frequency list was opened to generate a word frequency list of the chosen part of speech (Appendix 10). For example, in Appendix 8, the ‘CS’ (i.e. subordinating conjunction) was opened to generate the list of words and their frequency (Appendix 10). The 'CS' was also opened to generate the concordance lines of the words (Appendix 11). The concordance lines were then exported into Excel file (as in Appendix 12).

5.25 Processing of KWIC Output

In the Excel files, the concordance lines were examined and the mistagged ones were removed. The resulting list of the keywords for the POS tag was considered as the frequency list for the indicator. The manual inspection of the Wmatrix 3 output is important because the accuracy of the part-of-speech tagging is 96-98% while that of semantic tagging is 91% (Rayson, 2003, 2009a). The data analysis procedure in this study was influenced by the work of (Semino et al., 2013; Semino et al., 2015).

5.26 Corpus Analysis Procedure

5.26.1 Quantitative Analysis

Quantitative analysis was carried at two levels: Word frequency analysis and key word in context (KWIC) analysis.

Word frequency analysis. After the irrelevant tokens have been removed from the word frequency list, the list of the remaining words was considered as the raw frequency for the indicator concerned. The relative frequency of each token in the frequency list was calculated per 100 tokens. For example, assuming the students' word count for PBL group 1 (PBL1) is 30, 068; if the raw frequency for indicator 'because' is

45, the relative frequency was calculated as (45x100) ÷ 30068 to give 0.15.

Key Word in Context (KWIC) analysis. This concerns a more in-depth analysis of the indicator in the context of the surrounding words. This was carried out to gain a better understanding of the function of the indicator as informed by the other words around it. KWIC (or concordance) analysis procedure involved reading through each line of the concordance in the Wmatrix software 3, expanding the scope of the lines (i.e. left and right co-texts) as required or fully open the transcript in the Wmatrix 3 (Appendix 13). The keyword that is being read in context is indicated in red colour. On occasions, sections of the whole group transcript in word file were read to grasp the meaning of the indicator. To facilitate KWIC analysis, the columns of the excel file were labelled as Left co-text, key-word-in context, right co-text, function of the keyword and type of knowledge construction (Appendix 12). A coding scheme developed for each research question was applied to the categorisation and subcategorization of the indicator functions.

The development of the codes for KWIC analysis is based on an initial concept-driven and later data-driven strategies. It is almost impossible to create coding frame that is purely concept-driven or purely data-driven (Schreier, 2012). Main coding categories were based on the tagging of Wmatrix 3. The first step in the generation of the sub-category codes is concept-driven in that the categories were based on previous literature (e.g. Amidon, 1976; Barrie-Blackley, 1973; Beach and Anson, 2004; Bloom et al., 1980; Bolden, 2009; Carter and

McCarthy, 2006; Clark, 1971; Clark and Wilkes-Gibbs, 1986a; Dagher, 1992; Ehri and Galanis, 1980; Emerson and Gekoski, 1980; Feagans, 1980; French, 1988; Johnson and Chapman, 1980; Katz and Brent, 1968; McCabe and Peterson, 1985, 1988; Pappas et al., 2004; Peterson and McCabe, 1988; Schegloff, 1982; Short, 2004; Silva, 1991; Thompson and Mulac, 1991; Yaguchi, 2001). In the second step, emerging subcategory codes were generated based on the data and added to the initial codes.

5.26.2 Qualitative.

This is an extension of key-word-in-context (KWIC) analysis. It involves determining whether the knowledge construction indicator is representing autonomous construction or co-construction. The knowledge construction indicator that occurred within the statement of a participant is considered as autonomous knowledge construction indicator (Appendix 12). If the indicator occurred at the turn-initial (Cheepen, 2000; McCarthy, 2003), in response to contribution of previous speaker, it is classified as co-construction indicator. The analysis also involves providing extracts from the transcripts to demonstrate how the functions of the indicators were framed in the dataset. This provided a richer understanding of the functions of the commonly occurring indicators.

5.26.3 Statistical Analysis

The output of data analysis on the excel file was then prepared and exported into the statistical package for social sciences (SPSS) version 22 software (Appendix 14). The SPSS software was used for descriptive analysis of the frequency of each indicator and function types for each PBL group. The resulting frequency results were exported into the excel file for computing the overall raw frequency and relative frequency (NF) for the students' and facilitators' subcorpora. The overall raw and relative frequencies for the indicator functions were similarly calculated. Statistical significance of relative frequency variation was estimated using Log Likelihood (LL) calculator. The critical LL level was set at 6.63 equivalent to pv< 0.01. Overall relative frequency of the indicators, and the functions were reported as bar charts. Thus, indicators and functions with relative frequency of zero were not represented on the bar chart.

5.27 Reporting Research Findings

A major challenge that I faced in reporting the results of this research is the limitation of space for the massive data generated. Thus, the results of the analysis were reduced using the strategy found in the literature for corpus linguistics research. Gabrielatos and Baker (2008) reported only the interesting findings in their results because of limited space: "Due to issues of space, we cannot report all the results of our analysis in this paper, but instead have summarised some of the most interesting findings…" (2008: p. 17). Demmen et al. (2015) for similar reason set the threshold for including a token in the analysis at 10 or more frequencies in the combined data: "…and for reasons of space, we confine our analysis to only those which occurred ten or more times in any stakeholder group…." (2015: p. 213). Thus, in this study, indicators that are more than zero are included in the bar chart. The functions that are most or commonly occurring on relative frequency calculation are reported on the bar chart and explored further with qualitative analysis. The token frequencies on relative frequency analysis are divided into three categories which are: (1) most commonly occurring tokens occurred more than 0.1 per 100 tokens; (2) commonly occurring tokens have relative frequencies of 0.05 to 0.1 per 100 tokens; (3) and les commonly occurring tokens have relative frequencies of 0.04 and below per 100 tokens.

5.28 Rigor and quality criteria

Several steps were taken to ensure the validity and reliability of the study. The study was reviewed with my supervisor at intervals to ensure construct and content validity. The coding frame in the study is considered to have met the exhaustiveness criterion. Coding frame is a system for grouping the data into variables and sub-variables according to some defined criteria (Fruh, 2007; Holsti, 1969 cited in Schreier, 2012). The criterion of exhaustiveness was considered to have been met because of the categorisation of the research themes through software tagging and at least one subcategory in the coding frame was used. Saturation was considered to have been reached when no new subcategory emerged during data analysis.

The coding of functional subcategories was explicitly defined, and this coding scheme (Appendix 15) was used to guide key word in context

(KWIC) data analysis. Any emerging functional subcategories were similarly defined in an explicit way.

The size of the corpus is another quality criterion. This depends on the research objectives, and it is considered that the size of the corpus used in this study is large enough to allow drawing appropriate conclusion from the results of the research. The transcripts were also thoroughly read to correct transcription errors and unify spellings.

I underwent a training workshop, involving hands-on training in the use of Wmatrix 3 software for corpus linguistics analysis at the Department of Linguistics, University of Lancaster, UK (Appendix 16). The statistical analysis with Wmatrix 3 was reviewed with Dr Paul Rayson, who designed Wmatrix software (Appendix 17).

5.29 Ethical Issues

The ethical practices guiding research procedure with respect to ethical issues, including informed consent, voluntary participation, confidentiality and anonymity, and respect for the learning environment were observed during this research.

Informed Consent. All the students and the facilitators need to give a consent to be able to conduct the research. Therefore, all the students and the facilitators in year-one of the GEM programme were individually invited to participate in the study. The participants were given the information sheet (Appendix 1) about the project, and they signed a consent form (Appendix 2) for the audio and video recordings and for transcription of the audio recordings and the use of the transcripts for research purposes.

Voluntary participation. The participation in the study was voluntary. The participants were fully aware that they could exercise the right to withdraw from the study without giving any explanation and without any adverse effect to their status in the future.

Confidentiality & Anonymity. Participant confidentiality is very important in any research process. Video and audio files are very sensitive materials, because they contain a high level of personal and identifiable information that need to be respected and concealed. To safeguard the confidentiality and anonymity of the research participants, the video materials containing the images of the participants were kept in password-protected folders in the office of the supervisor, and they were only accessible to the researcher and the supervisor. These materials are not to be used in any conference, presentation, or educational materials without prior individual permission from the participants for such use. Audio files were, similarly, kept in secure condition and password protected in the office of the supervisor. Audio recordings were transcribed verbatim by a professional transcriber in line with the prior consent given by the participants. The responsibility for audio and video recording was given to the students. They were advised to stop the recordings (audio and video) when any confidential or sensitive issue was discussed. This was to further safeguard the confidentiality and anonymity of the participants. At no point was any part of the recordings shared with the student assessors or evaluators of the facilitators.

Respect for Research Site. Respect for the learning environment is an important aspect of an educational research (Creswell, 2012; Lincoln, 2009). Research conduct is an intrusive process on the research site or context. It is the responsibility of the researcher to ensure that research does not unduly interfere with the learning of the students. In line with this principle, the researchers did not stay with the students but handed over the responsibility of the recordings to the students. The study is observational in nature hence no alteration was made to the learning procedure or arrangement and there was no interference with the institutional culture.

Ethical Approval. The study is a part of a larger research project conducted by the Chief investigator and approved by the University of Nottingham Ethics Committee (Appendix 3). The number for the Ethics approval for the study is **D/9/2008.**

5.30 Summary of the chapter

This chapter summarises how the data were acquired and the problems associated with constructing the corpus for the study. It also describes the procedure for analysing students' verbal interactions and facilitators' contributions in eight tutorial groups to understand how the students jointly construct knowledge and how the facilitators' guide the process. A CL methodology, with Wmatrix 3 software, was employed for data extraction and knowledge construction analysed using lexicogrammatical analysis method. The statistical analysis and the study quality criteria are discussed. The results of data analysis are presented in the next two chapters.

CHAPTER 6

RESULTS OF STUDENTS' TALK ANALYSIS

6.1 General Overview of Results Presentation

The purpose of this study is to use corpus linguistics methodology and the Wmatrix 3 software analytic tool to explore how the GEM students at the Derby Medical School of the University of Nottingham collaboratively construct knowledge in problem-based learning tutorials and how the PBL facilitators guide this process of knowledge construction. Thus, two groups of participants are involved in this study.

The research findings are reported in two chapters, each describing the findings of the analysis of speech from these two groups of participants. The first of these chapters (i.e. the current chapter) presents the results of the analysis of the students' conversation while the second chapter (Chapter 7) presents the findings of the analysis of the facilitators' talk.

Male student participants were designated as M, female student participants as F, and the facilitators were labelled as T. Participants M1 to M6 represent male students and participants F1 to F7 make up the female students and both make up the participant group of students. Participants T1 to T8 make up the participant group of facilitators.

6.2 Results I: Students' Talk

A total of 63 students participated in the study, out of which 35 were male and 28 female. The PBL group distribution of the students is shown in Table 6.1. The number of students per group ranged between 7 and 10.

Table 6.1: Distribution of student participants

Participant								*PBL Group*	*Total*
	1	2	3	4	5	6	7	8	
Male	4	4	4	6	5	4	4	4	35
Female	3	3	3	3	3	6	3	4	28
Total	**7**	**7**	**7**	**9**	**8**	**10**	**7**	**8**	**63**

Students' subcorpus description. The students' subcorpus consisted of 210,077 words. This represents 88.33% of the study corpus of 237,820 words. The semantic tags and domains of the students' subcorpus were compared with the BNC sampler CG Educational spoken corpus. The relative frequencies of the five most frequent semantic tags and domains are presented in Table 6.2 (below). The figures in the table show that the contents of the students' discussions frequently related to anatomy, physiology, diseases, medicines, medical treatments, and cause-and-effect connections.

Table 6.2: Frequent semantic domains in students' subcorpus: Relative frequencies of the five most commonly occurring semantic tags and semantic domains in the students' subcorpus

	PBL1	*PBL2*	*PBL3*	*PBL4*	*PBL5*	*PBL6*	*PBL7*	*PBL8*
B1	1.72	1.98	2.05	3.80	3.38	3.52	2.78	1.46
B2-	1.02	1.04	1.66	1.96	2.57	1.81	2.17	2.08
B3	0.77	0.68	0.88	–	0.69	–	–	0.97
A2.2	0.81	0.74	–	1.26	1.50	1.10	1.20	0.87
Z4	2.80	3.45	2.60	2.06	2.15	2.13	3.01	2.89
F3	–	–	0.44	–	–	–	–	–
O1	–	–	–	0.45	–	0.64	–	–
A2.1+	–	–	–	–	–	–	0.74	–

B1 – anatomy & physiology; B2 – disease; B3 – medicines & medical treatment; A2.2 – cause & effect connection; Z4 – discourse bin; F3 – smoking & non-medical drugs; O1 – chemical substances in general; and A2.1+ – change.

The students' talk was analysed to answer the following research questions:

1. What are the frequencies and functions of the commonly occurring referring expression indicators in the students' talk?

2. What are the frequencies and functions of the commonly occurring shared knowledge indicators in the students' talk?

3. What are the frequencies and functions of the commonly occurring knowledge extension indicators in the students' talk?

4. What are the frequencies and functions of the commonly occurring knowledge enhancement indicators in the students' talk?

6.3 Results Presentation Approach

The presentation of the results is divided into two parts: (1) quantitative and (2) qualitative. For each question, the results of quantitative analysis are first presented, followed by a presentation of the results of the qualitative analysis.

Analysis of referring expression indicators

Quantitative analysis

The methodology described for the retrieval of indicators led to the identification of 2,325 reporting indicators in the students' subcorpus. The raw frequency of the indicators across the PBL groups is shown in Table 6.3 (below).

Word frequency analysis. The reporting indicators could be divided into three broad groups – verbal and mental indicators, and learning situation, and curricular materials. There are 1, 024 (44.04%) verbal reporting indicators and ‘say’, ‘says’, ‘saying’ and ‘said’ indicators are most frequently used. Mental reporting tokens in form of ‘thinking’, ‘thinks’, ‘think’, and ‘thought’ account for 982 (42.24%) of the indicators.

Table 6.3: Raw and relative frequency of referring indicators

	PBL1 RF NF	PBL 2 RF NF	PBL 3 RF NF	PBL 4 RF NF	PBL 5 RF NF	PBL 6 RF NF	PBL 7 RF NF	PBL8 RF NF	Total RF NF
Saying	15	20	31	13	21	8	17	47	172
	0.05	0.07	0.08	0.07	0.10	0.05	0.08	0.12	0.08
Say(s)	49	38	46	8	35	15	38	102	331
	0.16	0.13	0.13	0.05	0.16	0.10	0.19	0.26	0.16
Said	28	22	50	17	36	17	19	45	234
	0.09	0.07	0.14	0.10	0.17	0.11	0.09	0.12	0.11
Point	7	17	11	1	2	1	12	30	81
	0.02	0.06	0.03	0.01	0.01	0.01	0.06	0.08	0.04
Talking	10	8	20	7	10	2	10	32	99
	0.03	0.03	0.05	0.04	0.05	0.01	0.05	0.08	0.05
Talk(s)	4	3	4	0	2	3	5	21	42
	0.01	0.01	0.01	0.00	0.01	0.02	0.02	0.05	0.02
Talked	3	2	0	3	4	0	5	5	22
	0.01	0.01	0.00	0.02	0.02	0.00	0.02	0.01	0.01
Told	5	4	10	6	4	5	4	5	43
	0.02	0.01	0.03	0.03	0.02	0.03	0.02	0.01	0.02
Thinking	17	20	21	1	4	6	11	17	97
	0.06	0.07	0.06	0.01	0.02	0.04	0.05	0.04	0.05
Think(s)	95	119	166	80	47	50	68	141	766
	0.32	0.39	0.45	0.46	0.22	0.32	0.34	0.36	0.36
Thought	18	8	16	7	11	22	11	26	119
	0.06	0.03	0.04	0.04	0.05	0.14	0.05	0.07	0.06
Diagram(s)	5	9	2	2	4	4	3	1	30
	0.02	0.03	0.01	0.01	0.02	0.03	0.01	0.00	0.01
Notes	1	1	6	0	2	0	2	0	12
	0.00	0.00	0.02	0.00	0.01	0.00	0.01	0.00	0.01
Library	1	1	0	0	0	0	0	0	2
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Book	4	8	4	2	0	0	6	10	34
	0.01	0.03	0.01	0.01	0.00	0.00	0.03	0.03	0.02
Dictionary	0	0	0	4	0	9	0	1	14
	0.00	0.00	0.00	0.02	0.00	0.06	0.00	0.00	0.01

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

Table 6.3: Raw and relative frequency of referring indicators (cont.)

	PBL1 RF NF	PBL 2 RF NF	PBL 3 RF NF	PBL 4 RF NF	PBL 5 RF NF	PBL 6 RF NF	PBL 7 RF NF	PBL8 RF NF	Total RF NF
Article	1 0.00	0 0.00	2 0.01	0 0.00	0 0.00	1 0.01	2 0.01	14 0.04	20 0.01
Website	2 0.01	0 0.00	4 0.01	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	6 0.00
Graphs	0 0.00	4 0.01	4 0.01	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	8 0.00
Leaflets	0 0.00	0 0.00	2 0.01	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	2 0.00
Cards	0 0.00	0 0.00	2 0.01	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	2 0.00
Website	0 0.00	0 0.00	4 0.01	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	4 0.00
Lecture(s)	11 0.04	13 0.04	24 0.07	4 0.02	25 0.12	3 0.02	3 0.01	13 0.03	96 0.05
Reading/ read	11 0.04	12 0.04	15 0.04	0 0.00	5 0.02	7 0.05	5 0.02	30 0.08	85 0.04
Clinical skills	2 0.01	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	2 0.00
Workshop	2 0.01	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	2 0.00
Total	**291 0.96**	**309 1.02**	**444 1.21**	**155 0.89**	**212 1.00**	**153 1.00**	**221 1.10**	**540 1.39**	**2325 1.11**

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

Indicators relating to learning situations and materials were less common, accounting for 319 (13.72%) of the total indicators. Of the group of indicators pertaining to learning situation and curricular materials, ‘reading’, ‘read’ and ‘lecture(s)’ occurred more commonly. More than half of the total indicators (55.61%) were derived from PBL2, PBL3, and PBL8.

Of the verbal reporting indicators, six were consistently used across the PBL groups. The variation was statistically significant for each item at the level of LL 6.63, p< 0.01: ‘saying’ – LL 23.45; ‘say(s)’ – LL 66.31; ‘said’ – LL 16.35; ‘point’ – LL 12.81; ‘talking’ – LL 17.92; and ‘told’ – LL 10.25. All the mental indicators, i.e. ‘thinking’, ‘think(s)’, and ‘thought’, were consistently used across the PBL groups. The variation of use of each token across the PBL groups was statistically significant: ‘thinking’ – LL 16.78; ‘think(s)’ – LL 39.80; and ‘thought’ – LL 15.07. Only two learning indicators were consistently used in the PBL groups – diagram(s) and lecture(s). The variation of their use across the groups was also statistically significant: diagram(s) – LL 16.98, and lecture(s) – LL 47.22.

Figure 6.1 (below) shows the distribution of the relative frequencies of the most commonly occurring reporting indicators in the students’ subcorpus. A token was assumed to be relatively more common if the relative frequency was found to be greater than zero during the relative frequency computation.

Of the 27 indicators in the raw frequency analysis, two-thirds (19/27) were frequently occurring in the relative frequency analysis. Overall, the 19 types of reporting indicators had a relative frequency of 1.11.

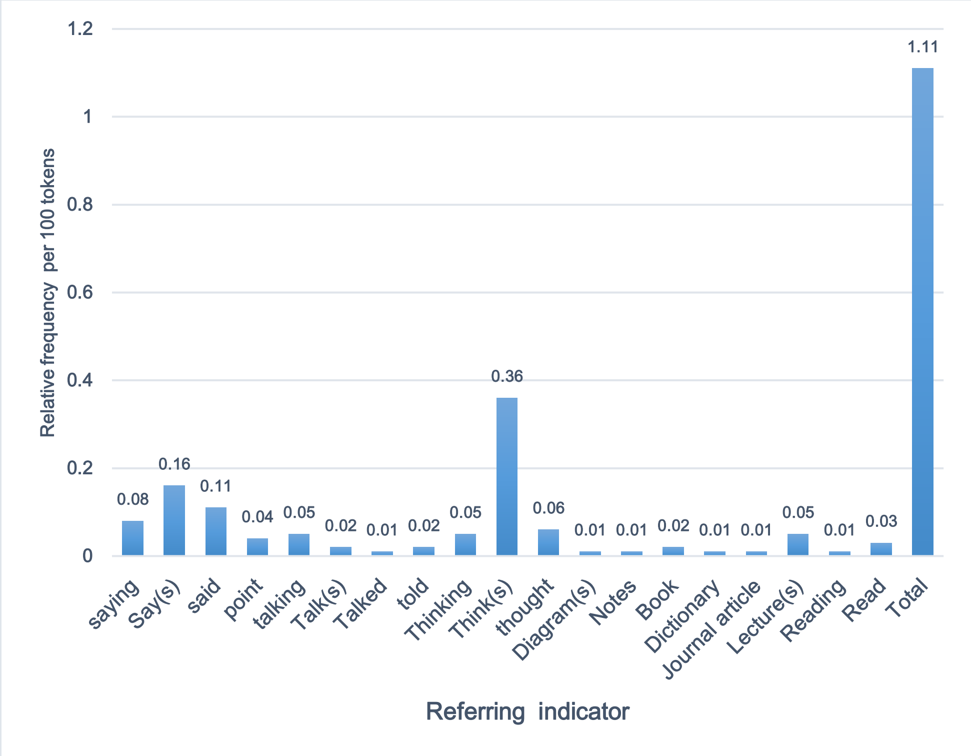


Figure 6.1: Frequency of common referring indicators: Relative frequency per 100 tokens of the commonly occurring referring expression indicators in the students' subcorpus

The 19 indicator types can be divided into three groups based on the relative frequencies. The most commonly occurring indicators 'think(s)', 'said', and 'say(s)' have relative frequencies more than 0.1 per 100 tokens. The indicator 'think(s)', i.e. 'think' and 'thinks', may suggest that the students were referring to ongoing mental processes; the indicator 'said' shows reporting of past speech; and 'say(s)', i.e. say and says, suggests reference to ongoing speech.

The commonly occurring indicators had relative frequencies between 0.05 and 0.1 per 100 tokens. This group of indicators consisted of 'saying' (0.08), 'thought' (0.06), 'talking' (0.05), and 'lecture(s)' (0.05). These indicators suggest that the students reported previous thought, ongoing talk, and made references to lectures. In case of the less commonly occurring group, the relative frequency of these indicators was less than 0.05 per 100 tokens. There were 11 items in this group (as shown in Figure 6.1).

The results of the frequency analysis suggest that students most commonly reported the ‘talk’ and ‘thoughts’ indicators, but less commonly referred to learning situations and materials. A KWIC analysis was conducted to understand the functions of these reporting tokens. The results are discussed in the next section.

KWIC analysis. The frequently occurring reporting expression indicators, i.e. 19 of 27, were explored in the context of the surrounding words to understand their functions. The results of the analysis are presented in Table 6.4 (below). A total of 982 mental functions, 1,016 verbal functions, and 278 learning functions were identified following the KWIC analysis, giving a total of 1,998 functions.

The figures in the table show that the students used two and half (40.35%) of the 1,016 verbal reporting indicators to refer to peer talk. Nearly one-fifth (17.32%) of the indicators were used to talk about the students’ own discussions. Less than one-tenth of the indicators were used for other functions: 7.87%, 8.37%, and 5.02% of them were used to refer to the teachers’ talks, group collection discussions, and doctors’ speech respectively.

Table 6.4: Functions of verbal and mental referring indicators: Raw and relative frequency of the functions of verbal and mental referring expression indicators per 100 tokens

	PBL 1 RF NF	*PBL 2 RF NF*	*PBL 3 RF NF*	*PBL 4 RF NF*	*PBL 5 RF NF*	*PBL 6 RF NF*	*PBL 7 RF NF*	*PBL 8 RF NF*	*Total RF NF*
Peer talk	57 0.19	71 0.24	79 0.21	14 0.08	31 0.15	15 0.10	34 0.17	109 0.28	410 0.20
Case information	9 0.03	6 0.02	5 0.01	6 0.03	9 0.04	2 0.01	3 0.01	5 0.01	45 0.02
Group talk	11 0.04	9 0.03	3 0.01	5 0.03	11 0.05	1 0.01	14 0.07	31 0.08	85 0.04
Concept/theory	3 0.01	3 0.01	0 0.00	0 0.00	0 0.00	0 0.00	1 0.00	2 0.01	9 0.00
Self-talk	16 0.05	13 0.04	21 0.06	3 0.02	17 0.08	4 0.03	34 0.17	68 0.18	176 0.08
Unnamed person	3 0.01	1 0.00	6 0.02	1 0.01	2 0.01	1 0.01	4 0.02	7 0.02	25 0.01
Other/impersonal	4 0.01	2 0.01	4 0.01	5 0.03	8 0.04	3 0.02	1 0.00	10 0.03	37 0.02
Teacher talk	10 0.03	0 0.00	13 0.04	5 0.03	34 0.16	12 0.08	0 0.00	6 0.02	80 0.04
Test result	1 0.00	0 0.00	2 0.01	0 0.00	1 0.00	0 0.00	1 0.00	0 0.00	5 0.00
Book/notes	3 0.01	2 0.01	6 0.02	3 0.02	0 0.00	8 0.05	2 0.01	28 0.07	52 0.02
Facilitator talk	1 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	2 0.01	3 0.01	6 0.00
Case talk	2 0.01	3 0.01	6 0.02	2 0.01	0 0.00	2 0.01	8 0.04	6 0.02	29 0.01

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

Table 6.4: Raw and relative frequency of the functions of verbal and mental referring expression indicators per 100 tokens (cont.)

	PBL 1 RF NF	PBL 2 RF NF	PBL 3 RF NF	PBL 4 RF NF	PBL 5 RF NF	PBL 6 RF NF	PBL 7 RF NF	PBL 8 RF NF	Total RF NF
Doctor talk	0	2	23	11	1	3	2	9	51
	0.00	0.01	0.06	0.06	0.00	0.02	0.01	0.02	0.02
Online resource	0	0	2	0	0	0	4	0	6
	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00
Peer thought	5	10	14	6	1	3	9	14	62
	0.02	0.03	0.04	0.03	0.00	0.02	0.04	0.04	0.03
Group thought	1	10	5	1	0	2	3	7	29
	0.00	0.03	0.01	0.01	0.00	0.01	0.01	0.02	0.01
Other's thought	2	0	3	2	0	0	0	1	8
	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00
Externalise	122	111	156	73	56	64	60	141	783
	0.41	0.37	0.42	0.42	0.26	0.41	0.30	0.36	0.37
Case thought	0	1	6	0	0	0	7	2	16
	0.00	0.00	0.02	0.00	0.00	0.00	0.03	0.01	0.01
Exhorting	0	13	18	2	2	8	8	10	61
	0.00	0.04	0.05	0.01	0.01	0.05	0.04	0.03	0.03
Expert thought	0	2	1	2	0	0	0	0	5
	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Unnamed person	0	0	0	1	1	1	2	2	7
	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.01	0.00
Impersonal	0	0	0	1	1	0	0	7	9
	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.02	0.00
Animated	0	0	0	0	1	0	1	0	2
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	250	259	373	143	176	129	200	468	1998
	0.83	0.86	1.01	0.82	0.83	0.83	1.00	1.21	0.95

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

Of the 982 commonly occurring mental indicators, nearly four-fifth (79.74%) were used to report the students' own ideas. Less than 10% of the indicators were used to refer to other types of thought. For example, 6.21%, 6.31% and 2.95% of the indicators were used to refer to exhorting, peer ideas, and collective group thinking respectively. Table 6.5 shows 278 commonly occurring learning indicators. About a quarter (25.90%) of these indicators were used to refer to previous lectures while more than a third (33.81%) were used to report self-study. Thirty-nine (14.03%) of the indicators were used to talk about conceptual maps. Less than 10% of the indicators were used to talk about other types of learning situations and curriculum materials. For example, 8.63% of them were used to refer to books, and 4.68% were used to refer to future lectures.

The distribution of the functions varied across the tutorial groups. Only six types of verbal functions consistently occurred across the tutorial groups: peer talk, self-talk, group talk, case information, talk about unnamed persons, and impersonal talk. The occurrence of these functional types varied across the tutorial groups. The variation was statistically significant for peer talk (LL 25.69), self-talk (LL 68.45), group talk (LL 19.04), case information (LL 16.50), and impersonal talk (7.58) but was not significant for talk about unnamed persons (LL 3.45).

Table 6.5: Functions of learning referring indicators: Raw and relative frequency of the functions of learning referring indicators per 100 tokens

	PBL 1 RF NF	PBL 2 RF NF	PBL 3 RF NF	PBL 4 RF NF	PBL 5 RF NF	PBL 6 RF NF	PBL 7 RF NF	PBL 8 RF NF	Total RF NF
Previous lecture	11 0.04	3 0.01	22 0.06	4 0.02	28 0.13	2 0.01	3 0.01	1 0.00	72 0.03
Future lecture	0 0.00	10 0.03	1 0.00	0 0.00	0 0.00	1 0.01	0 0.00	1 0.00	13 0.01
Prior self-study	13 0.04	12 0.04	21 0.06	3 0.02	1 0.00	7 0.05	5 0.02	32 0.08	94 0.04
Conceptual map	5 0.02	13 0.04	7 0.02	2 0.01	4 0.02	4 0.03	3 0.01	1 0.00	39 0.02
Lecture notes	1 0.00	1 0.00	6 0.02	0 0.00	2 0.01	0 0.00	2 0.01	0 0.00	12 0.01
Journal article	2 0.01	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	14 0.04	16 0.01
Book	0 0.00	1 0.00	9 0.02	3 0.02	0 0.00	9 0.06	0 0.00	2 0.01	24 0.01
Unclassified	0 0.00	0 0.00	3 0.01	0 0.00	1 0.00	0 0.00	2 0.01	0 0.00	6 0.00
Total	32 0.11	40 0.13	69 0.19	12 0.07	36 0.17	23 0.15	15 0.07	51 0.13	278 0.13

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

Only two of the mental function types – externalising one's own thinking and peer thought – consistently occurred across the PBL groups. The variation of externalisation was statistically significant (LL 49.44) while that of peer thought was not significant (LL 0.65).

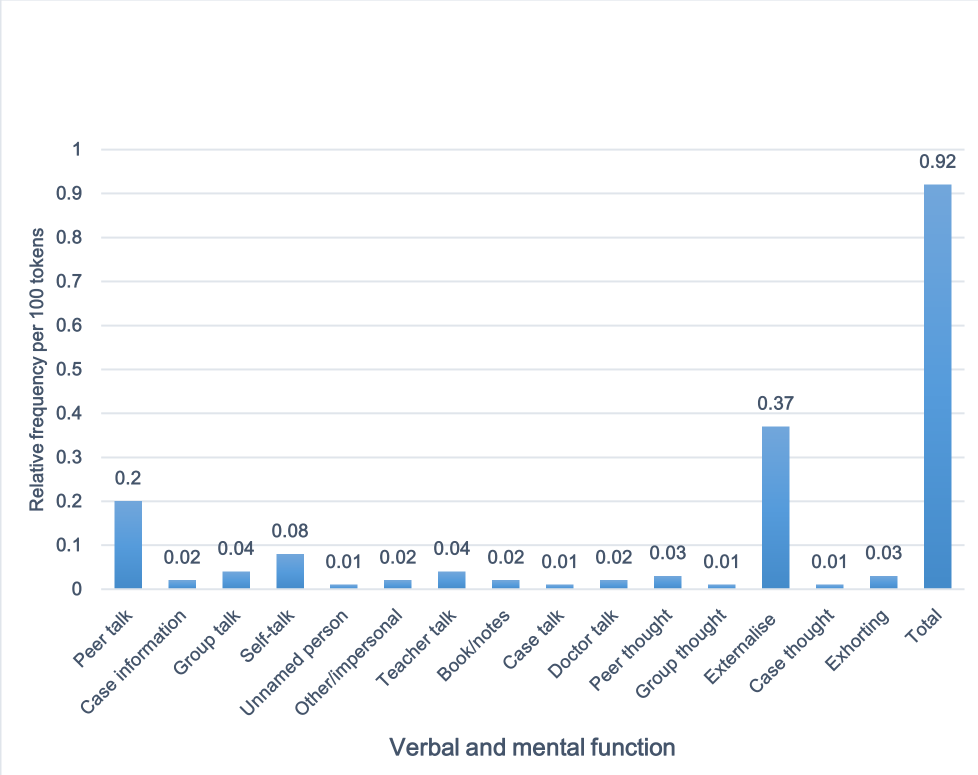


Figure 6.2: Frequent verbal and mental indicator function: Relative frequency of the commonly occurring verbal and mental referring indicator functions in the students' subcorpus per 100 tokens

Of the three consistent functional types in the learning situations and materials group, the cross-group variation of the indicators of ‘previous lecture’ (LL 82.90) and ‘concept maps’ (LL 21.95) was significant while that of ‘prior self-study’ was not significant (LL 3.58). The relative frequencies of the functions per 100 tokens were computed. Overall, the relative frequency of the commonly occurring verbal and mental functions per 100 tokens was 0.92 (Figure 6.2 above). Of the 14 verbal function types, 10 occurred frequently in the students’ subcorpus. Only references to peer talk occurred most commonly – more than 0.1 per 100 tokens. Similarly, self-talk reporting occurred commonly – 0.08 per 100 tokens. The remaining 8 functions were less commonly occurring with relative frequencies between 0.01 and 0.04.

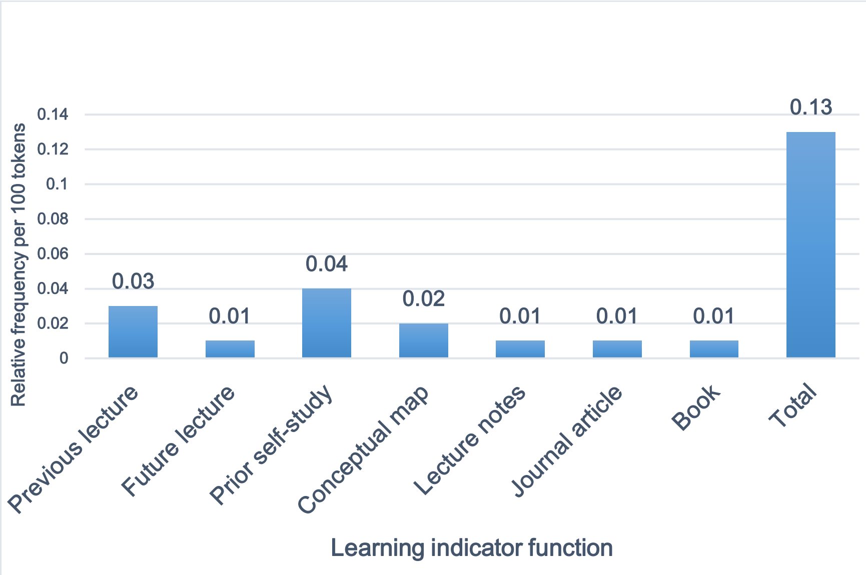


Figure 6.3: Frequent learning indicator function: Relative frequency of the most frequently occurring learning referring indicator function in the students' subcorpus

Of the 10 mental function types, five occurred commonly. Externalisation of one's own thinking was the most commonly occurring type – 0.37 per 100 tokens. The remaining mental function types were less commonly occurring with relative frequencies between 0.01 and 0.03 per 100 tokens (as shown in Figure 6.2).

Of the eight function types in the learning situation and materials category, seven were commonly occurring. All the seven functional types were less occurring with relative frequencies between 0.01 and 0.04 (Figure 6.3 above). These types included 'previous lecture', 'prior self-study', and 'conceptual map', which had relative frequencies between 0.02 and 0.04. Each of the other functional types – 'future lecture', 'lecture notes', 'journal article', and 'book' – had a relative frequency of 0.01 per 100 tokens.

The implications of the KWIC analysis are very significant. Scientific discourse is constructed by incorporating the ideas and statements of other people (Halliday and Matthiessen, 2004; Martin and White, 2007). This assumption is similar to the integration-oriented co-construction of Fischer et al. (2002) and of Weinberger and Fischer (2006). Similarly, Clark and Wilkes-Gibbs (1986b) considered referring as a collaborative process.

Although these results indicate the interactive actions of the students, a finer grasp of their experiences (which were indicated by the frequencies and the KWIC analysis) needed an in-depth analysis of extracts from the corpus. This required the use of the qualitative analytic technique, the results of which are presented in the following sections.

Qualitative analysis results.

In this section, the results of the qualitative analysis of the frequently occurring functions are discussed. This analytic approach involves an in-depth analysis of the functions in the context of a stretch of talk to understand how they were used in the talk's dataset. Only the functional types that had values greater than 10 on the raw frequency analysis and greater than zero on the relative frequency analysis were further explored via the qualitative analysis technique. The results are presented under verbal, mental, and learning categories.

Verbal reporting indicator function

Peer talk. The students used verbal reporting indicators to refer to the contributions of their peers. More specifically, these verbs were used to commend peers' positive contributions (extracts 1 and 5), clarify peers' contribution and seek explanation (extract 2), remember previous peers' discussions (extract 3), and integrate peer contributions into one's own knowledge construction (extract 4).

Extract 1:

M1 – Ash has just made a really good point, she said its consequences, they should have a value of consequences… (PBL1);

Extract 2:

M1 – And Rachel's... yeah, what were you asking about before? what were you asking about? (PBL3);

F3 – It is just that you said 'for this lady' as though normally you would not put them on an ACE inhibitor and I just wanted you to explain why. (PBL5);

Extract 3:

M2 – That's right. Yeah, so if it's all constricted then you've got ….

M3 – Yes, but do you remember what Gemma said about the dead space?

M2 – Yeah, yeah. (PBL2);

Extract 4:

M6 – Yes, it was like John was saying, if the kidney is getting all sorts of salts and proteins forced out of it then it will mess with a lot of stuff. However, in terms of the peripheral sites, it is quite a mix. (PBL4);

Extract 5:
F2 – Asha's transformation is complete she stood up and talked to us, this girl has come far, our work here is done (PBL6).

From the instances above, one can see that referring to peers' contributions seems to serve the purpose of maintaining social cohesion, establishing shared knowledge, and promoting collaborative knowledge construction.

Self-talk. Verbal reporting indicators were also used to refer to students' own contributions such as clarifying one's contribution (extract 1), clarifying misunderstanding (extracts 2 and 3), and expressing concessions (extracts 4 and 5).

Extract 1:
M3 – I thought you were talking about the long one thing.
M1 – I am talking about this one. It's…
M3 – The Purkinje fibre one?
M1 – Yeah. (PBL1)

Extract 2:
F2 – I am not saying it is not related, I am just trying to think through. (PBL2)

Extract 3:
M1 – No no no. I get that, I'm just saying, to prime it he would have to produce IGE to a virus, is that what you're saying? (PBL2)

Extract 4:

M1 – ...I would say I agree with you for the vast majority but there are some people who don't help themselves and don't try. Don't use their drive. (PBL3)

Extract 5:

M2 – I agree with you that there are x many causes to cardiac failure, but what I am trying to say is the causes and risks are going to explain the pathophysiology. (PBL6)

Examples (1) to (5) would suggest that students referred to their own contributions mainly for the purpose of attaining shared knowledge.

Group talk. This refers to a situation where expression verbs were used to denote collective statements or agreements. This could be in the form of clarifying a point of discourse (extract 1), recounting previous group discourses (extracts 2 and 3), regulating peers' verbal behaviours (extracts 4 and 5), and monitoring group discourse (extract 6).

Extract 1:

F2 – Did you say that?
M1 – No.
M3 – What are **we** talking about?
F2 – Palpitations. (PBL1)

Extract 2:

M2 – The drugs?
M3 – The classifications of...
M1 – Because we **talked** about what roughly atrial fibrillation is but this like in terms of our current... (PBL1)

Extract 3:

M4 – Yeah, are we going to do that thing that we did last week where we kind of **talk** about what is good about this case and all that? (PBL2)

Extract 4:

M2 – Ok, guys too much cross **talk** now. And cystic fibrosis hum is the other thing we have missed. (PBL2)

Extract 5:

M2 – Guys, talking over each other again.

Extract 6:

F3 – We are talking about palliative care as a whole and a lot of that is obviously for the elderly and terminally ill and they do know they are going to die, so have you looked into the psychological background, because that is what I would be interested in? (PBL6)

The examples above would suggest that students used verbals to refer to the group's social and cognitive regulation, shared knowledge building, and integration of prior knowledge into the new knowledge framework.

Unnamed person. Students also used verbal reporting indicators to report the speech of an unnamed person, one who could not be identified from the context of the discourse. In the dataset, students referred to a piece of advice (extract 1), a piece of knowledge (extracts 2 and 4) and a prior experience (extract 3) associated with this unidentified person.

Extract 1:

M4 – I was **told** to look at palliative care, which we have touched on before and also iron-deficient anaemia. (PBL6)

Extract 2:

M4 – This is a different one this time because we were **told** that the older you get, when you get to your 60's and stuff you're more prone to…. (PBL8)

Extract 3:

M4 – I don't know when exactly it was but they **told** the story when I went to shadow it. (PBL3)

Extract 4:

F2 – No but some cardio problems, like do you remember **them saying** if you have got congested heart failure you go to sleep and then you wake up because you can't breathe anymore. (PBL1)

Examples (1) to (4) suggest that students had integrated knowledge from these unnamed external sources (evident from the parts in bold in the above extracts) into their knowledge structures.

Impersonal reference. This aspect refers to specific events in which a student did not indicate any personal involvement, but got the information from the person who had been personally involved. The students used verbals to report the experiences of other people such as those related to ethical dilemma in clinical practice, as can be seen in extract 1 below.

Extract 1:

M2 – If you want a real ethical dilemma, one of the GPs **told** me that one of the patients he saw was depressed and on the verge of committing suicide, but he had cardiac failure. He did not tell him he had cardiac failure because he was depressed and he was going to commit suicide and he did not want anybody to know. What would you do in that circumstance? (PBL2)

Teacher talk. Students used verbal reporting indicators to recount the teaching of their lecturers. They used knowledge from the teachers to interpret test results (extract 1), recounted the teachers' opinion (extract 2), remembered learning in the lecture room (extracts 3 and 4), and regulated the scope of their discourse with the teachers' statements (extract 5).

Extract 1:

F2 – Do you remember the thing the ECG lady did when she said that thing like we have got some... (PBL1)

Extract 2:

F2 – John Frayne **said** it was rubbish.

M1 – Is it? I quite liked it...

F2 – He said you tap out the rhythm but I do not think that is any use at all to anyone. (PBL1)

Extract 3:

F2 – He said in terms of actions but did not Danny **say** that two thirds of blood is just through, passing… (PBL1)

Extract 4:

M4 – But I thought that he **said** that the atrium contracting just slightly tops up the blood that actually fills into the ventricles so why would you get…? (PBL1)

Extract 5:

F2 – That's the lecture, guys, that's the lecture we had when she **said** you might want to come back to this in a week's time rather than thinking too much about it now. (PBL3)

Extract 6:

M2 – He **said** we didn't have to go into all those various patterns, I mean I think it's quite useful. (PBL8)

Extracts 1 to 6 suggest that students used information from the teachers in the lecture rooms to construct their knowledge and regulate their cognitive activities.

Book/notes. The students referred to books to orally share knowledge (extracts 1 and 5), defend their position (extract 2), demand evidence (extract 3), and correct misconceptions (extract 4). They used lecture notes to resolve confusion about information (extracts 6 to 7) and share knowledge (extract 8).

Extract 1:

M1 – … I was reading around a lot and a lot of the literature was saying you can't classify asthma because everyone tries to classify it… (PBL3)

Extract 2:

M1 – We don’t want to give him diazepam though. Doesn’t that affect your respiratory drive as well? So that will lower the respiratory drive.

F1 – I read it in one of the books that said diazepam for breathlessness and to help them to sleep. (PBL5)

Extract 3:

F2 – You just made that up.

M1 – I did not.

F2 – Yes, you did. I do not believe you.

F3 – **Cite** your sources. (PBL6)

Extract 4:

F3 – The collapsing is such a generic term and it said in a couple of **books** that its use should be avoided as a medical term…. (PBL4)

Extract 5:

M4 – The dictionary says it is due to damaged heart valves, ventricular muscle or both. (PBL4)

Extract 6:

M1 – It was explicitly **said** in George Marbeza’s lecture notes that the inspiratory loop was relative to the… (PBL3)

Extract 7:

M1 – … looking at his notes he doesn’t refer to it as an obstructive respiratory disease or a restrictive respiratory disease, he’s just saying it’s an airway obstruction. (PBL3)

Extract 8:

F1 – And you get a fall in FEV_1. And it's obstructive co-existing with emphysema, that's what it says here. (PBL3)

Case information. To understand the case better and guide their discussion, students frequently referred to case information as illustrated in the two examples below (extracts 1 and 2).

Extract 1:

F2 – He doesn't drive much it says. (PBL7)

Extract 2:

F1 – It depends, what's his age, does it say?

M1 – No it not does say. (PBL7)

Doctor's talk. By referring to statements from the doctors, students were able to share the experience they had gained from the GP surgery placement with their peers and these experiences related to understanding the patient's perspective about their disease condition (extract 1), clinical decision-making strategies (extract 2), and diagnostic processes (extract 3).

Extract 1:

F2 – Actually my GP said that to somebody yesterday, had some woman come in to have…. what do you think it is and she came out with all this stuff. And she was wrong, but she was like she was really worried about. (PBL4)

Extract 2:

M4 – … I was speaking to my GP and he said in that instance you haven't necessarily solved the problem and that person will go away and come back to another GP in the surgery and another GP and it will keep happening… (PBL3)

Extract 3:
F2 – The GP also said to me yesterday, talking about the triad of signs, that she will always ask about shortness of breath, syncope… (PBL4)

The above examples highlight the fact that students referred to the doctors' talk to provide prior knowledge, which could be used to anchor new knowledge construction.

Case patient talk. Finally, the students used reporting expression indicators to refer to the case patients' speech. This could provide opportunities to understand the case problem and grasp the psychosocial issues surrounding the patients' health problems (extract 1).

Extract 1:
M4 – So he will say it hurts when I'm walking, but he wants to play bowls so he doesn't say oh it hurts when I'm playing bowls. (PBL7)

The qualitative analysis provided a richer insight into the functionality of reporting expression indicators. The reporting expression indicators were used to pursue social cohesion, share knowledge, promote social and cognitive regulation, construct knowledge, resolve knowledge-related confusions, and grasp psychosocial issues behind patients' clinical features.

Thought reference function. Students used thought referring indicators to refer to peer thought, group thought, case thought, externalisation of one's own thinking, and exhortation of thinking. The purpose of this reference was to achieve several objectives as illustrated below.

Externalise one's own thought. Thought referring indicators were used to report students' own thinking as they made their thinking visible to their peers. This served the purpose of generating hypothesis (extract 1), establishing shared understanding (extract 2), and expressing disagreement (extracts 3 and 4).

Extract 1:
M1 – I was thinking about carotid artery stenosis cos he's already got vascular disease pretty much everywhere else, leg, heart. (PBL7)

Extract 2:
M4 – Yeah. That's what I was thinking. Is it normal physiology which I think it is, or is it a physiology that happens if you have too much… (PBL7)

Extract 3:
M1 – Because I don't think varicose veins generally kill many people. (PBL7)

Extract 4:
M1 – I think that's a good idea yeah. (PBL8)

Peer thought. The students in the PBL groups referred to the thinking of their peers using referring mental verbs for the purpose of asking questions (extract 1), clarifying peers' thinking (extract 2), and expressing conflict of ideas (extract 3).

Extract 1:
M1 – So can you think why else you'd get a Pansystolic murmur with a… (PBL8)

Extract 2:
M1 – You're thinking about the stimulation and why you can't just tell yourself to breathe? (PBL3)

Extract 3:
F1 – You can't be thinking as a GP's thinking I am going to do this for the greater good, you have to treat your patient, and if more GPs treat their patients then this is going to contribute to public health later on. (PBL3)

Group thought. Cognitive verbs were also used to refer to the group's collective thinking. The reference was used to suggest topics for group consideration, as illustrated in extracts 1 to 3 below.

Extract 1:
M1 – ….can we think about the arteriosclerotic process and whether it's the same in the leg. (PBL7)

Extract 2:
F1 – Also are we going to think about coronary– and cardiac– arrest (PBL8)

Extract 3:
M1 – Yes, so the clinical stuff that we **think** we should probably look at… (PBL 2)

Case patient thought. Reporting cognitive indicators were also used to refer to the opinions of the case patient about his/her health condition as highlighted in the examples below. This may suggest students' attempt to understand case patients' expectations and psychosocial issues (extracts 1 to 3).

Extract 1:

F1 – Then we have what does he think is wrong with him if he's anxious about having surgery, cos his friend had his leg amputated. What does he think is wrong? (PBL7)

Extract 2:

F2 – ….well one of those is to try and get them to exercise… Obviously he thinks it's a risk. (PBL7)

Extract 3:

M1 – Varicose veins, which he thinks is causing his cramps. (PBL7)

Exhortation of thinking. Furthermore, students used mental reporting indicators to motivate and encourage their peers to think about issues in particular ways (extracts 1 and 2).

Extract 1:

M1 – Yes, but with so much blood it must take ages to get to that point, when you think about it, if you have all these peripheral places for the blood to sit anyway. (PBL5)

Extract 2:

F2 – … and if you think about forming these fatty streaks in various parts of the body and having different symptoms, for example now you're talking about legs and feet… (PBL7)

Summary

The thought referring indicators used commonly by students aimed at reporting their own, peer, collective, and case patient thinking. These were used to generate hypotheses, achieve understanding, define group task and action plan, and understand case patient expectations and psychosocial issues.

Learning situation reporting function

Learning situations represent another category of items that students referred to in their tutorial discussions. They were in the form of previous lectures, future lectures, and prior self-study sessions. A qualitative approach was deployed to understand the framing of these functions in the students' subcorpus.

Previous lecture. Data analysis showed that students applied information from previous lectures to understand concepts (extracts 1 and 2), jointly construct knowledge (extract 3), give explanations (extracts 4 to 7), and extol the usefulness of those lectures (extracts 8 and 10).

> **Extract 1:**
> F2 – But we had a lecture on this, race versus ethnicity versus culture and it's all… I can't remember which one is which but culture is to do with race plus ethnicity… (PBL2)

Extract 2:

F2 – As far as I could understand from the lecture that is the amount of damaged muscle you have in the heart decreases the ionotropy. (PBL5)

Extract 3:

F2 – Did you see the thing in the lecture, Friday's lecture...?

F1 – Which thing...?

F2 – I am being really specific, aren't I?

M4 – Yeah, you are.

F2 – About the Framingham risk assessment. That is what I was talking about last time and I could not remember what it was called.

M2–Refresh me on that.

F2 – It was to work out your risk of cardiovascular disease or accident, your ten-year risk factor based upon things like... (PBL1)

Extract 4:

F3 – There were two lectures; one was already explaining everything that happened...

M2 – We need to memorise...

F3 – … but the one before that I think was on like... when CO_2 increases, this happens and this means this and this means this. That's the second lecture and that's all under the bracket of Arterial Blood Gas analysis. (PBL3)

Extract 5:

F2 – He mentioned it briefly in his lecture, that if you've got a blockage here it's affecting inspiration and expiration which is why you get that funny shape… (PBL3)

Extract 6:

F2 – In the pathophysiology lecture we had the other day heart failure is not a disease in itself, it is caused by a number of different things. (PBL4)

Extract 7:

F2 – … the muscle is affected in those particular ways and in the lecture he talked about dilatation of the ventricle and that kind of thing. (PBL4)

Extract 8:

M4 – I think the lecture the other day on the ECGs has really helped me and I'm really excited about taking more ECGs myself, I'm starting to become a little bit more – cardiologist – it's not good is it? (PBL8)

Extract 9:

F2 – … yeah, good, very good lectures, I thought I understood ECGs and now I don't. (PBL8)

Extract 10:

F2 – … yesterday's lectures really, really helped everything and I really felt like it all made sense. (PBL1)

Future lecture. The students talked about future lectures during task planning, as illustrated by the extract 1 below.

Extract 1:

F2 – But Monday we've got lectures so we could maybe do an hour. (PBL2)

Prior self-study. The students talked about their previous study to refer to shared knowledge (extract 1), support peer explanation (extract 2), correct old teaching (extracts 3–5), and understand patients' psychosocial issues (extract 6).

Extract 1:

F2 – ... from what I've read, if it's a bacterial infection or infection of some sort then you're going to get a productive cough… (PBL2)

Extract 2:

M2 – … what Gemma said, you know with the problems at night, I read that when the COPD becomes really advanced then the respiratory problem or the breathlessness or respiratory problems exist even at rest…. (PBL3)

Extract 3:

F2 – We've been taught that the initial damage to vessels occurs through the pressure-related damage, but my reading last night was basically that the damage is done by lipids setting on the vessels and by oxidisation and free-radical… (PBL7)

Extract 4:

M2 – The other little note that I came across in my reading last night was that you know I think we've been told that what 50%, 75% occlusion is going to lead to something symptomatic, yeah? (PBL5)

Extract 5:

M2 – My reading said rather an old sort of figure said that mostly it's 20%, the end is 20% so that's a lot lower. (PBL6)

Extract 6:

F3 – I was reading some of those and they think that it's all their fault, like the way they've lived and the way they've eaten and now they're making their family suffer for it because they're in pain and can't walk. (PBL8)

Students referred to previous lectures and self-study sessions to construct new knowledge and share information with their peers. They referred to future lecture schedules as they talked about task plans.

Learning materials

The learning materials that the students referred to included conceptual maps and models, lecture notes, journal articles, and books. A qualitative analysis provides opportunities for evaluating how the materials were used.

Conceptual maps and models. The students used diagrams and models to share learning resources and explain concepts (extracts 1 and 2).

Extract 1:

M1 – Yeah, that is what I am asking. And what happens up here?
M2 – The potassium gates open up.
M1 – You get that, you get a plateau phase and then it drops off.
F3 – And where is this on the diagram?
F2 – They are just part cells, aren't they? (PBL1)

Extract 2:

F2 – There is a better diagram there; so that is your bundle of His, branches into left and right; that is your anterior fascicle, and Purkinje bits... (PBL2)

Lectures notes. Lecture notes were used by the students to guide their discussions and resolve conflicts of ideas, as exemplified in extracts 1 and 2 below.

Extract 1:
M2 – My **notes** say apicoposterior yeah, but I've put that in later I think. I've seen it written up in something. (PBL2)

Extract 2:
M4 – No there's a nerve underneath the bronchi; underneath the diaphragm… I've got it written **down.** It's called the parasympathetic reflux loop nerve endings underneath the lining of the bronchus. (PBL2)

Journal article. Students referred to journal articles to share knowledge (extract 1) and learning resources (extracts 2 and 3), as highlighted in the examples below.

Extract 1:
M4 – I don't…. kind of do but I read something called a variant in a journal about prinzmetal's angina. (PBL8)

Extract 2:
F1 – …I don't know if anyone read the article I sent around but there were a few women saying that it's difficult to get appointments and by the time you get an appointment the pain has gone so there's no point in going to the doctor anymore. (PBL8)

Extract 3:
F3 – But yeah I can send you those links for the papers if you want. (PBL7)

Book/booklet. Students talked about books for the purpose of resolving knowledge conflicts (extracts 1 and 5) and sharing learning resources (extracts 2 to 4), as illustrated in the examples below.

Extract 1:

M1 – We had one book… if two people come up with totally different ideas and you go right… (PBL2)

Extract 2:

F2 – There's some stuff that we can read up on that… (PBL4)

Extract 3:

M4 – There's a good book you know in the library (PBL2)

Extract 4:

M1 – You know the recommended ECG book; the blue one. Get in that book and there's a section on naming the waves depending on where they are and what comes before and what comes after that is really useful. (PBL8)

Extract 5:

M2 – … I'd check up on that cos I wasn't convinced about the femoral being the most common place and just a couple of vascular medicine, vascular surgery books…. and the most common sites are….. aortoiliac is the first one… (PBL7)

Summary

The results of the analysis showed that students were referring to the contributions of their peers and talking about previous teachings, thereby suggesting that they were integrating the statements of their peers and teachers in the process of forming their own knowledge.

They were also referring to their own contributions, which may suggest that they were clarifying their knowledge and understanding with their peers.

Furthermore, the results showed that students were externalising their own thinking. This could mean that students were putting forth their knowledge and sharing their ideas with their peers. The knowledge and ideas could then be discussed and negotiated with a potential for correction of misconceptions and learning. Besides, knowledge is reorganised when it is externalised (Huber 1987 cited in Weinberger and Fischer, 2006). The results showed that students referred to the ideas of their peers and encouraged their thinking through exhortation. This may suggest that students were asking for their peers' opinions and ideas. The reference which the students made to future lectures may relate to task planning and organisation. Besides, students referred to books and conceptual diagrams possibly as sources of knowledge to resolve case problems and as tools to build conceptual networks of the problem.

Question 2. Analysis of shared knowledge expressions

Quantitative analysis results

Quantitative analysis was carried out on two levels – word frequency analysis and KWIC analysis.

Word frequency analysis. The data retrieval strategy generated 3,437 shared knowledge indicators from the students' subcorpus (Table 6.6). There were 28 types of indicators. Only eight types of indicators were in double digits. Of the 8, ‘yeah’, ‘yes’, ‘no’, and ‘oh’ occurred more frequently. While ‘yes’, ‘no’ and ‘oh’ were consistently used across the eight PBL groups, ‘yeah’ only occurred in five of the eight groups.

Table 6.6: Frequency of shared knowledge indicators: Raw and relative frequency of shared knowledge indicators per 100 tokens

	PBL 1 RF NF	*PBL 2 RF NF*	*PBL 3 RF NF*	*PBL 4 RF NF*	*PBL 5 RF NF*	*PBL 6 RF NF*	*PBL 7 RF NF*	*PBL8 RF NF*	*Total RF NF*
Yeah	*223 0.74*	*324 1.07*	*336 0.91*	*0 0.00*	*0 0.00*	*0 0.00*	*252 1.26*	*418 1.08*	*1553 0.74*
Yes	*110 0.37*	*156 0.51*	*54 0.15*	*171 0.98*	*229 1.08*	*141 0.91*	*14 0.07*	*37 0.10*	*912 0.43*
No	*80 0.27*	*127 0.42*	*91 0.25*	*24 0.14*	*54 0.25*	*43 0.28*	*49 0.24*	*112 0.29*	*580 0.28*
Oh	*36 0.12*	*54 0.18*	*48 0.13*	*9 0.05*	*4 0.02*	*1 0.01*	*34 0.17*	*49 0.13*	*235 0.11*
Ah/a-h-ah	*6 0.02*	*16 0.05*	*18 0.05*	*4 0.02*	*3 0.01*	*1 0.01*	*11 0.05*	*4 0.01*	*63 0.03*
Wow	*3 0.01*	*0 0.00*	*1 0.0*	*0 0.00*	*0 0.00*	*0 0.00*	*2 0.01*	*2 0.01*	*8 0.00*
Yep	*1 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*1 0.00*
Hooo	*1 0.00*	*0 0.00*	*0 0.0*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*1 0.00*
Nope	*1 0.00*	*0 0.00*	*0 0.0*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*1 0.00*
Hello	*1 0.00*	*1 0.00*	*0 0.00*	*2 0.01*	*0 0.00*	*0 0.00*	*0 0.00*	*1 0.00*	*5 0.00*
Ssh	*1 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*1 0.00*
Oh God	*1 0.00*	*3 0.01*	*4 0.01*	*0 0.00*	*1 0.00*	*0 0.00*	*1 0.00*	*1 0.00*	*11 0.01*
Blah-blah	*1 0.00*	*0 0.00*	*0 0.00*	*3 0.02*	*0 0.00*	*0 0.00*	*0 0.00*	*2 0.01*	*6 0.00*
Um/hum	*0 0.00*	*10 0.03*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*1 0.00*	*4 0.01*	*15 0.01*
Uh-huh	*0 0.0*	*2 0.01*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*0 0.00*	*2 0.00*

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

Table 6.6: Raw and relative frequency per 100 tokens of shared knowledge indicators (cont.)

	PBL 1 RF NF	PBL 2 FR NF	PBL 3 RF NF	PBL 4 RF NF	PBL 5 RF NF	PBL 6 RF NF	PBL 7 RF NF	PBL8 RF NF	Total RF NF
Yah	1 0.00	2 0.01	1 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.0	3 0.00
Yea	1 0.00	2 001	1 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	3 0.00
Oh well	1 0.00	2 0.01	1 0.00	0 0.00	0 0.00	0 0.00	0 0.00	1 0.00	3 0.00
Crikey	1 0.00	2 0.01	0 0.00	1 0.01	0 0.00	0 0.00	0 0.00	0 0.00	3 0.00
Huh	0 0.00	1 0.00	0 0.00	0 0.00	0 0.00	0 0.00	1 0.00	0 0.0	2 0.00
Oh no	0 0.00	1 0.00	1 0.00	0 000	0 0.00	0 0.00	2 0.01	6 0.02	10 0.0
Oh my God	0 0.00	1 0.00	3 0.01	0 0.00	0 0.00	0 0.00	0 0.00	1 0.00	5 0.00
Oi	0 0.00	10.0 0	2 0.01	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	3 0.00
Och/gosh	0 0.00	0 0.00	3 0.01	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	3 0.00
Cheers	0 0.00	0 0.00	1 0.00	0 0.00	0 0.00	0 0.0	0 0.00	0 0.00	1 0.00
Oh dear	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	1 0.00	2 0.01	3 0.00
Oh really	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	1 0.00	0 0.0	1 0.00
Hey	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	2 0.01	1 0.00	3 0.00
Total	466 1.55	705 2.33	563 1.53	214 1.23	291 1.37	186 1.20	371 1.85	641 1.65	3437 1.64

RF – raw frequency; NF – normalised or relative frequency per 100 tokens

Figure 6.4 (below) shows the shared knowledge indicator types that occurred more than 10 times in the students' subcorpus. These seven indicator types accounted for 3,369 (98.02%) of the retrieved indicators. Four of the indicators had a frequency of more than 0.1 per 100 tokens and were considered as most frequent. The remaining three with relative frequencies between 0.01 and 0.03 were taken to be less common.

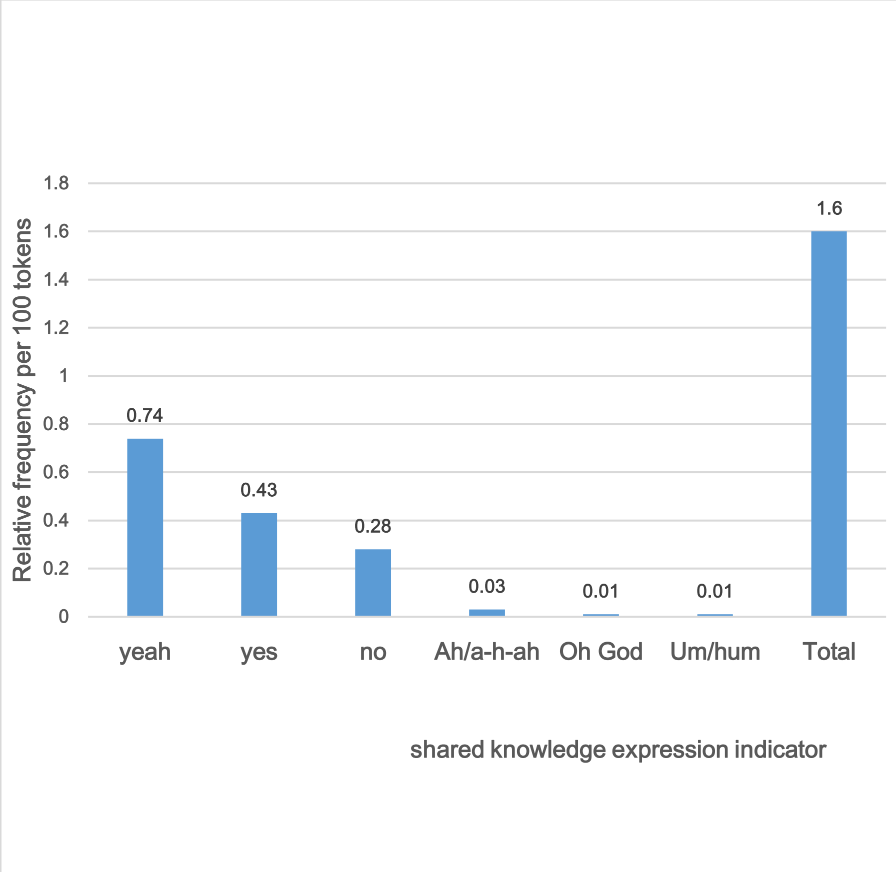


Figure 6.4: Frequent shared knowledge indicators: Relative frequency of the commonly occurring shared knowledge indicators in the students' subcorpus per 100 tokens

The commonly occurring shared knowledge indicators were divided into three structural categories: affirmation, negation, and no-lexical-content categories (Table 6.7 below). Although there was some variation in the use of the affirmation indicators across the PBL groups, it was not statistically significant (LL 2.09, $p > 0.01$). However, the variation of the negation and the nonlexical content indicators across the PBL groups was statistically significant – LL 39.16, $p < 0.01$ and LL 67.48, $p < 0.01$ respectively.

Table 6.7: Categories of shared knowledge indicators: Structural categories of the commonly occurring shared knowledge indicators

	PBL1 RF NF	***PBL2 RF NF***	***PBL3 RF NF***	***PBL4 RF NF***	***PBL5 RF NF***	***PBL6 RF NF***	***PBL7 RF NF***	***PBL8 RF NF***	***Total RF NF***
Affirmation	333 1.11	480 1.59	390 1.06	171 0.98	229 1.08	141 0.91	266 1.33	455 1.17	2465 1.17
Negation	80 0.27	127 0.42	91 0.25	24 0.14	54 0.25	43 0.28	49 0.24	112 0.29	580 0.28
Nonlexical	43 0.14	83 0.27	70 1.19	13 0.07	8 0.04	2 0.01	47 0.23	58 0.15	324 0.15
Total	456 1.52	690 2.29	551 1.50	208 1.19	291 1.37	186 1.20	362 1.80	625 1.61	3369 1.60

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

Figure 6.5 below shows that affirmation indicators were used four times more than the negation indicators while the nonlexical content indicators accounted for 10% of the commonly occurring indicators. Similarly, in each group, affirmation indicators accounted for more than 70% of the total shared knowledge indicators.

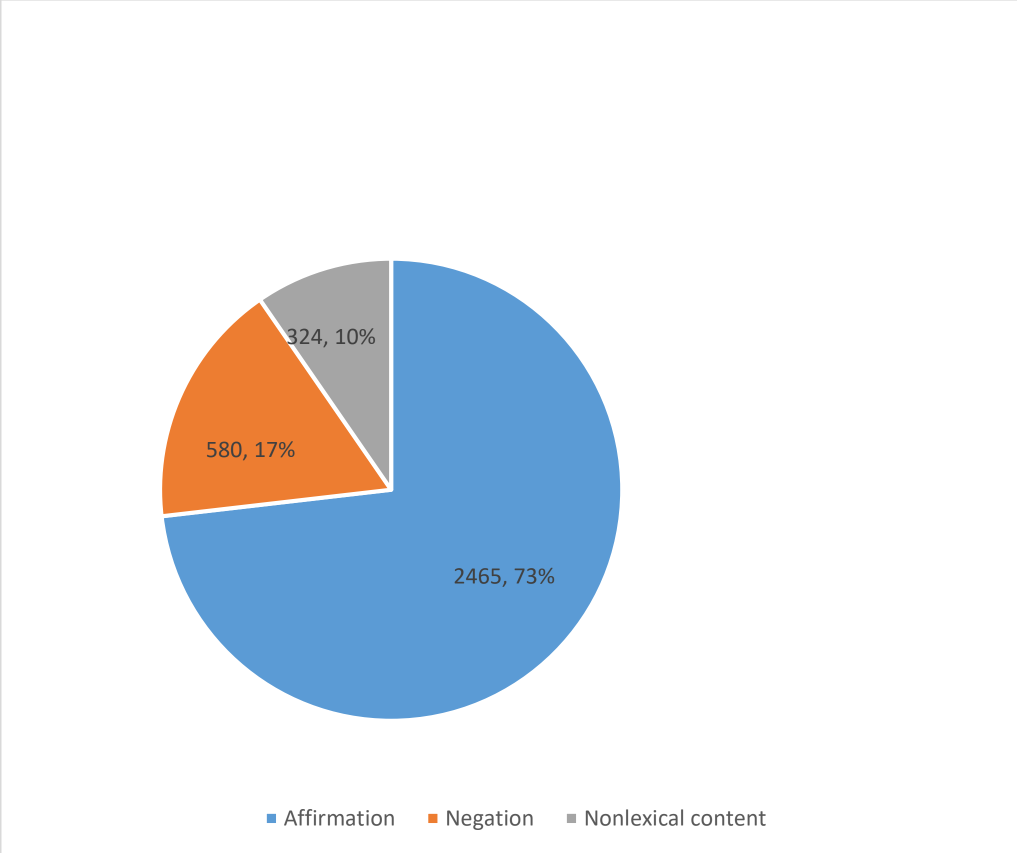


Figure 6.5: Structural categories of shared knowledge indicators: Proportions of structural categories of commonly occurring shared knowledge indicators in the students' subcorpus.

Key Word in Context (KWIC) analysis. The initial results from the word frequency analysis were examined at a deeper level with the KWIC analysis technique. The analysis involved the 3,369 commonly occurring shared knowledge indicators. The results of the analysis are shown in Table 6.8 (below).

Overall, the figures in the table show that the students most commonly used shared knowledge tokens to achieve integration-oriented shared knowledge. This function was closely followed by the agreement function while the acknowledgement and conflict-oriented functions ranked third. The results are next discussed according to the structural indicator categories.

The students most commonly used affirmation tokens to achieve integration-oriented shared knowledge. More specifically, the students used slightly less than half (42.31%) of the affirmation indicators for this purpose.

Table 6.8: Frequency of shared knowledge functions: Raw and relative frequency of shared knowledge functions per 100 tokens

	PBL1 RF NF	PBL2 RF NF	PBL3 RF NF	PBL4 RF NF	PBL5 RF NF	PBL6 RF NF	PBL7 RF NF	PBL8 RF NF	Total RF NF
Integration	123 0.41	177 0.59	127 0.35	96 0.55	140 0.66	80 0.52	106 0.53	194 0.50	1043 0.42
Agree	108 0.36	148 0.49	146 0.40	46 0.26	65 0.31	34 0.22	79 0.40	124 0.32	750 0.30
Acknowledge	70 0.23	97 0.32	66 0.18	17 0.10	18 0.08	10 0.06	49 0.24	103 0.27	430 0.17
Question	27 0.09	41 0.14	39 0.11	2 0.01	6 0.03	8 0.05	14 0.07	26 0.07	163 0.07
Repetition	5 0.02	17 0.06	12 0.03	10 0.06	0 0.00	9 0.06	18 0.09	8 0.02	79 0.03
Conflict	47 0.16	101 0.33	64 0.17	22 0.13	35 0.16	41 0.27	25 0.12	71 0.18	406 0.70
Negation	27 0.09	14 0.05	23 0.06	2 0.01	10 0.05	2 0.01	13 0.06	16 0.04	107 0.18
Question	0 0.00	3 0.01	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	3 0.00
Repetition	6 0.02	9 0.03	4 0.01	0 0.00	9 0.04	0 0.00	11 0.05	25 0.06	64 0.11
Orientation	39 0.13	62 0.21	54 0.14	8 0.05	5 0.02	1 0.01	37 0.18	46 0.12	252 0.78
Recall	4 0.01	8 0.03	12 0.03	5 0.03	3 0.01	1 0.01	8 0.04	7 0.02	48 0.15
Exclamation	0 0.00	3 0.01	4 0.01	0 0.00	0 0.00	0 0.00	1 0.00	1 0.00	9 0.03
Reactive	0 0.00	10 0.03	0 0.00	0 0.00	0 0.00	0 0.00	1 0.00	4 0.01	15 0.05
Total	456 1.52	690 2.29	551 1.50	208 1.19	291 1.37	186 1.20	362 1.80	625 1.61	3369 1.60

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

Nearly one-fifth (17.44%) of the affirmation indicators were used for acknowledgement; about one-third (30.43%) for agreement; less than one-tenth (6.61%) were used to preface questioning while less than one-twentieth (3.20%) were used for repetition. Almost three-quarters (70%) of the negation shared knowledge indicators were used for conflict-oriented sharing of knowledge. Nearly one-fifth (18.43%) of the indicators were used for simple negation while about one-tenth (11.03%) were used for mere repetitions. The students used more than three-quarters (77.78%) of the nonlexical content group of indicators to orientate themselves to ideas and information in the group discourse; more than one-tenth (14.81%) of the tokens were used to recall forgotten information while less than one-twentieth (4.63%) were used for reactive vocalisations.

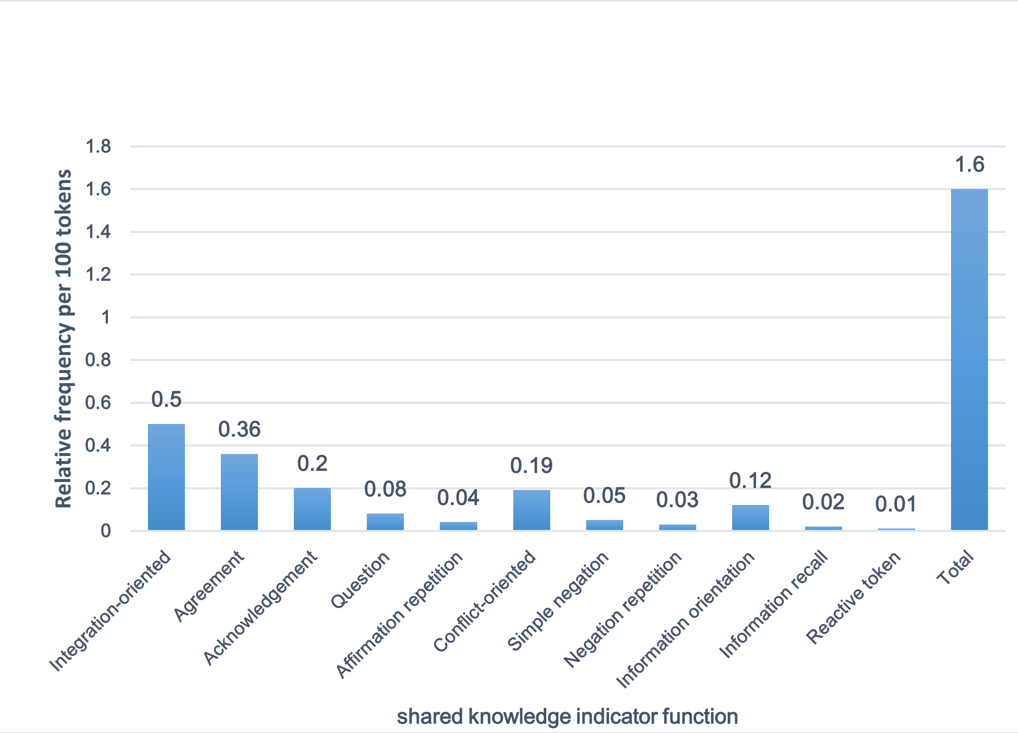


Figure 6.6: Frequent knowledge indicator functions: Relative frequency of the commonly occurring shared knowledge indicator function per 100 tokens

Figure 6.6 shows the commonly occurring shared knowledge functions in the students' subcorpus. There were 11 commonly occurring function types, and five most commonly occurring functional categories. Each had a relative frequency in excess of 0.1 per 100 tokens. Three of the leading functions – integration-oriented consensus, agreement, and acknowledgement were oriented towards achieving agreement between the participants.

This finding corroborates the finding of the word frequency analysis. Conflict-oriented consensus and information orientation functions were also most common. Only question-oriented affirmation and simple negation were commonly occurring while the students used less than 0.05 indicators to pursue each of the remaining functions.

The occurrence of the leading functional types across the PBL groups was statistically significant at the level of $p < 0.01$: integration-oriented shared knowledge building (LL 24.29), agreement (LL 45.06), acknowledgement (LL 8.03), conflict-oriented shared knowledge building (LL 73.32), and orientation to information (LL 57.05).

Further data exploration showed that addition, repetition and rephrasing, paraphrasing and causal explanation were the most frequent types of integration-oriented shared knowledge (see Figure 6.7 below).

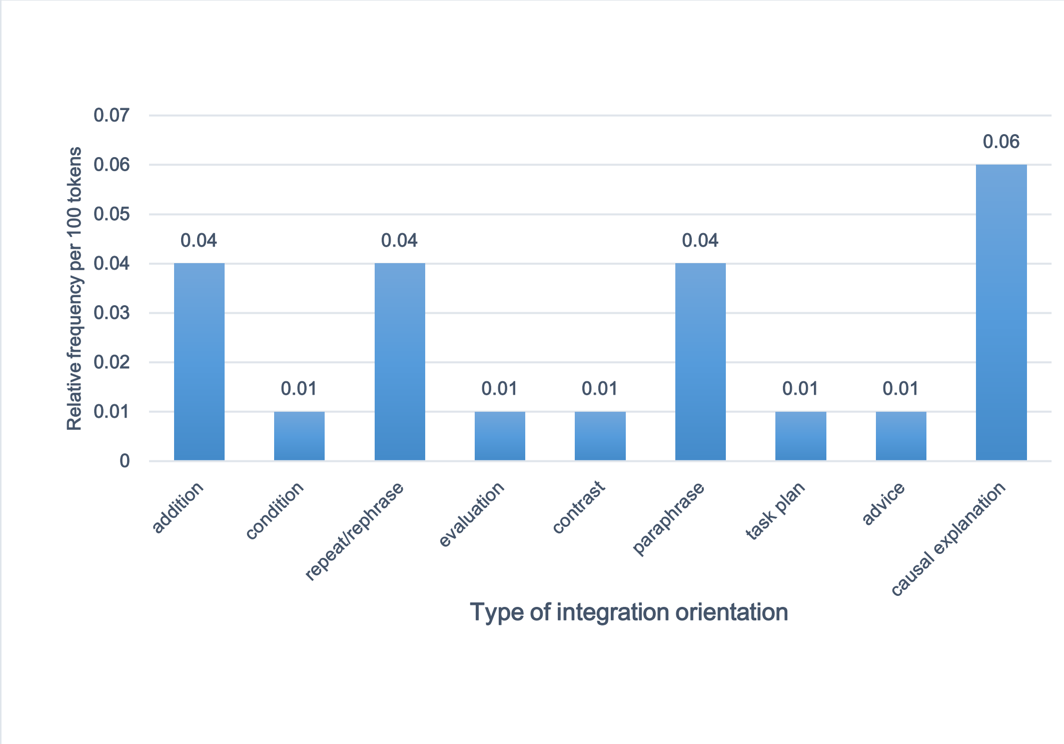


Figure 6.7: Affirmation elaboration types: Relative frequency of the commonly occurring affirmation elaboration types per 100 tokens

Similarly, conflict-oriented shared knowledge was most frequently achieved through correction while elaborative, contrast and causal conflict-oriented shared knowledge were less common (see Figure 6.8 below).

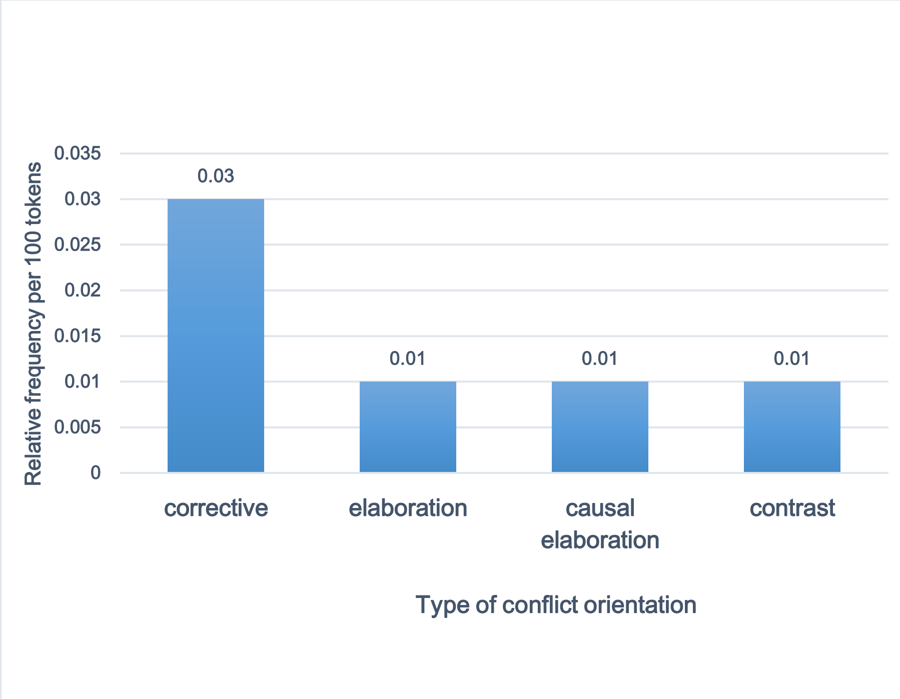


Figure 6.8: Negation elaboration types: Relative frequency of the commonly occurring negation elaboration types per 100 tokens

Qualitative analysis

In this section, I turn away from frequencies and form and function, and investigate more deeply how students used shared knowledge indicators by examining a longer stretch of discourse. This fine-grained analysis has shown the learning experiences that were marked with shared knowledge markers and the way in which co-construction was enacted among the learners.

Integration-oriented and conflict-oriented consensus building may be simple, elaborate or causal. Simple integration-orientated consensus building occurs when a speaker repeats, clarifies or paraphrases the contribution of the previous speaker. Elaborate integration-oriented consensus building occurs when a speaker adds non-causal information to the contribution of the previous speaker. The information may be a comparison, addition, example, or an alternative idea.

Integration-oriented shared knowledge. The students had been talking about causes of cough. Extract 1 is part of the exchanges regarding this point of discussion. In line 9, student M1 hypothesised that the cough may be due to congenital abnormality. This was paraphrased by F2 in line 10. In line 11, M1 endorsed this and extended the knowledge by speculating that the condition may have been undiagnosed. This stretch of discourse exemplifies simple integration-oriented shared knowledge construction.

Extract 1:

1. M4– Hold on.
2. F2 – Did you want congenital?
3. M3 – Yes o.
4. F3 – separate….
5. M2 – Am I going to…?
6. M1 – It could be anything.
7. M4 – Yes, I am trying to get that.
8. F1 – Keep going.
9. M1 – so a congenital abnormality actually comes too early like tetralogy or erm vascular…..
10. F2 – Is it congenital or has it not been there from birth?
11. M1 – Yeah, it might not have been identified.
12. F2 – But if he’s only had two episodes in six months…. symptoms (PBL2).

Causal elaboration. Extract 2 shows a combination of elaborate and causal integration-oriented consensus as is evident from its lines. In line 2, F1 was elaborating following the question in line 1. This was followed by a mechanistic contribution by M3 in line 3. The reason was again provided by M3 in line 5. Following a clarification question in line 6, M3 gave a further causal explanation.

Extract 2:

1. M2 – Is this stage one?
2. F1 – Yes, when the collagen starts forming then you get the plaque formation or something; I do not know what it is called. Lipid plaque.
3. M3 – Yes, lipid plaque and when it fibroses in the next bit – so you have your lipid plaque there, which is the foam cells and everything and then the smooth muscle proliferates over the top as well.
4. M4 – Smooth muscle?
5. M3 – Yes, because it is internalised –
6. F1 – The plaque grows down into the smooth muscle?
7. M3 – Yes, so if this is the smooth muscle, it is also on the outside as well and this is inside it, but what happens is that this smooth muscle hypertrophies as well, so it proliferates around the outside (PBL5).

Agreement. Extract 3 exemplifies agreement (as seen in lines 3 and 5 by M2). Both were confirmations to clarification requests.

Extract 3:

1. M2 – It says that he’s actually had something that would require antibiotics.
2. F2 – Has he had a chest infection?
3. M2 – He has yeah.

4. F2 – Two chest infections?
5. M2 – Yeah.
6. F2 – Both times?
7. M2 – He has. I have had a chest infection. (PBL2)

Conflict-oriented shared knowledge building. This mode of shared knowledge building is characterised by disagreeing and modifying or replacing the ideas and opinions of the learning partners (Weinberger and Fischer, 2006)

Extract 4:

M2 – This is basically what Vishnu (one of the students) said about the signs and symptoms. Vishnu, do you want to help me out? What crackles are they? It is the bronchi that crackle –

1. M4 – No, it is the weird noise that you get with breathing.
2. F3 – Is it like a snoring sound?
3. M1 – Okay, not crackles. That was a test. I did not get what pink frothy sputum could be.
4. F3 – Pulmonary oedema
5. M1 – Why is it pink? Is that blood in it?
6. F3 – No, it is protein. (PBL6)

Extract 5:

1. F3 – This goes down, but this is also activated by 7 and calcium, so it is a roundabout kind of thing and that goes into those two and 7 also activates… That goes to there and then this goes to 10, but it also goes past 9 to activate 9.
2. F5 – Twelve and then 11 and then 9 and then 10. Clearly this is not the order that they would assemble.
3. F3 – **No** both of these catalyse this reaction, so it is extrinsic and intrinsic coming together and then up there it is common.
4. M2 – Therefore 13 is catalysed, thrombin 2A, fibrin and 13A.
5. F3 – **No**, 13, once thrombin A is activated, along with thrombin 2A and fibrin this one gets activated to cross-linked fibrin, so this

one needs these two in order to make the fibrin cross-linked, which makes it strong.

6. M2 – Yes, I get it.

7. F5 – Would not that go that way then? I do not understand that.

8. F3 – You have thrombin and fibrin and then once that has happened these two can then make this happen, so until this is made this one will not be able to cross-link. The arrows might not be the best.

9. F5 – I am trying to work this out. This 13 here becomes 13A, but it needs fibrin to catalyse that.

10. F3 – **Yes** and thrombin 2A.

11. F5 **–** Thrombin 2A catalyses, so it does not go here, it goes there.

12. F3 – **Yes**. (PBL6)

Extracts 4 and 5 exemplify conflict-oriented shared knowledge building. In extract 4, the students were discussing abnormal breath sounds. In line 1, student M4 gave an elaborate correction of the misconception of the previous speaker. F3 asked a clarification question. In line 3, M1 said that he did not understand what the cause of the frothy sputum might have been and F3 in line 4 provided a reason. In line 5, M1 wondered whether blood was responsible for the pink colour of the sputum. F3 corrected this and said that it was protein rather than blood.

In extract 5, students were discussing clotting mechanisms. F2 gave an elaborated disagreement in line 3 to F5’s contribution in line 2. In line 4, M2 externalised his understanding which F3 disagreed with in line 5 and F3 gave an explanation. In line 6, M2 verbalised his new understanding. In line 7, F5 verbalised her confusion and F3 gave another extensive explanation in line 8. In line 9, F5 verbalised her new understanding. This was endorsed by F3 in line 10 with addition of the information that F5 had missed out. In line 11, F5 externalised her uptake of F3's contribution that the latter had endorsed in line 12. This extract shows how negotiations help to construct knowledge and build shared understanding.

Orientation to information. This occurs when a discourse partner orientates to the contribution of the discourse participants.

Extract 6:

1. M1 – I don't know cos I don't know the root cause of it.
2. M2 – Pain receding question mark I think will do me.
3. M1 – What was that? I didn't get it.
4. M2 – It's the pain receding part of the cramps and what's going on with that.
5. M1 – **Oh yeah**. Cos it stops and it goes away.
6. M2 – Yeah.
7. M1 – So it's obviously something to do with the demand he's placing on his muscles. (PBL7)

Extract 6 shows how reactive indicators were used to mark information awareness or orientation. In line 1, M1 verbalised his lack of understanding. M2 gave a response which M1 did not seem to comprehend (line 4). M2 repeated the explanation to which M1 responded by verbalising his new understanding in line 6. He prefaced his new understanding with 'oh yeah'. Again, his new understanding was endorsed by M2 in line 7.

Acknowledgement. This refers to mere recognition of peer contribution.

Extract 7:

M1 – Because each bronchopulmonary segment has a number.

F2 – Yeah. (PBL2)

Simple negation. In this case, only a negative token is used to respond to a peer contribution

Extract 8:

M3 – It is not 100%.

F3 – No. (PBL2)

Information recall. The indicator is used to mobilise previously acquired but forgotten information.

Extract 9:

M2 – We had Danny's lecture.

M3 – Oh I remember having it but I was not quite sure… (PBL4)

Preface/question token. This is a situation where the students used affirmation indicators as a token of a question.

Extract 7:

1. M1 – You might read in certain places; some physiologists believe that there is actually a communication of cells between here kind of like this bundle that makes it go across faster but it is debatable.
2. M2 – So the myocyte conduction is a bit slower than the bundle of this, **yeah?**
3. M1 – Yeah, myocytes yeah because the Bundle of His – well, we are getting down to that bit next year. (PBL1)

In extract 7, M2 asked a confirming question marked with the shared knowledge indicator ‘yeah’ in line 2. M1 responded to this by giving an elaborated response.

Summary

The results of the analysis of shared knowledge indicators showed remarkable findings. The results of the word frequency analysis tentatively pointed us to what happened in each tutorial group discussion. The result suggests that both agreement and disagreement tokens occurred in each of the PBL group discussions, but the students tended to use affirmation tokens four times more than they used negation tokens. This could mean that the students were agreeing, acknowledging, and confirming opinions, facts, and ideas and were cooperating to accomplish the group task with fewer disputes.

The high occurrence of the structural forms with lexical content and low occurrence of nonlexical content forms suggest that students engaged in discourse as one would expect in the PBL tutorials rather than merely listening as might happen in a traditional learning context. The low occurrence of nonlexical content forms also indicates that the students devoted less time to mere emotional reactions. More importantly, the relatively low frequency of nonlexical content indicators suggests the interactive nature of the group discussions and points to the amount of collaborative efforts that the students made. A collaborative process is considered an effortful process (Baker et al., 1999; Clark and Wilkes-Gibbs, 1986a).

The possible interactive actions deduced from the word frequency analysis were supported by the findings of the KWIC analysis. The dominant interactive patterns were integration-oriented and conflict-oriented consensus building. The findings of the KWIC analysis were quite significant. The social modes of co-construction describes the extent to which learners refer to contributions of their discourse partners and this has been linked to knowledge acquisition (Fischer et al., 2002; Weinberger and Fischer, 2006). Integration-oriented and conflict-oriented shared knowledge building could mean that students were contributing to each other's conversations and were constructing knowledge together. In order to attain common ground in a conflict-oriented discourse, students need to modify their point of view or provide alternative ideas and this develops knowledge (Weinberger and Fischer, 2006: p. 79). Operating at this level of discourse also imposes the need to reason on the students. Orientation to information and information recall are considered to be associated with knowledge growth (Schiffrin, 1987). The question category would suggest that shared knowledge markers were used to monitor understanding and confirm exchanges. The fewer simple negation tokens per 100 tokens may suggest less common contentious talk occurring in the students' discourse. The high relative frequency of agreement function may suggest quick consensus building or the presence of social contact.

The analysis of shared knowledge indicators shows that integration-oriented and conflict-oriented knowledge constructions were very frequent in the discussions of the students. The use of shared knowledge markers for acknowledgement and agreement suggests that students were attending to and confirming each other's contributions. The fewer occurrence of simple negation functions suggests that disputes were infrequent in the students' discourse.

Question 3. Knowledge extension indicators

This section reports on the main quantitative and qualitative results obtained from the analysis of the commonly occurring knowledge extension indicators. In the first section, the results of the word frequency and KWIC analyses are discussed. In the second section, the results of the in-depth qualitative analysis are presented. The results are meant to answer the third research question.

Quantitative analysis

A total of 6,520 knowledge extension indicators were retrieved by the Wmatrix 3 tool. Two approaches were employed to analyse these indicators: (1) word frequency analysis and (2) KWIC analysis.

Word frequency analysis. Table 6.9 shows the raw and relative frequencies of the knowledge extension indicators. There were 6,652

Table 6.9: Frequency of knowledge extension indicators: Raw and relative frequency of the knowledge extension indicators per 100 tokens in each PBL group

	PBL 1 RF NF	*PBL 2 RF NF*	*PBL 3 RF NF*	*PBL 4 RF NF*	*PBL 5 RF NF*	*PBL 6 RF NF*	*PBL 7 RF NF*	*PBL 8 RF NF*	*Total RF NF*
And	621	555	672	378	471	348	368	743	4156
	2.07	1.84	1.83	2.17	2.21	2.25	1.83	1.92	1.98
And stuff	5	2	4	5	0	6	8	8	38
	0.02	0.01	0.01	0.03	0.00	0.04	0.04	0.02	0.02
And other	3	0	1	2	5	0	2	2	15
	0.01	0.00	0.00	0.01	0.02	0.00	0.01	0.01	0.01
And everything	2	1	2	2	2	1	3	0	13
	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.0	0.01
And that is it	0	0	0	1	1	1	0	0	3
	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00
And what not	0	0	0	0	0	1	0	0	1
	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
And what have	0	0	0	0	0	0	0	1	1
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Or	158	165	143	73	105	77	75	147	943
			0.39	0.42	0.49	0.50	0.37	0.38	0.45
Or whatever	9	4	1	4	2	1	2	0	23
	0.03	0.01	0.00	0.02	0.01	0.01	0.01	0.00	0.01
Or anything	7	0	3	3	4	6	1	4	28
	0.02	0.00	0.01	0.02	0.02	0.04	0.00	0.01	0.01
Or so	1	3	0	0	0	0	1	2	7
	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00
But	191	2020	245	98	144	102	161	281	1424
	0.64	.67	0.81	0.56	0.68	0.66	0.81	0.72	0.68
Total	**997**	**932**	**1071**	**566**	**734**	**543**	**621**	**1188**	**6652**
	3.32	3.09	3.55	**3.25**	**3.45**	**3.51**	**3.09**	**3.06**	**3.17**

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

knowledge extension indicators, which were grouped into three main categories – additive, alternative, and adversative. There were 4,227 (63.54%) additive, 1,001 (15.05%) alternative, and 1,424 (21.41%) adversative indicators. The most frequent indicators were ‘and’, ‘or’, and ‘but’, and the three accounted for 98.06% of the total frequency.

Figure 6.9 (below) shows the relative frequency (per 100 tokens) of the commonly occurring indicators in the students’ subcorpus.

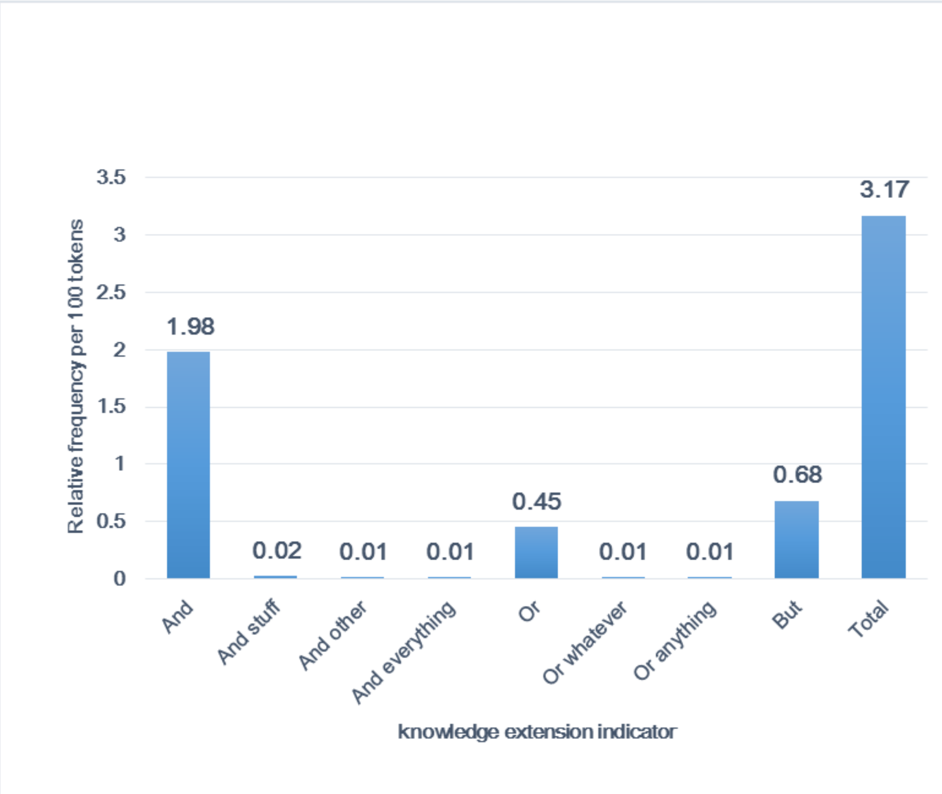


Figure 6.9: Frequent knowledge extension indicators: Relative frequency per 100 tokens of the commonly occurring knowledge extension indicators in the students' subcorpus

The students used close to 2 additive indicators per 100 tokens but used about 0.5 ‘or’ and ‘but’ indicators per 100 tokens. The remaining five indicators were less frequent, having relative frequencies between 0.01 and 0.02 per 100 tokens.

Table 6.10 (below) shows the raw frequency of the indicators used for knowledge construction and knowledge co-construction. The figures indicate that about 16% of the additive, 13% of the alternative, and nearly 33% of the adversative indicators were used for knowledge co-construction.

Table 6.10: Frequency of knowledge extension constructions: Raw and relative frequency of indicators used for knowledge construction and co-construction per 100 tokens

	PBL 1	*PBL 2*	*PBL 3*	*PBL 4*	*PBL 5*	*PBL 6*	*PBL 7*	*PBL 8*	*Total*
Construction	795	682	841	516	672	469	473	922	5370
	2.64	2.26	2.29	2.96	3.16	3.03	2.36	2.38	2.56
Co-construction	202	247	230	50	62	74	148	266	1279
	0.67	0.82	0.63	0.29	0.29	0.48	0.73	0.69	0.61
Total	997	932	1071	566	734	543	621	1188	6652
	3.32	3.09	3.35	3.25	3.45	3.51	3.09	3.06	3.17

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

The relative frequency per 100 of the tokens used for knowledge construction corroborated the raw frequency finding: additive and alternative indicators were less frequently used for knowledge co-construction compared to adversative indicators (see Figure 6.10

below). The relative frequency of the dominant indicators, i.e. of 'and', 'or', and 'but' varied significantly across the tutorial groups: 'and' – LL 129.96, $p < 0.01$; 'or' – LL 99.33, $p < 0.01$; and 'but' – LL 81.35, $p < 0.01$.

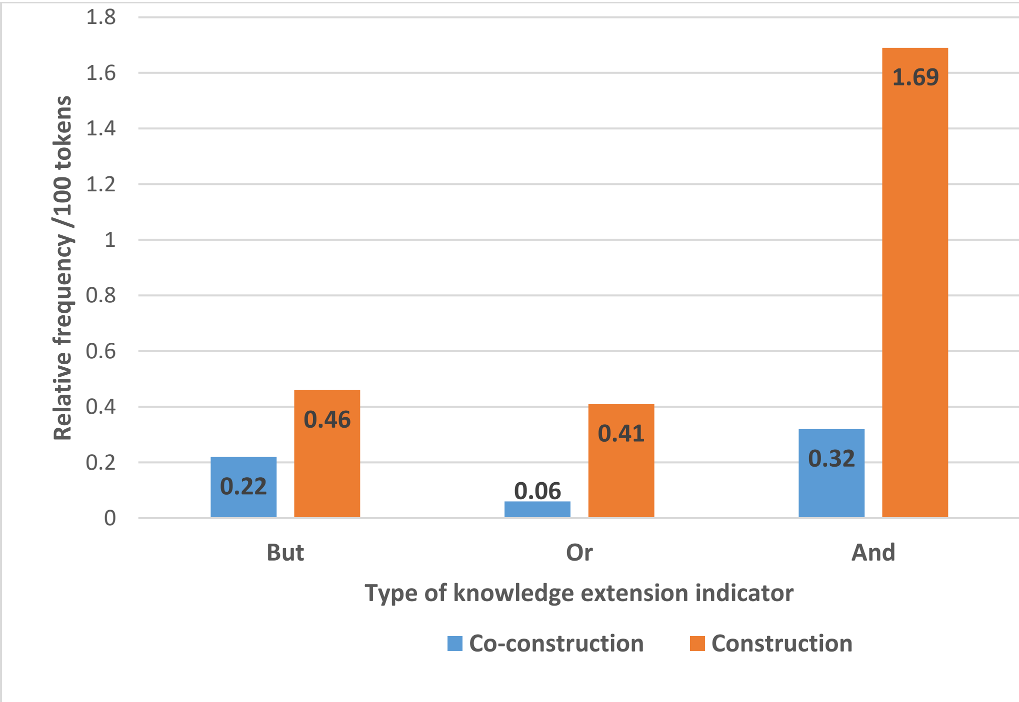


Figure 6.10: Extension indicator knowledge construction types: Relative frequency per 100 tokens of the knowledge construction type with knowledge extension indicator categories in the students' subcorpus

The frequency analysis gives a hint about the verbal interactions in the tutorials. These coordinating indicators suggest that students were continuing or extending, contrasting and possibly giving alternative actions and ideas in their dialogue. A KWIC analysis was conducted to understand the functions of these indicators. The results of the KWIC analysis are presented next.

KWIC analysis. A closer exploration of the commonly occurring words in the relative frequency analysis in the context of the surrounding words led to the identification of the functions of the knowledge extension indicators as highlighted in Table 6.11 (below).

Table 6.11: Frequency of knowledge extension functions: Raw and relative frequency of the knowledge extension indicator function per 100 tokens

	PBL 1 RF NF	***PBL 2 RF NF***	***PBL 3 RF NF***	***PBL 4 RF NF***	***PBL 5 RF NF***	***PBL 6 RF NF***	***PBL 7 RF NF***	***PBL 8 RF NF***	***Total RF NF***
Simple addition	393 1.31	240 0.79	315 0.86	165 0.95	227 1.07	161 1.04	211 1.05	420 1.08	2132 1.01
Temporal	86 0.29	80 0.26	56 0.15	39 0.22	45 0.21	35 0.23	41 0.20	98 0.25	480 0.23
Causal/condition	83 0.28	83 0.27	161 0.44	93 0.53	98 0.46	91 0.59	41 0.20	119 0.31	769 0.37
Elaborate	13 0.04	57 0.19	70 0.19	47 0.27	74 0.35	38 0.25	27 0.13	62 0.16	388 0.18
Contrast	17 0.06	33 0.11	30 0.08	18 0.10	7 0.03	3 0.02	10 0.05	16 0.04	134 0.06
Same word	7 0.02	2 0.01	0 0.00	2 0.01	0 0.00	9 0.06	4 0.02	9 0.02	33 0.02
Indefinite addition	14 0.05	16 0.05	27 0.07	20 0.11	10 0.05	12 0.08	23 0.11	24 0.06	147 0.07
Preface question	0 0.00	0 0.00	3 0.01	0 0.00	5 0.02	2 0.01	0 0.00	0 0.00	10 0.00
Other (additive)	18 0.06	47 0.16	17 0.05	3 0.02	12 0.06	4 0.03	23 0.11	5 0.01	129 0.06
Alternative question	66 0.22	640. 21	35 0.10	18 0.10	33 0.16	25 0.16	25 0.12	35 0.09	301 0.14
Alternative concept	78 0.26	82 0.27	94 0.26	36 0.21	65 0.31	41 0.27	36 0.18	70 0.18	502 0.24

Indefinite alternative	22 0.07	16 0.05	17 0.05	24 0.14	13 0.06	16 0.10	10 0.05	33 0.09	151 0.07

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

Table 6.11: Raw and relative frequency of the knowledge extension indicator function per 100 tokens (cont.)

	PBL 1 RF NF	***PBL 2 RF NF***	***PBL 3 RF NF***	***PBL 4 RF NF***	***PBL 5 RF NF***	***PBL 6 RF NF***	***PBL 7 RF NF***	***PBL 8 RF NF***	***Total RF NF***
Other (alternative)	8 0.03	7 0.02	1 0.00	2 0.01	0 0.00	2 0.01	7 0.03	13 0.03	40 0.02
Preface Elaboration	27 0.09	32 0.11	31 0.08	31 0.18	31 0.15	23 0.15	24 0.12	43 0.11	242 0.12
Contrast	77 0.26	94 0.31	84 0.23	28 0.16	70 0.33	43 0.28	53 0.26	114 0.29	563 0.27
Concession/denial	27 0.09	28 0.09	64 0.17	19 0.11	17 0.08	23 0.15	27 0.13	48 0.12	253 0.12
Corrective	18 0.06	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	4 0.01	22 0.01
Knowledge/reality	13 0.04	20 0.07	26 0.07	6 0.03	7 0.03	2 0.01	21 0.10	25 0.06	120 0.06
Conditional	11 0.04	11 0.04	16 0.04	10 0.06	11 0.05	8 0.05	16 0.08	28 0.07	111 0.05
Preface question	0 0.00	5 0.02	13 0.04	2 0.01	4 0.02	2 0.01	8 0.04	4 0.01	38 0.02
Other (adversative)	18 0.06	12 0.04	11 0.03	2 0.01	4 0.02	1 0.01	12 0.06	15 0.04	75 0.04
Total	996 3.31	929 3.08	1071 2.91	565 3.24	733 3.44	541 3.50	620 3.09	118 5 3.06	6640 3.16

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

The figures in the table show that most of the functions were in three digits. Additive indicators were used to mark simple, temporal, causal- condition, elaborate, contrast, and indefinite addition. Alternative indicators were used for offering alternative questions and ideas while adversative indicators were used to link elaborative, contrasting, concessional and conditional clauses to the previous ones.

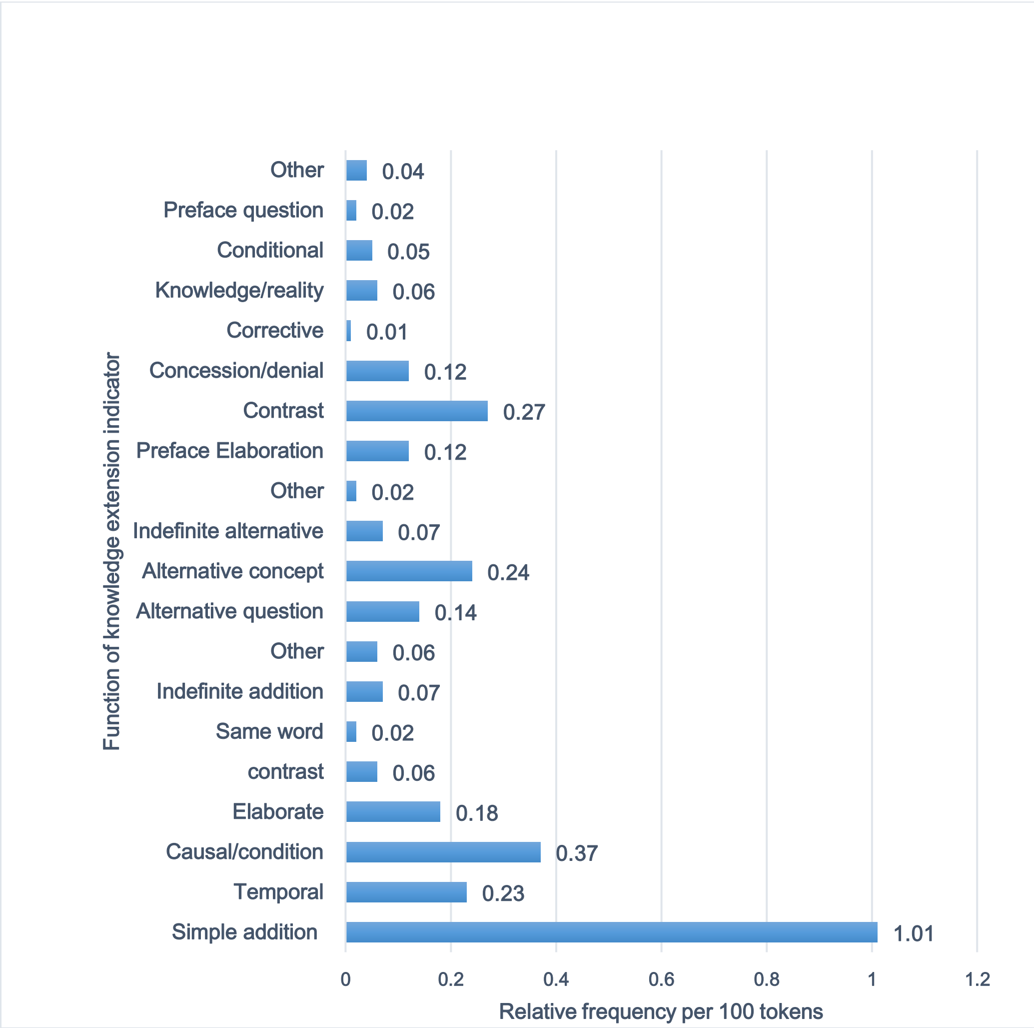


Figure 6.11: Frequent knowledge extension indicator function: Relative frequency per 100 tokens of the commonly occurring knowledge extension indicator functions in the students' subcorpus

Figure 6.11 (above) shows the relative frequencies of the commonly occurring functions of the knowledge extension indicators.

Simple addition, temporal sequence, and causal-conditional and elaborate additions are the most commonly occurring functions of additive indicators. The alternative indicators are mostly used for questioning and concept variation while elaboration, concession, and contrast are the most commonly occurring functions of adversative indicators. There are six commonly occurring functions with relative frequencies between 0.05 and 0.1. Five functions – same word repetition, other additive and alternative functions, adversative correction, and questioning – are less common.

The dominant functions of the additive indicators may suggest that the students were engaged in constructing knowledge by accumulating idea units, and elaborating ideas through cause-conditional relations and temporal relations. The high relative frequency of alternative concepts and questions, and the relatively lower use of alternative indicators for the co-construction of knowledge would suggest that alternative indicators were commonly used by the students to offer personal alternative questions and point of views. This in turn highlights the interactive nature of the discourse. The alternative questions may also reduce the cognitive load inherent in open-ended questions. Furthermore, elaboration, contrast, and concession are the most frequently occurring adversative functions, and these functions possibly occurred at the interpersonal level because of the high use of adversative indicators for knowledge co-construction. This usage would suggest that challenging each other, conceding positions, and elaborating ideas were the tasks that the students engaged in. The next section discusses the results of the qualitative analysis, showing how these functions were carried out in the dataset.

Qualitative analysis

This analysis involved an in-depth exploration of a stretch of the students' talk to understand the framing of the interactive actions with respect to the most commonly occurring functions of the knowledge extension indicators.

Example of construction and co-construction

1. M1 – Any other members of your extended family have asthma?
2. M2 – Actually yeah. My father has a history of nasal allergy.
3. M1 – So he's got rhinitis, interesting.
4. M2 – So mother has dry skin **and** he's got nasal polyps.
5. M1 – So he's got nasal polyps and rhinitis
6. M3 – **And** your mother's got dry skin **and** he's got eczema **and** asthma, so bloody hell…. (PBL2)

In this extract, M2 in line 4 used the indicator 'and' to link biomedical concepts together – an example of knowledge construction. In line 6, M3 linked his contribution to the previous peer's statement by using the turn initial discourse marker 'and'.

Additive indicators

Simple addition. In this extract, 'and' was used to join the pieces of biomedical concepts together and F3 used 'and' to link her contribution to her peer's previous statement.

Extract 1:

F3 – Yeah and his age as well. (PBL3);

M4 – **And** your mother's got dry skin **and** he's got eczema **and** asthma... (PBL2)

Temporal addition. The temporal function of 'and' is enacted when it is used either to stage activities or procedures. This is illustrated in the extract below. It is part of a long stretch of talk about task planning. 'And' was used to jointly deliberate on the sequence of the task.

Extract 1:

1. M4 – I wish we can make it more practical, **and then** we can do the physiology before the anatomy.
2. F2 – Maybe if we do anatomy before Friday **and then** we just do the physiology.
3. F1 – What about the anatomy on Friday **and then** the physiology on the Tuesday?
4. M4 – What is it?
5. F1 – Maybe do anatomy of Friday **and then** the physiology on the Tuesday.

Causal-conditional. Extracts 1 to 3 below describe the causal-conditional relationship between pairs of clauses. Extract 4 shows the conditional relationship established by 'and' prefacing depend.

Extract 1:

M1 – … you get bacterial pathogens going down there **and** causing like a zone of secondary bacterial infection same as the stuff we were doing on pneumonia. (PBL1)

Extract 2:

M4 – You constrict your bronchus, airways **and** it stops that gas going through. (PBL2)

Extract 3:
M3 – … it is carried from one point to another **and** it lodges in the other point… (PBL6)

Extract 4:
F2 – … you get some deep signal moving away and some moving toward **and** depends on where the leads are….. (PBL1)

Elaborative addition. Elaboration function means that 'AND' is used to connect an assertion to an idea unit which comments in some way on the assertion. The comments could be a qualification, clarification or interpretation. The extract below illustrates this function of 'AND'.

Extract 1:
M1 – Some people are just more aware of their heartbeat and it can be perfectly normal. (PBL1)

Extract 2:
F2 – But above that way you've got one tube **and** that's the trachea… (PBL2)

Extract 3:
F1 – The other thing is hypercoagulability **and** that is to do with the constituent makeup of the blood… (PBL6)

Extract 4:
M1 – **And** that's what I was meaning by looking at the classifications underneath coronary artery disease… (PBL8)

The above examples illustrate two explanatory functions of 'and'. In extracts 1 to 3, 'and' was used to link assertion to qualification/clarification, while in extract 4 it was used to connect interpretation to assertion. The finding highlights the use of ‘and’ to mark idea elaboration.

Elaboration may also be in form of contrast. The contrastive marking of an idea unit with ‘and’ signals that the upcoming idea unit differs from the frame of the assertion. The next extract demonstrates this function.

Extract 1:
F2 – I rung them with a chest infection **and** they told me to call an ambulance because I was probably having a heart attack **and** all I wanted was a GP I could go to.

Functions of alternative indicators

This section presents extracts of the most commonly occurring alternative indicator functions.

Alternative question. ‘Or’ also functioned to connect two questions that the students had offered for consideration during the tutorial discourse. The extract below illustrates the framing of the alternative question function of ‘or’ in the dataset.

Extract 1:
1. M4 – I know but you are actually asking us to look through... are we going to look through the actual test or the physiology? (PBL3)
2. F2 – Are they fine crackles or coarse? (PBL4)
3. M4 – Is your stuff also called Bundle of His or it’s the bundled branches of His?
4. F2 – Is that all arrhythmias or just arrhythmias?

These examples show that ‘or’ was most commonly used to offer choices of inclusive alternative questions. The alternative question would suggest that students tended to prefer interactive discourse mode for offering ideas and point of views, and for deliberation.

Alternative concept. The alternative concept function of ‘or’ has several uses, as illustrated in the extracts above. They show that students offered options where either of the options was equally acceptable (as in extracts 1 and 2). Alternative concepts were also offered in uncertain situations (as in extracts 3 and 4). ‘Or’ was also used to connect two alternatives that were mandatory (as in extract 5).

Extract 1:
1. M2 – So say the atrioventricular is myogenic conduction or myocyte conduction. (PBL1)

Extract 2:
2. F3 – Hyperuricaemia is gout **or** leads to gout, which is a big problem in old people anyway. (PBL5)

Extract 3:
3. M1 – I cannot remember if it is left or right atrium. (PBL1)

Extract 4:
4. M1 – Calcium exchange or whatever it is called. (PBL6)

Extract 5:
5. F1 – Because whether or not it’s cultural, he has a concealed weapon. (PBL3)

Functions of adversative indicators

The third category of knowledge extension indicators is adversative, which is marked by ‘but’. It is used to connect an assertion to a contrasting one in some way.

Adversative elaboration. In explanation, the second clause is used to comment on the first clause. The comment may be in the form of extension or restriction.

In the extract below, students were discussing a patient's symptoms. F1 asked whether it was worse in the afternoon. This was paraphrased by M1 in terms of disease periodicity. F1 followed this statement by saying that the daytime exacerbation might be related to the patient's occupation. M2 added to this idea by generating another reason for the exacerbation as ‘a build-up’. This represents an expansive discourse space in which causes of the patient’s symptoms are expanded by using ‘but’.

Extract 1:
F1 – Is it worse in the afternoon?
M1 – Short-term or long-term progression.
F1 – It might mean that because he’s a builder and he’s being… that’s why it’s worse by the afternoon.
M2 – It could mean that **but** it might also mean a build-up; some kind of bio-chemical, metabolic build-up of a problem at the end of the day.
M1 – Like if he’s slowly retaining more and more carbon dioxide you mean?
M2 – Exactly. Or his acidosis or whatever. (PBL4)

Contrastive function. In this extract, M4 was explaining the prevalence of heart failure. After he had mentioned the countries with low prevalence of heart failure, he contrasted the statement by mentioning countries with increasing incidence of heart failure.

> **Extract 1:**
> M4 – It is variable for different countries. You do not see many heart failure cases in places like Africa, China or East Asia, **but** in western countries it is increasing over time especially with age, which is probably the biggest risk factor for it. (PBL6)

> **Extract 2:**
>
> M4 – ... because ultimately you should get it on one of the other leads, I do not know which one it is now, you should get effectively the same sort of amplitude or deflection but in the opposite direction and it doesn't quite work and we get lower amplitude... (PLB1)

In extract 2, M4 was talking about ECG waves. He was saying that if a Q-wave is present on the ECG, it should be seen on one of the leads (though he could not remember which one) that has the same amplitude, but the wave should be on the opposite direction.

Extracts 1 and 2 show something interesting about how the students used words to construct. There was movement from the general to the specific in the ways phenomena were construed. The contrast was enacted at the level of specificity. In extract 1, the student moved from the level of general 'global variability of heart failure' to the specificity of contrasting the countries with higher incidence to those with increasing incidence of heart failure. The same occurs in extract 2 where the student did not stop with saying that the amplitudes of the waves were supposed to be the same, but the contrast was carried out at a deeper level to specify that the direction of one is different from that of the other. This could represent a fine-grained level of information processing.

Concession/denial of expectation. In case of concession, the second clause endorses the proposition expressed in the first clause but also claims that something that may be contradictory is equally true (Barth, 2000).

In the next extract, student M3 proposed that recurrent bronchitis was the cause of recurrent chest infections. He appealed to the collective prior knowledge as evidence. F1 challenged this idea indirectly by asking a verification question. M3 responded by endorsing F1's point of view, but at the same time he claimed that his own point was also possible.

Extract 1:

M3 – I'm sure that the recurring chest infections are just the bronchitis flaring up every now and again, that's what it is... Can you remember? I can't remember where we read it or whether it was in a workshop but you get an inflammation and it damages it and you get fluid trapped somewhere and that starts another one.

F1 – **But** does it always do that?

M3 – I guess it doesn't always have to be but it's likely. (PBL3)

In the extract below, M3 asked a verification question which M1 did not seem to understand. M3 ended his verification questions by stating his own position. Initially, M1 seemed to have agreed with M3, but then changed his mind to say that ‘less is produced but it is still produced’.

Extract 2:

M3 – …Do ACE inhibitors not cancel out angiotensin II as opposed to receptor, because in the lecture she said there were some other methods of angiotensin I being converted?

M1 – Yes, sorry?

M3 – It is the receptor one that blocks aldosterone completely? Do they both do all those things? I guess they do if you reduce it.

M1 – Yes, it does not block the receptor, sorry. I just mean that less is produced, **but** it is still produced through this way. Sorry, I should not have said that. (PBL6)

The concessional use of ‘but’ in the two extracts fulfils co-constructive and constructive functions. The co-construction function is shown by the turn initial position of ‘but’ in F1’s contribution. This contrast move could provide the opportunity for M3 to restructure his knowledge in order to attain a new understanding. ‘But’ serves the function of ‘although’ when it performs a concessional role. The use of ‘although’ makes one idea unit subordinate to the other, but the use of ‘but’ coordinates the two idea units and assigns equal status to them. By removing the inferior–superior dichotomy with respect to the two idea units, the scope for further negotiating the two ideas in the future remains open (Barth, 2000). It also implies that the two ideas are added to the knowledge deposits of the discourse participants. This is another indicator of an expansive stance in which all ideas are considered tentative and evolving. Furthermore, concession also indicates a cooperative and polite stance in discourse participants.

Denial of expectation. Denial of expectation implies that the second clause contradicts an expectation aroused by the first clause. The students were talking about improved treatment of heart failure, which has not translated to improvement in mortality despite improved forms of treatment. The link between expectation and its denial can be paraphrased as follows: ‘although we have become very good at treating heart failure, the mortality rate is still pretty high’. This sentence implies negated causal-consequence relation, suggesting that the student in question was engaging in clinical reasoning.

Extract 1:

M4 – We have become very good at treating heart disorders, **but** the mortality rate is still pretty high: half of patients diagnosed will be dead within five years and about 5% of all hospital admissions are linked to heart failure. It costs £1 billion a year to treat, so it is a pretty big hit on funding as well as on people’s lives. (PBL6)

In summary, knowledge extension indicators (coordinating conjunctions) were used for additive, alternative, and adversative functions. Beyond these uses, they performed subconjunctive functions. In-depth contextual analysis of their use suggests that they were used autonomously and collaboratively to construct knowledge through their capacity of idea expansion, negotiation, and elaboration.

Question 4. Knowledge enhancement construction

Knowledge enhancement indicators retrieved from the corpus were analysed at the level of word frequency and KWIC, and at the qualitative level. The results of the quantitative analysis are presented first.

Quantitative analysis

A total of 6,402 knowledge enhancement indicators were retrieved from the students' subcorpus.

Word frequency analysis. Table 6.12 (below) shows the raw frequencies of the knowledge enhancement indicators used by the PBL group. The commonly used indicators were 'that', 'because', 'if', 'so', 'as', and 'when'. These indicators were consistently used across the groups. The least commonly used indicators were 'since' and 'although'.

Table 6.12: Frequency of knowledge enhancement indicators: Raw and relative frequency of knowledge enhancement indicators per 100 tokens

	PBL1 RF NF	*PBL2 RF NF*	*PBL3 RF NF*	*PBL4 RF NF*	*PBL5 RF NF*	*PBL6 RF NF*	*PBL7 RF NF*	*PBL8 RF NF*	*Total RF NF*
Because	161 0.54	177 0.59	216 0.59	101 0.58	108 0.51	88 0.57	134 0.67	275 0.71	1260 0.60
If	177 0.59	169 0.56	217 0.59	126 0.72	167 0.78	103 0.67	143 0.71	307 0.79	1409 0.67
So	114 0.38	135 0.45	146 0.40	48 0.28	88 0.41	62 0.40	83 0.41	146 0.38	822 0.39
When	54 0.18	55 0.18	56 0.15	41 0.24	29 0.14	26 0.17	36 0.18	67 0.17	364 0.17
Where	19 0.06	15 0.05	14 0.04	13 0.07	4 0.02	2 0.01	9 0.04	16 0.04	92 0.04

Before	11 0.04	4 0.01	18 0.05	2 0.01	10 0.05	1 0.01	2 0.01	14 0.04	62 0.03
While	8 0.03	1 0.00	9 0.02	0 0.00	2 0.01	0 0.01	1 0.00	8 0.02	29 0.01
Until	6 0.02	4 0.01	5 0.01	1 0.01	0 0.00	5 0.03	1 0.00	6 0.02	28 0.01
Though	5 0.02	3 0.01	10 0.03	9 0.05	4 0.02	1 0.01	4 0.02	7 0.02	43 0.02
Although	3 0.01	3 0.01	2 0.01	2 0.01	2 0.01	3 0.02	1 0.00	2 0.01	18 0.01

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

Table 6.12: Raw and relative frequency per 100 tokens of knowledge enhancement indicators (cont.)

	PBL1 RF NF	***PBL2 RF NF***	***PBL3 RF NF***	***PBL4 RF NF***	***PBL5 RF NF***	***PBL6 RF NF***	***PBL7 RF NF***	***PBL8 RF NF***	***Total RF NF***
Unless	3 0.01	3 0.01	8 0.02	4 0.02	1 0.00	3 0.02	4 0.02	1 0.00	27 0.01
Once	2 0.01	4 0.01	1 0.00	0 0.00	5 0.02	5 0.03	2 0.01	10 0.03	29 0.01
Whereas	2 0.01	6 0.02	6 0.02	2 0.01	0 0.00	2 0.01	7 0.03	7 0.02	32 0.02
Like	2 0.01	6 0.02	4 0.01	3 0.02	1 0.00	1 0.01	1 0.00	8 0.02	26 0.01
Since	2 0.01	4 0.01	3 0.01	0 0.00	1 0.00	0 0.00	3 0.01	31 0.08	16 0.01
After	1 0.00	6 0.02	4 0.01	1 0.01	2 0.01	3 0.02	1 0.00	7 0.02	25 0.01
As	18 0.06	25 0.08	50 0.14	21 0.12	27 0.13	18 0.11	23 0.11	43 0.11	225 0.11
Than	15 0.05	8 0.03	27 0.07	6 0.03	13 0.06	11 0.07	6 0.03	19 0.05	105 0.05
That	233 0.77	201 0.67	253 0.69	140 0.80	182 0.86	164 1.06	181 0.90	333 0.86	1687 0.80
Whether	18	15	16	5	15	6	16	18	109

	0.06	0.05	0.04	0.03	0.07	0.04	0.08	0.05	0.05
Total	854	838	1065	525	661	504	658	1297	6402
	2.84	2.78	2.90	3.01	3.11	3.26	3.28	3.34	3.05

RF – raw frequency; and NF – normalised or relative frequency per 100 tokens

Figure 6.12 (below) shows the relative frequency per 100 tokens of the commonly occurring knowledge enhancement indicators in the students' subcorpus. Six indicators were most commonly occurring in the students' subcorpus with relative frequency in excess of 0.1 per 100 tokens each.

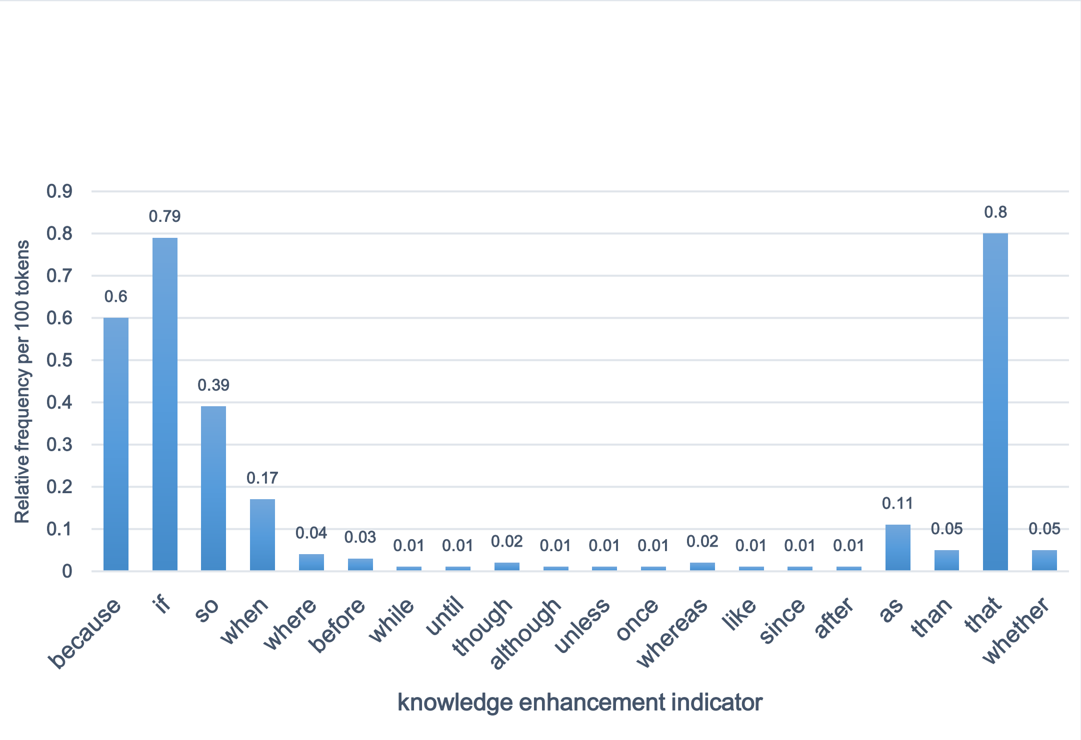


Figure 6.12: Frequent knowledge enhancement indicators: Relative frequency per 100 tokens of the commonly occurring knowledge enhancement indicators in the students' subcorpus

Overall, ‘that’ (0.8), ‘if’ (0.79), ‘because’ (0.6), ‘so’ (0.39), ‘when’ (0.17), and ‘as’ (0.11) were the most commonly occurring indicators in the students’ subcorpus.

Table 6.13 shows (below) that overall, three-quarters of the total indicators (4,780) were used for knowledge construction while about a quarter was used for knowledge co-construction. Specifically, across the PBL groups, between 16.94% and 29.24% of the indicators were used for co-construction while 70.76% to 83.06% of them were used for autonomous knowledge construction.

Table 6.13: Frequency of enhancement constructions: Raw and relative frequency of knowledge enhancement indicators used for knowledge construction and co-construction by PBL groups

	PBL 1 RF NF	*PBL 2 RF NF*	*PBL 3 RF NF*	*PBL 4 RF NF*	*PBL 5 RF NF*	*PBL 6 RF NF*	*PBL 7 RF NF*	*PBL 8 RF NF*	*Total RF NF*
Construction	661 2.20	593 1.96	762 2.07	387 2.22	549 2.58	390 2.52	503 2.51	935 2.41	4780 2.28
Co-construction	193 0.64	245 0.81	303 0.82	138 0.79	112 0.53	114 0.74	155 0.77	362 0.93	1622 0.77
Total	854 2.84	838 2.78	1065 2.90	525 3.01	661 3.11	504 3.26	658 3.28	1297 3.34	6402 3.05

Overall, the students used more than three-quarters (74.66%) of the indicators per 100 tokens for autonomous knowledge construction and about a quarter (25.34%) for co-construction. Across the PBL groups, between 0.53 and 0.93 indicator per 100 tokens were used for knowledge co-construction while between 1.96 and 2.58 indicators per 100 tokens were used for autonomous knowledge constructions.

Based on the results of word frequency analysis, certain predictions can be made. The high relative frequencies of ‘because’, ‘if’, and ‘so’ may be a clue to the reasoning and explanations going on in the students’ conversation. The results also indicate that the students used about one-quarter as many indicators for co-construction as they used for autonomous construction, thus suggesting that about one-third of the reasoning and explaining took place between discourse partners. A KWIC analysis was carried out to provide more insights into the interactive actions of the students. The results of this analytic approach are the subject of the next section.

KWIC analysis. The reading of the knowledge enhancing indicators in the context of the surrounding words led to the identification of their functional categories. The raw and relative frequencies of the functions are presented in Table 6.14 (below). There were 34 functional types of knowledge enhancing indicators. Each of the indicators is present in tens.

Table 6.14: Frequency of enhancement functions: Raw and relative frequency of the functions of knowledge enhancement indicators per 100 tokens

	PBL1 RF NF	PBL2 RF NF	PBL3 RF NF	PBL4 RF NF	PBL5 RF NF	PBL6 RF NF	PBL7 RF NF	PBL8 RF NF	Total RF NF
Correlation	4 0.01	8 0.03	7 0.02	0 0.00	5 0.02	7 0.05	8 0.04	7 0.02	46 0.02
Comparison	26 0.09	6 0.02	32 0.09	12 0.07	26 0.12	13 0.08	27 0.13	47 0.12	189 0.09
Report	72 0.24	38 0.13	61 0.17	25 0.14	72 0.33	66 0.43	60 0.30	76 0.20	470 0.22
Extension	25 0.08	60 0.20	110 0.30	14 0.08	77 0.36	46 0.30	74 0.37	126 0.32	532 0.25
Alternatives	8 0.03	15 0.05	0 0.00	3 0.02	10 0.05	4 0.03	12 0.06	0 0.00	52 0.02
Alternative	27 0.09	6 0.02	24 0.07	10 0.06	0 0.00	7 0.05	6 0.03	26 0.07	101 0.05
Possibility	41 0.14	35 0.12	22 0.06	27 0.15	10 0.05	30 0.19	15 0.07	22 0.06	202 0.10
Concession	0 0.00	6 0.02	12 0.03	12 0.07	6 0.03	3 0.02	7 0.03	12 0.03	58 0.03
Conditional	132 0.44	140 0.46	187 0.51	114 0.65	140 0.66	74 0.48	98 0.49	218 0.56	110 0.53
Animated	3 0.01	3 0.01	1 0.00	9 0.05	0 0.00	1 0.01	1 0.00	1 0.00	19 0.01
Tautological	2 0.01	3 0.01	5 0.01	0 0.00	0 0.00	0 0.00	2 0.01	5 0.01	17 0.01
Rationale	4 0.01	13 0.04	15 0.04	1 0.01	0 0.00	14 0.09	3 0.01	30 0.08	80 0.04
Existential	5 0.02	0 0.00	1 0.00	1 0.01	6 0.03	1 0.01	1 0.00	1 0.00	16 0.01
Verbal	8 0.03	0 0.00	0 0.00	4 0.02	0 0.00	0 0.00	2 0.01	0 0.00	14 0.01
Genetic	11 0.04	7 0.02	4 0.01	5 0.03	6 0.03	5 0.03	8 0.04	3 0.01	49 0.02
Epistemic	13 0.04	16 0.05	10 0.03	7 0.04	4 0.02	8 0.05	7 0.03	10 0.03	75 0.04

Hypothetical	14	0	10	0	25	0	3	10	62
	0.05	0.00	0.03	0.00	0.12	0.00	0.01	0.03	0.03
Functional	10	15	17	3	5	2	5	18	75
	0.03	0.05	0.05	0.02	0.02	0.01	0.02	0.05	0.04
Feature	37	50	62	31	38	32	35	77	362
	0.12	0.17	0.17	0.18	0.18	0.21	0.17	0.20	0.17
Mechanical	24	8	4	6	4	0	10	4	36
	0.08	0.03	0.01	0.03	0.02	0.00	0.05	0.01	0.02
Mechanistic	14	14	38	24	26	18	21	28	183
	0.05	0.05	0.10	0.14	0.12	0.12	0.10	0.07	0.09
Inference	51	78	65	20	32	20	23	80	369
	0.17	0.26	0.18	0.11	0.15	0.13	0.11	0.21	0.18
Psychological	15	13	28	5	4	2	19	38	124
	0.05	0.04	0.08	0.03	0.02	0.01	0.09	0.10	0.06
Consequence	64	29	76	26	60	38	34	71	398
	0.21	0.10	0.21	0.15	0.28	0.25	0.17	0.18	0.19
Durative	9	33	35	20	0	11	10	47	165
	0.03	0.11	0.10	0.11	0.00	0.07	0.05	0.12	0.08
Locative	11	26	17	10	1	6	16	21	108
	0.04	0.09	0.05	0.06	0.00	0.04	0.08	0.05	0.05
Stative	12	23	20	18	13	8	10	18	122
	0.04	0.08	0.05	0.10	0.06	0.05	0.05	0.05	0.06
Order	13	12	21	3	15	5	9	23	101
	0.04	0.04	0.06	0.02	0.07	0.03	0.04	0.06	0.05
Similarity	0	20	9	16	11	8	8	11	83
	0.00	0.07	0.02	0.09	0.05	0.05	0.04	0.03	0.04
Simultaneity	63	14	38	8	21	10	14	51	219
	0.21	0.05	0.10	0.05	0.10	0.06	0.07	0.13	0.10
Mathematical	0	2	18	1	0	0	4	21	46
	0.00	0.01	0.05	0.01	0.00	0.00	0.02	0.05	0.02
Contrast	0	9	6	4	0	3	2	5	29
	0.00	0.03	0.02	0.02	0.00	0.02	0.01	0.01	0.01
Instantaneity	0	3	0	0	0	8	3	1	15
	0.00	0.01	0.00	0.00	0.00	0.05	0.01	0.00	0.01
Behavioural	0	12	18	0	3	0	0	21	54
	0.00	0.04	0.05	0.00	0.01	0.00	0.00	0.05	0.03

The figures in the tables show that conditional, extension, report, consequence, and inference functions have high frequencies while verbal, instantaneity, existential, tautological, and animated functions have the least frequencies. Of the 34 type of functions, only 16 were consistently used across the PBL groups, while 18 functions were not consistently used.

Figure 6.13 (below) shows the relative frequency of the commonly occurring functions of the knowledge enhancement indicators in the students' subcorpus. The indicators can be grouped into three categories. There were 6 most commonly occurring indicators with relative frequencies more than 0.1 per 100 tokens. Ten indicators fell into the commonly occurring categories with relative frequencies of 0.05 to 0.1 per 100 tokens. Eighteen indicators were less commonly occurring with a relative frequency of less than 0.05 per 100 tokens.

The high relative frequencies of conditional, extension, report, consequence, inference, and feature specification functions may mean that the students were mobilising prior experience, defining phenomena and constructing explanatory knowledge in their discussions. However, an in-depth qualitative analysis of the use and function of the most commonly occurring knowledge enhancing indicators was carried out to make sense of the quantitative results. The discussion of the results of the qualitative analysis is the subject of the next section.

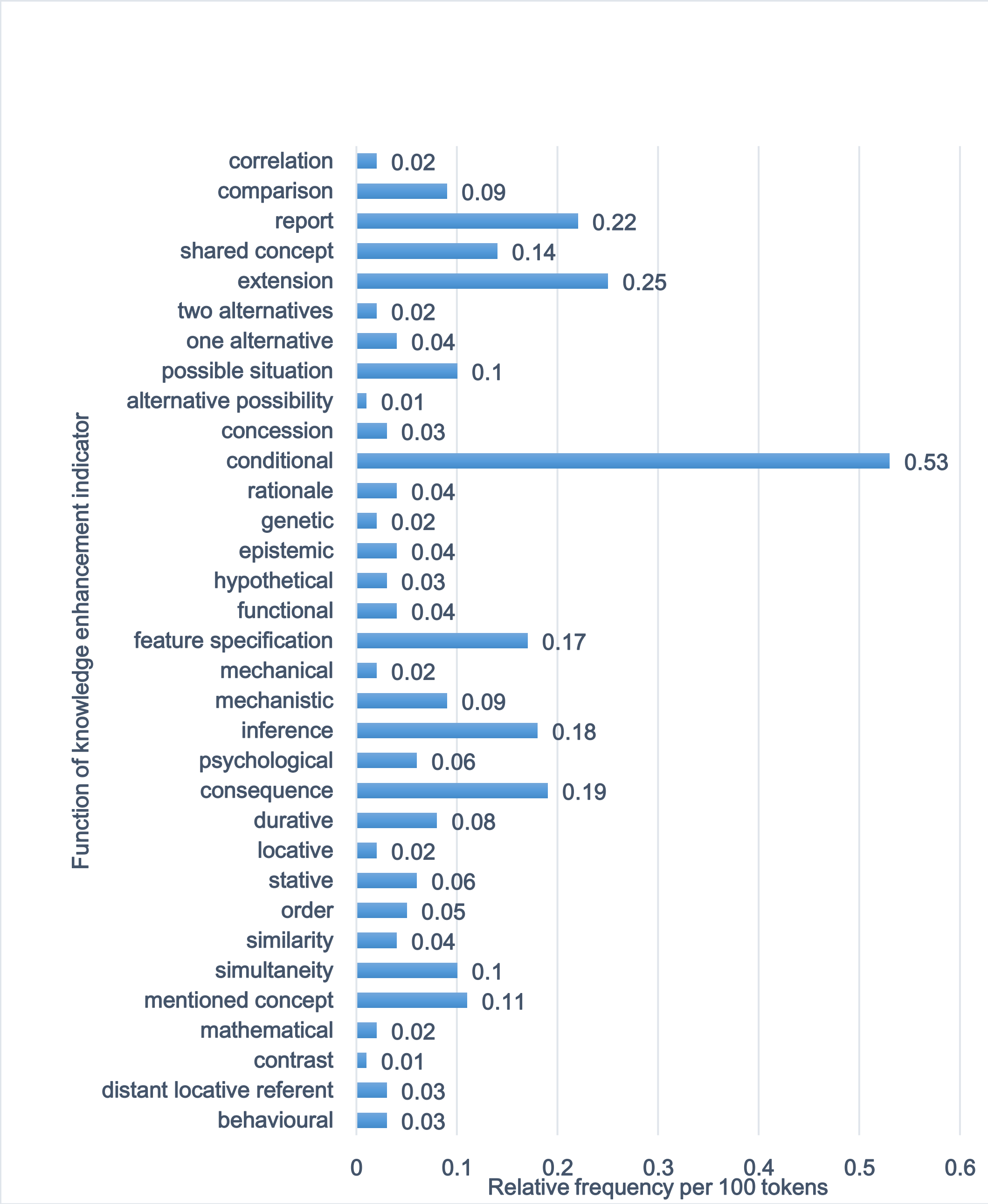


Figure 6.13: Frequent knowledge enhancement indicator functions: Relative frequency per 100 tokens of the commonly occurring functions of knowledge enhancement indicators in the students' subcorpus

Qualitative analysis results

The KWIC analysis generated many functional categories (Figure 6.13). The most commonly occurring functions were selected for further exploration with qualitative approach.

Report. Qualitative exploration of the report function showed that students used ‘that’ to report relational, cognitive, and material processes. This function was widespread across the tutorial groups.

Extract 1:

M1 – But there is also the hygiene bits, the hygiene theory as well which says that if you’re not exposed to enough viruses and bacteria then your immune system is more likely to move towards the TH2. (PBL2)

Extract 2:

F2 – Well look, Ben thought that he knew it so he stood up and...
M1 – I thought I knew it.
M1 – Yeah. Who knows the most about this anatomy? (PBL4)

Extract 3:

M1 – And they’ve proved that if you do that it reduces your likelihood of stopping and staying stopped. (PBL3)

Extract 4

M1 – You might read in certain places, some physiologists believe that there is actually a communication of cells between here kind of like this bundle that makes it go across faster but it is debatable. (PBL5)

Extract 5:

F1 – I remember there's a word that's a mnemonic…

M1 – I still don't know if you should call it a mnemonic because it's not particularly… is it? (PBL7)

Extract 6:

F1 – What's making…? I know it says smokers' cough but we don't know that he smokes.

F2 – Yeah so shall we say, 'Are you a smoker or passive smoker?' (PBL5)

Extract 7:

M3 – If she had leg cancer and it was blocking something here, you would get unilateral pitting oedema, would you not?

M1 – I am not convinced that we know exactly what the difference is between pitting oedema and non-pitting. (PBL4)

These extracts suggest that students evoked theories to make sense of the case problem (extract 1); utilised peer knowledge when discussing issues (extract 2); allowed research evidence to influence their discussions (extract 3); discussed diverse professional viewpoints (extract 4); used a cognitive tool in the form of mnemonics to aid their information retrieval (extract 5); and critically evaluated propositions (extracts 6 and 7).

Extension. Another most frequent function of knowledge enhancement indicators is extension. This function is majorly related to the indicator 'that'. In this instance, the students amplified idea units using 'that' to add information specificity.

Extract 1:

F2 – … there is a tool called the something calculator…. it is like a tool **that** GPs use to calculate your risk of having a cardiac episode of some description and calculate your five year risk and ten year risk. (PBL1)

Extract 2:

M1 – Absolutely. It could be **that** he has concurrent angina as well.

F2 – Could it be muscular? Maybe some sort of muscle problem that's making it harder for him to breath.

M1 – Absolutely (PBL3)

Extract 3:

F2 – **...** and it is just a response **that** our body seems to go into pretty easily for a number of different reasons, the usual – smoking, being hugely overweight, too much salt, drinking too much, stress. (PBL4)

The above extracts illustrate the extension function of 'that' in the dataset. It was used to construct knowledge by adding specificity and precision to previous information.

Feature specification. Feature specification was one of the most commonly occurring function of knowledge enhancement indicators with a relative frequency of 0.17 per 100 tokens. The following extracts highlight the framing of the feature specification in the dataset.

Extract 1:

F2 – So you end up dehydrated **because** you go to the bathroom more to get rid of urine (PBL2

Extract 2:

M3 – Are you drinking your pee **because** it is sweet? (PBL2)

Extract 3:

M4 – Yes, but if you have heart failure **because** you have had an MI or because you have hypertension and you have – (PBL4)

Extract 4:

M1 – And he's got hyper-responsiveness of the bronchi **because** he's having exercise induced asthma. (PBL2)

Extract 5:

F1 – It might mean that **because** he's a builder and he's being… that's why it's worse by the afternoon. (PBL3)

In these extracts, feature specification was used to provide reason or explanation for the assertion of the previous clause. Students provided functional reasons for the clinical feature of a disease (extract 1); goal-oriented reasons for human actions (extract 2); pre-existing reasons for disease presence (extract 3); implicit mechanistic reasons for observed symptoms (extract 4); and mechanical reasons for disease periodicity (extract 5).

Conditional function. Conditional function is the most frequently occurring function in the dataset with the highest relative frequency of 0.53 per 100 tokens. The extracts below illustrate the framing of conditional functions in the dataset.

Extract 1:

F1 – He might have had a cocoa before he went to bed.

F2 – No but some cardio problems, like do you remember them saying **if** you have got congested heart failure you go to sleep and then you wake up because you can't breathe anymore. And I wonder if palpitations are related to why he's not sleeping.

Extract 2:

M1 – If you have got a very high ventricular rate, **if** you are symptomatic from it you use rate control to reduce their symptoms… (PBL1)

Extract 3:

M1 – **… if** you're not exposed to enough viruses and bacteria then your immune system is more likely to move towards the TH2. (PBL2)

Extract 4:

M2 – **If** too much fluid builds up whereby the lymphatic system cannot deal with it, that is when oedema starts to take place and it happens for various reasons. (PBL6)

Extract 5:

M2 – **… if** you have hypertension and you have a very high capillary hydrostatic pressure in the arterial end, you get a lot of fluid being pushed into the interstitial space and not enough being drained back to compensate. (PBL6)

These extracts illustrate that the students used the 'if' knowledge enhancement indicator to link assertions to their conditions. The logical relations suggest that diseases, and their manifestations and treatments are dependent on situation and circumstance.

Inferential function. This function indicates that the knowledge enhancement indicators were used to link logical deductions to their premises.

Extract 1:
F2 – Syncope.
F3 – No, she called for help, **so** she is not unconscious. (PBL8)

Extract 2:
F2 – … because with any heart block you have decreased output anyway, **so** it can lead to dizziness… (PBL5)

Extract 3:
M2 – The simple definition is, out of place. Atopic means out of place.
F2 – **So** it means out of place allergic reaction.
M2 – Yes. (PBL2)

Extract 4:
M1 – If you have ischaemia or maybe an infarction or something, that would cause damage to a certain part of the heart, which is then causing the heart failure, **so** that is not necessarily the specific cause of the heart failure. (PBL4)

These extracts illustrate how students used the ‘so’ connective to link logical deductions to their premises.

Consequential function. In this instance, an idea unit is expanded by linking it with its consequence via the use of a knowledge enhancement indicator.

Extract 1:

M5 – Angiotensin also acts on the heart to promote myocyte and fibroblast activity, **so** it remodels and you get a bit of a messed up heart. Of course, the liver produces angiotensin. (PBL5)

Extract 2:

F2 – You pee the glucose out because your… is it pancreas?
M2 – Your pancreas can't get rid of it **so** you pee it out.
F2 – **So** you end up dehydrated because you go to the bathroom more to get rid of… (PBL1)

In these extracts, physiological mechanisms and organ functions were linked to their consequences with the use of 'so'.

Possible situation. The 'if' knowledge enhancement indicator was also used to link a possible situation to a previous assertion. The function is illustrated by the following extracts (extracts 1 and 2).

Extract 1:

F2 – Do you mind if we go through remodelling and then we will come back to that? (PBL5)

Extract 2:

M4 – I'm just wondering if anybody knows. (PBL3)

Simultaneity. Simultaneity relates to the temporal function of the indicators. It relates to two or more actions, events or processes occurring at the same time. The following extracts illustrate this function.

Extract 1:

M2 – De Musset's sign is head nodding **while** your heart beats, which I find brilliant. (PBL5)

Extract 2:

F1 – Okay. Have you had any pain in your legs **when** you have been doing any exercise? (PBL7)

Extract 3:

F2 – Has anyone ever had palpitations **when** you are in the bath and it is too hot? (PBL1)

Comparison function. The comparison function of the knowledge enhancement indicators is achieved in the dimensions of similarity (extracts 1 and 3), quantity (extract 2), and functionality (extract 4).

Extract 1:

F3 – They are, yes. They are the **same** as potassium-sparing, but they are in a different category. (PBL5)

Extract 2:

F2 – What is this sickle cell disease that is…?
M3 – more prevalent in Africa
F2 – Yeah, than it is here. (PBL2)

Extract 3:

F2 – Briefly, the definition is, as you said, a syndrome not a diagnosis. (PBL5)

Extract 4:
M1 – Once aldosterone enters the cell it binds to its receptor and does the same thing as the steroids from the previous module: (PBL5)

Mechanistic function. In this instance, the knowledge enhancement indicator links a previous assertion to the following clause, which provides the physiological explanation for the assertion. Extracts 1 and 2 below illustrate this function in the dataset.

Extract 1:
M1 – People with respiratory diseases like, well, lots of different respiratory diseases have headaches when they wake up **because** their gas exchange or their ventilation changes at night. (PBL2)

Extract 2:
F2 – It does not just spread all the way down the ventricles **because** you need the ventricles to contract from the bottom of the bloods. I love that toothpaste analogy. I love the toothpaste analogy, it is great. (PBL8)

Summary

In summary, the commonly occurring knowledge enhancement indicators were ‘because’, ‘than’, ‘that’, ‘so’, ‘if’, ‘when’, ‘as’, and ‘whether’. The indicators were frequently used for conditional, consequential, causal, inferential, reporting, extension, and feature specification functions. About one-third of the knowledge enhancement indicators were used for knowledge co-construction. The qualitative analysis provided a richer insight into the dataset. It was found that students co-constructed explanatory case models by linking causes to clinical symptoms and other disease features through collaborative discussions of the case.

6.4 Summary of the Chapter

The students used referring expression indicators to externalise their own thoughts, and to refer to peer, collective, tutor, and patient thought and talk. They were used to generate hypothesis, construct knowledge, define group task, and understand case patient's expectation and psychological issues. The referenced educational resources – which included lectures, lecture notes, journal articles, textbooks, and concept maps – were used to share learning materials, explain concepts, guide discussions, and resolve idea conflicts.

Integration-oriented and conflict-oriented knowledge constructions were very frequent in the discussions of the students. The agreement and acknowledgement indicators suggest that students were attending to and confirming each other's contributions. The presence of fewer simple negation indicators might suggest that disputes were infrequent in the students' discourse.

The knowledge extension indicators were used for additive, alternative, and adversative functions. The also performed subconjunctive functions. They were used autonomously and collaboratively to construct knowledge because of their capacity for idea expansion, negotiation, and elaboration. Adversative indicators were often used for knowledge co-construction.

The knowledge enhancement indicators were often used for conditional, consequential, causal, inferential, reporting, extension, and feature specification functions. About one-third of the knowledge enhancement indicators were used for knowledge co-construction. The students tended to engage in co-construction of explanatory case models by linking causes to clinical symptoms and other disease features through collaborative discussions of the case.

CHAPTER 7

RESULTS II: ANALYSIS OF FACILITATORS' TALK

7.1 Introduction

This is the second chapter of the presentation of results. It presents the findings of the analysis of the facilitators' talk. Participants in the group of the facilitators were labelled T1 to T8. There were three females and five male facilitators, and each of the facilitators guided each of the PBL tutorial groups.

7.2 Facilitators' Subcorpus Description

The facilitators' subcorpus consisted of 27,743 words. The semantic tags and domains of the facilitators' subcorpora were compared with the BNC sampler CG Educational spoken corpus. The relative frequencies of the five most frequent semantic tags and domains are presented in Table 7.1. The semantic domains in the table show that the facilitators frequently discussed anatomy, physiology, diseases, medicines, medical treatments, and cause-and-effect connections.

Table 7.1: Semantic domains in facilitators' talk: Relative frequencies of the five most commonly occurring semantic tags and domains per 100 tokens in the facilitators' subcorpus

	PBL1	PBL2	PBL3	PBL4	PBL5	PBL6	PBL7	PBL8
Z4	1.26	1.88	1.76	1.95	4.18	1.91	2.19	1.34
B1	2.17	1.88	–	1.66	1.85	2.13	1.83	–
B2-	0.99	0.90	–	1.14	1.94	1.64	1.79	1.31
A2.2	1.16	–	–	–	–	–	1.33	–
B3	0.08	–	1.17	–	0.94)	1.03		1.05
O4.5	–	0.65	–	–	–	–		
A15+	–	0.41	–	–	–	–		
F3	–	–	1.17	–	–	–		
A15-	–	–	0.37	–	–	–		
T2-	–	–	0.81	–	–	–		
A13.6	–	–	–	0.76	–	–		
Z1	–	–	–	0.10	–	–	0.05	0.09
K5.2	–	–	–	–	0.26	–		
A13						0.15		
A4.1								1.00

B1 – anatomy & physiology; B2 – disease; B3 – medicines & medical treatment; A2.2 –cause-and-effect connection; Z4 – discourse bin; F3 – smoking & non-medical drugs; O4.51 – texture; A15+ – safe; A15 – danger; T2 – Time: ending; A13.6 – degree: diminishers; Z1 – proper names; K5.2 – games; A13 – degree; and A4.1 – kinds, groups, and examples in general.

The analysis of the facilitators' talk was aimed to answer the following research questions:

1. What are the frequencies and functions of the question indicators in the facilitators' contributions?

2. What are the frequencies and functions of the commonly occurring directive expression indicators?

3. What are the frequencies and functions of the commonly occurring probability expression indicators in the facilitator's talk?

7.3 Results Presentation Approach

The approach to result presentation in this case was similar to that adopted for the presentation of the students' results. It is divided into two parts: quantitative and qualitative. For each question, the results of the quantitative analysis are presented first followed by a presentation of the qualitative analysis results.

Question 1: Facilitators' Questions

Quantitative analysis results

The quantitative analysis consists of word frequency and KWIC analyses.

Word frequency analysis. Table 7.2 (below) shows the pattern of tokens relating to the questions asked by the facilitators. A closer examination of the results shows striking findings. The figures in the table suggest wide variation in the use of the question indicators across the PBL groups. Of a total of 418 question indicators, about one-fifth occurred in PBL7. The 'what' question indicator accounted for nearly one-third (28.95%) of the total 418 indicators. Of the 35 types of indicators, three – 'what', 'why', and statement tokens – were consistent across the groups. Only 28.57% of the indicators had a total raw frequency of more than 10, indicating that they were the most commonly occurring indicators. The 10 most commonly occurring indicators were 'does', 'what', 'do', 'why', 'how', 'is', 'are', 'would', 'did', and statements (Figure 7.1 below).

Table 7.2: Frequency of facilitators' question indicators: Raw frequency of the indicators of the questions asked by the facilitators

	PBL1	PBL2	PBL3	PBL4	PBL5	PBL6	PBL7	PBL8	Total
Does	7	2	0	1	1	2	3	0	16
Could	2	1	1	1	0	0	1	0	6
What	35	3	1	15	18	8	34	7	121
Do	5	1	0	3	12	4	6	0	31
Statement	3	1	2	2	1	2	6	10	27
Can	2	1	0	0	1	0	2	0	6
Why	7	1	2	1	1	2	8	2	24
How	7	6	2	1	0	1	10	0	27
Is	2	0	0	8	5	8	9	3	35
Are	1	5	2	2	5	3	2	0	20
Which	2	0	0	0	1	0	1	0	4
Have	1	0	1	3	0	2	0	0	7
Or	1	0	0	0	0	1	1	1	4
When	1	0	0	0	2	0	0	0	3
Doesn't	1	0	0	0	1	0	0	2	4
Has	1	0	1	0	0	0	0	0	2
Would	1	2	0	0	1	5	3	1	13

Table 7.2: Raw frequency of the indicators of the questions asked by the facilitators (cont.)

	PBL1	PBL2	PBL3	PBL4	PBL5	PBL6	PBL7	PBL8	Total
Shall	0	3	0	1	3	1	0	0	8
Will	0	1	0	1	0	1	0	0	3
Was	0	0	1	0	0	0	2	0	3
Isn't	0	0	1	0	3	1	0	1	6
Yeah	0	0	2	0	1	0	2	0	5
Where	0	0	0	1	0	1	2	4	8
Did	0	0	0	5	5	2	6	1	19
Aren't	0	0	0	0	1	0	0	0	1
Wasn't	0	0	0	0	1	1	0	0	2
Didn't	0	0	0	0	1	0	0	0	1
Hasn't	0	0	0	0	1	0	0	0	1
Right	0	0	0	0	1	0	1	0	2
Don't	0	0	0	0	0	2	1	0	3
Haven't	0	0	0	0	0	2	0	0	2
Anything else	0	0	0	0	0	1	0	0	1
Who	0	0	0	0	0	0	3	0	3
Were	0	0	0	0	0	0	1	0	1
Total	79	27	16	42	69	50	103	32	418
	1.35	2.20	1.17	2.00	2.24	1.52	1.41	0.91	1.51

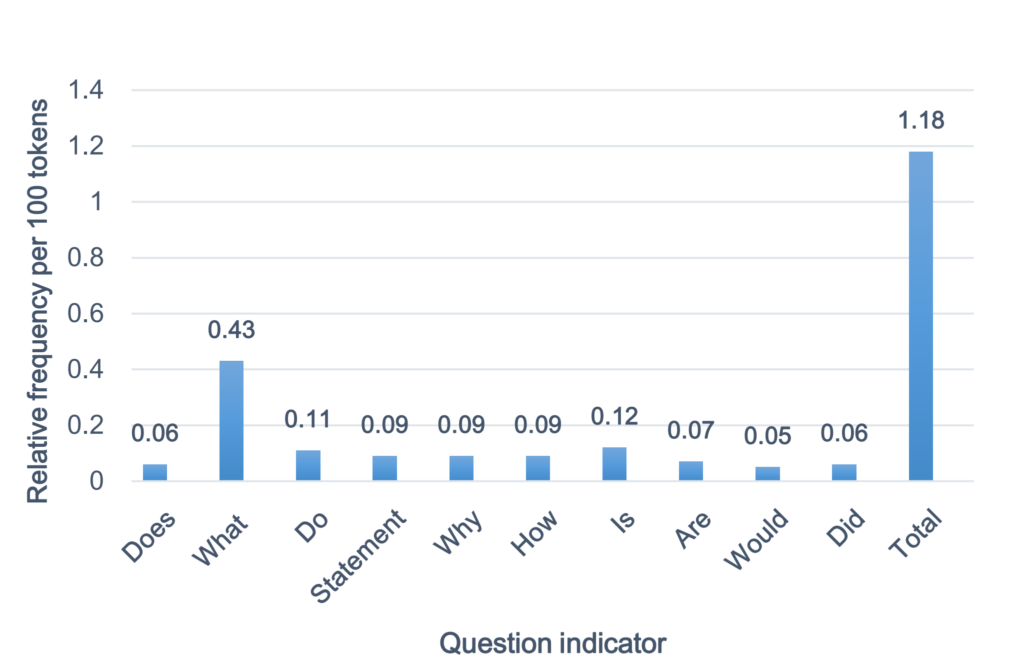


Figure 7.1: Frequent facilitator's question indicators: Relative frequency of the commonly occurring facilitators' question indicators per 100 tokens

These indicators accounted for 332 (79.43%) of the sum total of 418 indicators. The high occurrence of these indicators may mean that the facilitators asked checking, confirming, information-seeking, and explaining questions. A deeper KWIC analysis was needed to uncover how the question indicators were used and what functioned they performed. The results of the KWIC analysis are discussed next.

KWIC analysis. A deeper analysis of the most frequently occurring question indicators was carried out. The analysis involved identifying the use of the indicators in the context of the surrounding words. The results are shown in Table 7.3 (below).

Table 7.3: Frequency of functions of facilitators' questions: Showing raw frequency of the functions of the facilitators' question indicators

	PBL1	PBL2	PBL3	PBL4	PBL5	PBL6	PBL7	PBL8	Total
Explanatory	20	5	3	3	2	4	19	3	59
Definition	9	0	0	2	1	2	1	0	15
Completion	13	0	0	0	1	2	5	3	24
Recall	3	0	0	6	17	5	20	4	55
Verification	7	11	3	14	27	12	23	10	107
Clarification	1	1	1	2	0	0	2	0	7
Seek opinion	1	0	0	1	0	1	0	0	3
Quantification	1	3	1	0	0	0	2	0	7
Judgemental	2	1	0	0	0	0	1	0	4
Exemplification	1	0	0	0	0	0	0	0	1
Suggestion	1	0	1	2	1	9	5	3	22
Rhetorical	5	0	0	2	0	1	4	0	12
Incomplete	1	0	0	0	0	0	0	0	1
Refocusing	0	0	0	2	0	0	0	0	2
Transferability	0	0	0	1	0	0	0	0	1
Comparison	0	0	0	1	0	0	2	1	4
Doubt	0	0	0	0	0	1	0	0	1
Comprehension	0	0	0	0	0	0	2	0	2
Total	65	21	9	36	49	37	86	24	327

The figures of the raw frequencies of the question function in Table 7.3 show that PBL facilitators consistently asked verification and explanation questions in the tutorial groups, but the frequency of verification questions was nearly twice of that of the explanatory questions. Overall, about one-third (32.72%) of the question indicators were used to verify students' information and ideas. Information recall questions were the third most common questions, accounting for 16.82% of the total.

Table 7.4 below shows the raw and relative frequencies of the facilitators' question categories and their functions. The figures show more lower-order questions – 217 (67.18%) – as against higher-order questions – 70 (21.67%). The table shows that facilitators used 0.25 higher-order questions per 100 tokens compared to 0.78 lower-questions per 100 tokens, suggesting that lower-order questions were three times as frequent as higher-order questions. The findings are similar to those of Profetto-McGrath et al. (2004) in case of nursing PBL, where a higher proportion of lower-order questions were found, but contradict those of Hmelo-Silver and Barrows (2008) in case of medical problem-based learning, where higher order questions were frequently asked by the facilitators.

Table 7.4: Frequency of facilitators' question categories: Relative frequency per 100 tokens of the facilitators' questions categories and their functions in the facilitators' subcorpus

	Function	*Normalised frequency/100 tokens*
Higher order (n = 70)	Explanatory	0.21
	Judgemental	0.01
	Opinion seeking	0.01
	Comparison	0.01
Lower order (n = 217)	Concept definition	0.05
	Concept completion	0.09
	Information recall	0.20
	Verification	0.39
	Clarification	0.03
	Quantification	0.03
	Comprehension	0.01
Other (n = 12)	Rhetorical	0.04
Process question (n = 24)	Suggestion	0.08
	Refocusing	0.01
Total		1.16

The questions were further explored to understand the facilitation functions that they performed. Figure 7.2 shows that the facilitators often used the questions to stimulate elaboration, to elicit response, and to prompt. These techniques, which were meant to raise critical awareness, were deployed in response to insufficient explanation, knowledge gaps or inconsistent thinking, and were meant to raise the group discussion to a higher cognitive level (Gilkison, 2003).

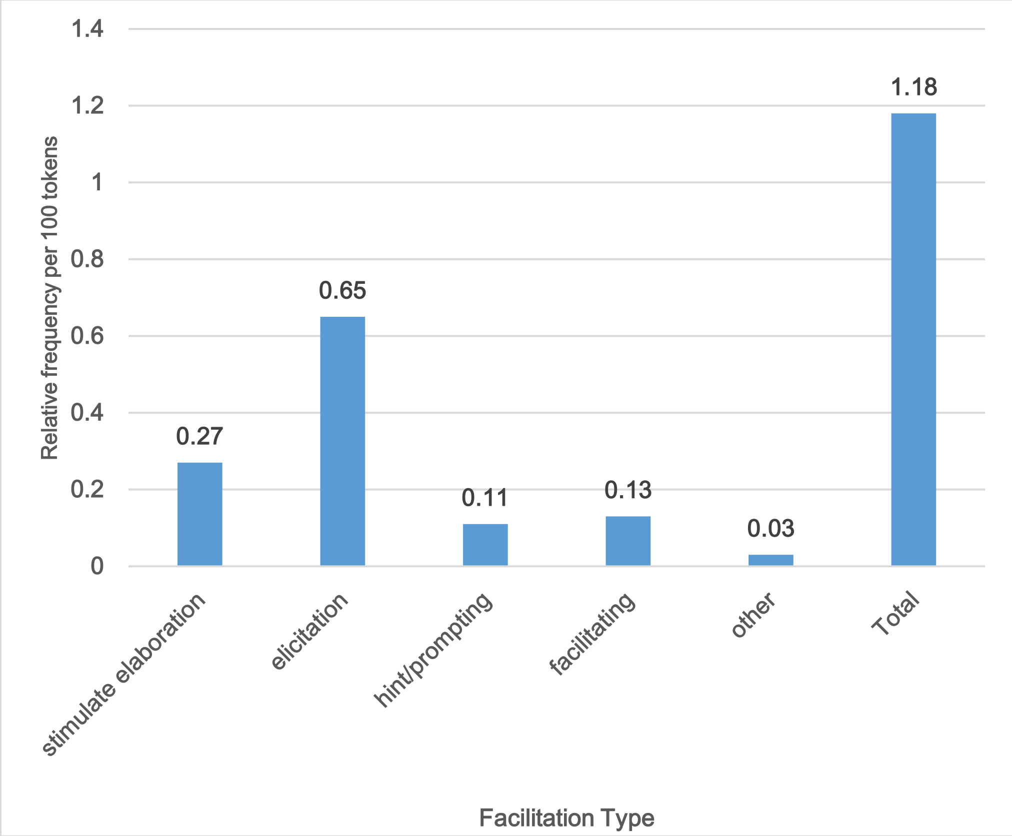


Figure 7.2: Frequent question-related facilitation techniques: Relative frequency of the techniques in which the facilitation questions were used per 100 tokens

To gain a deeper and richer understanding of the interactive features of the discourse that surrounded these questions, an in-depth analysis of the text segments was carried out. The results of the qualitative analysis are discussed next.

Qualitative analysis

This section presents the results of the qualitative analysis of the questions asked by the facilitators in the PBL tutorials. Only the functions that had raw frequencies of 10 and above were further explored with the qualitative approach. These functions were seven in number. The results are presented according to the functional categories.

Explanatory function. Extracts 1 and 3 exemplify how the facilitators' questions to stimulate elaboration promoted knowledge co-construction among the students. However, in extract 2, the question failed as the students did not respond to the question from the facilitator and there was no further evidence of discussion relating to it in the transcript. Questions in this category required students to provide cause-effect connections or to provide simple elaboration.

Extract 1:

T1 – Does anybody know anything about blood sugar and going to the bathroom in the middle of the night?
M3 – Diabetes you can be peeing more.
M1 – Why do you do that? Why do you wee in the night?
F2 – You pee the glucose out because your… is it pancreas?
M2 – Your pancreas can't get rid of it so you pee it out.
F2 – So you end up dehydrated because you go to the bathroom… (PBL1)

In this extract, the facilitator asked a question regarding the connection between blood sugar and going to the bathroom in the night. The students elaborated the connection and they built on the facilitator's question by asking themselves further questions. Different students contributed to the elaboration, thereby suggesting knowledge co-construction.

Extract 2:

T2 – More importantly… why do we exhale after every swallow?

M1 – Oh god!
F3 – Shall I put that in as well?
M1 – Yes, yes, the end bit that’s the profile so that’s just ready to be chopped so you can see it. They’re you go, that’s perfect. Well done. (PBL2)

In extract 2, the facilitator asked an elaboration stimulating question, but this question was never answered and there was further talk relating to it.

Extract 3:
1. T4 – Why do you want to do these?
2. M3 – Why would you do a chest x-ray, to check the lungs?
3. F3 – I suppose to ensure she is not anaemic, for a start, because although the signs are saying that she has heart failure there is nothing to say she is also not anaemic, so you would want to check things like that.
4. T4 – What would you do with the Us and Es? What is that about?
5. F3 – Looking at renal function, also sodium levels and –
6. M2 – Yes, retention of salt and water.
7. F3 – Yes, sodium (PBL3)

In extract 3, the question of the facilitator (line 1) stimulated knowledge co-construction. F3 (line 3) gave a causal elaboration building upon the M3’s contribution (line 2). Another question by the facilitator (line 4) similarly stimulated knowledge co-construction as F3 (line 5), M2 (line 6), and F3 (line 7) built upon each other’s contribution to answer it. In extract 5, the question stimulated the students to build causal models for the case patient symptoms.

Extract 4: Judgemental

T1 – … if the guy has atrial fibrillation which we have said that he does, does that in anyway influence your decision to do carotid sinus massage via doing this some other way and if so why? (PBL1)

Extract 5:

T4 – Is there anything in particular about doing housework that might make it something more likely to make her faint?
M1 – Funny positions. Maybe she has some postural hypotension if she was bending over and putting her head under the thing.
F1 – She was probably breathing in dust and polish and stuff.
F3 – At least her maxillary sinuses will be nice and clear.
M1 – Yes, that is a good point, maybe it is something to do with that.
T4 – Is there anything from other cases that you have done so far that you would like to go over again that you think might be relevant?
F1 – Shall we ask her if she has any arrhythmias and any palpitations, because sudden collapse could be arrhythmia? (PBL4)

Concept definition. In extracts 1 and 2, the facilitator asked for factual knowledge regarding the definition of bruit and fibrillation. The exchanges between the student and the facilitator followed the initiation-response-evaluation (IRE) pattern. This model is considered appropriate for factual knowledge recall.

Extract 1:

T1 – So what is a bruit?
M1 – Bruit is a sound you hear over the carotid.
T1 – Bruit is a sound and you hear it over the carotid, why?

M1 – Atherosclerosis of the carotid and you can displace something. (PBL1)

Extract 2:

T1 – What is fibrillation?

M1 – Chaotic electrical activity.

T1 – What is it?

M1 – Chaotic electrical activity.

T1 – Lack of organised electrical activity. (PBL1)

Concept completion. In extracts 1 and 2, the facilitator asked concept completion questions. The facilitator used hints to stimulate the recall of factual knowledge.

Extract 1:

1. T1 – And one of them has potentially no relevance whatsoever to physiology but if you basically put blood on a slide it does what?
2. F3 – Contact.
3. F2 – Clot. (PBL2)

Extract 2:

1. T1 – Because the palpitations will then usually resolve when you do what?
2. F1 – Rest.
3. T1 – Rest or otherwise do what to the patient?
4. F2 – Do anything that is going to get his heart to slow down.
5. T1 – So like beta-blockers. (PBL6)

Information recall. Extracts 1 to 4 illustrate information-seeking questions asked by the facilitators.

Extract 1:

1. T8 – What determines the baseline in addition to electrical activity?
2. F2 – Muscle tremor.
3. T8 – Muscle tremor, shivering, moving around–. (PBL1)

Extract 2:

T7 – … which molecule is the most important molecule in terms of this guy's problem here? (PBL7)

Extract 3:

T6 – … what else in this pattern leads you in the direction of heart failure? (PBL6)

Extract 4:

T1 – When you say the idea behind it, what do you mean? (PBL7)

Verification. In extracts 1 and 2 below, the facilitator asked verification questions to ascertain the correctness of the students' contributions.

Extract 1:

1. T4 – … but having collapsed what other things might be about collapse? Is collapse and dizziness the same thing?
2. M1 – We are not getting these today, are we? We are not moving. (PBL4)

Extract 2:

1. T4 – Therefore, pitting results when there is an ordinary amount of protein, just too much water in the tissue?
2. M6 – Usually, yes.
3. T4 – Is that how it works? (PBL4)

Rhetorical. Rhetorical questions are not meant to be answered, but in this context, they were asked by the facilitator to give examples or illustrate a point of view (extract 1).

> **Extract 1:**
> T1 – And I think the other thing is, and as you raised the question about the medications. Are there medications that have come up in their history that you want to go over in terms of thinking about whether they would be appropriate what they tell you if the patient is taking them, you know, what is Digoxin? What does it do? Would you expect that it would wear off? (PBL6)

Suggestion. In this instance, the facilitator put forward an idea for the students' consideration. The idea could be accepted or rejected. In the extract 1, the students rejected the facilitator's suggestion.

> **Extract 1:**
> M1 – Yeah, okay.
> T1 – Do you guys want to take five minutes to briefly hit rate control versus rhythm control before going around or did you want to skip that and just go around?
> M1 – It would be useful but we do not have time to I do not think. (PBL1)

In summary, 418 question indicator markers were retrieved from the facilitators' subcorpus across 18 PBL tutorial groups. There were 10 most frequently occurring question indicator types and they accounted for about 80% of the indicators. Lower-order questions were frequently used, with a relative frequency of 0.78 per 100 tokens. The facilitators used 0.39 indicators per 100 tokens to request verification; 0.21 per 100 tokens to ask explanatory questions; and 0.20 per 100 tokens for information recall. The questions were used for content and process facilitation. The results of the analysis of the facilitators' stance expressions are presented next.

Analysis of stance expressions

Stance refers to a range of different ways in which speakers temper their verbal interactions (Eggins, 2005; Eggins and Slade, 1997). Stance expressions can convey many different kinds of personal feelings and opinions, including a speaker's attitudes towards information and their certainty about its veracity (Biber, 2006). Two kinds of stance were analysed in this study: (1) one that dealt with tempering the categorical nature of the information that the facilitator exchanged; and (2) one that tempered the directness with which the facilitator sought to act upon or influence the students. The indicators of these two types of stance are analysed in the following two questions.

Question 2. Directive expression

The indicators of directive expression were analysed at two levels: quantitative and qualitative.

Quantitative analysis

The quantitative analysis consisted of word frequency analysis and KWIC analysis.

Word frequency analysis. A total of 174 directive expression indicators were retrieved from the facilitators' dataset. Overall, only six ('should', 'have to', 'need', 'supposed', 'would' and 'can') of the 15 directive indicator types occurred in tens (Table 7.5 below).

Table 7.5: Frequency of directive expression indicators: Raw frequency of the directive expression indicators in the facilitators' talk

	PBL1	*PBL2*	*PBL3*	*PBL4*	*PBL5*	*PBL6*	*PBL7*	*PBL8*	*Total*
Should	5	0	3	0	1	2	9	4	24
Have got to	2	0	0	0	0	0	0	0	2
Have to	7	1	0	1	3	7	3	5	27
Has to	1	0	0	0	0	1	1	0	3
Needs	0	1	1	0	1	0	0	0	3
Needed	0	0	0	0	0	0	1	1	2
Need	2	5	0	3	1	5	2	4	22
Supposed	7	0	0	0	1	0	2	2	12
Must	1	0	0	0	0	2	0	1	4
Having to	0	0	0	0	0	1	1	0	2
Had to	0	0	0	0	0	2	0	0	2
Would	8	2	0	1	3	4	25	1	44
Could	2	1	1	2	0	0	2	0	8
Can	3	2	0	1	2	1	5	0	14
'D	0	1	0	0	0	0	2	2	5
Total	38	13	5	8	12	25	53	20	174

The distribution of the relative frequency of the indicators is shown in Figure 7.3 (below). The facilitators used more than 0.1 per 100 tokens of two types of indicators; 0.05–0.1 per 100 tokens of two; and less than 0.05 per 100 tokens of 10 types of indicators. These indicators tentatively indicated that the facilitators talked about some degree of obligation, necessity, and possibly requirement. A contextual examination of the indicators yielded the functions of the indicators which are discussed below.

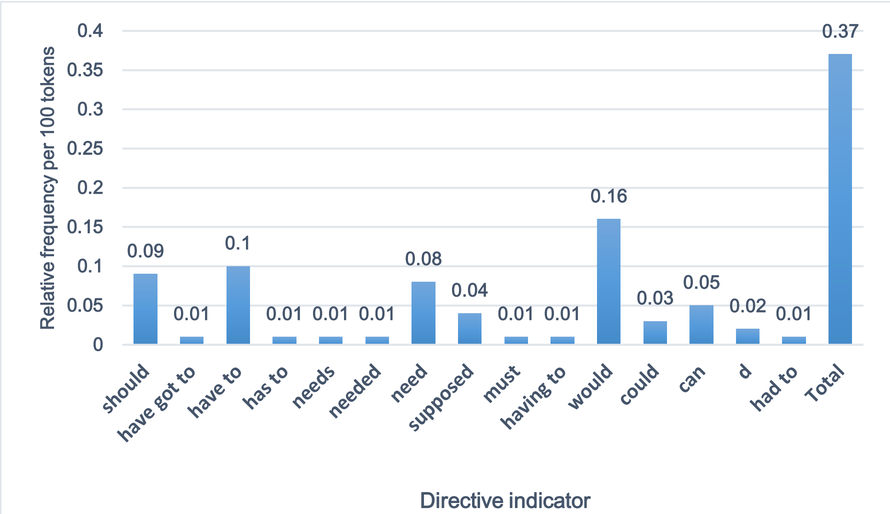


Figure 7.3: Frequent facilitator's directive expression indicators: Relative frequency of the directive expression indicators in the facilitators' subcorpus per 100 tokens

KWIC analysis. The directive expression indicators were used to achieve 20 types of functions out which six were more frequent and occurred in tens in the raw frequency analysis. The pattern of the functions is shown in Table 7.6 below.

Table 7.6: Frequency of directive expression functions: Raw frequency of the directive expression indicators' functions in the facilitators' talk

	PBL1	PBL2	PBL3	PBL4	PBL5	PBL6	PBL7	PBL8	Total
Expectation	11	2	0	0	1	2	6	9	31
Unnecessary	1	1	0	1	1	3	0	0	7
Self-compulsion	1	0	0	0	2	5	0	0	8
Necessary process	3	0	0	0	0	0	0	0	3
Requirement	2	3	1	1	1	1	4	1	14
Duty bound	0	0	0	0	0	1	0	2	3
Self-exhortation	1	0	0	0	1	0	0	0	2
Exhorting	4	0	1	0	0	3	2	3	13
Appropriateness	1	0	1	0	1	0	3	0	6
Logical estimation	0	1	1	0	0	1	2	2	7
Necessity	0	0	0	1	0	3	2	0	6
Unexpected	1	0	0	0	0	0	0	0	1
Denial of own statement	0	0	0	1	0	0	0	0	1
Likely a joke	0	0	0	0	0	1	0	0	1
Intention	4	1	0	1	2	0	14	1	23
Indirect question	5	4	1	2	2	4	13	1	32
Indicate willingness	2	0	0	0	0	1	7	1	11
Seek permission	1	0	0	1	1	0	0	0	1
Give permission	1	0	0	1	1	0	0	0	3
Indirect request	0	1	0	0	0	0	0	0	1
Total	38	13	5	8	12	25	53	20	174

Only 15 of the functions were commonly occurring in the facilitators' subcorpus (Figure 7.4 below). These commonly occurring functions were grouped into three categories. A frequency of more than 0.1 indicators per 100 tokens was used to mark expectation and indirect questioning. A frequency of between 0.05 and 0.10 indicators per 100 tokens was used to preface intention, exhortation, and requirement functions. The other 10 functions had a relative frequency of 0.04 and below per 100 tokens.

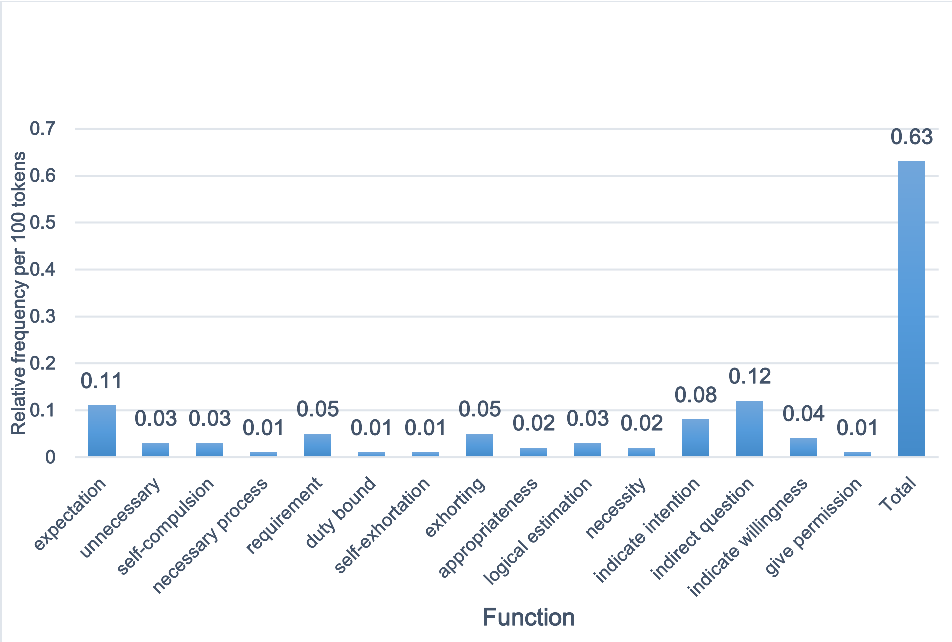


Figure 7.4: Frequent directive expression indicator functions: Relative frequency per 100 tokens of the commonly occurring functions of the directive expression indicators in the facilitators' subcorpus

Table 7.7 below shows how the facilitators used the directive expression indicators to guide the PBL discussions. Three facilitation techniques (i.e. learning process guidance, directing learning and raising awareness to critical curriculum issues) occurred in tens.

Table 7.7: Frequency of directive facilitation techniques: Raw frequency of facilitation techniques for the use of directive expression indicators

	PBL1	PBL2	PBL3	PBL4	PBL5	PBL6	PBL7	PBL8	Total
Facilitating Learning	9	6	2	2	5	10	8	6	48
	0.15	0.49	0.15	0.10	0.16	0.30	0.11	0.17	0.17
Directing Learning	16	4	1	2	3	12	21	14	73
	0.27	0.33	0.07	0.10	0.10	0.36	0.29	0.40	0.26
Intra-personal Behaviour	4	0	0	0	1	0	0	0	5
	0.07	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.02
Interaction/ Accountability	0	0	0	0	0	1	1	0	2
	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00	0.01
Awareness	8	3	1	3	3	2	23	0	43
	0.14	0.24	0.07	0.14	0.10	0.06	0.31	0.00	0.15
Other	1	0	1	1	0	0	0	0	3
	0.02	0.00	0.07	0.05	0.00	0.00	0.00	0.00	0.01
Total	38	13	5	8	12	25	53	20	174
	0.65	1.06	0.37	0.38	0.39	0.76	0.73	0.57	0.63

The three facilitation techniques were commonly used by the facilitators (Figure 7.5below). Overall, content facilitation was more frequently used than process facilitation. The most frequently occurring functions of the directive expression indicators and facilitation techniques were further explored with qualitative analysis. The results are described in the next section.

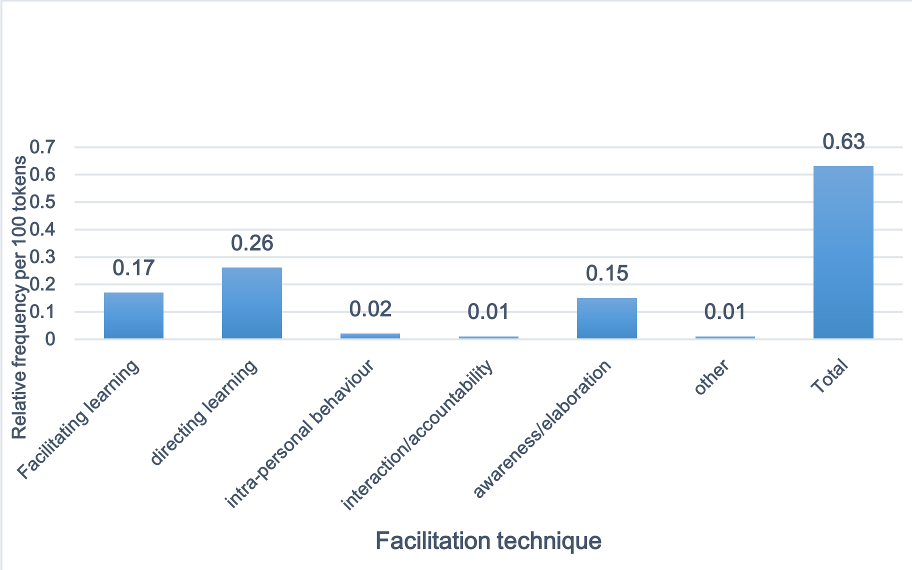


Figure 7.5: Frequent directive expression-related facilitation techniques: Relative frequency per 100 tokens of the frequently occurring facilitation techniques in which directive expressions indicators were used in the facilitators' subcorpus

Qualitative analysis

The results of the qualitative analysis are discussed under two headings: (1) functions of directive expression indicators and (2) facilitation techniques associated with the use of directive expression indicators.

Functions of directive expression indicators

This section discusses the results of the qualitative approach to the analysis of the functions of directive expression indicators.

Expectation. The directive expression indicator was used to express expected normal and abnormal results (extract 1), expected clinical examination steps (extract 2), integrated knowledge (extract 3), expected knowledge (extract 4), and expected learning points (extract 5).

Extracts 1:

T1 – … you can try and know what is different about this thing I am looking at versus what it **should** look like… (PBL1)

T1 – But as everybody seems to agree you can estimate where it **should** be. (PBL1)

T8 – Troponin T is 0.14 and that's micrograms per litre. And it **should** be below 0.03. (PBL8)

Extract 2:

T1 – There is a key step you are **supposed** to do before you do carotid sinus massage.

Extract 3:

T7 – … I think we've had a whole long discussion today about how these things, actually all, **should** interrelate… (PBL7)

Extract 4:

T6 – It is obviously an important point that you have picked up and if you are considering giving somebody Warfarin that is something you absolutely **need to** know. (PBL6)

T2 – … if you want to swallow safely then you **need to** know what consistency you **need to** reproduce. (PBL2)

Extract 5:

T8 – This is going to be kind of above and beyond what you **needed** to get for this particular case. We will come back to heart failure so don't panic too much about the **need to** understand this, because you **need to** get to the next case. (PBL8)

Requirement. Data analysis also showed that the facilitators used directive expression indicators to talk about the requirement for learning issues (extract 1), after-hospital care for the patient (extract 2), and learning requirements (extract 3).

Extract 1:
T7 – We're on target to finish on time but then someone decided we needed an extra three learning areas. (PBL7)

Extract 2:
T3 – He was given dietary advice, and was advised, as Asher said, to stop drinking alcohol immediately. Occupational health needs were assessed. (PBL2)

Extract 3:
T2 – … having studied some of these areas your intention will be come to the group to discuss any area that you feel **you need** further clarification on. (PBL3)

Exhortation. In several instances in the dataset, the facilitators used directive expression indicators to preface advice, either in the form of self-exhortation (extract 1), self-inclusive exhortation (extract 2), or in the form of advice to the students (extracts 3 to 5).

Extract 1:
T5 – To be honest, I am not sure what happens in the specific vessels. I would have to go and look at it. (PBL5)

Extract 2:
T7 – … I said maybe we should just explore regulation of blood pressure in greater depth… (PBL7)

Extract 3:

T7 – … you should let your thought process dictate what questions you ask and don't get frustrated if you don't have that information. (PBL7)

Extract 4:

T8 – … the next case there will probably be at least five learning objectives per day so it will be a bit heavier then this case is. So I think you have to accept to prepare for that for all of you. (PBL8)

Extract 5:

T1 – You guys have got to remember it is multiple cells. (PBL1)

Mark intention. In this instance, an indicator of directive expression was used to mark the speaker's intention.

Extract 1:

T3 – I just figured I would give you the full benefit of what you have for his… (PBL3)

Extract 2:

T1 – … there are three other points that I would make on the EKG interpretation… (PBL1).

Preface question. In this case, an indicator of directive expression was used to mark a question.

Extract 1:

T7 – The other question then is would his symptoms have been having intermittent obstructions to blood flow, would that potentially cause varicose veins? (PBL7)

Extract 2:

T1 – So what would be the consequences of having too little calcium or too little potassium or too much of either? (PBL1)

Facilitation Techniques

This section provides examples of facilitation techniques where directive expression indicators were used in the dataset. Three frequently occurring facilitation techniques – facilitating learning, directing learning, and raising critical awareness (Figure 7.5) – are discussed.

Facilitating group process

This involves behaviours such as suggesting, refocusing, summarising, providing feedback, and evaluating on the facilitators' part (Gilkison, 2003).

Extract 1:

T6 – I think we **should** pass over to James, because these questions all seem to be very related to the pathophysiological process. (PBL6)

Extract 2:

T4 – You may not **need to** make it a learning topic as in 'what is a collapse?' (PBL4)

Extract 3:

T8 – We will come back to heart failure so don't panic too much about the **need to** understand this, because you **need to** get to next case. (PBL8)

Extract 4:

M1 – Khattab, you are very good because I was ready to change my answer with your look.

T6 – Maybe there is a **need to** trust your own reasoning.

In these extracts, the facilitators marked their suggestions (extract 1 and 2), refocusing attempts (extract 3), and feedback (extract 4) with directive indicators. The use of directive indicators was toned down with indirective devices such as maybe (extract 4) and may (extract 2).

Directing learning. This occurs when the facilitator dispenses facts, information, opinions, or ideas to the students or provides answers to the students' questions (Gilkison, 2003), as can be seen in the extracts below.

Extract 1:

T1 – Atrial fibrillation does not **have to** be fast and in most patients it is not fast but when it becomes fast that is when you tend to see these patients acutely and that is usually what? Does anybody know? (PBL1)

Extract 2:

T1 – What has to happen to get the contraction? You **have to** have a coupling between the electrical impulse and the physical contraction. And so anything you can do inside the cells that messes up the electrical depolarisation of the plasma membrane to contractual apparatus connectivity will change the relationship between the ECG and the actual contraction. And so that's another level at which you can have problems (PBL1)

Extract 3:

T6 – Yes, I had to spray it under my tongue and it made me a bit dizzy, so I do not use it unless I have to. (PBL6)

Raising critical awareness. The exchanges under this category include elicitation, and prompting and elaboration stimulation in response to inadequate explanation, gaps in students' knowledge or inconsistencies in their thinking (Gilkison, 2003). In the extracts below, prompting and elicitation questions are marked with directive indicators.

Extract 1:

T6 – Do you **need to** do any tests at all?
M2 – I would like to do an ECG. (PBL6)

Extract 2:

T1 – But if you were thinking that it was related to dynamic changes in his blood pressure what kind of assessment **must** you consider doing on him? (PBL1)

Summary

The directive expression indicators were frequently used to mark the expected learning behaviours of students and the obligation of doctors, and to intimate students about curriculum requirements. The facilitators used the indicators to mark their own self-exhortation and preface advice to the students. They were commonly deployed for directing learning and guiding the learning process.

Q3. Probability expression

The indicators of probability stance were analysed at two levels: quantitative and qualitative.

Quantitative analysis

The quantitative analysis consists of word frequency analysis and key word in context (KWIC) analysis.

Word frequency analysis. A total of 27 types of probability indicators were retrieved from the dataset. Their raw frequencies are shown in Table 7.8 above. Only 9 indicators occurred in tens.

Table 7.8: Raw frequency of probability expression indicators

	PBL1	PBL2	PBL3	PBL4	PBL5	PBL6	PBL7	PBL8	Total
Perhaps	5	0	0	1	0	0	16	1	23
Maybe	5	0	0	2	2	4	12	5	30
Would	27	4	6	5	16	5	18	4	85
Could	15	0	1	0	3	2	17	4	42
Can	18	2	1	3	3	2	27	5	61
Potentially	11	0	0	0	0	0	5	4	20
Potential	2	0	0	0	1	0	2	3	8
Might	2	0	2	12	3	5	5	2	31
May	3	2	0	4	1	4	19	5	38
Probably	6	0	0	0	0	4	16	4	30
Sure	3	1	1	0	1	1	0	0	7
Make sure	2	0	0	0	0	0	0	1	3
Likely	3	0	1	1	0	0	3	0	8
Clearly	1	0	1	0	0	1	0	1	4
Possibly	0	0	0	0	0	1	1	1	3
Possible	0	3	1	0	1	0	2	0	7
Clarification	0	2	0	0	0	0	0	0	2
Clarifying	0	0	0	0	0	1	0	0	1
Clear	0	0	0	0	0	1	0	0	1
Certainly	0	0	0	0	0	2	0	0	2
Clarify	0	0	0	1	0	0	0	0	1

Chance	0	0	0	0	0	1	0	0	1
Making sure	0	0	0	1	0	0	0	0	1
Unlikely	0	0	1	0	0	0	0	0	1
Definitely	0	0	1	0	0	0	0	0	1
Ca	0	1	2	0	0	0	0	0	3
'D	0	0	4	0	0	0	2	1	7
Total	103	15	22	30	31	34	145	41	421

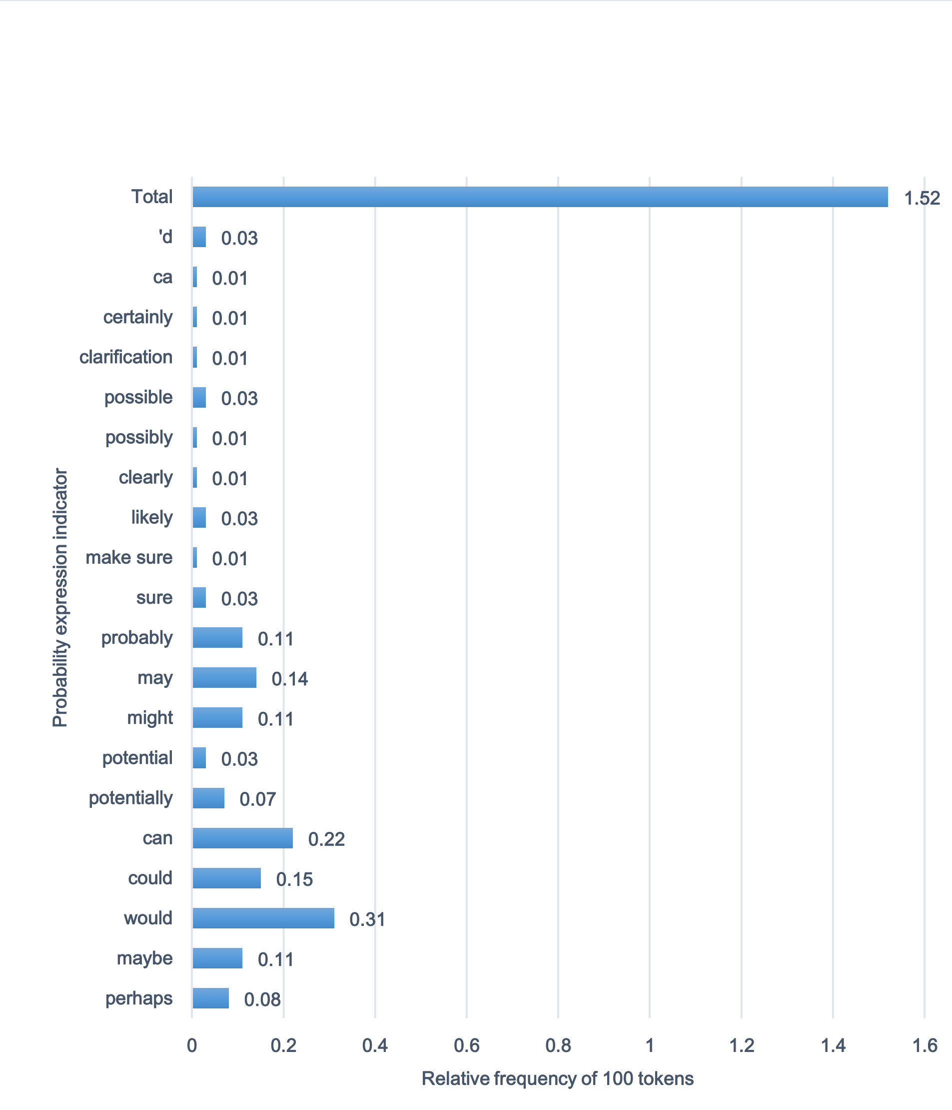


Figure 7.6: Frequent probability expression indicators: Relative frequency per 100 tokens of the frequently occurring probability expression indicators in the facilitators' subcorpus

In Figure 7.6 (above), 20 frequently occurring probability indicators in the facilitators' subcorpus are shown. The facilitators used more than 0.1 indicators per 100 tokens for seven types of indicators. Two indicator types had relative frequency between 0.05 and 0.10 per 100 tokens. Of the remaining 11 indicator types, the facilitators used between 0.01 and 0.04 indicators per 100 tokens in the subcorpus.

KWIC analysis. The raw frequency of the functions of the probability indicators is shown in Table 7.9 (below). Only four types of functions (possibility, prediction, hedging, and logical deduction) occurred in tens.

Table 7.9: Raw frequency of the functions of probability indicators

	PBL1	*PBL2*	*PBL3*	*PBL4*	*PBL5*	*PBL6*	*PBL7*	*PBL8*	*Total*
Possibility	73	7	5	16	9	19	122	33	284
Prediction	8	2	3	1	10	0	14	4	42
Hedging	14	3	4	6	1	4	0	0	32
Uncertainty	4	1	1	0	0	0	0	0	6
Uncertain	1	0	0	0	0	0	0	0	1
Certainty	3	0	2	0	0	3	0	1	9
Make certain	0	2	0	2	0	2	0	1	7
Impossible	0	0	2	0	0	0	0	0	2
Deduction	0	0	3	3	5	4	6	2	23
Disagreement	0	0	1	1	0	1	0	0	3
Unlikely	0	0	1	0	0	0	0	0	1
Likelihood	0	0	0	1	0	0	3	0	4
Agreement	0	0	0	0	5	1	0	0	6
Check certainty	0	0	0	0	1	0	0	0	1
Total	103	15	22	30	31	34	145	41	421

Figure 7.7 below shows the 11 frequently occurring probability functions in the facilitators’ subcorpus.

Overall, the indicators had a cumulative relative frequency of 1.52 indicators per 100 tokens. Three types of probability indicator (possibility, prediction, and hedging) had a frequency of more than 0.1 per 100 tokens. Facilitators used fewer than 0.1 per 100 tokens of one indicator. Seven indicator types were less commonly occurring and the facilitators used between 0.01 and 0.03 of these indicators per 100 tokens.

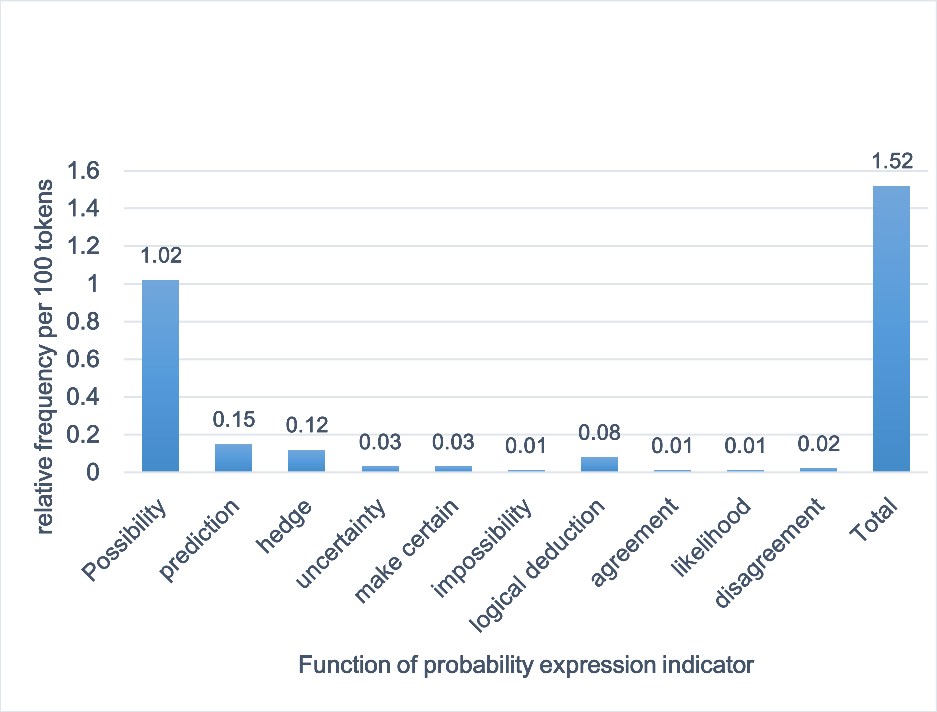


Figure 7.7: Frequent probability expression indicator functions: Relative frequency of the commonly occurring functions of probability expression indicators in the facilitators' subcorpus per 100 tokens

Table 7.10 shows how the indicators were deployed for PBL facilitation.

Table 7.10: Frequency of probability expression facilitation techniques: Raw and relative frequency of the facilitation techniques in which probability indicators were used by the facilitators across the PBL groups

	PBL1 RF NF	PBL2 RF NF	PBL3 RF NF	PBL4 RF NF	PBL5 RF NF	PBL6 RF NF	PBL7 RF NF	PBL8 RF NF	Total RF NF
Directing	67 1.15	9 0.73	4 0.29	8 0.38	18 0.58	13 0.39	76 1.04	12 0.34	211 0.76
Facilitating	20 0.34	4 0.33	9 0.66	13 0.62	7 0.23	15 0.46	44 0.60	29 0.82	144 0.52
Questioning	15 0.26	2 0.16	7 0.51	7 0.33	6 0.19	2 0.06	25 0.34	0 0.00	64 0.23
Reflection	2 0.03	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	2 0.01
Total	103 1.76	15 1.22	20 1.47	28 1.33	31 1.00	30 0.91	145 1.98	41 1.17	414 1.49

Overall, the figures show that facilitators used 0.76 probability expression indicators per 100 tokens for directing group process compared to 0.52 probability expression indicators per 100 tokens used for group process facilitation. The pattern is shown graphically in Figure 7.8 below.

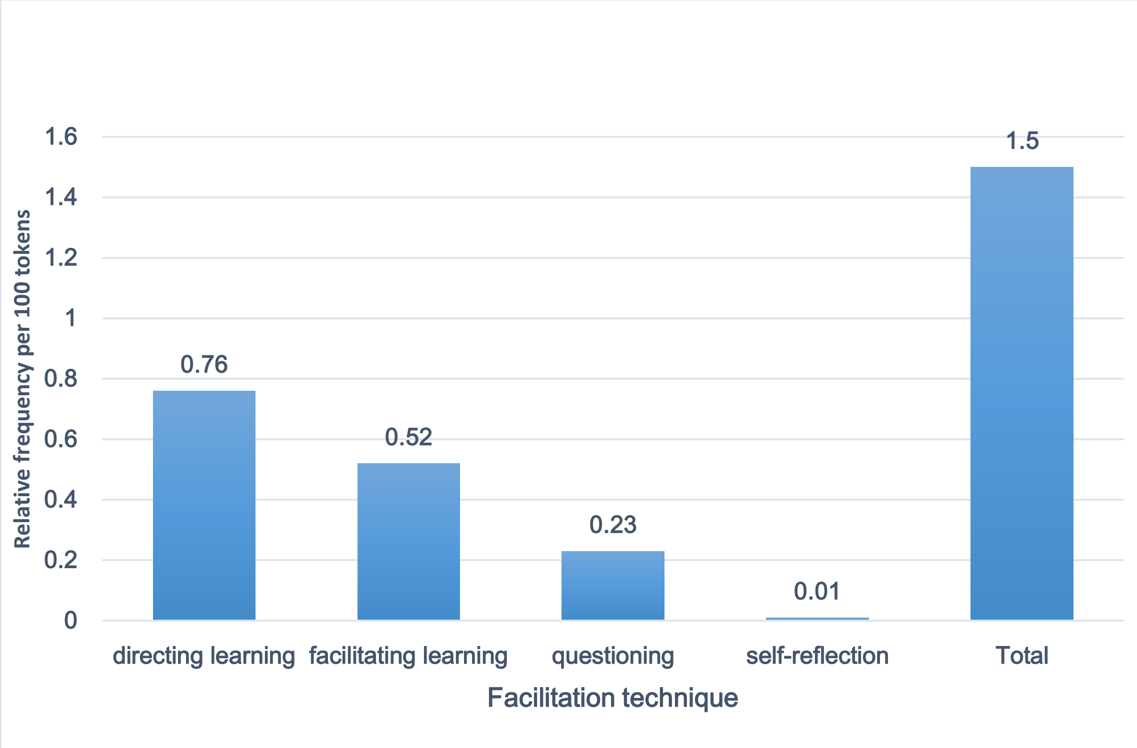


Figure 7.8: Frequent probability expression-related facilitation techniques: Relative frequency per 100 tokens of the frequently occurring facilitation techniques in which probability expression indicators were used in the facilitators' subcorpus

In four groups (PBL1, PBL2, PBL5, and PBL7), the facilitators used more probability indicators (0.58–1.15) per 100 tokens to direct learning than to facilitate group process (0.23–0.60). In another four groups (PBL3, PBL4, PBL6, and PBL8), the facilitators used more probability indicators (0.46–0.82) per 100 tokens to facilitate group process than to direct learning (0.29–0.39) (as shown in Table 7.10). It is possible that contextual factors influenced the facilitators' group behaviours. However, overall, the facilitators frequently dispensed information more than they facilitated group process.

Qualitative analysis

The qualitative analysis was carried out to describe the four frequently occurring functions with extracts from the transcripts.

Possibility. Possibility indicators were used by the facilitators to mark the degree of certainty of the students' point of views and were used in the context of responding to the students' questions (extract 1), offering feedback to students (extract 3), refocusing tutorial discourse (extract 4), offering suggestions (extract 5), and asking questions (extract 6). Extracts 1 and 2 describe the two types of responses found in the dataset. In extract 1, the facilitator acted as the case patient responding to the students' questions about symptoms and signs. This occurred in PBL4 to PBL6. In extract 2, the facilitators provided answers to content questions asked by the students. Provision of feedback occurred in almost all the groups, as exemplified in extract 3.

Extract 1:
T4 – Once in a while on a special occasion maybe, not very often, but a sherry here and there.
F3 – What is your diet like at the moment?
T4 – I have never been a big eater. I do not eat very much, mostly a bit of toast. I do not have much of an appetite.
F2 – Have you ever had rheumatic fever when you were younger?
T4 – I do not think so. (PBL4)

Extract 2:
T8 – When you say are they alright, so yes they all have a role. So I think what you're asking when you say are they alright is which

ones are the most important. I think that's a very difficult thing. If you look at people with diabetes, you'll often see that there is variability in how things occur and that's probably because of the multi-factorial nature. (PBL8)

Extract 3:

T3 – … I was very impressed by the way that you said that because you didn't have personal knowledge of this smoking cessation process, that conversation was helpful because it would help you to understand the patients, I thought that was a really nice thing to say. (PBL3)

Extract 4:

T8 – I was just trying to interject myself into the conversation to see if we wanted to steer away from the health beliefs and maybe onto a different topic so that we could get to history taking some time in the next ten, fifteen– (PBL8)

Extract 5:

T1 – The suggestion then would be you guys want to look at mechanisms under the line of pathology, perhaps you actually want to explore the normal situation as well. (PBL1)

Extract 6:

T7 – … do you think perhaps the anatomy of the veins and the arteries is **perhaps** more pathological? (PBL7)

T4 – … is there another thing that is integral to catching your breath that **might** be related to blood and stuff? (PBL4)

Prediction. This pertains to the situation where the probability indicators were used to make statements about what would happen in the future. The extracts below illustrate this function.

Extract 1:

T2 – But the other thing is you mentioned education but didn't mention that he **would** need appointments with the practice nurse in the short term and long term. (PBL2)

Extract 2:

T3 – … that conversation was helpful because it **would** help you to understand the patients, I thought that was a really nice thing to say. (PBL3)

Extract 3:

T7 – The pain's becoming more frequent, what **would** that potentially have to do with arterial blockage? (PBL7)

In each of these extracts, the facilitator used 'would' to mark a future occurrence or event.

Indefiniteness. This aspect refers to the linguistic elements used to weaken the assertions of academic discourse as being hard and indisputable facts.

Extract 1:

T1 – I think that is actually a good point from a clinical perspective because my take **would be** this bathroom thing is a red herring, it is clearly the heart that woke him up and it is a thyroid issue but if we take the case at face value and turn out... (PBL1)

Extract 2:

T1 – So what would be examples of beta blockers? (PBL8)

Extract 3:

T3 – … I don't know whether it's worth having a summary table.

T3 – **I would** find it useful.

M4 – It's not a bad idea actually.

M2 – The disease, the features and what the causes are type of table?

T3 – It's just a question; I'm not saying that you should. (PBL3)

In these extracts, the facilitator marked their points of view (extracts 1 and 3) and questions (extract 2) with probability markers.

Logical deduction. The probability indicators were also used to preface logical deductions from available evidence. The logical deduction function of the probability indicators is illustrated in extracts 1 to 3 below.

Extract 1:

T4 – I had a bit of a faint, I think, and when I fell over my husband decided it **would** be a good idea to get an ambulance. (PBL4)

Extract 2:

T5 – I am a bit elderly for that; my knees **would** not let me. (PBL5)

Extract 3:

T7 – So maybe an underlying theme here **would** be vascular disease or vascular disorder under causes with arthrosclerosis potentially being one, physical blockage being another… (PBL7)

Facilitation techniques

Directing learning. This occurs when the facilitator dispenses facts, information, opinions, or ideas to the students or provides answers to the questions asked by the students (Gilkison, 2003), as can be seen in the extracts below.

> **Extract 1:**
> T1 – …so you can have two to one, three to one; you can have a two to one followed by a three to one. Which would be so you can have a regular rhythm with a flutter, but you could also have a regularly irregular rhythm with a flutter. Because you could have one conduction every second wave and one conduction every third wave, those are just the most common ones that you'll see. And so if it is irregularly irregular, it is typically going to be... (PBL1)

Facilitating learning. This aspect involved behaviours such as suggesting (extract 2), refocusing (extract 3), summarising (extract 1), and providing feedback (extract 4) on the facilitators' part (Gilkison, 2003).

> **Extract 1:**
> T5 – Encapsulate it in two different ways for normal things to look at in terms of mechanisms. You have identified something behind you there as being potentially related to a cardiac problem that is causing palpitations. (PBL5)

> **Extract 2:**
> T1 – … so perhaps looking at normal conduction of the heart would help understand pathological situations. (PBL1)

Extract 3:

T4 – I know where you are going. Hold that thought; it might be useful in the next case. (PBL4)

Extract 4:

M2 – Yeah, that is the answer but it is obviously…

T4 – Well potentially it is an answer. (PBL 4)

Raising awareness. The exchanges under this category involved questioning (extracts 1 and 2) in response to inadequate explanation, gaps in the students' knowledge or inconsistencies in their thinking (Gilkison, 2003).

Extract 1:

T5 – What tests would you do? (PBL5)

Extract 2:

T1 – And then how would you detect a conduction problem between the atrium and the ventricles? (PBL1)

In summary, 27 types of probability indicators were retrieved. Nine of these indicators occurred 10 times or more on the raw frequency list. The probability expression indicators were most frequently functioned as possibility, prediction, logical deduction, and hedging markers. The probability expression indicators were more frequently used for directing learning than for facilitating the learning process.

7.4 Summary of the chapter

The question indictors were most frequently used for asking verification, explanatory, and information-recall questions, and lower-order questions predominated in this regard. The directive expression indicators were most frequently used to mark expectations. The directive expressions were frequently used to facilitate and direct learning as well as raise critical awareness. The probability expression indicators were frequently used to mark possibility, prediction, and hedging. Facilitators frequently used probability expression indicators for directing and facilitating learning, and for questioning. Discussion and conclusions are discussed in the next chapter.

CHAPTER 8

DISCUSSIONS, RECOMMENDATIONS AND CONCLUSIONS

8.1 Introduction

Research in problem-based learning has evolved from concept confirmation through outcome studies to process research. The analysis of verbal interactions in the tutorials is a major part of the process-oriented research strategy. Although the main tool for learning in a PBL tutorial is language, only a few studies have recorded, transcribed, and analysed PBL discourses. These studies have analysed limited data using analytic techniques that are time-consuming, error-prone, and not based on a linguistic framework. The exceptions to date have not analysed collaborative knowledge construction in PBL tutorials. The present study has investigated collaborative knowledge construction in PBL tutorials and contributed to understanding knowledge construction in the graduate entry problem-based learning curriculum at the Derby Medical School, UK. This chapter discusses and concludes the findings of the study in terms of the following research questions:

1. What are the frequencies and functions of the commonly occurring reporting expression indicators?

2. What are the frequencies and functions of the commonly occurring shared knowledge indicators?

3. What are the frequencies and functions of the commonly occurring knowledge extension indicators?

4. What are the frequencies and functions of the commonly occurring knowledge enhancement indicators in the students' talk?

5. What are the frequencies and functions of the facilitators' questions in the problem-based learning tutorial?

6. What are the frequencies and functions of the commonly occurring facilitators' directive expression indicators?

7. What are the frequencies and functions of the commonly occurring probability expression indicators evident in the facilitators' talk?

8. How can the result of the research be used to improve the problem-based learning process?

As demonstrated in the review of the literature, studies that have analysed language data derived from PBL tutorials have not been based on an explicit linguistic theoretical framework. This corpus-based study has investigated collaborative knowledge construction in case of a PBL curriculum in the UK. By using the lexicogrammatical approach of the functional linguistic theory, the study implemented the corpus linguistics methodology to provide rich data on students' knowledge construction practices within some sociocultural theoretical perspectives. The verbal interactions of first year GEM students' collaborative knowledge construction were explored. In this chapter, the constructivist framework that underpins PBL and medical expertise is first presented. The aim is that this will provide a context for the discussion and interpretation of the research findings. Next, a discussion of the knowledge construction processes of the medical students and the facilitation of these processes by the facilitators is presented. This is followed by the presentation of the implications of the research findings for the PBL tutorial process and research. The chapter concludes by highlighting the significance of the research and its contributions, along with presenting the limitations of the study.

8.2 Constructivist Medical Knowledge Development

The key principle of constructivism which underpins PBL is that learning is equivalent to knowledge construction (Dennick, 2016). Students use previous knowledge to make sense of new experience and the new knowledge is then used to restructure prior knowledge. According to Dennick (2016), through the interaction of an individual with the world, meaning is extracted to construct a coherent, consistent, and elaborate view of the world. The new view of the world could be described as a cognitive (mental) model or schema. The process of creating the new mental model is similar to hypothetico-deductive reasoning. This process occurs in a mutual constitutive manner in that the mental model is used to interpret experience and experience leads to modification or restructuring of the mental model.

The importance of collaboration in learning lies in the fact that collaboration may create interactive activities (e.g. explanation, disagreement, and mutual regulation), which may trigger extra-cognitive mechanisms (e.g. knowledge elicitation and internalisation) that would not occur in a lone learning situation.

Medical education is a form of knowledge construction process in that it aims to create the medical cognitive model (or schema) that a clinician can use to view and interpret the patients' problems and diseases. A schema is an internal model that enables individuals to construct interpretations, representations, and perceptions of situations (Glaser, 1985). Thus, a mental model is a form of prior knowledge (Charlin et al., 2000; Dennick, 2016).

The problem-based learning curriculum provides optimum conditions for the training of doctors in that its underlying constructivist framework aligns with the theoretical basis of medical practice. The principles of PBL are well suited to a dynamic medical cognitive model and the collaborative condition in which PBL takes place can help challenge and correct students' misconceptions. The construction of knowledge in PBL follows the hypothetico-deductive model.

The constructivism framework specifies how this process should progress and this is done mainly through the use of language. Units of knowledge are related together with the use of connective linguistic devices such as coordinating and subordinating conjunctions. By analysing the language used in the process of knowledge construction, one can have an idea of whether or not the ideals of the PBL constructivist framework have been followed. Thus, the main research question for this study could be restated as follows: 'Did the medical students in the Derby GEM School construct knowledge according to the constructivist philosophy of PBL and did the facilitator guide the process of knowledge construction according to the principles of PBL facilitation?' In the sections that follow, the findings of the study are discussed with respect to these two questions.

8.3 Use of Referring Expressions

The referring expression indicators retrieved from the dataset fall into three categories: verbal and mental indicators, and learning materials and situations. The indicators were used to report personal experience, self-inclusive collective experience, others' experience, and experience with learning materials and situations. A qualitative analysis provided a richer insight into the functionality of these expression indicators. They were used to pursue social cohesion, share knowledge, promote social and cognitive regulation, jointly construct knowledge, resolve knowledge-related confusions, and grasp psychosocial issues behind patients' clinical features. Prior knowledge was used to make sense of the case scenario on one hand and on the other hand to refer to the discourse partners' talk within the tutorial, thereby suggesting co-construction of knowledge. The use of referring expressions is in line with the constructivist principle of problem-based learning. The amount of prior knowledge determines the amount of new material that can be learned (Schmidt, 1993). The positive effects of prior knowledge activation and elaboration on learning are well documented. Bransford and Johnson (1972) have shown that comprehension and recall of information is poor when students fail to activate the appropriate prior knowledge. Yew et al. (2011) have shown that previous knowledge promotes subsequent learning and recall. Moust et al. (1986) have also demonstrated that elaboration promotes the recall of information. The benefits of using conceptual maps for learning have also been documented: De Leng and Gijlers (2015) have shown that the use of diagrams and conceptual maps fosters knowledge construction and helps students to relate one concept to the other. Similarly, in a randomised study, Veronese et al. (2013) demonstrated that the use of concept maps helped students to integrate physiological mechanisms and identify their knowledge gaps. While the two studies on the benefits of conceptual maps were based on the subjective reporting of the research participants, the present study presents explicit reports of how conceptual maps were used to promote learning in PBL tutorials. The study provides evidence of how students integrate each other's statements and ideas. The appropriation of the ideas and opinions of others into one's own construction and statement of knowledge is described as intertextuality (Bakhtin et al., 1981; Koschmann, 1999; Tannen, 2007). Moreover, the study has described how the participating students used textbooks and lecture notes to resolve idea conflicts and guide tutorial discussions.

8.4. Shared Knowledge and Knowledge Construction

The next important finding of this study relates to the use of shared knowledge indicators. This aspect has been variously described as 'common ground' (Mercer, 2000), 'grounding' (Baker et al., 1999;

Clark and Schaefer, 1987; Clark and Brennan, 1991), 'shared knowledge' (Puntambekar, 2006), and 'conceptual convergence' (Oliveira and Sadler, 2008).

The markers studied as indicators of shared knowledge in this project have been variously labelled. For example, McCarthy (2003) has described them as 'small interactional response tokens'; Clancy et al. (1996) has called them 'reactive tokens', while Yngve 1970 cited in Young and Lee (2004) has named them 'backchannel responses'. These markers have affective and interactive meanings and are used to reveal intersubjectivity through conversation; they also function as discourse monitoring and co-construction devices (Gardner, 2007; McCarthy, 2003; Young and Lee, 2004).

In his study on small interactive response tokens in everyday conversation, McCarthy (2003) described several types of response tokens: (1) response token occupying the whole response move after which the turn reverts to another speaker (e.g. 'yeah', and 'yes'); (2) response move with minimal accompaniment after which the turn is taken over by another speaker (e.g. 'yeah exactly', and 'yes that's right'); (3) response tokens prefacing expanded response moves (e.g. 'yes because she has asked that question before'); and (4) negated response tokens. This classification scheme has significant implication for the findings of this research.

Three types of affirmative shared knowledge types – integration-oriented, agreement, and acknowledgement, which are similar to expansion responses, response move with minimal accompaniment, and response token occupying the whole response move respectively in the above scheme – are prevalent in the corpus. The high frequency of integration-oriented convergence may suggest that the students co-constructed knowledge by taking up each other's views points and building on them. They built on each other's contributions through paraphrasing, evaluation, addition, and causal explanation. Judging by the frequency of agreement tokens, they used affirmative tokens to confirm their peers' ideas and monitor their discussions. The frequency of acknowledgement tokens may also suggest that the students enjoyed some degree of social contact as they listened to and monitored each other's contribution, and provided appropriate support.

Although low in frequency, the use of reactive tokens (e.g. 'oh', 'aha', and 'oh yes') is remarkable. Data analysis demonstrated that the tokens were used to preface orientation to new information or to mark information recollection. The common characteristic of these response tokens is that they 'externalise a presumed inward state' (Goffman 1978:794 cited in Wilkinson and Kitzinger (2006) and Heritage (1984) have stated that they signify a change in the state of knowing. Data analysis demonstrated that the tokens were used to mark students' orientation towards information (e.g. 'Oh you think he's got an infection…') or to mark information recall (e.g. 'Oh I remember having it…'). Schiffrin (1987) has claimed that when an individual orientates to new information (through phrases such as 'oh yes'), they replace old idea with new knowledge. The implication of this finding is that the students constructed new knowledge when they orientated themselves to new information and came to a shared understanding with the speaker. They also used reactive tokens to mark information recall. In either of these two instances, shared knowledge is established.

Another remarkable finding relates to the use of negative reactive tokens. The results of the analysis of the shared knowledge indicators also showed that conflict-oriented convergence was frequent in the tutorial group talk. Data analysis showed four types of conflict-oriented shared knowledge – corrective, elaborative, and contrasting. (Jasinskaja, 2013) has stated that elaboration and correction refer to situations where two utterances describe the same situation.

Students' readiness to disagree with discourse partners and challenge viewpoints is central to the process of gathering shared knowledge (Baker et al., 1999; Oliveira and Sadler, 2008; Swain, 2011). Incompatibility of ideas and viewpoints between learning partners can lead to cognitive conflict, which may trigger a search for new information and experiences to achieve cognitive equilibrium. This will lead the student to present more ideas and rationales, and integrate the other's information, leading to the achievement of a conceptual restructuring. It would appear that in this study, the supporting environments of the groups enabled the students' readiness to disagree with the result that the conflicts were resolved constructively through different strategies such as correction, concession, and rephrasing. The unelaborated divergence of the students' contributions mainly related to responding to verification and confirmation questions.

Finally, the students took over the administration of the group and there is evidence that they regulated their conversational behaviours.

8.5 Extension as Knowledge Construction

Idea units and experiences are often represented as grammatical clauses or groups (Halliday and Matthiessen, 2004; Schiffrin, 1987). According to Halliday and Matthiessen (2004), these grammatical elements are logically linked together to form complex clauses. Halliday and Hassan (1976) have described the semantic structuring of the text segments as text formation, which is carried out through cohesion devices. In this context, text is taken to mean both written and spoken discourse (Halliday and Hassan, 1976). When coordinating connectives are used, Halliday describes the process as extension (Halliday and Matthiessen, 2004). The formation of texts as described above is similar to the formation of medical cognitive models. The mental cognitive representation of a diagnostic problem occurs in the form of the illness script into which patient attributes, signs, and symptoms of disease, and the knowledge underlying the links are structured (Charlin et al., 2000; Collard et al., 2009b; Dennick, 2016; McMillan, 2010). The nodes of this script are similar to grammatical forms and its logical linkages are similar to grammatical connectives. This section of the discussion will focus on the extension mode of knowledge construction.

Three coordinating connectives were analysed in this study: additive (i.e. ‘and’), alternative (i.e. ‘or’), and adversative (‘but’). In this study, ‘and’ was the most commonly occurring connective followed by ‘but’ and ‘or’. This finding is in line with previous literature. Schiffrin (1987) found that 1,002 idea units in her corpus were marked with ‘and’ compared to 440 marked by ‘but’ and 53 marked by ‘or’. Chafe (1988) found that ‘but’ was one-fifth as common as ‘and’ and that ‘or’ was one-sixteenth as common as ‘and’ in the corpus of his research participants. In this study, the predominant use of additive knowledge extension indicators for simple addition and the low frequency of their use for co-construction would suggest that additive indicators were mostly used for intrapersonal cumulative knowledge construction. Alternative indicators were mainly used to mark alternative questions and concepts. This finding, along with the one recording the low use of knowledge extension indicators for co-construction (only 13%), would suggest that alternative indicators were commonly used in an interactive fashion with the students offering ideas to their peers for consideration. The use of ‘or’ to mark questions also suggests an interactive and friendly social environment in that questions were asked as opinions and options presented to colleagues for deliberation. Because of the supportive social environment, the students were not afraid to volunteer their ideas and doing so reduced the cognitive load of open questions on their colleagues. Adversative indicators were used more than the other two indicators for knowledge co-construction. The most frequent use of adversative indicators was to contrast the assertion in previous clauses. Putting these two findings together, one could conclude that the students actually challenged each other’s ideas.

Apart from the coordinating functions, additive and adversative indicators were also used for other functions. Additive indicators were used to mark temporal sequence, contrast, causal-conditional relationships, and non-causal elaboration while adversative indicators were used to mark elaboration, concession, condition, and knowledge-vs-reality relationships. This implies that these coordinators were used to perform the role of subordinating conjunctions. Experts in language studies have recognised the advantages of this mode of coordinator use. Spoken language places great cognitive demand on discourse partners because a large amount of data needs to be produced and processed within a short period of time. The listener cannot return to previous talk to process them without interrupting the discourse and the speaker's flow of thought. Thus, paratactic connectives are said to be useful in simplifying the processing of high-paced spoken discourse as it avoids the use of subordinating connectives, which are more difficult to process (Beaman 1984 cited in Barth, 2000; Chafe, 1988). Paratactic connectives are used to link clauses of equal grammatical status while subordinating connectives join clauses of unequal grammatical status. The use of coordinating indicators in knowledge co-construction suggests that the discourse partners assign equal status to each other's contributions. This provides opportunities for idea negotiation and debate. Where an idea is being contrasted, the use of the coordinating connective 'but' preserves the claims of the initial speaker as equal status is assigned to both contributions, implying that the previous idea is not considered to be inferior (Barth, 2000). The use of alternative coordinating connectives is also important. Within the Piagetian tradition, knowledge growth is considered to be dependent on coordinating alternative viewpoints and the use of coordinating connectives between the ideas of discourse partners balances power relations (Howe, 2009). The implication of this mode of knowledge construction is that it keeps the discourse space open and expansive, and allows for free exchanges of ideas and viewpoints with the potential result of constructing an expanded knowledge base.

8.6 Enhancement as Knowledge Construction

The second group of devices for structuring idea units together is called subordinating conjunctions (Halliday and Hassan, 1976; Halliday and Matthiessen, 2004). In this study, the most prevalent knowledge enhancement indicators in the students' talk were 'because', 'if', 'so', 'when', 'as', 'whether', 'than', and 'that'. These indicators were used for explanation and reasoning, and for defining phenomenon and mobilising prior knowledge. About a third of the indicators were used for knowledge co-construction. It is also interesting that students evoked theories, supported their explanations with evidence, evaluated propositions critically, and encoded information in mnemonics to aid future retrieval. This would suggest that the students were engaged in the co-construction of an explanatory case model linking causes to clinical symptoms and in collaborative reasoning that involves making deductions and predictions.

The results of the study provide an understanding about how discourse in PBL tutorials provides support for the development of students' biomedical knowledge models. This is very important because medical practice, by its very nature, seeks to explain medical problems and medical treatment strategies as conditional and situational. Through discourse, the students are being initiated into the problem representation and understanding frameworks of experts in the medical profession. The importance of social interaction and language use in the development of conceptual knowledge has long been recognised. Vygotsky has said that intermental social activity promotes intramental cognitive development (Vygotsky, 1930/1978). Halliday has conceived that knowledge acquisition is synonymous with language learning (Gee, 1994; Halliday, 1978; Wells, 1994).

The educational literature has empirical evidence highlighting the positive effects of collaborative explanation and reasoning on knowledge development. Wegerif et al. (1999) demonstrated that collaborative explanation and reasoning among eight- and nine-year-old children in small-group exploratory discourse improved their performance on assessment tests. Mercer et al. (2004) also showed, in a comparative study, that children who used language to learn science through reasoning and explanation gained significantly better scores in science tests compared to the control group.

In a small group setting, the discourse partners may seek explanation or give explanation to others. The relative benefit of these approaches to explanation has been documented in the literature. In a study by Chi et al. (1994), the performance of 14 grade four students who were prompted to give explanations after reading a text on circulatory system was compared with that of a control group, which had not received the prompt. The students' understanding was assessed by asking them to answer very complex questions and by conducting a mental model analysis of the self-explanation protocols. The results showed that students who were more proficient at explaining achieved the correct mental model of the circulatory system, whereas unprompted students and low explainers did not. Webb (1989), in a literature review on explanation and learning, has demonstrated that giving elaborate explanations positively influences students' achievement whereas receiving elaborate explanations only has few connections to academic achievement.

In this study, the discussions of the students aligned with collaborative knowledge construction as they collaboratively constructed a causal representation of the biomedical knowledge model of the problem case. Feature specification was used to provide reasoning or explanation for hypotheses in line with how medical professionals make diagnoses based on the disease attributes in the illness script (Charlin et al., 2000). Other types of explanations were advanced to explain the problem case. Students provided physiological explanation for observed features, consequent explanations for case patient behaviours, antecedent reasons for disease presence, mechanistic reasons for observed symptoms, and physical reasons for disease periodicity. The results of the study are in line with that of

Hmelo-Silver and Barrows (2008) and Frederiksen (1999) who studied explanatory knowledge construction of medical students in PBL settings. However, in these studies, very limited data was manually analysed using the discourse analysis technique.

8.7 Facilitators' Questions and Knowledge Construction

The role of the facilitator is central to the success of PBL tutorial sessions. One of the strategies that facilitators use to guide learning in PBL tutorials is by asking questions. Such questions play an important role in determining the nature of discursive activities in the PBL tutorials. For an inquiry-oriented learning situation like the PBL, the questions that are well suited for reflective thinking and expansive conceptual knowledge building need to be open ended and not accompanied by evaluation.

Analysis of the facilitators' dataset showed that there were questions that were indirect, and stimulated co-construction and reflection. On the other hand, there were questions that conformed to the IRE model of questioning. More than two-thirds of the questions asked by the facilitators were fact-recall questions, while only about one-fifth demanded speculative, inferential, and evaluative thinking. Moreover, less than 10% of the questions were process-oriented, suggesting that the majority of the questions were content-based. The findings are in line with the literature on the relative prevalence of higher-order and lower-order questions in the classroom. Cotton (2001), in a literature review on classroom questioning, found that 60% of the questions asked in the classroom were lower-order questions and 20% were higher-order questions. Profetto-McGrath et al. (2004) classified the types of questions asked by the facilitators in a nursing PBL tutorial; they found that 78.5% of the questions were lower-order questions while only 5.5% demanded reflective, synthetic, and evaluative responses. Similarly, Phillips and Duke (2001) and Hsu (2007) have found that clinical nursing teachers ask between 65% and 90% lower-order questions. The findings in this study contradict the findings of Hmelo-Silver and Barrows (2008) where the facilitator asked mainly open-ended and meta-category questions, which stimulated the students to monitor their progress and focus on self-directed learning. However, the findings in Hmelo-Silver and Barrows (2008) study may not be generalisable as the facilitator in the study was Barrow himself who was highly experienced in PBL facilitation and had worked within the orthodox PBL curriculum. Research findings on the relationship between the cognitive level of teachers' questions and students' academic achievement have been mixed. Some studies have found that higher-order questions are superior to lower-order ones; some have found no difference between these two types of questions, while some other studies have found the opposite to be true (that lower-order questions are superior to higher-order ones) (Cotton, 2001).

However, medical practice demands information analysis, synthesis, and evaluation, and doctors' critical thinking skills to help develop their problem-solving abilities as they function in complex clinical situations. Thus, theory seems to suggest that these set of skills would be better supported by higher-order questioning. The meagre use of process-oriented questions in this study would suggest that the facilitation of the tutorials was content-driven rather than process-based.

8.8 Facilitators' Stance Expressions and Knowledge Construction

One of the key requirements of effective PBL facilitation is that the facilitators need to refrain from directing group discussion by dispensing information to the students (Al-Shawwa, 2005; Azer, 2005; Neville, 1999). They also need to exhibit interactive behaviours that expands the group discussion as a scientific community (Carter and McCarthy, 2006; Howe, 2009).

The findings in this study show that the facilitators frequently used directive expression indicators to mark the expected learning behaviours of the group participants and to preface doctors' obligation. At other instances, the facilitators used self-exhortation indicators to preface advice to students and intimate them about curriculum requirements. It is important to note how the directive expression indicators were deployed to facilitate PBL interactions.

Data analysis showed that the directive expression indicators were most frequently used to direct learning mainly in the area of dispensing information and ideas. This would suggest that the facilitators engaged in mini-lectures in the tutorials, although this occurrence varies from group to group. While some tutors used more facilitation techniques, others used subject matter to facilitate: in four groups, the tutors used more directing techniques; in three, the tutors used more facilitating techniques; and in one group, the tutor balanced facilitation with directing.

The findings are in line with the literature regarding the contextual behaviour of facilitators. Schmidt (1994) has shown that facilitators adopt directing roles when students have insufficient prior knowledge or when the case structure is less clear.

Similarly, the findings of the analysis of the probability expression indicators showed striking observations. The facilitators often used probability expression indicators as possibility markers when responding to students' questions, offering feedback, refocusing tutorial discourse, offering suggestions, and asking questions. These indicators were also used to mark anticipated events and to preface logical deductions from explicit and implicit evidence. These indicators were more commonly used to temper the content information they dispensed to the students and less commonly for process facilitation.

Problem-based learning tutorials, like other communal modes of scientific learning, engage students in academic discourse, which concerns exchanging viewpoints and conclusions drawn from evidence rather than hard indisputable facts. The language of the tutor who is responsible for guiding the group needs to reflect this ideal. The use of such discourse-expanding language widens the scope of discourse and helps build robust networks of cognitive medical models to support the students' future professional practice. Such verbal behaviour also tones down the power relations between the facilitators and the students. By marking their contributions with probability tokens, the facilitators model the students in the act of scientific discourse. Moreover, such verbal behaviour may be an invaluable humanistic trait in their future professional practice.

In summary, I have demonstrated – using lexicogrammatical analysis – how conceptual medical knowledge is constructed. Knowledge is a representation of experience, and it is constructed and co-constructed with others by mobilising prior experience and various cognitive tools. Facilitators are quite central to students' knowledge construction as they provide conditions for expansive discourse by asking questions and using linguistic devices that could regulate power relations and temper their contributions. This type of conversation behaviour has the potential to expand the tutorial discourse space and could lead to an expanded causal model of knowledge construction.

8.9 Emerging Pedagogic Issues

The results of this research raise several pedagogic issues regarding the link between students' knowledge construction and PBL facilitation. One of the key findings that emerged from the study is a lack of alignment between facilitation practice and the constructivism philosophy of PBL.

While curricula are mainly concerned with ***the what*** of education, pedagogy, on the other hand, is concerned with ***the why*** and ***the how*** of education. A curriculum may be very relevant and interesting but what makes it work is how it is presented and delivered in the classroom (Nieto, 2010). This section discusses the facilitation behaviour of the PBL facilitators on the basis of two main pedagogical theoretical frameworks. The findings in this study suggest that the facilitators conducted the PBL tutorials using a mixture of behavioural and constructivist theories. However, the two theories are significantly different in the way instructions based on them are organised and facilitated.

Behavioural learning theory concerns externally observable behaviour, which is shaped through conditioning, reinforcement, and punishment (Gould, 2012). The instructional practice based on behavioural theory differs. The planning of instruction under behaviourist perspective consists of stating the externally observable behaviours that students are expected to attain by the end of the instruction. Education researchers have continued to show that educational deliveries in the classrooms vary little from the traditional methods (Neville, 1999; Nieto, 2010; Perrone, 1998; Wiske, 1998a). Findings from this study suggest that tutors in some tutorial groups employed facilitation styles that involved drilling, initiation-response-evaluation (IRE) exchanges, and asking students lower-order questions that required students to display their knowledge ('The correct answer is yes you feel better'). The extract below shows the types of IRE discourse exchanges noted in the dataset. For example:

T1 – Because the palpations will usually resolve when you do what?

F1 – Rest
T1 – Rest or otherwise do what to the patient?
F2 – Do anything that is going to get his heart to slow down.
T1 – So like beta-blockers.

Furthermore, about 67% of the questions asked by the facilitators were of the lower-order variety, while higher-order questions constituted about 22%. In several instances, data analysis suggests that some of the facilitators dominated the discourse and transmitted information to the students. The ideal of constructivist pedagogy discourages facilitators from dispensing information to students and overdirecting their exploration of the subject matter. For example:

T8 – Because you have a very short window in which to actually act with the treatments, you just assume that it's an MI to begin with, because the sooner that you get started with things like thrombolytics, the better the outcomes. The management protocol has specifically built into it the ability to do some of that differential while also being very aggressive with the treatment.
M1 – Yeah it makes sense. It's quite–
T8 – It doesn't sound like you guys have a difficulty understanding the difference between angina and MI, but I think some folks do just because if you look at the management plan, it's almost as if you're treating both of them the same way and that's the reason, because if it is an MI, you'd rather get that going early, and if it's angina you don't need to do the thrombolytics.

Another striking finding of the study is the impulse for coverage noted in the data from the tutorial groups. Data analysis showed that facilitators were under pressure to cover the case under discussion. This forced the facilitators to overdirect the tutorial process and transmit knowledge to the students. For example:

> T1 – I am impressed that we have got as much work done as we did. That was the other thing I was going to say, I mean part of it was I was also trying to push to wrap up most of the case last time if you had not noticed.

Recognising that this facilitation technique violates the facilitation principle of PBL, the facilitator apologised to the students for overdirecting them.

> T1 – I am glad you guys did not feel like I talked too much. I was not really happy with saying quite as much today and was feeling kind of bad about it I think part of the way through there so I am glad that at least that seemed to go fairly well.

Education experts of the constructivist tradition assert that the depth of information processing is more important than superficial coverage of the broad curriculum content (Gould, 2012; Perrone, 1998; Wiske, 1998b). They have recommended more thorough inquiry around a smaller number of critical ideas, concepts, and themes that must be studied in depth and suggested that teachers need to resist the urge to put everything into any single course (Perrone, 1998).

Another striking finding of the study relates to the high frequency of mini-lectures such that the tutorial discussions were mere aggregations of mini-lectures given by the students to each other. In one group, a student spoke about 1,900 words without interruption from their peers. Unlike the behaviourist approach, which sees learning as the accumulation of knowledge and focuses on external response to stimuli, constructivist theory is less concerned about the actual stimuli and the resulting response. Rather, it concentrates on the happenings between the two events, considers learning as internal to individuals, and focuses on internal processes that connect external responses to stimuli (Gould, 2012). Thus, the constructivist approach to learning focuses on understanding and meaning, and mere accumulation of knowledge is not considered as learning (Marlowe and Page, 2005). The analogy of a jigsaw puzzle is very useful in seeing how learning occurs through the constructivist approach. The individual pieces of a jigsaw puzzle have very little meaning in themselves. However, as pieces of a jigsaw puzzle are fitted together, a picture emerges such that every piece of the puzzle that is added changes the perception of the emerging picture. The jigsaw puzzle becomes meaningful as the whole picture is seen when the pieces of the puzzle relate to and fit with each other.

Thus, individual pieces of knowledge have limited meaning in themselves and so, accumulating more and more knowledge does not constitute learning. The accumulation of knowledge without meaningful relationship of ideas and knowledge pieces is termed coverage, while the structuring of knowledge with the emergence of meaning and understanding is termed knowledge construction. The pedagogic process in the setting of constructivist theory involves helping students to form whole pictures and meaningful linkages between pieces of knowledge, and not to encourage the mere accumulation of knowledge (Gould, 2012). The guided creation of students' understanding, which variously implies 'the ability to think and act flexibly with what one knows' (Perkins, 1998: p. 40); and the 'ability to creatively use presented information to solve transfer problems' (Mayer, 1989: p. 43), is the important feature of PBL facilitation. This involves an active learning process where students can self-direct, openly explore, explain, build reasoned arguments, apply knowledge to solve problems, and link knowledge to prior knowledge and outside school experiences (Ritchhart et al., 2011, 1998). Findings in this study reveal several examples of pedagogic behaviours that are compatible with the constructivist instructional process.

In all the tutorial groups, students related their knowledge to prior experiences such as previous PBL cases ('Similar to that woman the other day who you did a history for'); learnings from previous lectures ('In the pathophysiology lecture we had the other day heart failure is not a disease in itself...'); previous job experience ('… when I was working in respiratory, the Head of the Department drew these out'); and general life experiences ('I have been through immigration from the States and Canada lots of times with someone with a dagger...').

Furthermore, the facilitators encouraged the students to make links with their previous knowledge ('What is oedema? What do we already know about it?'; and 'Is there anything from other cases that you have done so far that you would like to go over again that you think might be relevant?'). The students were not just assimilating information; they exhibited intellectual scepticism towards information presented in the tutorial (‘Yes you did. I do not believe you’; and ‘Cite your source’), and were critical of their lecturers (‘I’m not very confident about Dr X’s notes to be honest; it was all a bit of a shambles, wasn’t it?’). However, the participation of the facilitators in the tutorials varied. Analysis of the tutorial talk data suggests that some facilitators exhibited a hands-off facilitation style while some were overtly directive. The students were disillusioned with the hands-off facilitation style because their confusion relating to the case was not resolved, but applauded detailed and deep discussion of the material of the case. For example:

> M1 – I am very much enjoying the way that it is much more down to the nitty gritty now, looking at the smaller level of what these things are in terms of physiology.
> M2 – It is not as vague as the respiratory stuff in my opinion.
> M1 – Yeah, I agree.
> M2 – You can get a grip on details a bit more.

Lower-order questions of the knowledge display types were predominant across the tutorial groups, but some facilitators did ask referential questions (‘I don’t know and I am wondering whether it’s worth having a table...’; and ‘I do not know really. I am just making sure we establish what the relationship is’). Furthermore, the findings in this research showed that the students planned their task and the facilitators encouraged them to reflect on their performances at the end of each case, which was in line with the constructivist perspective.

The examples above indicate that learning processes in the PBL tutorials studied were variably conducted by the facilitators. A large number of studies suggest that the teachers' instructional methods are influenced by many factors (Elmore and Burney, 1997; Handa and Herrington, 2003; Mumtaz, 2000; Perrone, 1998). Therefore, pedagogic issues require holistic conceptualisation. The conceptualisation suggests that the learning process and the teacher-learner relationship in the classroom are influenced by complex interrelating teacher knowledge and beliefs, institutional culture, social processes, and political agenda in the wider society (Waring and Evans, 2015). Thus, facilitators' behaviour needs to be seen beyond what occurs in the tutorial groups, and this notion aligns with the view of educational researchers that several conditions are required for a constructivist learning process to thrive (Hmelo-Silver and Barrows, 2008).

The knowledge and belief of the facilitator are important determining factors in how they conduct the learning process in PBL tutorials. First, there seems to be a great deal of confusion and disputes about what PBL facilitation is and how its complexity is sufficiently captured within any definition. It has been observed that PBL curriculum becomes highly variegated as it gets planted in different cultures and institutions (Barrows, 1986; Karen et al., 1998). Thus, PBL facilitation is likely to mean different things to different people. As no one confronts reality with an empty mind, the facilitators and policy makers may also approach the notion of facilitation from very different perspectives and conceptual positions. Consequently, although facilitation takes place in all tutorial groups, the details of the facilitation may be very different.

Second, facilitators come to the PBL tutorials with different views about what learning is and how it should take place. These perspectives influence how the tutors facilitate PBL tutorials. However, a clear understanding of the learning theories, and the alignment of theory with practice has the potential of improving PBL facilitation. Teaching generally is a complex task (Elmore and Burney, 1997; Marlowe and Page, 2005; Waring and Evans, 2015) and classroom practice may not fit perfectly into the boxes that theory provides; however, learning theories can provide benchmarks against which classroom practice can be compared and analysed (Gould, 2012). This can provide a more informed view of facilitation and a basis for reflection on practice, which in turn can lead to a more effective PBL facilitation. Learning theory provides a framework against which to analyse and reflect on facilitation experiences, and identify possible ways in which things may be done differently the next time (Gould, 2012).

For example, one issue that features in the analysed tutorial data relates to the coverage of the case material. Emphasis on the coverage of the curriculum content relates to behaviourism theory. This is because instructional delivery is linear in nature without requiring the activation of prior knowledge. This is different from the constructivist approach where the curriculum is designed in a spiral fashion with cases carefully designed such that they relate to each other. Activation of prior knowledge, which is a key feature of constructivist learning theory, ensures that aspects of the previous cases are revisited. What is important from the constructivist perspective is not coverage but the depth of the processing of curriculum material and the establishment of links between concepts.

Third, a sound understanding of the professional nature of teaching is important for teacher education and consequent classroom practice. How teachers perceive professionalism within immediate and wider school roles affects how professionalism is enacted in classroom practice (Waring and Evans, 2015). Higgs and Titchen (2001 cited in Waring and Evans, 2015: p. 1) have conceived professional teaching holistically as a combination of 'people-centred and interactive processes, accountability and professional standards, practice wisdom, professional artistry, openness to knowledge growth and practice development and engagement in professional journeys towards expertise.'

Pedagogic expertise requires a blend of science, craft and art (Pollard 2010 cited in Waring and Evans, 2015). Teaching is a complex enterprise (Elmore and Burney, 1997; Lieberman and Miller, 1992) and a journey that requires commitment to continuous pedagogic and subject matter development, theory-guided reflection on practice, and research-informed pedagogic actions (Waring and Evans, 2015). Consequently, teachers need to be adaptable and possess the capacity to align their classroom practice with the theory underlying their curriculum.

Fourth, teachers do not change their classroom teaching methods just because a professional teaching program or a new curriculum tells them to do so (Guskey, 2002). The most powerful evidence that a teacher is successful in their teaching is their students' achievements in examinations (Lieberman and Miller, 1992). Generally, the government and the society measure the success of schools and teachers by the achievements of the students in examinations. It is known that learning is intimately related to the assessment system (Biggs and Tang, 2011). What is assessed and how it is assessed inform the students and the teachers about what is worth learning and how it should be learned (Biggs, 1996; Biggs and Tang, 2011). Thus, teachers teach in manners and ways that help their students to succeed in examinations. The implication is that if the assessment system is based on the behaviourism framework (e.g. objective tests, recognition tests, and recall tests), the teachers are more likely to adopt teaching methods that ensure students' success in the tests regardless of the theoretical framework that underlines the curriculum.

Finally, the institutional culture in which teaching occurs also determines how facilitation is conducted in the tutorial groups. Teaching methods that require students to demonstrate understanding and meaning are less likely to succeed in the presence of teacher-proof programs where teachers are without greater autonomy individually and collectively (Perrone, 1998). To this end, (Novak, 1994) presents the following quote by Linda Darling-Hammond:

> We have never invested in teachers in this country. We've created a system in which we regulate heavily and tell teachers what to do. We create teacher-proof curriculum. We pay teachers relatively little, and then we spend enormous sums creating a superstructure of supervisors to look over their shoulders

In some instances, the institution's administrative structure may pay lip service to the constructivist theoretical framework while there is a deep allegiance to behaviourist ideals. This could be because of fear of loss of influence or because there is no definite empirical evidence demonstrating that the constructivist instructional method is superior to the behaviourist approach.

8.10 Recommendations

The analysis of a large body of dataset in this study has contributed to understanding how knowledge is collaboratively constructed during GEM students' problem-based learning tutorials. The pedagogic issues that emerged from the findings of the study need to be seen within a wider context of interrelated factors in which daily PBL tutorial activities occur. The research findings can provide educational policy makers and teaching staff with useful information about students' learning and PBL facilitation.

8.10.1 Pedagogic recommendations

Generally, the constructivist perspective suggests that current knowledge builds upon previous learning; learning involves deriving understanding and meaning; meaning is dependent upon establishing relationships; and relationships are stored internally as cognitive structures (Gould, 2012). The facilitators are expected to be helping the students to achieve these features by using pedagogic methods that make these features visible. These pedagogic methods can take the following points into account:

1. The teachers deliberately request students to connect new learning with previous knowledge and life experiences, e.g. 'How is this related to what you have encountered in previous cases?'

2. An important aspect of understanding something is the ability to notice its parts and features, and describe it in detail. Thus, facilitators can help build understanding of a concept by requesting a description of its part and features, for example, 'How would you describe heart failure?' or 'What in your opinion counts as heart failure?'

3. The facilitators can encourage the students to make knowledge construction visible by requesting students to explain and elaborate. Halliday (1993) has observed that language learning is synonymous with learning: '… ontogenesis of language is at the same time the ontogenesis of learning' (p. 93); '…. language is the essential condition of knowing, the process by which experience becomes knowledge' (p. 94); and '…learning is learning to mean, and to expand one's meaning potential' (p. 113). Language is the means of building conceptual knowledge, which in turn serves the purpose of making sense of reality. That is why language is used 'to interpret experience by organising it into meanings' (Halliday, 1993: p. 94). If the medical conceptual model is a potential determinant of professional expertise and the structuring of the content knowledge is more important than mere aggregation of concepts (Charlin et al., 2000; McMillan, 2010), then how the medical knowledge is structured during tutorial discourse becomes very important. The structuring of ideas can be promoted by asking higher-order questions that urge students to explain and elaborate on their ideas.

4. The cases used in the tutorial need to be carefully designed so that the cases link together in some ways. This will ensure that a new case is linked to some aspects of the previous cases through activation of prior knowledge. Consequently, the impulse for superficial coverage at the expense of deep knowledge processing may be reduced or removed.

5. The assumptions underlying the ideas of students can be made visible by asking them to provide justification for their ideas. For instance, students may be asked to justify their hypothesis: 'Why do you think A causes B?'

6. During tutorial discussion, information is processed in short-term (or thinking) memory, which has limited capacity to hold information (Ormrod, 2004). The facilitators can help the students to hold more information in the working or thinking (short-term) memory by requesting students to categorise their ideas or crystallise talk into main ideas. For example, students are more likely to remember the causes of a disorder when they are grouped in some way (e.g. according to anatomical structure) than if the causes are learnt without the use of such categorisation (Baddeley, 1999).

7. The use of directive language devices has the potential to constrain conversation and diminish the construction of conceptual models (Davis et al., 1992; Eagle et al., 1992; Neville, 1999). The use of nondirective language devices and possibility markers can expand the discourse space and temper power relations between the students, and between the facilitators and the students (Dolmans et al., 2002; Webb, 2009). This emancipating situation can allow the free exchange of ideas, which can expand the mental conceptual knowledge of the students. It also reassures the silent participants who are afraid to contribute for fear of criticism.

8. The establishment of a corpus that exemplifies expert facilitator knowledge construction for use in guiding medical students' talk in the PBL tutorials.

8.10.2 Administrative recommendations

What the teacher does in the classroom may be the tip of the iceberg of the regulatory and administrative conditions influencing their behaviours. The facilitators need administrative support to function optimally. The following suggestions may be fruitful in this regard:

1. The facilitators need administrative support and some degree of freedom to perform their facilitation duties as professionals. They need to be supported in their continuous professional development. The complexity of facilitation needs to be understood and there is a need to see the facilitators as learners who are striving to attain expertise in facilitation.

2. The continuous professional development programs of the facilitators may involve tutorial visits by external educationists and provision of feedback to improve the facilitation process.

3. A collaboration of the facilitators aiming at quality improvement can be established, whereby colleagues observe each other's tutorial session to see how often the students provide extended and elaborate causal explanations and to evaluate the facilitator's performance and provide feedback. Collier again quotes Linda Darling-Hammond in this regard: 'Teaching is definitely a team sport'. Elsewhere in the same source, Hammond states the following:

> You always have a way to have partners help you solve your problem, to help you improve your practice. This allows you to be more effective with students. If you're more efficacious, you're more satisfied with your carrier (Collier, 2011: p.13)

4. The facilitators may form a regular discussion group characterised by openness and congeniality where facilitation challenges are discussed and solutions are provided by the members (Lieberman and Miller, 1992). An expert in facilitation may be present in such meetings to offer professional advice. A teaching community provides a forum where teachers can share knowledge and expertise such that the facilitation practices in the tutorial groups are products of group deliberation and consensus. To this end, (Wei et al., 2009: p.16) states the following:

> In line with other research on professional development, collegial, job-embedded models of support appeared to have more effect on practice than traditional workshop models of training.

8.11 Contributions to Knowledge

This study contributes to our understanding of interactive knowledge construction in graduate entry medical PBL tutorials. Previous studies have used subjective, laborious, and error-prone approaches to analyse knowledge construction by medical students in PBL tutorial settings. This has only afforded researchers the opportunity to analyse limited datasets.

To the best of my knowledge, this is the first study to apply the lexicogrammatical analysis approach to collaborative knowledge construction in the graduate entry medical PBL curriculum. Along with the corpus linguistics methodology, Wmatrix 3 was used to retrieve datasets from the corpus consisting of the video recordings of full PBL cycles of eight PBL groups, and the data was analysed using the lexicogrammatical approach and qualitative techniques. The functions of the referring, shared knowledge, and knowledge construction indicators along with those of the facilitators' questions and stance expression indicators were analysed. The findings from the study provided insights into how the GEM students could jointly construct medical knowledge under the guidance of the facilitator.

Another contribution that this study has made to the existing body of knowledge is the objective documentation of how the voices of others were integrated into the students' knowledge construction process. The study has also documented, in an objective way, the use of coordinating conjunctions to fulfil the functions of subordinating conjunctions. As has been highlighted previously, this also fulfils interactive and cognitive economic purposes. Similarly, as far as I know, linguistic modality forms were used as markers of facilitation directiveness and nondirectiveness for the first time in this research.

Although the use of diagrams has been documented in problem-based learning tutorials (De Leng and Gijlers, 2015; Veronese et al., 2013), these studies have largely been based on subjective reporting. In this study, the students' conversations were analysed to identify how the diagrams were used in the tutorials. I have also objectively documented how lecture notes and books were used by the students to help tutorial discussions.

8.12 Research Limitations

Certain limitations need to be considered when evaluating and applying the results of this study. These include the following:

1. The first obvious limitation of this study relates to its generalisability as it reflects the practice in a single institution. The aim of the research is theoretic generalisation and not statistical generalisation. Replication of the study in another context and curriculum is therefore advocated.

2. Participating in and consenting to the study was voluntary. The data collection for the study was done on two occasions. Several groups declined to participate in the study based on personal reasons and some participants who had previously consented dropped out of the study. Those who participated in the study were possibly more motivated. It is difficult to know whether the tutorial process of the groups that did not participate differed from those that did participate.

3. Language is a vast field in that there are many ways by which the same meaning may be conveyed. This study has analysed data from the structural perspective. The same meaning can also be constructed from non-structural perspectives. This limitation needs to be considered when drawing conclusions from this study.

4. The analysis of the functions of the indicators was done by a single person and no inter-coder reliability was observed.

8.13 Future Research Directions

Many opportunities exist for further research on problem-based learning using the lexicogrammatical approach:

1. Future studies to examine the linguistic indicators of the objectives of each phase of the tutorial cycle are desirable.

2. Future studies may examine the correlation between facilitators' personal characteristics and the indicators of directiveness and nondirectiveness.

3. Studies may also correlate the frequency of directive and nondirective indicators in the tutorial corpus with a questionnaire survey of the students' perceptions.

4. More research involving several tutorial groups is needed with respect to the lexicogrammatical indicators of the students' questions.

5. More lexicogrammatical research is needed to explore other types of knowledge construction connectives.

6. Future studies may explore the relationship between the frequency of facilitators' directive expression indicators and the students' rating of the facilitators' directiveness.

7. Further research is needed to explore the relationship between the frequency of students' knowledge construction indicators and their scores on achievement test.

8.14 Conclusions

Comparative outcome research on PBL curriculum has not shown that it is superior to traditional curriculum in terms of students' achievements. Small sample sizes and error-prone, laborious analytic techniques have been identified as significant limitations of process-oriented studies on PBL.

The aim of this study was to explore how GEM students in Derby collaboratively construct knowledge in PBL tutorials and how the PBL facilitators guide this process. To overcome some of the limitations of previous studies, the Wmatrix 3 online software tool was used within a corpus linguistics methodology to retrieve indicators linked to research questions from the compiled corpus of the PBL tutorial talk. Data analysis was carried out using the lexicogrammatical approach of the functional linguistics framework. Quantitative findings from the study were enriched with illustrative examples of the tutorial participants' talk. Several interesting findings emerged from the study.

First, the study demonstrated the feasibility of using the Wmatrix 3 tool as part of the corpus linguistics methodology to analyse the participants' talk over varying PBL tutorials.

Second, the participating medical students anchored knowledge construction on their prior knowledge, externalised their ideas, and jointly constructed knowledge with their peers through the extension and enhancement of ideas. However, the connective indicators were more frequently used for autonomous knowledge construction than for knowledge co-construction. The PBL tutorial facilitators most frequently asked lower-order questions and were more direct in their tutorial facilitation.

Third, several pedagogic issues that emerged from the research findings suggest that some facilitation moves in the tutorials were based on behaviourist perspectives: (1) impulse towards material coverage; (2) greater number of lower-order questions; (3) presence of IRE discourse pattern; (4) presence of many knowledge display questions; (5) presence of asynchronous mini-lecture presentation among the students; and (6) knowledge transmission by the facilitators. The pedagogic issues need to be seen in the wider context of the facilitators' knowledge and beliefs, institutional culture, administrative control, and assessment system.

The pedagogic issues that emerged from the study are potentially important for PBL curriculum and its facilitation. These issues support the need to open 'the black box' of PBL curriculum to understand the actual activities and learning processes that mediate the relationship between the PBL curriculum and its cognitive outcomes. Such understanding can explain why curriculum implementation may not fulfil its promises and may demonstrate the need to align learning theory with facilitation methods. It also provides the basis for giving pertinent advice and recommendations about how to conduct PBL tutorials.

Fourth, based on the emerging pedagogic issues, pertinent pedagogic advice about how to conduct PBL tutorials and administrative suggestions on how to further develop the facilitative expertise of the facilitators was provided.

Finally, several research limitations associated with generalisability, irregular nature of participation, issues of consent, and the nature of linguistic study were highlighted. Many opportunities for further research were also described.

BIBLIOGRAPHY

Aarnio M, Lindblom-Ylanne S, Nieminen J, et al. (2013) Dealing with conflicts on knowledge in tutorial groups. *Advances in Health Science Education Theory & Practice* 18: 215 – 230.

Aarnio M, Lindblom-Ylanne S, Nieminen J, et al. (2014) How do tutors intervene when conflicts on knowledge arise in tutorial groups? *Advances in Health Science Education Theory & Practice* 19: 329 – 345.

Aarts J. (1991) Intuition-based and observation-based grammars. In: Aijmer K and Altenberg B (eds.) *English Corpus Linguistics: Studies in honour of Jan Svartvik.* London and New York: Longman, 44 – 62.

Adolphs S. (2006) *Introducing Electronic Text Analysis: A Practical guide for language and literary studies,* London and New York: Taylor and Francis Group.

Adolphs S and Knight D. (2012) Building a spoken corpus: What are the basics? In: McCarthy M and O'Keeffe A (eds.) *Routledge Handbook of Corpus Linguistics.* Abingdon, Oxon and New York: Routledge: Taylor and Francis Group, 39 – 52.

Ahmed Z. (2014) Problems of group dynamics in problem based learning sessions. *Journal of Ayub Medical College Abbottabad* 26: 230 – 234.

Al-Drees AA, Khalil MS, Irshad M, et al. (2015) Students' perception towards the problem based learning tutorial session in a system-based hybrid curriculum. *Saudi Medical Journal* 36: 341 – 348.

Al-Shawwa LA. (2005) Tutors' characteristics and their effects on efficacy and commitment to problem-based learning. INDIANA UNIVERSITY.

Albanese M. (2000) Problem-based learning: why curricula are likely to show little effect on knowledge and clinical skills. *Medical Education* 34: 729 – 738.

Albanese MA and Mitchell S. (1993) Problem-based Learning: A Review of Literature on Its Outcomes and Implementation Issues. *Academic Medicine* 68: 52 – 81.

Alexa M. (1997) *Computer Assisted Text Analysis Methodology in the Social Sciences,* Mannheim Germany: ZUMA.

Alexa M and Zuell C. (2000) Text analysis software: Commonalities, Differences and Limitations: The results of a review. *Quality and Quantity* 34: 299 – 321.

Alexa M and Züll C. (1999) *A review of software for text analysis,* Mannheim, Germany: Zentrum für Umfragen, Methoden und Analysen (ZUMA). Available at: http://www.zuma-mannheim.de/ ZUMA. Accessed: 27 November 2014.

Alexandra H. (2013) Computer-aided content analysis of digitally enabled movements. *Mobilization: An International Quaterly* 18: 367 – 388.

AlHaqwi AI. (2014) Learning outcomes and tutoring in problem based-learning: how do undergraduate medical students perceive them? *International Journal of Health Sciences* 8: 125 – 132.

Allen D, Duch B and Groh S. (1996) The power of problem-based learning in teaching introductory science courses. *New Directions for Teaching and Learning* 1996: 43 – 52.

Amidon A. (1976) Children's understanding of sentences with contingent relations: why are tempotral and conditional connectives so difficult? *Journal of Experimental Child Psychology* 22: 423 – 437.

Anthony L. (2005) AntConc: Design and development of a freeware corpus analysis toolkit for the technical writing classroom. *In IEEE International Professional Communication Conference Proceedings. Available at:* http://uibk.ac.at/tuxtrans/docs/antConc.pdf*. Accessed: 16 December, 2014.*

Anthony L. (2013) A critical look at software tools in corpus linguistics. *Linguistic Research* 30: 141 – 161.

Applefield J, Huber R and Moallem M. (2001) Constructivism in theory and practice: Toward a better understanding. *The High School Journal* 84: 35 – 53.

Archer D, Wilson A and Rayson P. (2002) Introduction to the USAS category system. Benedict project report. Available at:

http://ucrel.lancs.ac.uk/usas/usas%20guide.pdf. Accessed: 30 January, 2015.

Arksey H and O'Malley L. (2005) Scoping studies: Towards a methodological framework. *International journal of social research methodology* 8: 19 – 32.

Armstrong E. (1999) A hybrid model of problem-based learning In: Boud D and Feletti G (eds.) *The challenge of problem-based learning (p. 137 - 148).* 2nd ed. London: Taylor and Francis Ltd.

Azer S. (2001) Problem-based learning: A critical review of its educational objectives and the rationale for its use. *Saudi Medical Journal* 22: 299 – 305.

Azer S and Azer D. (2015) Group interactions in problem-based learning tutorials: a systematic review. *European Journal of Dental Education* 19: 194 – 208.

Azer SA. (2005) Challenges facing PBL tutors: 12 tips for successful group facilitation. *Medical Teacher* 27: 676 – 681.

Baddeley A. (1999) *Essentials of human memory,* Philadelphia: Psychology Press.

Baker M, Hansen T, Joiner R, et al. (1999) The role of grounding in collaborative learning tasks. In: Dillenbourg P (ed.) *Collaborative learning and computational approaches.* Oxford: Elsevier Sciences Ltd, 31 – 63.

Bakhtin M. (1986) *Speech genres and other essays,* Austin: University of Texas Press.

Bakhtin M, Emerson C and Holquist M. (1981) *The dialogic imagination: four essays by M.M. Bakhtin,* Austin: University of Texas Press.

Barnbrook G. (1996) *Language and computers: A practical introduction to the computer analysis of language,* Edinburgh: Edinburgh University Press.

Barrett T and Moore S. (2011a) An introduction to problem-based learning. In: Barrett T and Moore S (eds.) *New approaches to problem-based learning: Revitalising your practice in higher education.* New York and London: Routledge: Taylor and Francis Group, 3 – 17.

Barrett T and Moore S. (2011b) Students maximising the potential of the problem-based learning tutorials: Generating dialogic knowing In: Barrett T and Moore S (eds.) *New approaches to problem-based learning :revitalising your practice in higher education.* New York and Abingdon Oxon: Routledge.

Barrie-Blackley S. (1973) Six-year-old children's understanding of sentences adjoined with time adverbials. *Journal of Psycholinguistic Research* 2: 153 – 165.

Barrows HS. (1983) Problem-based, self-directed learning. *The Journal of the American Medical Association* 250: 3077 – 3080.

Barrows HS. (1985) *How to design a problem-based curriculum for the preclinical years,* New York: Springer Publishing Company.

Barrows HS. (1986) A taxonomy of problem-based learning methods. *Medical Education* 20: 481 – 486.

Barrows HS. (1988) *The Tutorial Process,* Southern Illinois University: Springfield

Barrows HS. (1992) *The Tutorial Process,* Southern Illionis University School of Medicine: Springfield.

Barrows HS. (1996) Problem-based learning in medicine and beyond: A brief overview. *New Directions for Teaching and Learning* 1996: 3 – 12.

Barrows HS and Tamblyn RM. (1980) *Problem-based learning: An approach to medical education,* New york: Springer Publishing Company.

Barth D. (2000) "that's true, although not really, but still".: Expressing concession in spoken English. In: Couper-Kuhlen E and Kortmann B (eds.) *Cause-condition-concession-contrast: Cognitive and discourse perspectives.* Berlin and New York: Walter de Gruyter, 411 – 437.

Bazerman C. (2004) Intertextuality: How texts rely on other texts. In: Bazerman C and Prior P (eds.) *What writing does and how it does it: An introduction to analyzing texts and textual practices.* New York: Routledge, 83 – 94.

Beach R and Anson C. (2004) Stance and intertextuality in written discourse. In: Shuart-Faris N and Bloome D (eds.) *Uses of* *intertextuality in classroom and educational research.* Connecticut: Information Age Publishing Inc, 251 – 277.

Benko A and Lanyi C. (2009) History of artificial intelligence *Encyclopedia of information science and technology.* IGI Global, 1759 – 1762.

Bereiter C. (1994) Implications of postmodernism for sciences, or, science as progressive discourse. *Educational Psychologist* 29: 3 – 12.

Berkson L. (1993) Problem-based learning: Have the expectations been met? *Academic Medicine* 68: S79 – S87.

Biber D. (2006) *University Language: A corpus-based study of spoken and written registers,* Amsterdam and Philadelphia: John Benjamins Publishing Company.

Biber D, Conrad S and Reppen R. (1998) *Corpus Linguistics: Investigating language structure and use,* Cambridge: Cambridge University Press.

Biggs J. (1996) Enhancing teaching through constructive alignment. *Higher Education* 32: 347 – 364.

Biggs J and Tang C. (2011) *Teaching For Quality Learning At University (Society for Research Into Higher Education)* Maidenhead Berkshire: Open University Press.

Bloom L, Lahey M, Hood L, et al. (1980) Complex sentences: aquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language* 7: 235 – 261.

Bloome D and Egan-Robertson A. (2004) The social construction of intertextuality in classroom reading and writing lessons. In: Shuart-Faris N and Bloome D (eds.) *Uses of Intertextuality in Classroom and Educational Research.* Connecticut: Information Age Publishing Inc, 17 – 64.

Boelens R, De Wever B, Rosseel Y, et al. (2015) What are the most important tasks of tutors during the tutorials in hybrid problem-based learning curricular? *BMC Medical Education. Available at:* http://www.biomedcentral.com*. 15:84 DOI 10.1186/s12909-015-0368-4. Accessed: 30 May, 2015*.

Bolden G. (2009) Implementing incipient actions: The discourse marker 'so' in English conversation. *Journal of Pragmatics* 41: 974 – 998.

Boud D and Feletti G. (1999) What is problem-based learning? In: Boud D and Feletti G (eds.) *The challenge of problem-based learning.* London: Taylor and Francis Group, 15 – 16

Bourke J. (2005) The grammar we teach. *Reflections on English Language Teaching* 4: 87 – 97.

Bowling A. (2009) *Research methods in health: iInvestigating health and health services,* Berkshire and New York: Open University Press.

Bransford JD and Johnson MK. (1972) Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of verbal learning and verbal behavior* 11: 717–726.

Braun V and Clarke V. (2006) Using thematic analysis in psychology. *Qualitative research in psychology* 3: 77–101.

Carter R and McCarthy M. (2006) *Cambridge Grammar of English: A comprehensive guide Spoken abd written English grammar and usage,* Cambridge Cambridge University Press

Chafe W. (1988) Linking intonation units in spoken English. In: Haiman J and Thompson S (eds.) *Clause combining in grammar and discourse.* Amsterdam and Philadlphia: John Benjamins Publishing Company, 1 – 27.

Charlin B, Tardif J and Boshuizen H. (2000) Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. *Academic Medicine* 75: 182 – 190.

Charmaz K. (2006) *Constructing grounded theory: A practical guide through qualitative analysis,* London, Thousand Oaks and New Delhi: Sage Publications.

Cheepen C. (2000) Small talk in service dialogues: the conversational aspects of transactional telephone talk. In: Coupland J (ed) *Small Talk (Language In Social Life)* New York: Routledge, 288 – 312.

Chi MTH, De Leeuw N, Chiu MH, et al. (1994) Eliciting self-explanations improves understanding. *Cognitive Science* 18: 439 – 477.

Chng E, Yew EH and Schmidt HG. (2011) Effects of tutor-related behaviours on the process of problem-based learning. *Advances in Health Science Education Theory & Practice* 16: 491– 503.

Chng E, Yew EH and Schmidt HG. (2015) To what extent do tutor-related behaviours influence student learning in PBL? *Advances in Health Science Education Theory & Practice* 20: 5 – 21.

Chomsky N. (1965) *Aspects of Theory of Syntax,* Cambridge and Massachusetts: The MIT Press.

Clancy P, Thompson S, Suzuki R, et al. (1996) The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of Pragmatics* 26: 355 – 387.

Clark E. (1971) On the aquisition of the meaning of before and after. *Journal of Verbal Learning and Verbal Behaviour* 10: 266 – 275.

Clark E and Schaefer E. (1987) Collaborating on contributions to conversations. *Language and Cognitive Processes* 2: 19 – 41.

Clark H and Schaefer E. (1989) Contributing to discourse. *Cognitive Scince* 13: 259 – 294.

Clark H and Wilkes-Gibbs D. (1986a) Referring as a collaborative process. *Cognition* 22: 1 – 39.

Clark HH and Brennan SE. (1991) Grounding in communication. *Perspectives on socially shared cognition* 13: 127 – 149.

Clark HH and Wilkes-Gibbs D. (1986b) Referring as a collaborative process. *Cognition* 22: 1 – 39.

Coates A, Carpenter B, Case C, et al. (2011) Text detection and character recognition in scene images with unsupervised feature learning. *In Intl. Conference on Document Analysis and Recognition (ICDAR), 2011*. Available at: http:www.cs.stanford.edu/acoates/papers/coatesetal_icdar_2011.pdf. Accessed: 21 April, 2015.

Cobb P. (1994) Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher* 23: 13 – 20.

Cobb P. (2002) Theories of knowledge and instructional design: a response to Colliver. *Teaching and Learning in Medicine* 14: 52 – 55.

Cobb P and Yackel E. (1996) Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist* 31: 175 – 190.

Cockburn A and Dale T. (1997) CEVA: A tool for collaborative video analysis *Proceedings of GROUP 97 35-42.* New York, NY: ACM.

Cohen D and Crabtree B. (2006) Qualitative research guidelines project. Available at: http://www.qualres.org/pdf/. Accessed: 9 August 2014.

Cohen L, Manion L and Morrison K. (2011) *Research Methods in Education,* London: Routledge Taylor and Francis Group.

Collard A, Gelaes S, Vanbelle S, et al. (2009a) Reasoning versus knowledge retention and ascertainment throughout a problem-based learning curriculum. *Medical Education* 43: 854 – 865.

Collard A, Gelaes S, Vanbelle S, et al. (2009b) Reasoning versus knowledge retention and ascertainment throughout a problem based learning curriculum. *Medical Education* 43: 854 – 865.

Collier L. (2011) The need for Teacher Communities: An interview with Linda Darling-Hammond. *The Council Chronicle*.

Colliver J. (2002) Constructivism: The view of knowledge that ended philosophy or a theory of learning and instruction? *Teaching and Learning in Medicine* 14: 49 – 51.

Colliver JA. (2000) Effectiveness of problem-based learning curricula: research and theory. *Academic Medicine* 75: 259 – 266.

Conrad S and Biber D. (2000) Adverbial marking of stance in speech and wrtiting. In: Hunston S and Thompson G (eds.) *Evaluation in text: Authorial stance and the constructio of discourse.* Oxford & New York: Oxford University Press, 56 – 73.

Cotton K. (2001) Classroom questioning. *School improvement research series. Available at:https://*www.learner.org/workshops/socialstudies/pdf/.../6.ClassroomQuestioning.pdf*. Accessed: 10 July, 2016.*

Couto LB, Bestetti RB, Restini CB, et al. (2015) Brazilian medical students' perceptions of expert versus non-expert facilitators in a (non) problem-based learning environment. *Medical Education Online* 20: 268 – 293.

Creswell J. (2003) *Research Design: Qualitative, Quantitative and Mixed Method Approaches,* Thousand Oaks, CA: Sage Publications Ltd.

Creswell JW. (2009) *Research Design: Qualitative, Quantitative, and mixed methods approaches,* London: Sage Publications Ltd.

Creswell JW. (2012) *Educational Research: Planning, conducting, and evaluating quantitative and qualitative research,* Boston MA: Pearson Education Inc.

Creswell JW. (2013) *Qualitative inquiry and research design: Choosing among five approaches,* London: Sage Publications Ltd.

Creswell JW and Clark VLP. (2011) *Designing and conducting mixed methods research,* London: Sage Publications Ltd.

Curet MJ and Mennin SP. (2003) The effect of longterm vs shortterm tutors on the quality of the tutorial process and student performance. *Advances in Health Science Education Theory & Practice* 8: 117 – 126.

Da Silva A. (2013) Clinical reasoning development in medical students: An educational and transcultural comparative study. *School of Medicine.* UK: University of Nottingham.

Da Silva AL and Dennick R. (2010) Corpus analysis of problem ‐based learning transcripts: an exploratory study. *Medical Education* 44: 280 – 288.

Dagher Z. (1992) Verbal Explanations Given by Science Teacher: Their nature and implications. *Journal of Research in Science Teaching* 29: 361 – 374.

Das M, Mpofu D, Hasan MY, et al. (2002) Student perceptions of tutor skills in problem ‐based learning tutorials. *Medical Education* 36: 272 ‐ 278.

Davis WK, Nairn R, Paine ME, et al. (1992) Effects of expert and non-expert facilitators on the small-group process and on student performance. *Academic Medicine* 67: 470 – 474.

Davis WK, Oh MS, Anderson RM, et al. (1994) Influence of a highly focused case on the effect of small-group facilitators' content expertise on students' learning and satisfaction. *Academic Medicine* 69: 663 – 669.

De Grave W, Boshuizen H and Schmidt H. (1996) Problem-based learning: Cognitive and metacognitive processes during problem analysis. *Instructional Science* 24: 321 – 341.

De Grave W, Boshuizen H and Schmidt HG. (2001) Effects of problem-based discussion on studying a subsequent text: A randomised trial among first year medical students. *Instructional Science* 29: 33 – 44.

De Grave WS, Dolmans DHJM and Van Der Vleuten CPM. (1999) Profiles of effective tutors in problem-based learning: scaffolding student learning. *Medical Education* 33: 901 – 906.

De Leng B and Gijlers H. (2015) Collaborative diagramming during problem based learning in medical education: Do computerized diagrams support basic science knowledge construction? *Medical Teacher* 37: 450 – 456.

Demmen J, Semino E, Demjen Z, et al. (2015) A computer-assisted study of the use of violence metaphors for cancer and end of life by patients, family careers and health professionals. *International Journal of Corpus Linguistics* 20: 205 – 231.

Dennick R. (2015) Theories of learning: constructive experience. In: Matheson D (ed) *An introduction to the study of education.* 4th ed. London and New York: Routledge: Taylor and Francis Group, 36 – 63.

Dennick R. (2016) Constructivism: reflections on twenty five years teaching the constructivist approach in medical education. *International Journal of Medical Education* 7: 200 – 204.

Dennis J. (2003) Problem-based learning in online vs. face-to-face environments. *Education for Health* 16: 198 – 209.

Dickson R, Cherry M and Boland A. (2013) Carrying out a systematic review as a master's thesis. In: Boland A, Cherry G and Dickson

R (eds) *Doing a systematic review.* London,California, New Delhi and Singapore: Sage Publications Ltd, 1 - 16.

Diemers AD, Wiel MW, Scherpbier AJ, et al. (2015) Diagnostic reasoning and underlying knowledge of students with preclinical patient contacts in PBL. *Medical Education* 49: 1229 – 1238.

Dillenbourg P. (1999) What do you mean by 'collaborative learning'? In: Dillenbourg P (ed) *Collaborative-learning: Cognitive and Computational Approaches.* Oxford: Elsevier, 1 – 19.

Dochy F, Gijbels D, Raes E, et al. (2014) Team learning in education and professional organisations. In: Billett S, Harteis C and Gruber H (eds.) *International handbook of research in professional and practice-based learning.* New York and London: Springer, 987 – 1020.

Dolmans DH, Gijselaers WH, Moust JH, et al. (2002) Trends in research on the tutor in problem-based learning: conclusions and implications for educational practice and research. *Medical Teacher* 24: 173 – 180.

Dolmans DH and Ginns P. (2005) A short questionnaire to evaluate the effectiveness of tutors in PBL: validity and reliability. *Medical Teacher* 27: 534 – 538.

Dolmans DH and Schmidt HG. (2006) What do we know about cognitive and motivational effects of small group tutorials in problem-based learning? *Advances in Health Sciences Education* 11: 321 – 336.

Dolmans DH and Wolfhagen IH. (2005) Complex interactions between tutor performance, tutorial group productivity and the effectiveness of PBL units as perceived by students. *Advances in Health Science Education Theory & Practice* 10: 253 - 261.

Dolmans DH, Wolfhagen IH and Schmidt HG. (1996) Effects of tutor expertise on student performance in relation to prior knowledge and level of curricular structure. *Academic Medicine* 71: 1008 – 1011.

Dolmans DH, Wolfhagen IH, Schmidt HG, et al. (1994) A rating scale for tutor evaluation in a problem-based curriculum: validity and reliability. *Medical Education* 28: 550 – 558.

Dolmans DHJM, Wolfhagen IHAP, Van Der Vleuten CPM, et al. (2001) Solving problems with group work in problem ‐based learning: hold on to the philosophy. *Medical Education* 35: 884 ‐ 889.

Donner R and Bickley H. (1993) Problem-based learning in American medical education: an overview. *Bulletin Medical Library Association* 81: 294 – 298.

Dornyei Z. (2007) *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies,* Oxford: Oxford University Press.

Duch B. (2001) Models of problem-based instruction in undergraduate courses. In: Duch B, Groh S and Allen D (eds.) *The power of problem-based learning: A practical "How To" for teaching undergraduate courses in any discipline.* Virginia: Stylus Publishing, LLC, 39 – 46.

Duek J. (2000) Whose group is it, anyway? Equity of student discourse in problem-based learning (PBL). In: Evensen DH and Hmelo C (eds.) *Problem-based learning: A research perspective on learning interactions.* New Jersey and London: Lawrence Erlbaum Associates Publishers, 75 – 107.

Duit R and Treagust D. (1998) Learning in science - from behaviorism towards social constructivism and beyond. In: Fraser B and Tobin K (eds.) *International handbook of science education* London: Kluwer, 3 – 25.

Eagle CJ, Harasym PH and Mandin H. (1992) Effects of tutors with case expertise on problem-based learning issues. *Academic Medicine* 67: 465 – 469.

Eggins S. (2005) *Introduction to systemic functional linguistics,* London & New York: Continuum Internation Publishing Ltd.

Eggins S and Slade D. (1997) *Analysing Casual conversation,* London: Equinox.

Ehri L and Galanis A. (1980) Teaching children to comprehend propositions conjoined by "Before" and "After". *Journal of Experimental Child Psychology* 30: 308 – 324.

Elmore R and Burney D. (1997) Investing in Teacher: Staff development and instructional improvement in community school

District, New York City. National Commission on Teaching and America's Future, New York, NY and Consortium for Policy Research in Education, Philadelphia, PA: Office of Educational Research and Improvement (ED), Washington, DC, 4 – 61.

Elo S and Kyngäs H. (2008) The qualitative content analysis process *Journal of Advanced Nursing* 62: 107–115.

Emerson H and Gekoski W. (1980) Development of comprehension of sentences with "Because" or "If" *Journal of Experimental Child Psychology* 29: 202 – 224.

Engel CE. (1999) Not just a method but a way of learning. In: Boud D and Feletti G (eds) *The challenge of problem-based learning.* London: Routledge, 21 – 31.

Evensen DH and Hmelo CE. (2000) Introduction. In: Evensen DH and Hmelo CE (eds.) *Problem-based learning: A Research Perspective on Learning Interactions.* Mahwah NJ: Lawrence Erlbaum Associates, Inc.

Evison J. (2012) What are the basics of analysing a corpus? In: O'Keeffe A and McCarthy M (eds.) *The Routledge Handbook of Corpus Linguistics.* Abingdon, Oxon and New York: Routledge: Taylor and Francis Group, 122 – 135.

Exley K and Dennick R. (2004) *Small group teaching: Tutorials, Seminars and beyond:key guides for effective teaching in higher education,* London and New York: Routledge Falmer.

Faidley J, Evensen D, Salisbury-Glennon J, et al. (2000) How are we doing? Methods of assessing group processing in a problem-based learning context. In: Evensen D and Hmelo C (eds.) *Problem-based learning: A Research Perspective on Learning Interactions.* New Jersey and London: Lawrence Erlbaum Associates, Publishers, 109 – 135.

Fairclough N. (2003) *Analysing discourse: Textual analysis for social research,* London and New York: Taylor and Francis Group.

Fan J. (2014) Thinking styles’ socialization and their roles in student development. *Faculty of Education.* Hong Kong: The University of Hong Kong.

Feagans L. (1980) Children's understanding of some temporal terms denoting order, duration, and simultaneity. *Journal of Psycholinguistic Research* 9: 41 – 55.

Fischer F, Bruhn J, Gräsel C, et al. (2002) Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction* 12: 213 – 232.

Fischer F and Mandl H. (2005) Knowledge convergence in computer-supported collaborative learning: The role of external representational tools. *Journal of the Learning Sciences* 14: 405 – 441.

Flowerdew J. (1993) Concordancing as a tool in course design. *System* 21: 231 – 244.

Fontaine L. (2013) *Analysing English Grammar: A systemic Functional Introduction,* Cambridge: University Press.

Fosnot C. (1989) *Enquiring teachers, enquiring learners: A constructivist approach for teaching,* New York: Teachers College Press.

Fosnot C. (1993) Rethinking science education: A defence of Piagetian constructivism. *Journal of Research in Science Teaching* 30: 1189 – 1201.

Fosnot C and Perry R. (1996) Constructivism: A psychological theory of learning. In: Fosnot C (ed.) *Constructivism: Theory, perspectives, and practice.* New York: Teachers College Press, 8 – 33.

Fraser B. (1998) The birth of a new journal: Editor's introduction. *Learning Environments Research* 1: 1 – 5.

Frederiksen CH. (1999) Learning to reason through discourse in a problem ‐based learning group. *Discourse Processes* 27: 135 ‐ 160.

French L. (1988) The development of children's understanding of "Because" and "So". *Journal of Experimental Child Psychology* 45: 262 – 279.

Frey L, Botan C and Kreps G. (1990) *Investigating communication: An introduction to research methods,* Boston, London, Toronto, Sydney, Tokyo and Singapore: Allyn and Bacon.

Gabrielatos C and Baker P. (2008) Fleeing, Sneaking, Flooding A Corpus Analysis of Discursive Constructions of Refugees and Asylum Seekers in the UK Press, 1996-2005. *Journal of English Linguistics* 36: 5 – 38.

Gardner R. (2007) The right connections: Acknowledging epistemic progression in talk. *Language in Society* 36: 319 – 341.

Garside R. (1987) The CLAWS word-tagging system. In: Garside R, Leech G and Sampson G (eds.) *The computational Analysis of English: a corpus-based approach.* Harlow, Essex: Longman Group UK Limited, 30 – 41.

Gee JP. (1994) Discourses: Reflections on M.A.K.Halliday's "Toward a Language-Based Theory of Learning". *Linguistics and Education* 6: 33 – 40.

Geelan D. (1997) Epistemological anarchy and the many forms of constructivism. *Science & Education* 6: 15 – 28.

Gentleman R, Huber W and Carey V. (2008) *Supervised machine learning,* New York: Springer.

Gijbels D, van Den Bossche P and Loyens S. (2013) Problem-based learning. In: Hattie J and Anderman E (eds.) *International Guide to Student Achievement.* New York and London: Routledge, 382 – 384.

Gijselaers WH. (1996) Connecting problem ‐based practices with educational theory. *New Directions for Teaching and Learning* 1996: 13 ‐ 21.

Gijselaers WH. (1997) Effects of contextual factors on tutor behavior. *Teaching and Learning in Medicine* 9: 116-124.

Gilkison A. (2003) Techniques used by "expert" and "non-expert" tutors to facilitate problem-based learning tutorials in an undergraduate medical curriculum. *Medical Education* 37: 6 – 14.

Glaser R. (1985) The Nature of Expertise. Occasional Paper No. 107. National center for research in vocational education: Ohio state University, 3 – 26.

Glasersfeld E. (2000) Piaget’s legacy: Cognition as adaptive activity. *Understanding representation in the cognitive sciences*: 283 – 287.

GMC. (1993) Tomorrow's Doctors: Recommendations on Undergraduate medical Education. London: General Medical Council Education Committee.

Gould J. (2012) *Learning Theory and Classroom Practice in the Lifelong Learning Sector,* London, California, New Delhi and Singapore: Sage Publications Ltd.

Greene J, Caracelli V and Graham W. (1989) Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis* 11: 255 – 274.

Greene J and Curucelli V. (1994) Defining and describing the paradigm issue in mixed-method evaluation. *New Directions for Evaluation* Summer: 5 - 17.

Greeno J. (1997) On claims that the answer wrong questions. *Educational Researcher* 26: 5 - 17.

Gries S. (2009) What is Corpus Linguistics? *Language and Linguistics Compass* 3: 1–17.

Groves M, Rego P and O'Rourke P. (2005) Tutoring in problem-based learning medical curricula: the influence of tutor background and style on effectiveness. *BMC Medical Education* 5: 20. Available at: http://www.biomedcentral.com/1472-6920/5/20. Accessed: 12 November, 2015.

Guba E and Lincoln Y. (1994) Competing paradigms in qualitative research. In: Denzin N and Lincoln Y (eds.) *Handbook of Qualitative Research.* Thousand Oaks, CA: Sage Publications, 105 – 117.

Gukas ID, Leinster SJ and Walker R. (2010) Verbal and nonverbal indices of learning during problem-based learning (PBL) among first year medical students and the threshold for tutor intervention. *Medical Teacher* 32: 5 – 11.

Guskey T. (2002) Professional development and Teacher change. *Teachers and Teaching: Theory and Practice* 8: 381 – 391.

Hak T and Maguire P. (2000) Group process: The black box of studies on problem-based learning. *Academic Medicine* 75: 769 – 772.

Halela M and Fagerholm H. (2008) *Tracing the roles of the PBL tutor: A journey of learning*: HAAGA-HELIA University of Applied Sciences. Available at: http://www.haaga-helia.fi/sites/default/files/Kuvat-ja-liitteet/Palvelut/Julkaisut/pbl.pdf. Accessed: 10 March, 2016.

Halliday M. (1975) *Learning how to mean: Explorations in the development of language,* London, Caulfield East and Maryland: Edward Arnold Ltd.

Halliday M. (1978) *Language as social semiotic: The social interpretation of language and meaning,* Oxon: Hodder Arnold.

Halliday M. (1991a) Corpus studies and probabilistic grammar. In: Aijmer K and Altenberg B (eds.) *English Corpus Linguistics: Studies in honour of Jan Svartvik.* London and New York: Longman, 30 – 43.

Halliday M. (1991b) Language as system and language as instance: The corpus as a theoretical construct. In: Startvik J (ed.) *Nobel Symposium 82.* Stockholm, Sweden: Walter de Gruyter & Co., 61 – 78.

Halliday M. (2002) *Linguistic studies of text and discourse,* London and New York: Continuum.

Halliday M and Hassan R. (1976) *Cohesion in English,* Oxon & New York: Taylor & Francis Group.

Halliday MAK. (1993) Towards a Language-Based Theory of Learning. *Linguistics and Education* 5: 93 – 116.

Halliday MAK. (2006) Afterwords. In: Thompson G and Hunston S (eds.) *System and Corpus: Exploring Connections.* London: Equinox Publishing Ltd, 293 – 299.

Halliday MAK and Matthiessen C. (2004) *An Introduction to functional Grammar,* London: Hodder Education.

Handa B and Herrington A. (2003) Mathematics teachers' beliefs and curriculum reform. *Mathematics Education Research Journal* 15: 59 – 69.

Hanson W, Creswell J, Clark V, et al. (2005) Mixed methods research designs in counseling psychology. *Journal of Counseling Psychology* 52: 224 – 235.

Harden RM and Davis MH. (1998) The continuum of problem-based learning. *Medical Teacher* 20: 317 – 322.

Hardie A. (2012) CQPweb - combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics* 17: 380 – 409.

Hassan R. (2002) Semiotic mediation, language and society: Three exotripic theories - Vygotsky, Halliday and Bernstein. *Second International Basil Bernstein Symposium: Knowledges, pedagogy and society.* Cape Town.

Hendry GD, Ryan G and Harris J. (2003) Group problems in problem-based learning. *Medical Teacher* 25: 609 – 616.

Heritage J. (1984) A change of state token and aspects of its sequential placement. In: Atkinson J and Heritage J (eds.) *Structures of social action.* Cambridge, UK: Cambridge University Press, 299 – 345.

Hewitt J and Scardamalia M. (1998) Design principles for distributed Knowledge building processes. *Educational Psychology Review* 10: 75 – 96.

Hmelo-Silver C and Eberbach C. (2012a) Learning theories and problem-based learning. In: Bridges S, McGrath C and Whitehill T (eds.) *Problem-based learning in clinical education: The next generation.* London and New York: Springer, 3 – 17.

Hmelo-Silver CE. (2002) Collaborative ways of knowing: Issues in facilitation. International Society of the Learning Sciences, 199 – 208.

Hmelo-Silver CE. (2003a) Analyzing collaborative knowledge construction: multiple methods for integrated understanding. *Computers and Education* 41: 397 – 420.

Hmelo-Silver CE. (2003b) Facilitating collaborative knowledge construction. *System Sciences (HICSS '03), 36th Annual International Conference.* Hawaii IEEE Computer Society Washington, DC.

Hmelo-Silver CE. (2004) Problem-based learning: What and how do students learn? *Educational Psychology Review* 16: 235 – 266.

Hmelo-Silver CE and Barrows HS. (2006) Goals and strategies of a problem-based learning facilitator. *Interdisciplinary Journal of Problem-based Learning* 1: 4. Available at: http://dx.doi.org/10.7771/1541-5015.1004. Accessed: 10 February, 2016.

Hmelo-Silver CE and Barrows HS. (2008) Facilitating collaborative knowledge building. *Cognition and Instruction* 26: 48 – 94.

Hmelo-silver CE and Eberbach C. (2012b) Introduction. In: Bridges S, McGrath C and Whitehill TL (eds.) *Problem-based learning in clinical education: The nexr generation.* Springer.com: Springer.

Holtz M. (2011) Lexico-grammatical properties of abstracts and research articles. PhD Thesis. Technische Universität Darmstadt. Systemic Functional Linguistics Online Theses: http://www.isfla.org/Systemics/Print/Theses.html. Accessed: 6 November, 2015.

Houlden R, Collier C, Frid P, et al. (2001) Problems identified by tutors in a hybrid problem-based learning curriculum. *Academic Medicine* 76: 81.

Howe C. (2009) Expert support for group work in elementary science: The role of consensus. In: Schwarz B, Dreyfus T and Hershkowitz N (eds.) *Transformation of knowledge through classroom interaction.* London and New York: Routledge, 93 – 103.

Howe K. (1988) Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher* 17: 10 – 16.

Howley I, Mayfield E and Rosé C. (2013) Linguistic Analysis Methods for Studying Small Groups. In: Hmelo C, Chinn C, Chan C, et al. (eds.) *The international handbook of collaborative learning.* New York & Oxon: Routledge.

Hsu LL. (2007) Conducting clinical post ‐conference in clinical teaching: a qualitative study. *Journal of Clinical Nursing* 16: 1525 ‐ 1533.

Huitt W and Hummel J. (2003) Piaget's theory of cognitive development. *Educational psychology interactive* 3. Available at: http://www.edpsycinteractive.org/topics/cognition/piaget.html. Accessed: 21 July, 2015.

Hung W, Jonassen D and Liu R. (2008) Problem-based learning. In: Jonassen D, Spector M, Driscoll M, et al. (eds.) *Handbook of Research in Educational Communications and Technology.* 3rd ed. New York: Lawrence Erlbaum Associates 485 – 506.

Hurk MMvd, Dolmans DHJM, Wolfhagen IHAP, et al. (2001) Quality of student-generated learning issues in a problem-based curriculum. *Medical Teacher* 23: 567 – 571.

Hutchinson SA. (1986) Education and grounded theory. *Journal of Thought*: 50 – 68.

Imafuku R. (2007) A case study of a medical PBL tutorial: Tutor and student participation. In: Marriott H, Moore T and Spencer-Brown R (eds.) *Learning discourses and the discourse of learning.* Melbourne, VC:Monash: University ePress, 111 – 117.

Imafuku R. (2012) Japanese first-year PBL students' learning processes: A classroom discourse analysis. In: Bridges S, McGrath C and Whitehill T (eds.) *Problem-Based Learning in Clinical Education: The next Generation (Innovation and Change in Professional Education). Dordrecht, Heidelberg, London and New York. Springer,* 153 – 170.

Imafuku R, Kataoka R, Mayahara M, et al. (2014) Students' experiences in interdisciplinary problem-based learning: A discourse analysis of group interaction. *The Interdisciplinary Journal of Problem-based Learning* 8: 1 – 18.

Jasinskaja K. (2013) Corrective elaboration. *Lingua* 132: 51 – 66.

Jennings P. (2013) PBL but not as we know it: an ethnography of the practice and facilitation of 'problem-based learning' within a hybrid graduate-entry medical programme in England. *School of Medicine.* Available at: http://eprints.nottingham.ac.uk/14438/1/602367.pdf. Accessed: 12 July, 2016. : University of Nottingham.

Jin J. (2014) Understanding silence in problem-based learning: a case study at an English medium university in Asia. *Clinical Linguistics & Phonetics* 28: 72-82.

John-Steiner V and Meehan T. (2000) Creativity and collaboration in knowledge construction. In: Lee C and Smagorinsky P (eds.) *Vygotskian perspectives on literacy research: constructing meaning through collaborative inquiry.* Cambridge. Cambridge University Press, 31 – 48.

Johnson H and Chapman R. (1980) Children's judgement of and recall of causal connectives: A developmental study of "Because," "So," and "And". *Journal of Psycholinguistic Research* 9: 243 – 259.

Johnson R and Onwuegbuzie A. (2004) Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher* 33 14 – 26.

Johnson R, Onwuegbuzie A and Turner L. (2007) Toward a Definition of Mixed Methods Research. *Toward a Definition of Mixed Methods Research* 1: 112 – 133.

Kalyuga S. (2009) Knowledge elaboration: A cognitive load perspective. *Learning and Instruction* 19: 402 – 410.

Kamin CS, O'Sullivan PS, Younger M, et al. (2001) Measuring critical thinking in problem-based learning discourse. *Teaching and Learning in Medicine* 13: 27 – 35.

Karen C, Penny M and Hansen B. (1998) The many faces of problem-based learning: A framework for understanding and comparison. *Medical Teacher* 20: 323 – 330.

Kassab S, Abu-Hijleh MF, Al-Shboul Q, et al. (2005) Student-led tutorials in problem-based learning: educational outcomes and students' perceptions. *Medical Teacher* 27: 521 – 526.

Katz E and Brent S. (1968) Understanding connectives. *Journal of Verbal Learning and Verbal Behaviour* 7: 501 – 509.

Kaufman DM and Holmes DB. (1998) The relationship of tutors' content expertise to interventions and perceptions in a PBL medical curriculum. *Medical Education* 32: 255 – 261.

Kennedy G. (1998) *An introduction to Corpus Linguistics,* London and New York: Longman.

Khalid K, Kunz R, Kleijnen J, et al. (2011) *Systematic Reviews to Support Evidence-based Medicine,* London: CRC Press.

Kindler P, Grant C, Kulla S, et al. (2009) Difficult incidents and tutor interventions in problem-based learning tutorials. *Medical Education* 43: 866 – 873.

Kingsbury M and Lymn J. (2008) Problem-based learning and larger student groups: Mutually exclusive or compatible concepts - a pilot study. *BMC Medical Education. Available at:* http://www.biomedcentral.com/1472-6920/8/35*. Accessed: 13 March, 2015*.

Koester A. (2012) Building small specialised corpora. In: McCarthy M and O'Keeffe A (eds.) *Routledge Handbook of Corpus Linguistics.* Abingdon, Oxon and New York: Routledge: Taylor and Francis Group, 66 – 79.

Kok J, Boers E, Kosters W, et al. (2009) ARTIFICIAL INTELLIGENCE - Artificial intelligence: Definitions, trends, techniques and cases. *Encyclopedia of life support systems (EOLSS).* Available at: http://www.eolss.net/sample-chapters/c15/e6-44.pdf. Accessed: 12 December, 2015.

Kolmos A and Graaff E. (2007) History of problem-based and project-based learning. In: Kolmos A and Graaff E (eds.) *Management of change: Implementation of problem-based and project-based learning in engineering.* Rotterdam The Netherlands: Sense Publishers, 1 – 8.

Kondracki N, Wellma N and Amundson D. (2002) Content analysis: Review of methods and their applications in nutrition education. *Journal of Nursing Education Behaviour* 34: 224 – 230.

Koschmann T. (1999) Toward a dialogic theory of learning: Bakhtin's contribution to understanding learning in settings of collaboration. *Proceedings of the 1999 conference on Computer support for collaborative learning.* International Society of the Learning Sciences, 308 – 314.

Koschmann T and Evensen DH. (2000) Five readings of a single text: Transcript of a video analysis session. *Problem-based learning: A research perspective on learning interactions*: 137 – 166.

Koschmann T, Glenn P and Conlee M. (1997) Analysing the emergence of a learning issue in a problem-based learning meeting. *Medical Education Online.Available at:* http://med-ed-online.nrt/index.php/meo/article/view/4290*. 10 Accessed June, 2016.*

Koschmann T, Glenn P and Conlee M. (2000) When is a problem-based tutorial not tutorial? Analyzing the tutor's role in the emergence of a learning issue. In: Evensen D and Hmelo C (eds.) *Problem-based Learning: A Research Perspective on Learning Interactions.* New Jersey and London: Lawrence Erlbaum Associates, Publishers, 53 – 74.

Koschmann T and MacWhinney B. (2001) Opening the black box: why we need a PBL talkbank database. *Teaching and Learning in Medicine* 13: 145 – 147.

Kotsiantis S. (2007) Supervised machine learning: A review of classification techniques. *Informatica* 31: 249 – 268.

Krippendorff K. (1989) Content analysis. In: Barnouw E, Gerbner G, Schramm W, et al. (eds.) *International encyclopedia of communication.* New York: Oxford University Press. Available at:http://repository.upenn.edu/asc_papers/226. Accessed: 17 September, 2014.

Krippendorff K. (2004) *Content analysis: An Introduction to its Methodology* Thousand Oaks, London and New Delhi: Sage Publications Inc.

Krippendorff K. (2013) *Content Analysis: An Introduction to its Methodology,* London: Sage Publications Ltd.

Kwan C. (202) Problem-based learning and the teaching of Pharmacology. *Naunyn-Schmiedeberg's Archives of Pharmacology* 366: 10 – 17.

Lee G-H, Lin Y-H, Tsou K-I, et al. (2009) When a problem-based learning tutor decides to intervene. *Academic Medicine* 84: 1406 – 1411.

Lee W, Wong F and Mok E. (2004) Problem-based learning: ancient Chinese educational philosophy reflected in a modern educational methodology. *Nurse Education Today* 24: 136 – 144.

Leech G. (1987) General Introduction. In: Garside R, Leech G and Sampson G (eds.) *The Computational Analysis of English: A Corpus-based Approach.* New York: Longman Inc., 1 – 15.

Leech G. (1991) The state of the art in corpus linguistics. In: Aijmer K and Altenberg B (eds.) *English Corpus Linguistics: Studies in honour of Jan Svartvik.* London and New York: Longman, 8 – 29.

Leech G. (1992a) 100 million words of English: The British National Corpus. *Language Research* 28: 1 – 13.

Leech G. (1992b) Corpora and theories of linguistic performance. In: Svartvik J (ed.) *Trends in Linguistics Studies and Monographs 65: Directions in Corpus Linguistics Proceedings of Nobel Symposium 82.* Stockholm Sweden: Mouton de Gruyter, 105 – 122.

Leech G. (1992c) Trends in Linguistics: Studies and Monographs 65. In: Svartvik J (ed.) *Directions in Corpus Linguistics: Nobel Symposium 82.* Stockholm: Mouton De Gruyter, 105 –122.

Leech G. (2000) Grammars of spoken English: New outcomes of corpus-oriented research. *Language Learning* 50: 675 – 724.

Levon E. (2010) Organising and processing your data: The nuts and bolts of quantitative analysis. In: Litosseliti L (ed.) *Research Methods in Linguistics.* London, New Delhi, New York and Sydney: Bloomsbury Academics, 68 – 92.

Lieberman A and Miller L. (1992) *Teachers - Their World and Their Work: Implications for School Improvement* Washington DC: Teachers College Press.

Lincoln Y. (2009) Ethical practices in qualitative research. In: Mertens D and Ginsberg P (eds.) *The Handbook of Social Research Ethics.* Thousands Oaks, London, New Delhi and Singapore: Sge Publications Ltd, 150 – 169.

Lincoln YS and Guba EG. (1985) *Naturalistic inquiry*: Sage Publications, Inc.

Lincoln YS, Lynham SA and Guba EG. (2011) Paradigmatic controversies, contradictions, and emerging confluences, revisited. In: Denzin NK and Lincoln YS (eds.) *The Sage Handbook of Qualitative Research.* London: Sage Publications Ltd, 97 – 128.

Lindquist H. (2009) *Corpus Linguistics and the Description of English,* Edinburgh: Edinburgh University Press.

Liu C and Matthews R. (2005) Vygotsky's philosophy: Constructivism and its criticisms examined. *International Education Journal* 6: 386 – 399.

Lloyd-Jones G, Margetson D and Bligh J. (1998) Problem-based learning: a coat of many colours. *Medical Education* 32: 492 – 494.

Lobb D, Inman D and Butler R. (2004) Problem-based learning in reproductive physiology. *Journal of Midwifery and Womens Health* 49: 449 – 453.

Lustigman L and Berman R. (2016) Form and function in early clause-combining. *Journal of Child Language* 43: 157 – 185.

Lycke KH. (2002) Inside PBL groups: observation, confirmations and challenges. *Education for Health* 15: 326 – 334.

Maclellan E and Soden R. (2004) The importance of epistemic cognition in student-centred learning. *Instructional Science* 32: 253 – 268.

Macnamara J. (2005) Media content analysis: Its uses, benefits and best practice methodology. *Asian Pacific Public Relations Journal* 6: 1 – 34.

Margetson D. (2003) The elements of problem-based learning. *International Journal of Medicine* 5: 177 – 181.

Marlowe B and Page M. (2005) *Creating and sustaining the constructivist classroom,* California, London and New Delhi: Cowin Press: Sage Publications Ltd.

Martin JR and White PRR. (2007) Engagement and Graduation: Alignment, solidarity and the construed reader. *The Language of Evaluation: Appraisal in English.* Basingstoke, Hampshire: Palgrave Macmillan.

Martin ND, Nguyen K and McDaniel MA. (2016) Structure building differences influence learning from educational text: Effects on encoding, retention, and metacognitive control. *Contemporary Educational Psychology* 46: 52 – 60.

Matthews P. (1997) Problems with Piagetian constructivism. *Science & Education* 6: 105 – 119.

Matthiessen C. (1999) The system of transitivity: An exploratory study of test-based profiles. *Functions of Language* 6: 1 – 51.

Matthiessen C. (2006) Frequency profiles of some basic grammatical systems: an interim report. In: Thompson G and Hunston S (eds.) *System and Corpus: Exploring connections.* London and Oakville: Equinox, 103 – 142.

Maudsley G. (1999) Roles and responsibilities of the problem based learning tutor in the undergraduate medical curriculum. *British Medical Journal* 318: 7184.

Maudsley G, Williams EM and Taylor DC. (2008) Problem-based learning at the receiving end: a 'mixed methods' study of junior medical students' perspectives. *Advances in Health Science Education Theory & Practice* 13: 435 – 451.

Mayer R. (1989) Models for Understanding. *Review of Educational Research* 59: 43 – 64.

McCabe A and Peterson C. (1985) A naturalistic study of the production of causal connectives by children. *Journal of Child Language* 12: 145 – 159.

McCabe A and Peterson C. (1988) A comparison of adult's versus children's spontaneous use of because and so. *Journal of Genetic Psychology* 149: 257 – 268.

McCarthy M. (2003) Talking Back: “Small” Interactional Response Tokens in Everyday Conversation. *Research on Language and Social Interaction* 36: 33 – 63.

McCorduck P, Minsky M, Selfridge O, et al. (1977) *History of artificial intelligence.* Morgan Kaufmann Publishers Inc.

McEnery T and Gabrielatos C. (2006) English corpus linguistics. In: Aarts and McMahon A (eds.) *A handbook of English Linguistics.* Malden, Oxford and Carlton: Blackwell Publishing Ltd, 33 – 71.

McEnery T and Hardie A. (2012) *Corpus Linguistics: Theory, Method and Practice,* Cambridge: Cambridge University Press.

McEnery T and Wilson A. (2001) *Corpus Linguistics: An Introduction,* Edinburgh: Edinburgh University Press.

McEnery T, Xiao R and Tono Y. (2006) *Corpus-based Language Studies: An advanced resource book,* London: Routledge Taylor and Francis Group.

McMillan WJ. (2010) Teaching for clinical reasoning–helping students make the conceptual links. *Medical Teacher* 32: 436 – 442.

Mercer N. (1995) *The guided construction of knowledge: Talk amongst teachers and learners,* Clevedon, Philadelphia, Toronto, Sydney and Johannesburg: Multilingual Matters Ltd.

Mercer N. (2000) *Words and minds: How we use language to think together,* London & New York: Routledge: Taylor & Francis Group.

Mercer N. (2004) Sociocultural discourse analysis: analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics* 1: 137 – 168.

Mercer N, Dawes L, Wegerif R, et al. (2004) Reasoning as a scientist: ways of helping children to use language to learn science. *British Educational Research Journal* 30: 359 – 377.

Mercer N and Littleton K. (2007) *Dialogue and the development of children's thinking: A sociocultural approach,* Oxford: Routledge.

Mergenthaler E. (1996) Computer-assisted content analysis. In: Zuell C, Harkness J and Hoffmeyer-Zlotnik J (eds.) *Text Analysis and Computers Conference.* Mannheim, Germany: Zentrum für Umfragen, Methoden und Analysen (ZUMA), 3 – 32.

Mertens D. (2010) *Research and evaluation in education and psychology: Integrating diversity with quatitative, qualitative and mixed methods,* Thousand Oaks, London, Singapore and New Delhi: Sage Publications,Inc.

Meyer C. (2002) *English Corpus Linguistics: An Introduction,* Cambridge and New York: Cambridge University Press.

Monroe B and Schrodt P. (2008) Introduction to the special issue: The statistical analysis of political text. *Political Analysis* 16: 351 – 355.

Morley G. (2000) *Syntax in functional grammar: an introduction to lexicogrammar in systemic linguistics,* London and New York: Continuum.

Morrison J. (2004) Where now for problem based learning? *THE LANCET* 363: 174.

Morse J. (1991) Approaches to qualitative-quantitative methodological triangulation. *Nursing Research* 40: 120 – 123.

Moust J, Schmidt HG, De Volder M, et al. (1986) Effects of verbal participation in small group discussion on learning. Available at: http//www.repub.eur.nl/pub/2797/18263.pdf. Accessed: 5 June, 2016.

Moust JH, De Volder ML and Nuy HJP. (1989) Peer teaching and higher level cognitive learning outcomes in problem-based learning. *Higher Education* 18: 737 – 742.

Moust JHC, Schmidt HG, Volder ML, et al. (1987) Effects of verbal participation in small-group discussion on learning. In: Richardson J (ed.) *Student Learning (Society for Research into Higher Education).* Open University Press, 147 – 154.

Mumtaz S. (2000) Factors affecting teachers' use of information and communications technology: a review of the literature. *Journal of Information Technology for Teacher Education* 9: 319 – 342.

Nesselhauf N. (2005) Corpus Linguistics: A practical introduction. Available at: http://www.as.uni-heidelberg.de/personen/Nesselhauf/files/Corpus%20Linguistics%20Practical%20Introduction.pdf. Accessed: 25 December, 2014.

Neufeld VR and Barrows HS. (1974) The" McMaster Philosophy": an approach to medical education. *Academic Medicine* 49: 1040 – 1050.

Neville AJ. (1999) The problem-based learning tutor: Teacher? Facilitator? Evaluator? *Medical Teacher* 21: 393 – 401.

Neville AJ. (2009) Problem-based learning and medical education forty years on. *Medical Principles and Practice* 18: 1 – 9.

Neville AJ and Norman GR. (2007) PBL in the undergraduate MD program at McMaster University: Three iterations in three decades. *Academic Medicine* 82: 370 – 374.

Niebles J, Wang H and Fei-Fei L. (2008) Unsupervised learning of human action categories using spatial-temporal words. *International Journal of Computing Vision* 79: 299 – 318.

Nieto S. (2010) *Language, Culture and Teaching: Critical perspectives,* Abingdon Oxon and New York: Routledge.

Norman G and Schmidt HG. (1992) The psychological basis of problem-based learning: a review of the evidence. *Academic Medicine* 67: 557 – 565.

Norman GR and Schmidt HG. (2000) Effectiveness of problem ‐based learning curricula: theory, practice and paper darts. *Medical Education* 34: 721 – 728.

Novak C. (1994) Interview with Linda Darling-Hammond. *TECHNOS Quarterly* 3: 2.

O'Grady G, Yew E, Goh K, et al. (2012). Introduction and Overview. In: O'Grady G, Yew E, Goh K, et al. (eds.) *One-day,One-problem: An Approach to Problem-based Learning. Dordrecht, Heildelberg, London and New Yor.* Springer.

O'Loughlin M. (1992) Rethinking science education: Beyond piagetian constructivism toward a sociocultural model of teaching and learning. *Journal of Research in Science Teaching* 29: 791 – 820.

O'Loughlin M. (1993) Some further questions for Piagetian constructivists: A reply to Fosnot. *Journal of Research in Science Teaching* 30: 1203 – 1207.

Oakes M. (1998) *Statistics for corpus linguistics,* Edinburgh: Edinburgh University Press.

Oliveira A and Sadler T. (2008) Interactive patterns and conceptual convergence during student collaborations in science. *Journal of Research in Science Teaching* 45: 634 – 658.

Oostdijk N. (1991) *Corpus linguistics and the automatic analysis of English,* Amsterdam and Atlanta: Rodopi.

Ormrod J. (2004) *Human Learning,* Columbus, OH: Pearson Prentice Hall

Osborne J. (2010) Arguing to learn in science: The role of collaborative, critical discourse. *Science* 328: 463 – 466.

Packer M and Greco-Brooks D. (1999) School as a site for the production of persons. *Journal of Constructivist Psychology* 12: 133 – 149.

Packer MJ and Goicoechea J. (2000) Sociocultural and constructivist theories of learning: Ontology, not just epistemology. *Educational Psychologist* 35: 227 – 241.

Papinczak T, Tunny T and Young L. (2009) Conducting the symphony: a qualitative study of facilitation in problem-based learning tutorials. *Medical Education* 43: 377 – 383.

Pappas C, Varelas M, Barry A, et al. (2003) Dialogic inquiry around information texts: The role of intertextuality in constructing scientific understandings in Urban Primary classrooms. *Linguistics and Education* 13: 435 – 482.

Pappas C, Varelas M, Barry A, et al. (2004) Dialogic inquiry around information texts: The role of intertextuality in constructing scientific understandings in Urban primary classrooms. In: Shuart-Faris N and Bloome D (eds.) *Uses of Intertextuality in Classroom and Educational Research.* Connecticut: Information Age Publishing Inc, 93 – 145.

Pastirik P. (2006) Using problem-based learning in a large classroom. *Nurse Education in Practice* 6: 261 – 267.

Peets AD, Cooke L, Wright B, et al. (2010) A prospective randomized trial of content expertise versus process expertise in small group teaching. *BMC Medical Education* 10: 1. Available at: http://www.biomedcentral.com/1472-6920/10/70. Accessed: 10 July, 2016.

Perkins D. (1998) What is Understanding? In: Wiske M (ed.) *Teaching for Understanding: Linking researh with practice.* San Francisco, CA: Jossey-Bass: A Wiley Company, 39 – 57.

Perrone V. (1998) Why do we need a pedagogy of understanding? In: Wiske M (ed.) *Teaching for understanding: Linking research with pratice.* San Francisco, CA: Jossey-Bass: A John Wiley Company, 13 - 38.

Peterson C and McCabe A. (1988) The connective AND as a discourse glue. *First language* 8: 19 – 28.

Phillips D. (1995) The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher* 24: 5 – 12.

Phillips N and Duke M. (2001) The questioning skills of clinical teachers and preceptors: a comparative study. *Journal of Advanced Nursing* 33: 523 – 529.

Piaget J. (2011) *The child conception of the world,* London: Lightning Source

Piaget J, Brown T and Thampy K. (1985) *Equilibration of Cognitive Structures: The Central Problem of Intellectual Development,* Chicago: University Press.

Polit DF and Beck CT. (2012) *Nursing Research: Generating and assessing evidence for nursing practice,* Philadelphia: Lippincott Williams and Wilkins.

Popping R. (2000) *Computer-assisted Text Analysis,* London, Thousand Oaks and New Delhi: Sage Publications Ltd.

Profetto-McGrath J, Smith KB, Day RA, et al. (2004) The questioning skills of tutors and students in a context based baccalaureate nursing program. *Nurse Education Today* 24: 363 – 372.

Puntambekar S. (2006) Analyzing collaborative interactions: divergence, shared understanding and construction of knowledge. *Computers and Education* 47 332 – 351.

Rajasekar S, Philominathan P and Chinnathambi V. (2013) Research Methodology Online. Availableat: http://arxiv.org/pdf/physics/0601009.pdf. Accessed: 9 August, 2014.

Rasinger S. (2010) Quantitative methods: Concepts, frameworks and issues. In: Litosseliti L (ed.) *Research Methods in Linguistics.* London, New Delhi, New York and Sydney: Bloomsbury Academic, 49 – 67.

Rayson P. (2003) WMatrix: A statistical method and software tool for linguistic analysis through corpus comparison. *Computer Science.* Avaialble at: http://ucrel.lancs.ac.uk/people/paul/publications/phd2003.pdf. Accessed: 20 December, 2014: Lancaster University.

Rayson P. (2008) From key words to key semantics domains. *International Journal of Corpus Linguistics* 13: 519 – 549.

Rayson P. (2009a) Key domains and MWE extraction using Wmatrix. *Workshop at Aston Corpus Summer School. Available at:*http://ucrel.lancs.ac.uk/people/paul/publications/AstonWorkshop09PRslides.pdf*. Accessed: 23 May, 2016*

Rayson P. (2009b) Wmatrix: a web-based corpus processing environment, Computing Department, Lancaster University. Available at: http://ucrel.lancs.ac.uk/wmatrix/. Accessed: 30 January, 2015.

Regehr G, Martin J, Hutchison C, et al. (1995) The effect of tutors ̓ content expertise on student learning, group process, and participant satisfaction in a problem ‐based learning curriculum. *Teaching and Learning in Medicine* 7: 225 – 232.

Reichardt C and Rallis S. (1994) *The Qualitative-Quantitative Debate: New Perspectives. ,* San Francisco: Jossey-Bass.

Remedios L, Clarke D and Hawthorne L. (2008) The silent participant in small group collaborative learning contexts. *Active Learning in Higher Education* 9: 201 – 216.

Reppen R. (2012) Building a corpus: What are the key considerations. In: McCarthy M and O'Keeffe A (eds.) *Routledge Handbook of Corpus Linguistics.* Abingdon, Oxon and New York: Routledge: Tyalor and Francis Group, 31 – 37.

Reznich CB and Werner E. (2004) Facilitators' influence on student PBL small group session online information resource use: a survey. *BMC Med Educ* 4: 9. Available at: http://www.biomedcentral.com/1472-6920/4/9. Accessed: 3 December, 2015.

Richardson V. (2003) Constructivist pedagogy *Teachers College Record* 105: 1623 – 1640.

Ritchhart R, Church M and Morrison K. (2011) *Making thinking Vissible: How to promote engagement, understanding and independence for all Learners,* San Francisco, CA: Jossey-Bass: A Wiley Company.

Ritchhart R, Wiske M, Buchovecky E, et al. (1998) How does Teaching for Understanding look in practice? In: Wiske M (ed.) *Teaching for Understanding: Linking research with practice.* San Francisco, CA: Jossey-Bass: A Wiley Company, 122 – 158.

Roberts C, Lawson M, Newble D, et al. (2005) The introduction of large class problem-based learning into an undergraduate medical curriculum: an evaluation. *Medical Teacher* 27: 527 – 533.

Robson C. (2011) *Real World Research: A resource for users of social research methods in applied settings,* Chichester: John Wiley & Sons Ltd.

Rogoff B. (2003) *The Cultural Nature of Human Development,* Oxford: Oxford University Press.

Rojas-Drummond S and Mercer N. (2003) Scaffolding the development of effective collaboration and learning. *International Journal of Educational Research* 39: 99 – 111.

Romito LM and Eckert GJ. (2011) Relationship of biomedical science content acquisition performance to students' level of PBL group interaction: are students learning during PBL group? *Journal of Dental Education* 75: 653 – 664.

Roschelle J. (1992) Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences* 2: 235 – 276.

Rossman G and Wilson B. (1991) Numbers and words revisited: Shamelessly eclctic *Educational ResearchAssociation.* Washington, DC AGENCY Office of Educational Research and Improvement, 1 – 18.

Russell S and Norvig P. (2010) *Artificial intelligence: A modern approach,* Upper Saddle River, NJ: Pearson Education, Inc.

Sadighi F and Bavali M. (2008) Chomsky's Universal Grammar and Halliday's Systemic Functional Linguistics: An appraisal and a compromise. *Journal of Pan-Pacific Association of Applied Linguistics* 12: 11 – 28.

Sampson E. (1981) Cognitive Psychology as ideology. *American Psychologist* 36: 730 – 743.

Sampson G and McCarthy D. (2004) Introduction. In: Sampson G and McCarthy D (eds.) *Corpus Linguistics: Readings in a Widening Discipline.* London and New York: Continuum,1 – 8.

Savery J. (2006) Overview of problem-based learning: Definitions and distinctions. *Interdisciplinary Journal of Problem-based Learning* 1: 9 - 20.

Savery JR and Duffy TM. (1996) Problem based learning: An instructional model and its constructivist framework. *Constructivist learning environments: Case studies in instructional design*: 135 – 148.

Savin-Baden M. (2000) *Problem-based learning in higher education: Untold stories,* Bukingham: The Society for Study in Higher Education & Open University Press.

Schegloff E. (1982) Discourse as an interactional achievement: Some uses of "Uh huh" and other things that come between sentences. In: Tannen D (ed.) *Analysing Discourse: Text and Talk.* Washington, DC: Georgetown University Press, 71 – 93.

Schiffrin D. (1987) *Discourse markers* Cambridge: University Press.

Schmidt H, van der Arend A, Kokx I, et al. (1994) Peer versus staff tutoring in problem-based learning. *Instructional Science* 22: 279 – 285.

Schmidt HG. (1983) Problem-based learning: rationale and description. *Medical Education* 17: 11 – 16.

Schmidt HG. (1993) Foundations of problem ‐based learning: some explanatory notes. *Medical Education* 27: 422 – 432.

Schmidt HG. (1994) Resolving inconsistencies in tutor expertise research: Does lack of structure cause students to seek tutor guidance? *Academic Medicine* 69: 656 – 662.

Schmidt HG. (2012) A Brief History of Problem-based Learning. In: O'Grady G, Yew EHJ, Goh KPL, et al. (eds.). *One-Day, One-Problem*: An Approach to Problem-based Learning. Dordrecht, Heidelberg, London and New York. Springer, 21-40.

Schmidt HG, Arend A, Moust J, et al. (1993) Influence of tutors' subject-matter expertise on student effort and achievement in problem-based learning. *Academic Medicine* 68: 784 – 791.

Schmidt HG and Boshuizen HPA. (1992) Encapsulation of biomedical knowledge. *Nato ASI Series for Computer & System Sciences* 97: 265 – 265.

Schmidt HG and Moust J. (1995) What makes a tutor effective? A structural-equations modeling approach to learning in problem-based curricula. *Academic Medicine* 70: 708 – 714.

Schmidt HG and Moust JH. (2000) Factors affecting small-group tutorial learning: A Review of research. In: Evensen D and Hmelo C (eds.) *Problem-based Learning: A Research Perspective on Learning Interactions.* New Jersey and London: Lawrence Erlbaum Associates, Publishers, 19 – 52.

Schmidt HG, Rotgans JI and Yew EH. (2011) The process of problem ‐ based learning: what works and why. *Medical Education* 45: 792 ‐ 806.

Schreier M. (2012) *Qualitative content analysis in practice,* Los Angeles, London, New Delhi, Singapore and Washington DC.: SAGE Publications Ltd.

Scollon R, Yung V, Tsang W, et al. (2004) Voice, appropriation, and discourse representation in a student writing task. In: Shuart-Faris N and Bloome D (eds.) *Uses of intertextuality in classroom and education research.* Connecticut: Information Age Publishing, 173 – 199.

Sebastiani F. (2002) Machine learning in automated text categorization. *ACM Computing Surveys* 34: 1– 47.

Semino E, Demmen J and Koller V. (2013) Using Wmatrix for metaphor identification and analysis. Available at: http://ucrel.lancs.ac.uk/melc/workshop_jan2013/MELC_workshop_Jan14_Using_Wmatrix_for_metaphor_analysis.pdf. Accessed: 11March, 2016.

Semino E, Hardie A, Koller V, et al. (2015) A computer-assisted approach to the analysis of metaphor variation across genres. Available at:http://www.research.lancs.ac.uk/portal/en/publications/a-computerassisted-approach-to-the-analysis-of-metaphor-variation-across-genres(75851a0b-3cd4-4f5e-a13e-8528970689f2).html. Accessed: 29 April, 2016.

Sendag S and Odabasi H. (2009) Effects of an online problem-based learning course on content knowledge acquisition and critical thinking skills. *Computers and Education* 53: 132 – 141.

Shami M and Verhelst W. (2007) An evaluation of the robustness of existing supervised machine learning approaches to the classification of emotions in speech. *Speech Communication* 49: 201 – 212.

Short K. (2004) Researching intertextuality within collaborative classroom learning environments In: Shuart-Faris N and Bloome D (eds.) *Uses of Intertextuality in Classroom and Educational Research.* Connecticut: Information Age Publishing Inc, 373 – 396.

Silva M. (1991) Simultaneity in children's narratives: the case of when, while and as. *Journal of Child Language* 18: 641 – 662.

Silver M and Wilkerson LA. (1991) Effects of tutors with subject expertise on the problem-based tutorial process. *Academic Medicine* 66: 298 – 300.

Skinner V, Braunack-Mayer A and Winning T. (2012) Getting on with each other: PBL group dynamics and function. In: Bridges S, McGrath C and Whitehill T (eds.) *Problem-Based Learning in Clinical Education: The next Generation (Innovation and Change in Professional Education). Dordrecht, Heidelberg, London and New York:* Springer, 189-205.

Slavin RE. (1996) Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology* 21: 43 – 69.

Smith A and Humphreys M. (2006) Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods* 38: 262 – 279.

Staarman J, Aarnoutse C and Verhoeven L. (2003) Connecting discourses: Intertextuality in a primary school CSCL practice. *International Journal of Educational Research* 39: 807 – 816.

Stahl G. (2000) A model of collaborative knowledge building. In: Fishman B and O'Connor-Divelbiss S (eds.) *Fourth International Conference of the Learning Sciences.* Mahwah, NJ: Erlbaum, 70 – 77.

Stahl G. (2003a) Knowledge negotiation in synchronous learning networks. *Paper presented at the Hawaii International Conference on Systems Sciences (HICSS2003).* Hawaii, USA.

Stahl G. (2003b) Meaning and interpretation in collaboration *Computer Support for Collaborative Learning (CSCL).* Bergen, Norway: Kluwer Academic Publishers. Available at: http://www.cis.drexel.edu/faculty/gerry/cscl/papers/ch20.pdf. Accessed: 3 January, 2015.

Stahl G. (2006) Building collaborative knowing elements of a social theory of CSCL. *Group Cognition: Computer Support for Building Collaborative Knowledge (Acting with Technology).* Cambridge, Massachusetts and London: The MIT Press, 303 – 330.

Steele DJ, Medder JD and Turner P. (2000) A comparison of learning outcomes and attitudes in student ‐versus faculty ‐led problem ‐based learning: an experimental study. *Medical Education* 34: 23 ‐ 29.

Sterling S and Centre S. (2004) Online problem-based learning. *Action Research Exchange* 3. Available at: http://www.academia.edu/3765808/Online_Problem-Based_Learning. Accessed: 22 April, 2014.

Svinicki MD. (2007) Moving beyond “it worked”: The ongoing evolution of research on problem-based learning in medical education. *Educational Psychology Review* 19: 49 – 61.

Swain M. (2000) The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In: Lantolf J (ed.) *Sociocultural theory and second language learning.* Oxford and New York: Oxford UNiversity Press, 97 – 114.

Swain M. (2011) The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In: Swain M, Kinnear P and Steinman L (eds.) *Sociocultural theory in second language education: an introduction through narratives.* Bristol: Multilingual Matters Tesxtbooks, 97 – 114.

Tannen D. (2007) *Talking voices: Repetition, dialogue and imagery in conversational discourse,* Cambridge: Cambridge University Press.

Tashakkori A and Creswell J. (2007) Editorial: The new era of mixed methods. *Journal of Mixed Methods Research* 1: 3 – 7.

Taylor C and Gibbs G. (2010) "Preparing data". *Learning qualitative data analysis on the web. Online: QDA Web Site* *[onlineqda.hud.ac.uk/Intro_QDA/preparing_data.php]. Accessed: 29 January, 2015*.

Teasley SD, Fischer F, Weinberger A, et al. (2008) Cognitive convergence in collaborative learning. *8th ICLS* International Society of the Learning Sciences, 360 – 367.

Teddlie C and Tashakkori A. (2009) *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioural sciences,* Los Angeles, London, New Delhi, Singapore and Washington DC: Sage Publications Ltd.

Teubert W and Čermáková A. (2007) *Corpus Linguistics: A short introduction,* London and New York: Continuum.

Thomas D. (2006) A general inductive approach for analysing qualitative evaluation data. *American Journal of Evaluation* 27: 237 – 246.

Thompson G. (2004) *Introduction to systemic functional linguistics,* London: Arnold Publications.

Thompson G and Hunston S. (2000) Evaluation: An introduction. In: Hunston S and Thompson G (eds.) *Evaluation in text: Authorial stance and the construction of discourse.* Oxford & New York: Oxford University Press, 1 – 27.

Thompson G and Hunston S. (2006) System and corpus: Two traditions with a common ground. In: Thompson F and Hunston S (eds.) *System and Corpus: exploring connections.* London: Equinox Publishing Ltd, 2 – 14.

Thompson S and Mulac A. (1991) The discourse conditions for the use of the complementizser that in conversational English. *Journal of Pragmatics* 15: 237 – 251.

Thorsten D, Thorsten P and Schmieder C. (2012) Manual (on) transcription. Transcription conventions, software guides and practical hints for qualitative researchers. 2nd English Edition. Marburg. Available at: http://www.audiotranscription.de/english/transcription-practicalguide.htm. Accessed: 29 January, 2015.

Tomasello M. (1999) *The cultural origins of human cognition,* Harvard: Harvard University Press.

Tomasello M, Carpenter M, Call J, et al. (2005) Understanding and sharing intentions: the origin of cultural cognition. *Behavioral and Brain Sciences* 28: 675 – 735.

Trebits A. (2009) Conjunctive cohesion in English language EU documents - A corpus-based analysis and its implications. *English for Specific Purposes* 29: 199 – 210.

Trochim WMK and Donnelly JP. (2008) *The research methods knowledge base,* Mason OH: Atomic Dog Publishing Inc.

Van Berkel HJM and Dolmans DHJM. (2006) The influence of tutoring competencies on problems, group functioning and student achievement in problem-based learning. *Medical Education* 40: 730 – 736.

Van Blankenstein F, Dolmans D, Van der Vleuten C, et al. (2013) Elaboration during problem-based group discussion: Effects on recall for high and low ability students. *Advances in Health Sciences Education* 18: 659 – 672.

Van Boxtel C, Van der Linden J and Kanselaar G. (2000) Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction* 10: 311– 330.

Van Dijk T. (1977) *Text and context: Explorations in the semantics and pragmatics of discourse.,* London and New York: Longman Group UK Limited.

Varelas M and Pappas C. (2006) Intertextuality in Read-Alouds of Integrated Science–Literacy Units in Urban Primary Classrooms: Opportunities for the Development of Thought and Language. *Cognition and Instruction* 24: 211 – 259.

Vernon TD and Blake RL. (1993) Does problem-based learning work? A meta-analysis of evaluative research. *Academic Medicine* 68: 550 – 563.

Veronese C, Richards JB, Pernar L, et al. (2013) A randomized pilot study of the use of concept maps to enhance problem-based learning among first-year medical students. *Medical Teacher* 35: 1478 – 1484.

Viera A and Garrett J. (2005) Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine* 37: 360 – 363.

Visschers-Pleijers A, Dolmans H, De Leng B, et al. (2006a) Analysis of verbal interactions in tutorial groups: a process study. *Medical Education* 40: 129 – 137.

Visschers-Pleijers AJ, Dolmans DH, de Grave WS, et al. (2006b) Student perceptions about the characteristics of an effective discussion during the reporting phase in problem-based learning. *Medical Education* 40: 924 – 931.

Visschers-Pleijers AJ, Dolmans DH, Wolfhagen IH, et al. (2004) Exploration of a method to analyze group interactions in problem-based learning. *Medical Teacher* 26: 471 – 478.

Visschers-Pleijers AJ, Dolmans DH, Wolfhagen IH, et al. (2005a) Development and validation of a questionnaire to identify learning-oriented group interactions in PBL. *Medical Teacher* 27: 375 – 381.

Visschers-Pleijers AJ, Dolmans DH, Wolfhagen IH, et al. (2005b) Student perspectives on learning-oriented interactions in the tutorial group. *Advances in Health Science Education Theory & Practice* 10: 23 – 35.

Voloshinov VM, L Titunik, IR. Transl.). (1973) *Marxism and the philosophy of language,* Cambridge, MA: Harvard University Press.

von Glasersfeld E. (1989) Cognition, construction and teaching. *Synthese* 80: 121 – 140.

von Glasersfeld E. (1995) A Constructivist Approach to Teaching. In: Steffe L and Gale J (eds.) *Constructivism in Education.* New Jersey: Lawrence Erlbaum Associates, Publishers, 3 – 16.

Vygotsky L. (1934/1986) *Thought and Language* London: MIT Press.

Vygotsky LS. (1930/1978) *Mind in society: The development of higher psychological processes,* Massachusetts: Harvard University Press.

Wallis S and Nelson G. (2001) Knowledge discovery in grammatically analysed corpora. *Data Mining and Knowledge Discovery* 5: 305 – 335.

Walton HJ and Matthews M. (1989) Essentials of problem ˌbased learning. *Medical Education* 23: 542 ˈ 558.

Waring M and Evans C. (2015) *Understanding Pedagogy: Developing a critical approach to teaching and learning,* Abingdon, Oxon and New York: Routledge.

Webb NM. (1989) Peer interaction and learning in small groups. *International Journal of Educational Research* 13: 21 – 39.

Webb NM. (2009) The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology* 79: 1 – 28.

Weber R. (1984) Computer-aided content analysis: A short primer. *Qualitative Sociology* 7: 126 – 147.

Weber R. (1990) *Basic Content Analysis,* London: Sage Publications Ltd.

Wegerif R, Mercer N and Dawes L. (1999) From social interaction to individual reasoning: an empirical investigation of a possible socio-cultural model of cognitive development. *Learning and Instruction* 9: 493 – 516.

Wei R, Darling-Hammond L, Andree A, et al. (2009) *Professional learning in the learning profession: A status report on teacher development in the United States and abroad* Dallas, TX: National Staff Development Council.

Weinberger A and Fischer F. (2006) A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education* 46: 71 – 95.

Wells G. (1990) Talk about text: Where literacy is learned and taught. *Curriculum Inquiry* 20: 369 – 405.

Wells G. (1994) The Complementary Contributions of Halliday and Vygotsky to a “Language-Based Theory of Learning”. *Linguistics and Education* 6: 41 – 90.

Wells G. (1998) Some questions about direct instruction: Why? To whom? How? and When? *Language Arts* 76: 27 – 35.

Wells G. (1999) Language and education: Reconceptualising educationa as dialogue. *Annual Review of Applied Linguistics* 19: 135 – 155.

Wells G. (2000) Dialogic inquiry in education: Building on the legacy of Vygotsky. In: Lee CD and Smagorinsky P (eds.) *Vygotskian perspectives on literary research: Constructing meaning through collaborative inquiry.* Cambridge: University Press, 51 – 85.

Wells G. (2007) Semiotic mediation, dialogue and the construction of knowledge. *Human Development* 50: 1 – 42.

Wells G and Arauz R. (2006) Dialogue in the classroom. *The Journal of the Learning Sciences* 15: 379 – 428.

Wertsch J. (1985) *Vygotsky and the social formation of mind,* Cambridge, Massachusetts and London: Harvard University Press.

Wertsch J. (1991) *Voices of the Mind: a Sociocultural Approach to Mediated Action* Harvard: Harvard University Press.

Wertsch J and Toma C. (1995) Discourse and learning in the classroom: A sociocultural Approach. In: Steffe L and Gale J (eds.) *Constructivism in Education.* New Jersey: Lawrence Erlbaum Associates Publishers, 159 – 174.

White M and Marsh E. (2006) Content Analysis: A flexible methdology. *Library Trends* 55: 22 – 45.

Wilkinson S and Kitzinger C. (2006) Surprise as an interactional achievement: Reaction tokens in conversation. *Social psychology quarterly* 69: 150 –182.

Wiske M. (1998a) The importance of Understanding. In: Wiske M (ed.) *Teaching for Understanding: Linking research with practice.* San Francisco, CA: Jossey-Bass: A Wiley Company, 1 – 12.

Wiske M. (1998b) What is teaching for Understanding? In: Wiske M (ed.) *Teaching for Understanding: Linking research with practice.* San Francisco, CA: Jossey-Bass: A Wiley Company, 61 – 86.

Yaguchi M. (2001) The function of the non-deictic that in English. *Journal of Pragmatics* 33: 1125 – 1155.

Yew EH, Chng E and Schmidt HG. (2011) Is learning in problem-based learning cumulative? *Advances in Health Sciences Education* 16: 449 – 464.

Yew EHJ and Schmidt HG. (2009) Evidence for constructive, self-regulatory, and collaborative processes in problem-based learning. *Advances in Health Sciences Education* 14: 251 – 273.

Yew EHJ and Schmidt HG. (2011) What students learn in problem-based learning: a process analysis. *Instructional Science*: 40:1-25.

Yin RK. (2009) *Case Study Research: design and methods,* London: Sage Publications Ltd.

Yoon C. (2011) Concordancing in L2 writing class: An overview of research and issues *Journal of English for Academic Purposes* 10: 130 – 139.

Yoshioka T, Suganuma T, Tang AC, et al. (2005) Facilitation of problem finding among first year medical school students undergoing problem-based learning. *Teaching & Learning in Medicine* 17: 136 – 141.

Young RF and Lee J. (2004) Identifying units in interaction: Reactive tokens in Korean and English conversations. *Journal of Sociolinguistics* 8: 380 – 407.

APPENDICES

Appendix 1. Description of the reviewed studies

	Citation	Purpose	Research type	Design	Discipline	Data source	Data type	Analysis method	Analysis technique	Key findings and weakness
1	Aarino et al. 2013 Finland	To explore types of conflicts of knowledge and methods of dealing with them in the PBL tutorial groups	Exploratory	Observational	Medical and dental four groups (n = 33)	(1) Videotaping of discussions in phase 2 of the cycle	(1) Video transcripts of conflict episodes	Coding system	Conflict episodes located on video with software. Coding done manually	There were 43 conflict episodes accounting for 7.6% of the time, factual conflicts 58%, and conceptual conflicts 42%. Conceptual conflicts last longer and are more elaborated. Men's duration of conflicts was 28 seconds. Elaboration was mostly done individually rather than collaboratively. Date was limited. Analysis was time consuming and error prone.

2	Aarino et al. 2014 Finland	To explore how facilitators help students resolve conflicts	Explorator y	Observati onal	Medical and dental four groups (n = 33)	(1) Videotapi ng of discussio ns in phase 2 of the cycle	(1) Video transcript s of conflict episodes	Coding system	Conflict episodes located on video with software. Coding done manually	Tutor interventions were 92. Conflict episodes were 43. Tutors intervened in 24; resolved 13. Methods of intervention involved (1) confirming (2) explanation (3) questions (rarely). Tutor often resolved factual conflicts, but rarely resolved conceptual conflicts. Dataset was small-sized. Analysis technique was time consuming and error-prone.
3	Duek 2000 US	To explore equity of participation in group discourse	Explorator y case study	Ethnograp hic	Medical	(1) Participan t observati on (2) audiotapi ng	(1) Field notes (2) Interview transcript s	Grounde d theory	Manual coding	Group participation was uneven and subject to dominance, social exclusion, and sometimes conflicts. There were discussion dominators and hypercontributors, withdrawing. Tutors felt that the groups were doing well

						students' interviews (3) audiotaping facilitator interviews				generally; they only said that some students talk more than others and never stated this as a problem.
4	Skinner et al. 2012 Australia & Ireland	What is the nature of group dynamics for the students? How do students understand group dynamics to shape group functions?	Exploratory case study (five groups)	Ethnographic	Cross-site study in two dental schools	(1) Participant observation (2) Individual student's interviews audiotaped	(1) Field notes (2) Interview transcripts	Grounded theory	Manual coding	In all the five case studies, the common finding in all the groups was that social relationship determines working interactions and influences how group members are engaged during PBL discussions.

5	Jin 2012 Japan	To explore silence in PBL tutorial	Exploratory case study	Ethnographic	Dental	(1) Questionnaires (2) interviews (3) participant observation (4) stimulated recall test	(1) Questionnaire scores (2) interview transcripts (3) field notes (4) audio-record of spoken discourse	Grounded theory and critical discourse analysis (CDA + GT)	(1) SPSS; (2) Soundscriber; (3) NVivo software	Students' silence performs specific roles in group communication and learning. Silence was seen not as a verbal disengagement, but as a productive resource, a collaborative practice, a platform to deal with conflicting understanding and a signal of shifting power relations. The generalisability of the study may be limited because of small sample size.
6	Remedios et al. 2008 Australia	To examine the experience of silent participants in PBL context	Exploratory case study	Ethnographic	Physiotherapy	(1) Participant's observation (2) semi-structured interview	(1) Participant observation notes (2) interview transcript	Grounded theory	Constructive grounded theory	Silence participation in PBL is based on several factors: learning style, motivation, preparation for the session, cultural literacy, language, concerns with face, and group dynamics. Silence did not preclude learning.

						(3) videotaping of tutorial interaction (4) video-stimulated recall interviews	s (3) transcripts of stimulated interviews			
7	Regehr et al. 1995 Canada	To investigate the influence of content expertise on tutorial process and students' satisfaction	Analytic	Cross-sectional survey	Medical	(1) Participant observation for 20 minutes. (2) Tutors categorised their level of expertise	(1) Number of students and tutor utterance. (2) Questions asked by the tutors	Quantitative	Statistical measures of significance	There were no difference in the process variables and students' satisfaction between the two groups.

							and the students. (3) Expert categories			
8	Yew and Schmidt 2009 Singapore	Evidence for constructive, self-regulative, and collaborative processes in PBL	Descriptive		Biomedical science	Audio recording of verbal interaction for three phases	Verbatim transcript of tutorial talk	Coding system	Manual coding	Learning activities were 53.3% collaborative episodes, 27.2% self-directed learning, and 15.7% constructive. Generalisability was limited: only one group was analysed.
9	Koschmann et al. 1997 Carbobdale USA	Analysing emergence of learning issues in medical PBL tutorials	Descriptive		Medical	Videotaping of tutorial segment and field notes	Videotaping tutorial talk Field notes	Conversation analysis	Manual coding	Evidence of collaborative knowledge negotiation and construction and questioning was provided, as well as tutor scaffolding. Small size (one case) and small segments of tutorial interaction

10	Hurk et al. 2001 Netherland	To rate quality of learning issues using a coding system	Descriptive		Medical	Tutorial talk	Talk transcript	Coding system	Manual coding	21% of all learning issues for one problem and 32% for the other problem scored high on all characteristics. 5% and 2% of learning issues scored low on all characteristics. Most learning issues were formulated non-concisely and ambiguously. There was no validation of the assessors' rating.
11	Lee et al. 2009 Taiwan	To explore when and how a tutor intervenes in PBL tutorial	Descriptive		(1) Medical (2) Nursing (3) Clinical psychology	(1) Videotaping of tutorial talk (2) participant observation	(1) Transcripts of video-stimulated recall (2) field notes	Grounded theory	Manual coding	There were 636 episodes in 40 videotaped tutorial discussions. The themes identified: (1) tutorial group process; (2) quality of discussion; (3) quality and quantity of materials discussed, including related to amount, datedness, and source.

						(3) audiotaping of stimulated recall				
12	David et al. 1994 US	To investigate the effect of content expertise on learning process and students' satisfaction	Analytic	Observational	Medical	(1) Questionnaire on tutor's expertise, tutorial process, and satisfaction (2) Participant observation (3) tutor self-rated	(1) Questionnaire scores (2) Field notes on number of questions asked by students and tutor and other utterances	Quantitative	Statistical comparison	There was no difference between content expert and non-content expertise group.

						expertise level				
13	Lycke 2002 Norway	To understand how PBL process corresponds to the theory of PBL	Descriptive		Medical	Videotaping and written reports of tutorial interactions	Video of tutorial talk and written reports	Narrative	Narrative	Discussions by the three groups aligned with PBL process in general, but tutor behaviours differed. One tutor was directive while the other two facilitated. Explanations and elaborations were present in the students' discourse.
14	Schmidt et al. 1994 Netherland	Effect of peer vs staff tutoring in PBL tutorial	Analytic	Cross-sectional survey	Health sciences	Tutor performance rating	Rating scores	Quantitative	Statistical with test of significance	Students felt that staff tutors were more knowledgeable, made more relevant contributions, and asked more stimulating questions than peer tutors. Peer tutors were more congruent for year 1 students.
15	Eagle et al. 1992 Canada	Effects of tutor expertise on (1) number	Analytic	Observational	Medical	Students' written report of number of	Students' tutorial notes.	Mixed Quantitative	Manual coding and	Expert tutored group generated statistically more significant learning issues and spent more time on self-study. Non-expert

		of learning issues and (2) study time				learning issues and time spent on self-study. Tutor interview.	Interview notes.		statistical with software	tutors with experienced in facilitation and read around the case before the tutorial achieved results, similar to expert tutors. Data was self-report by the students.
16	Schmidt et al. 1993 Netherland	Effect of tutor expertise on student effort	Analytic	Cross-sectional survey	Medical (n = 1120) in four curriculum years and tutors (n = 152)	Students' tutorial notes and questionnaire	Students' tutorial notes and questionnaire	Quantitative	Statistical with software	Expert-tutored group spent significantly more time on self-study. Expert tutors made use of content and process facilitation skills. Subject questionnaire report.
17	Silver and Wilkerson, 1991	Influence of tutor expertise on tutorial process	Analytic		Medical	Audiotaping of tutorial interactions and	Direct analysis of audiotaped talk and	Quantitative	Statistical with software	Content experts were more directive, spoke often and for longer periods, suggested topics for discussion, and dispensed knowledge. Results

	USA					tutors' self-rating of expertise	tutors' self-rating report			are based on small datasets of two sessions of two cases and self-rating of tutors' expertise.
18	Groves et al. 2005 Australia	Influence of tutor background and style on effectiveness	Analytic		Medical	Validate questionnaire	Questionnaire	Quantitative	Statistical with software	Clinically qualified used content to facilitate and were more socially congruent. Staff tutors emphasised assessment and established and maintained collaborative atmosphere. Both content and process expertise were needed. Weakness: subjective reporting.
19	Curet and Mennin 2003 USA	Effect of long-term vs. short-term tutors on PBL process	Analytic		Medical	Questionnaire	Questionnaire	Quantitative	Statistical with software	Long-term tutors were perceived to be statistically superior to short-term tutors in terms of cognitive, metacognitive, and motivational behaviours. There

										was voluntary participation and use of self-rated questionnaire.
20	Couto et al. 2015 Brazil	Influence of tutor expertise on tutorial process	Analytic	Cross-sectional survey	Medical	Questionnaire	Questionnaire	Quantitative	Statistical with software	Content experts were considered statistically superior to non-experts: building knowledge, guiding learning process, achieving cognitive learning, generating learning goals, and motivating self-study. The study was based on subjective data.
21	Peets et al. 2010 Canada	Influence of tutor expertise on tutorial process	Analytic	Randomised trial	Medical	Questionnaire	Questionnaire	Quantitative	Statistical with software	Process experts significantly performed better in all items of tutorial process than content experts. The study was based on self-rated questionnaire
22	Imafuku 2007 Australia	To study students and tutor participation in PBL	Exploratory	Ethnographic Case study	Medical Year 3	(1) Students' follow-up interview (2)	(1) Interview transcript (2) tutorial talk	Mixed methods FDA and descriptiv	Manual coding	A gap between practice and curriculum was detected: facilitator was directive and used IRE discourse patterns. Students felt that tutor's role in

		tutorial discussion		1 PBL group		audiotaping tutorial talk (3) participant observation	transcript (3) participant observation notes	e statistics	and descriptive statistics	the tutorial was helpful for their learning. Weakness: only one group was analysed.
23	Imafuku et al. (2014) Gifu University Japan	To explore learning experiences of interdisciplinary learning in PBL setting	Exploratory study	Ethnographic case study. Three PBL groups	Medical year 3; Dentistry year 3; Pharmacy year 3; Nursing year 2 Occup. health year 2; Physio year 2	(1) Video-recoding of PBL sessions (2) Students' reflective e-portfolio on learning experiences	(1) Video transcripts (2) reflective journal transcripts	Discourse analysis (video transcripts) and grounded theory (reflective journal)	Manual coding	Analysis showed elaboration and co-construction; questions and answers; confirmation/clarification request; agreement and disagreement. There was evidence of development of professional identity. Analysis technique was time consuming and laborious.

24	Imafuku 2012 Private University Japan	To explore Japanese first-year PBL learning processes	Exploratory study	Ethnographic case study one group of four students	Medical year 1; Nursing year 1; Pharmacy year 1 (two students)	(1) Video-recording of PBL sessions (2) participant observation (3) post-PBL session students' interview	(1) Video-recording transcripts (2) transcripts of interviews (3) field notes	(1) Discourse analysis (DA) for video transcript; (2) grounded theory (GT) for interview transcripts	Manual coding	The learning moves included elaboration, co-construction, knowledge negotiation, and shared understanding. Discourse participation was uneven. Factors affecting learning process included (1) prior learning experience and fear about communication; (2) identity as a marker of professionalism and motivation; (3) students' perceptions of learning in PBL and social relationship with peers and positioning in the group. Generalisability was low.
25	Papinczak et al. 2009	To explore facilitation techniques	Exploratory case study	Case study	Medical	(1) Students' written	(1) Students' written	Grounded theory	Manual coding	Themes: (1) tutor and students were confused about the role; (2) sensitive issue

	Australia	of tutors as perceived by students				reports (2) in-depth interviews of two students out of 295 (42.4% of the cohorts) who participated.	reports (2) transcripts of the in-depth interviews.			management. There was variation in the effectiveness of dealing with sensitive issues; (3) tutor style. Tutors used directive and non-directive approaches. Students reported lack of balance in scaffolding learning with some tutors dispensing information and some others leaving students to wander astray. Only students' views were reported; reasons for tutor behaviours are unknown. Analysis technique was laborious and error prone.
26	Gilkison 2003 UK	Exploration of techniques used by medical and	Exploratory	Case study	Medical	(1) Participant observation (2)	(1) Participant observation notes	Mixed methods Descriptive and	Manual coding and descriptive statistics	Tutors from the two backgrounds used the same techniques to (1) raise students' awareness; (2) facilitate group process; and

		non-medical facilitators				Audiotaping tutorial talk (3) semi-structure interviews of students and tutors	(3) tutorial talk transcript (3) interview transcripts	grounded theory		(3) direct students' learning. Medical tutor questioned the students; non-medical tutors wanted students to question each other. Non-medical tutors used more process facilitation. Sample size was small. Incomparable groups were studied. Analysis was laborious and error prone. No full transcripts were analysed.
27	Yoshioka et al. 2005 Taiwan	To evaluate the effects of intervention to facilitate case-based problem	Experimental Intervened (n= 89) Non-intervened (n = 95)		Medical	Interventions: (1) problem finding lecture; (2) encouragement; (3)	Outcome measures: (1) number and categories of	Quantitative	Statistical comparing problems extracted by both groups	Intervened group generated significantly more problems than non-intervened group. The intervened group was able to generate more questions in a greater number of specified categories.

		finding in PBL tutorial				nonverbal reinforce ment	problems extracted			
28	Chng et al. 2011 Singapor e	Influence of tutor behaviours on PBL process	Analytic	Observati onal	Health science s year 2 (n = 223); tutors (n = 7)	(1) Validated tutor behaviour questionn aire (2) concept recall test	(1) Questionn aire scores; (2) recall test scores	Quantitati ve	Statistics with software	(1) Social congruence had most effect on students' learning; (2) cognitive congruence and subject matter expertise did not affect learning. Self-reported with questionnaire. No sample size calculation was carried out.
29	Davis et al. 1994 USA	To investigate whether highly focused case could remove the effect of content expertise on	Analytic	Observati onal	Medical	(1) Participan t observati on of group interactio ns (2) Students' rating of	(1) Students-tutors' interactio n scores (2) Students' experienc e scores	Quantitati ve	Statistical with group comparison	(1) There were more tutor-directed activities in content expert group, but 62% of the activities were student directed; (2) there was no difference in students' satisfaction or achievement between the two groups. The case focus and amount of facilitator's training on the case

		students' learning and satisfaction				the experienc e				influenced tutor's behaviour and students' learning. There was no validation of interaction coding. Students' ratings were subjective.
30	Maudsle y 2008 UK	To explore how tutors conceptualis e their students' integrated learning agenda	Explorator y	Cross-sectional survey	Medical	(1) Telephon e interviews (n = 34)	(1) Interview transcript s	Qualitativ e	Grounded theory	All tutors responded. Almost all tutors identified the structure-function theme of the learning agenda. 50% differentiated structure-function theme from other three themes. Only 41% articulated the public health-based theme adequately without confusion, difficulty, or antagonism. The findings suggested PBL insecurities outside their comfort zone. The study was based on subjective data.

31	Dolmans et al. 1996 Netherland	Effects of tutors' expertise on students' performance in the context of varying curriculum structure and students' prior knowledge	Analytic	Cross-sectional survey	Medical tutorial group (n = 135); tutors (n = 119)	(1) Tutors' level of expertise self-rating (2) Students' ratings of prior knowledge and degrees of structure of curriculum materials	(1) Level of expertise (2) Level of students' prior knowledge (3) degrees of curriculum structure	Quantitative	Statistical analysis of variance	There was no significant difference between expert and non-expert groups in terms of students' performance on test of achievement. The expertise-structure and expertise-prior knowledge interaction effects were not statistically significant. Expert tutors did not compensate for the lack of structure or inadequate prior knowledge.
	De Grave et al. 1999	To define profiles of effective tutors	Theory construction	Cross-sectional survey	Medical tutors (n = 67)	Tutor intervention	Questionnaire scores	Quantitative	Descriptive statistics	Effective tutors' profiles involved elaboration, directing learning process, integration of knowledge, stimulating

	Netherland					questionnaire				interaction and individual accountability. Subjective rating.
32	Gijselaers 1997 Netherland	To study the effects of contextual factors and departmental affiliation on tutor behaviours	Analytic	Cross-sectional	Medical	(1) Students' completed questionnaire on tutor behaviours and aspects of the course	Questionnaire scores	Quantitative	Statistical analysis of correlations and variance	Tutor behaviour was related to departmental affiliation. Stability and generalisability of tutor behaviour was low. Subjective questionnaire ratings.
33	Dolmans and Wolfhagen 2005	To investigate how tutor performance, tutorial group	Theory construction	Cross-sectional survey	Medical	(1) Validate questionnaire	(1) Questionnaire scores	Quantitative	Statistical modelling	Tutor performance differed across different levels of group productivity. Group productivity differed across different levels of tutor performance. Both group productivity and tutor

	Netherland	productivity, and the effectiveness of tutorial unit interact with each other								performance had an impact on the effectiveness of PBL unit. Weakness: students' subjective ratings.
34	Kindler et al. 2009 Canada	To categorise difficult incidents and interventions that skilled tutors use in response and determine the effectivenes	Descriptive	Exploratory	Medical	(1) Semi-structured interviews	Interview transcript	Qualitative thematic	Manual coding	94 difficult incidents were recorded. Incidents were divided into two: (1) individual student and (2) group dynamics. There were 142 interventions categorised as feedback in the tutorial, feedback outside the tutorial, and student or group intervenes. Interventions worked with individual problems, but no success was observed with group dynamics problems.

		s of the responses								
35	Boelens et al. 2015 Belguim	To explore the most important task of hybrid PBL tutor in the tutorial	analytic	Cross-sectional survey	Medical	Validated questionnaire	Questionnaire scores	Quantitative	Structural modelling	1) Stimulation of active SDL enhances case quality and group functioning. (2) Stimulated collaborative learning did not affect group functioning. Subjective rating.
36	Berkel and Dolmans 2006 Netherland	(1) To investigate the effects of tutors' competencies on students' learning, group functioning, and	Analytic	Cross-sectional survey	Medical	Validated questionnaire	Questionnaire scores	Quantitative	Structural modelling	1) Stimulation of active and constructive learning, SDL, and collaborative learning by tutors enhanced the quality of the problems and group functioning. (2) The quality of the problems fostered group functioning, which had a positive effect on students' achievements. Self-reported with questionnaire.

		achievements								The hypothesis was tested only on one dataset. The model showed interrelation of variables but did not provide how.
37	Musal et al. 2004 Turkey	To investigate differences in year 1 and year 3 students' perceptions about self-study and reporting processes; (2) what were the lengths of students' self-study	Analytic	Cross-sectional survey	Medical	(1) Questionnaire on students' perceptions of self-study times and learning resources	(1) Questionnaire data on students' self-study times and learning resources	Quantitative	Statistics with regression analysis	Searching and preparatory phases of the self-study process significantly influenced breadth and depth of tutorial discussions during the reporting phase. The study was based on self-report by the students.

		times and their usage of learning resources?								
38	Moust et al. (1986) Netherland	To investigate the effect of elaboration on recall	Descriptive	Experimental	Year 2 medical (n = 24) in four groups	Free recall test and cued recall test	(1) Free recall test transcript; (2) cued recall test transcript	Quantitative	Descriptive statistics	(1) Free recall and stimulated recall scores were the same for silent and vocal students. The nature of elaboration of the silent students are unknown. The sample size was small.
39	Chng et al. 2015 Singapore	To determine the extent to which tutor behaviours influence students' learning	Analytic	Observational	Health sciences (n = 77); tutors (n = 4)	(1) Validated rating questionnaire (2) concept recall test	(1) Questionnaire scores (2) concept recall test scores	Quantitative	Statistical analysis of variance (ANCOVA)	(1) Tutor behaviour affected recall after problem analysis phase, but not after SDL and reporting phase; (2) tutor behaviours had an effect on average students, but not on stronger and weaker students. Weakness: small sample size; subjective tutor assessment;

										perceptions of the tutors unknown.
40	Das et al. 2002 USA	To investigate how students perceive tutor skills	Analytic	Cross-sectional survey	Medical	Questionnaire	(1) Questionnaire scores; (2) tutor feedback	Quantitative	Statistical analysis	Students rated the tutors as having average to outstanding skills overall. Students' and tutors' views diverged on tutor behaviours: students expected more tutors' support, but tutors emphasised self-directed learning. There is a possible religious and cultural divergence.
41	Kassab et al. 2005 Bahrain	To investigate the functions of student and faculty in PBL tutorials	Analytic	Random allocation	Medical	Questionnaire	Questionnaire scores	Quantitative	Statistical analysis	Student tutors were perceived as superior to faculty tutors in providing feedback and in understanding students' difficulties. Tutorial atmosphere, decision-making, and support for the group was better under student tutors. The ratings were subjective

										and tutors' opinions were unknown.
42	Kaufman and Holmes 1998 USA	(1) Do content experts differ from non-experts in the extent to which they dispense case knowledge? (2) Are tutors who dispense knowledge rated differently by the students?	Analytic	Cross-sectional survey	Medical	(1) Questionnaire (2) tutor's self-rating of expertise	(1) Questionnaire scores (2) expertise categories	Quantitative	Statistical measures of significance	Less than 50% of the said they almost never dispense knowledge. The tendency to present/explain case content correlated positively with tutor's level of expertise. Tutors who did not dispense knowledge rated PBL more highly. There was no difference in students rating of the two groups of tutors.

43	De Grave et al. 1999 Netherland	To validate tutor intervention profile	Theory construction	Cross-sectional survey	Medical	Questionnaire	Questionnaire scores	Quantitative	Statistical confirmatory factor analysis	Effective tutors' intervention profile were stimulation of elaboration, directing learning process, integration of knowledge, and stimulation of interaction and individual accountability.
44	Dolmans and Ginns 2005 Netherland	A short questionnaire to evaluate effect of tutors on PBL	Theory construction	Cross-sectional survey	Medical	Questionnaire	Questionnaire scores	Quantitative	Statistical confirmatory factor analysis	Effective PBL tutor behaviours supported active, contextual, and constructive learning, fostered self-directed learning, promoted collaboration among the students, were aware of their own limitations, and were very passionate about tutor role. Subjective questionnaire rating.
45	**Reznich and Werner 2004**	To investigate facilitator's encouragem	Analytic	Cross-sectional survey	Medical	(1) Student questionnaire rating	Questionnaire scores	Quantitative	Statistical with measures	Students who used online information resources rated their facilitators as more encouraging. The opposite

	USA	ent of use or non-use of internet on students' use of online information resources				of tutor on internet talk (2) Tutor questionnaire to rate internet use encouragement talk			of significance	rating was recorded for students who did not use online information resources.
46	Carlo et al. 2003 USA	Effects of motivational and cognitive factors on productivity of PBL tutorial groups	Analytic	Cross-sectional survey	Medical year 1 (n = 115)	(1) Student completed group productivity questionnaire; (2) tutor	(1) Students questionnaire score; (2) tutor questionnaire score	Quantitative	Statistical analysis	General scores for motivation, cohesion, interaction, and elaboration were higher than sponging and withdrawal. Productivity positively correlated with motivation, cohesion, interaction, and elaboration, but negatively correlated with withdrawal and

						completed group functioning questionnaire				sponging. Unproductive group scored low on motivation and elaboration while productive/highly productive group scored low on sponging and withdrawal. All female groups were productive; 3/4rth male groups were productive and 1/4rth unproductive. 52% of tutors' report said some students were not free to express their ideas in the group; in many groups, a few students dominated the discussion; all students said they shared tasks equally, but only 50% said they completed the tasks effectively. In all male groups, tutors reported problems of absenteeism and lateness.

47	Hendry et al. 2003 Australia	To explore most common PBL group problems in a medical GEM programme, and those that hinder students' learning	Descriptive	Cross-sectional survey	Medical year 1 and 2 students and tutors	(1) Students' questionnaire; (2) tutor questionnaire	(1) Questionnaire scores (2) Questionnaire scores	Quantitative	Descriptive statistics	Five most common overserved were psychosocial, disorganised tutorial process, quite student, lateness/absenteeism, and dominant student. Five least common were frustration with tutor's lack of content expertise, personality clash, superficial learning, group shortcuts, bullying. Most common problems that hindered learning were dominant students, disorganised tutorial process, group/students expressing frustration about tutor's lack of content expertise, personality clashes, superficial study of the problem, group rushing

										through the tutorial, and bullying. Students' most common responsive problems included group short-cuts, frustration with tutor, group rushes, bullying, and psychosocial disparagement. Tutor's most responsive five problems included frustration about tutor, superficial group engagement, disorganised tutorial process, group short-cuts, group rushes. Tutors were mostly well prepared for superficial group engagement, group short-cuts, group rushes, disorganised group engagement, and psychosocial disparagement. Weakness: subjective rating.

48	Visschers-Pleijers et al. (2004) Netherland	(1) To learn about knowledge elaboration; (2) To pilot analysis technique of Van Boxtel	Exploratory	Observational study	Medical: Year 1 (n = 2); year 2 (n = 1)	Videotaping of tutorial discourse	Transcript of tutorial discourse	Qualitative functional discourse analysis	Manual coding	(1) Cognitive interactions were present. (2) Cognitive interactions were analysable. Sample size was small. Analysis technique was laborious and error prone.
49	Visschers-Pleijers et al. (2006a) Netherland	(1) How much time was spent on different types of interactions? (2) How were the types of interactions distributed over the meeting?	Descriptive	Observational	Medical year 2 four tutorial groups.	Videotaping of tutorial talk	Direct software assisted analysis	Mixed methods	Analysis method: Quantitative direct analysis (without transcription) with computer software that quantify codes.	(1) Learning oriented-interaction made up 80% of interactions: cumulative reasoning (63%); exploratory questions (10%); and handling knowledge conflict (7%). Generalisability was limited.

50	Visschers-Pleijers et al. (2006b) Netherland	(1) Factors affecting effective discussions in PBL tutorial	Exploratory	Cross-sectional survey	Medical: year 1 (n = 23); year 2 (n = 25)	Audiotaping of focus group interview	Transcript of interview	Qualitative	Grounded theory analysis: qualitative-manual coding with computer software(ATLAS.ti)	Characteristics of effective discussion include: (1) explanatory discussion (asking, giving, and receiving); (2) knowledge integration and application; (3) different learning content discussion; (4) guiding and monitoring content and process. Analysis procedure was laborious and error prone
51	Visschers-Pleijers et al. (2005) Netherland	(1) To develop and validate an instruments to assess the quality of learning-oriented interactions	Analytic	Cross-sectional survey	Medical year 2 (n = 240)	Student completed Questionnaire	Questionnaire scores	Quantitative	Quantitative: factor and regression analysis	Exploratory questions and cumulative reasoning explained together 26% of the variance of the tutorial group productivity. Self-reported questionnaire

52	Visschers-Pleijers et al. (2005) Netherland	To measure students' perspectives on the occurrence and desirability of three interaction types in PBL tutorial	Descriptive	Cross-sectional survey	Medical students year 2 (n = 240)	Questionnaire	Questionnaire scores	Quantitative	Descriptive statistics	Three interaction types were present: exploratory reasoning had lowest score; cumulative reasoning had the highest score; and handling of conflicts was intermediate. The desirability scores for exploratory questions and cumulative reasoning were higher than occurrence scores, suggesting a need for improvement. Weakness: subjective self-reporting.
533	Ahmed (2014) Pakistan	To determine (1) frequency of conflicts; and (2) influence of	Analytic	Cross-sectional survey	Medical (n = 100) Tutor (n = 17)	Student completed questionnaire	Questionnaire scores	Quantitative	Statistical analysis	Conflicts: (1) dominancy; (2) personality clash; (3) quite students; (4) lack of commitment; (5) lateness; (6) absenteeism. Subjective rating.

		conflicts on learning								
54	Blankenstein et al. (2013) Netherland	(1) To assess the effect of elaboration on students' recall; (2) Effect of elaboration on students' cognitive ability	Analytic	Experimental	Health sciences (n = 67)	(1) Pre-intervention test; (2) Post-intervention test	Transcript tests	Quantitative	Statistical: One-Way ANOVA	Pre-test and Post-test recall. There was no difference in recall between the two groups. High-ability students outperformed low-ability students on recall tests. Data was small-sized. There was no blinding information and no randomisation information.
55	Hmelo-Silver and Barrows (2008)	(1) How is CKB achieved, and what are the characteristics? (2) How does the	Descriptive	Observational	Medical One group (n = 5) and master facilitator	Video-recording of tutorial discussion	Transcripts of tutorial talk	Functional discourse analysis	Mixed methods (manual coding and quantitative reporting)	(1) Group engaged in CKB; (2) The facilitator guided the students by asking open-ended metacognitive questions; (3) Students asked high-level questions and co-constructed ideas.

	New Jersey USA	facilitator provide affordances for CKB discourse? (3) What characterises the interactions within the group: between the facilitator and the students and among the students?								Dataset was small sized (one PBL group and two sessions). Manual coding was time consuming and error prone. There was low generalisability of results because the tutor was an expert.
56	AlHaqwi (2014)	1) Explore students' opinions	Analytical	Cross-sectional survey	Medical (n = 174)	Questionnaire	Questionnaire scores	Quantitative	Descriptive & Chi square	PBL developed (70%) (1) students' knowledge; (2) presentation skills; (3) team-

	Saudi Arabia	about PBL learning outcomes. (2) Explore students' views about the role of tutors and qualities of effective tutors								work abilities; (4) accepting peer criticisms. 75% indicated that tutor role was essential; 58% said the role was clear and well defined; 63% preferred peers as tutors; 80% preferred content and process experts. Self-reported with questionnaire. There was no information on questionnaire validation.
57	AL-Drees et al. (2015) Saudi Arabia	(1) To evaluate students' perceptions of PBL sessions	Analytic	Cross-survey	Medical year 1 (n = 167) and year 2 (n = 108)	Questionnaire	Questionnaire scores	Quantitative	One-Way ANOVA	PBL was helpful (1) to understand basic science concepts; (2) increase knowledge of basic sciences; (3) encouraging SDL and collaborative learning; (4) improving decision making. 54.5% said students were not well trained for the PBL. 25.1%

										said tutors were well prepared. Learning resources included (1) internet 93.1%; (2) lecture notes 76.7%; (3) books 64.4%. Self-reported with questionnaire. There was no evidence of validation.
58	Diemers et al. (2015) Netherland	(1) To explore the development of and transfer of knowledge with real patients	Analytic	Experimental longitudinal Intervention: Use of real patients in PBL	Medical year 3 (n = 13; 16.25% of the cohort) all females	Audiotaping of tutorial talk	Tutorial talk transcript; Recall talk	Mixed methods	Coding with ATLAS.ti and ANOVA	(1) Diagnostic accuracy increased; (2) Case processing time decreased; (3) Students used biomedical and clinical knowledge during diagnostic reasoning. The quality of pathophysiological explanation increased. The results were inferior for transfer. Sample size was small. It may not be generalisable because of the sample size and the gender included.

										It was difficult to establish the cause-effect relationship because there was no control group. Power of the sample was not calculated.
59	Romito et al. (2011) USA	To assess the relationship between biomedical science content acquisition and students' level of PBL group interaction	Descriptive study	Observation study	Dental students year 1 and 2	Students and facilitators' Questionnaire Recall and Application tests quizzes	Application test scores Recall test scores Questionnaire rating scores	Quantitative	Descriptive statistics	Year-one had higher assessment scores. Recall scores were higher than application scores. Assessment scores were correlated with interaction scores for year 1 group, but not for year 2. There was no relationship between knowledge application score and students' interaction scores. The study was based on self-reported data. The

										questionnaire used was not validated.
60	Da Silva and Dennick 2010 UK	To explore the feasibility of an analytic tool	Analytic Exploratory	Exploratory Case study	Medical one group GEM year 1 (n = 7) Tutor (n = 1)	Video-recording of tutorial talk	Video transcripts	Quantitative Corpus linguistics	Wmatrix software	Compared frequencies of technical words and of reasoning, explaining, and questioning episodes between sessions. Findings were related to the objectives of each of the tutorial phase. Sample size was small.
61	Gukas et al. (2010) UK	To explore verbal and non-verbal indices of learning in medical students' PBL tutorials	Descriptive	Observational	Medical year 1 (n = 50)	Participant observation with observation tool	Records of observations	Quantitative	Descriptive	Exploratory questioning, cumulative reasoning, and handling of conflicts were present. Common non-verbal indices of learning were silence.

62	Yew et al. (2011) Singapore	To examine whether the learning in PBL tutorial is cumulative	Analytic	Experimental cohort	218 health science students	Pre-intervention essay test; pre- and post-intervention recall test	Essay test scores Recall test scores	Quantitative	Statistical modelling	There was evidence that knowledge at previous phase of PBL affected learning in subsequent phases.
63	Collard et al. (2009) Belgium	To explore maturation increase in biomedical reasoning capacity in comparison with factual knowledge retention	Analytic	Cross-sectional	Medical students year 3 (n = 35); year 4 (n = 20); year 5 (n = 25) Year 6 = 24	Factual recall test Reasoning test	Factual recall test scores Reasoning test scores	Quantitative	Statistical analysis ANOVA	Years 5 and 6 had higher scores on reasoning test than years 3 and 4. Year 3 had higher scores on factual test than years 4–6. A positive correlation between factual test scores and reasoning test scores was noted in years 3 and 4. In each group, the ascertainment degree scores were higher for correct than for incorrect responses.

										Reasoning was present early in the curriculum, and it increased with curriculum years while factual knowledge decreased.
64	De Leng and Gijlers (2015) Netherland	To examine how collaborative diagramming affects discussion and knowledge construction when learning complex basic science topics in	Descriptive	Cross-sectional survey	Medical year 1 students (n = 70) in seven tutorial groups	Questionnaire video-recording of tutorial discussions Informal focus group interviews with tutors	Questionnaire score Interview transcripts Global direct video analysis Transcript of tutors video stimulated recall		Video was analysed with a software program TRANSANA 2013.	Students and tutors felt that CM promoted knowledge construction and their discussions. Students reported that CM helped them to structure knowledge and to relate one concept to the other. Tutors felt that CM stimulated interactions and fostered focus and detail of discussions. Weakness: self-reported questionnaire

		medical education								
65	Verones e et al. (2013) USA	Randomised controlled trial of the use of concept mapping to enhance learning in medical PBL	Analytical	Experime ntal study	Year 1 medical and dental students (n = 172)	Student questionn aire survey Audiotape Tutor interviews Exam scores	Student questionn aire scores Interviews transcript s Exam scores	Mixed methods	Statistical analysis of survey scores and manual coding of interview transcripts	82.6% planned to use CM in the future. More students in the CM planned to continue using it than in no CM group (p = 0.02); Qualitative analysis showed association of CM knowledge integration of physiological mechanisms, challenging students, knowledge of material and identification of knowledge gap. CM group performed statistically better than group without CM. Data was subjective report.
66	McLean 2003	Students' and facilitator	Analytic	Cross-sectional survey	Medical students	Facilitator questionn aire	Facilitator questionn	Mixed methods	Descriptive statistics	Many facilitators experienced problems in the tutorials for not giving content knowledge. This

	South Africa	perceptions of facilitation skills and roles in year medical year 1 PBL				Students open comments	aire scores Students comments text		and narratives	tended to improve with PBL experience. Many facilitators saw that students as colleagues are willing to be role models, but not mentors. The comments of the students were generally positive for the facilitators, but were critical of inadequate facilitators in terms of curriculum implementation. Subjective reports.
67	Faidley et al. (2000) USA	To explore whether two assessment methods would provide more viable means of assessing group	Observation	Case study	Year-one medical students (n = 20)	Learning team survey rating Non-participant rating of group behaviours on video	Results of ratings	Mixed methods	Mixed descriptive statistics and narrative	The salient behaviours of the tutorial participants fell into four categories: (1) teacher dominated; (2) student negotiated; (3) single student dominant or cautiously interactive; and (4) male dominant or aggressively interactive. Students' report in three groups diverged from

		processes in four PBL tutorials				segments (7 minutes at 6 intervals; total 42 minutes.					non-participants’ observation. Non-participant video rating explained students’ subjective assessment. Weakness: Small- sized data. The observation did not span the entire length of group interactions

Appendix 2. Research Participant Information Sheet

1. Research Project Title

Collaborated knowledge construction in problem-based learning – A corpus-based study

2. Introduction

We are a group of researchers consisting of two PhD students and a supervisor from the School of Medicine, Medical Education Unit, University of Nottingham.

3. Invitation paragraph

You are being invited to participate in a research project. Before you decide whether to participate in the project, it is important for you to understand why the research is being conducted and what it will involve. Please take some time to read the following information carefully. You are free to ask us questions if there is anything that is not clear or if you would like more information. Please take time to decide whether you wish to take part in the project. Thank you for reading.

4. What is the purpose of the project?

The study is a research project conducted by two postgraduate students for the PhD degree programme at University of Nottingham. The study seeks to provide information on how problem-based learning (PBL) students construct knowledge together when they talk in the tutorials. It also seeks to provide information about how the contributions of the facilitators help the students construct knowledge. It is anticipated that the study will provide useful information and ideas on how the process of problem-based learning can be improved.

5. Why have I being chosen?

You have been invited for the study as the member of year one graduate-entry medical students in Nottingham Medical School Derby, with a view that you might be interested in taking part in the study. This does not mean that you must take part.

6. Do I have to take part in the project?

No. Your participation in the project is entirely voluntary. You are not obliged to take part. You have been approached as a member of year one class in graduate-entry medical school with the view that you might be interested in taking part. This does not mean that you necessarily must take part. If you do not wish to take part, you do not have to give any reason, and we would not contact you again. Similarly, if you do agree to take part, you are free to withdraw at any time during the project if you change your mind without prejudice or negative consequences.

7. What do I have to do if I take part?

If you are interested in taking part in the project, you will be asked to complete the attached response slip and return it to the researcher who will visit your tutorial group. You will be required to complete a consent form for the audio and video recordings; for the transcription of the audio recordings; and for the use of the transcripts for research purposes. You will be required to be audio- and video-recorded during the tutorials. If you decide not to participate in the study, I would like you to return the slip to indicate this, but no further contact will be made.

8. What are the possible disadvantages and risks of taking part?

The only disadvantage of taking part in the study relates to the fact that you might not feel comfortable being audio- and video-recorded.

9. What are the possible benefits of taking part?

The study may not benefit you directly. However, the study may generate useful information that would help us provide advice that might improve the process of problem-based learning and its facilitation. By participating in the research, you will be contributing to the progress of problem-based learning curriculum in Derby Graduate Entry Medical School.

10. What if something goes wrong?

If you have wish to raise a complaint about the researcher or the process of the research, you can contact the research supervisor, Prof. Reg Dennick, via the contact below. However, if you feel that your complaint is not dealt with to your satisfaction, you can contact the school's research ethics coordinators via the contact details provided below.

11. Will my taking part in the project be kept confidential?

The audio and video recordings and transcripts will be kept strictly confidential. You will not be identifiable in the transcripts or any reports or publications of the study. The video and audio materials and the transcripts will be stored in password-protected files in the computers in the supervisor's office and will only be accessible to the researchers and the supervisor. These materials will not to be used in any conference, presentation, or educational materials without prior individual permission from the participants for such use. Audio recordings will be sent for transcription in password-protected files. The responsibility for audio and video recordings will be given to the students who will be advised to stop the recordings (audio and video) when any confidential or sensitive issue will be discussed. The recordings will not be shared with the student assessors or evaluators of the facilitators.

12. What type of information will be sought from me, and why is the collection of this information relevant for achieving the research project's objectives?

The conversation of the students and the facilitators will be recorded during the tutorial sessions. The talk of the participants is what the research is interested in exploring.

13. Will I be recorded, and how will the recorded media be used?

There will be audio and video recording of the talk of the participants during the tutorial sessions. The audio recording will be transcribed, and the video recording will be used to identify individual's talk during the tutorials for the purpose of analysis only.

14. What will happen to the results of the research project?

The results of the research will be presented as a PhD thesis and published in a journal of medical education. You will not be identified in any report or publication.

15. Who has ethically reviewed the project?

The research has been ethically approved by the education school's ethics procedures of University of Nottingham. The University of Nottingham Research Ethics Committee monitors the application and delivery of the university's ethics review procedure across the university.

16. Contacts for further information

Olukayode Matthew Tokode, School of Medicine, Medical Education Unit, University of Nottingham, UK. Tel: +44(0)7438238918, email: mcxot2@nottingham.ac.uk

Prof. Reg Dennick, School of Medicine, Medical Education Unit, University of Nottingham, UK. Tel: +44(0)115523, email: reg.dennick@nottingham.ac.uk

School of Education Research Ethics Coordinator (educationresearchethics@nottingham.ac.uk)

Thank you for taking part in this research.

Appendix 3. Research Participant Consent

Project title ……………………………………………………………….

Researcher's name ……………………………………………………….

Supervisor's name..……………………………………………………

- I have read the Participant Information Sheet, and the nature and purpose of the research project has been explained to me. I understand and agree to take part.
- I understand the purpose of the research project and my involvement in it.
- I understand that I may withdraw from the research project at any stage, and that this will not affect my status now or in the future.
- I understand that while the information gained during the study may be published, I will not be identified and my personal results will remain confidential. (If other arrangements have been agreed in relation to identification of research participants, this point will require amendment to accurately reflect those arrangements.)
- I understand that I will be audio/video recorded during the tutorials.
- I understand that data will be stored as password-protected files in the designated computer in the supervisor's office at the

School of Medicine, University of Nottingham. The researcher and the supervisor would have access to the data for the purpose of data analysis and result writing.

- I understand that I may contact the researcher or supervisor if I require further information about the research, and that I may contact the research ethics coordinator of the School of Education, University of Nottingham, if I wish to make a complaint relating to my involvement in the research.

Signed ………………………………… (research participant)

Print name …………………………………… Date ……………

Contact details

Researcher: (…………………………………..)

Supervisor: (……………………………………)

School of Education Research Ethics Coordinator

(educationresearchethics@nottingham.ac.uk)

Appendix 4. Ethics Consent Letter for the Research

Direct line/e-mail
+44 (0) 115 8231063
Louise.Sabir@nottingham.ac.uk

The University of Nottingham

Faculty of Medicine and Health Sciences

Medical School Research Ethics Committee
Division of Therapeutics & Molecular Medicine
D Floor, South Block
Queen's Medical Centre
Nottingham
NG7 2UH

Tel: +44 (0) 115 8231063
Fax: +44 (0) 115 8231059

10th December 2009

Dr Reg Dennick
Assistant Director of Medical Education
Medical Education Unit
B Floor, Medical School
QMC Campus
Nottingham University Hospitals
NG9 2TR

Dear Dr Dennick

Ethics Reference No: D/9/2008 - ***Please quote this number on all correspondence***
Study Title: Monitoring the development of technical language and clinical reasoning in problem based learning groups of graduate entry medical students.
Lead Investigator: Dr Reg Dennick, Assistant Director of Medical Education
Co Investigators: Ana Linda Da Silva, Matthew Olukayode Tokode, PhD Students, School of Community Health Sciences.

Thank you for your letter dated 11th November 2009 requesting the following amendment to the above study:

- Addition of Matthew Olukayode Tokode, PhD Student.

These have been reviewed and are satisfactory and the study amendment is approved.

Approval is given on the understanding that the Conditions of Approval set out below are followed.

Conditions of Approval

You must follow the protocol agreed and any changes to the protocol will require prior Ethic's Committee approval.

This study is approved for the period of active recruitment requested. The Committee also provides a further 5 year approval for any necessary work to be performed on the study which may arise in the process of publication and peer review.

You promptly inform the Chairman of the Ethic's Committee of

(i) Deviations from or changes to the protocol which are made to eliminate immediate hazards to the research subjects.

(ii) Any changes that increase the risk to subjects and/or affect significantly the conduct of the research.

(iii) All adverse drug reactions that are both serious and unexpected.

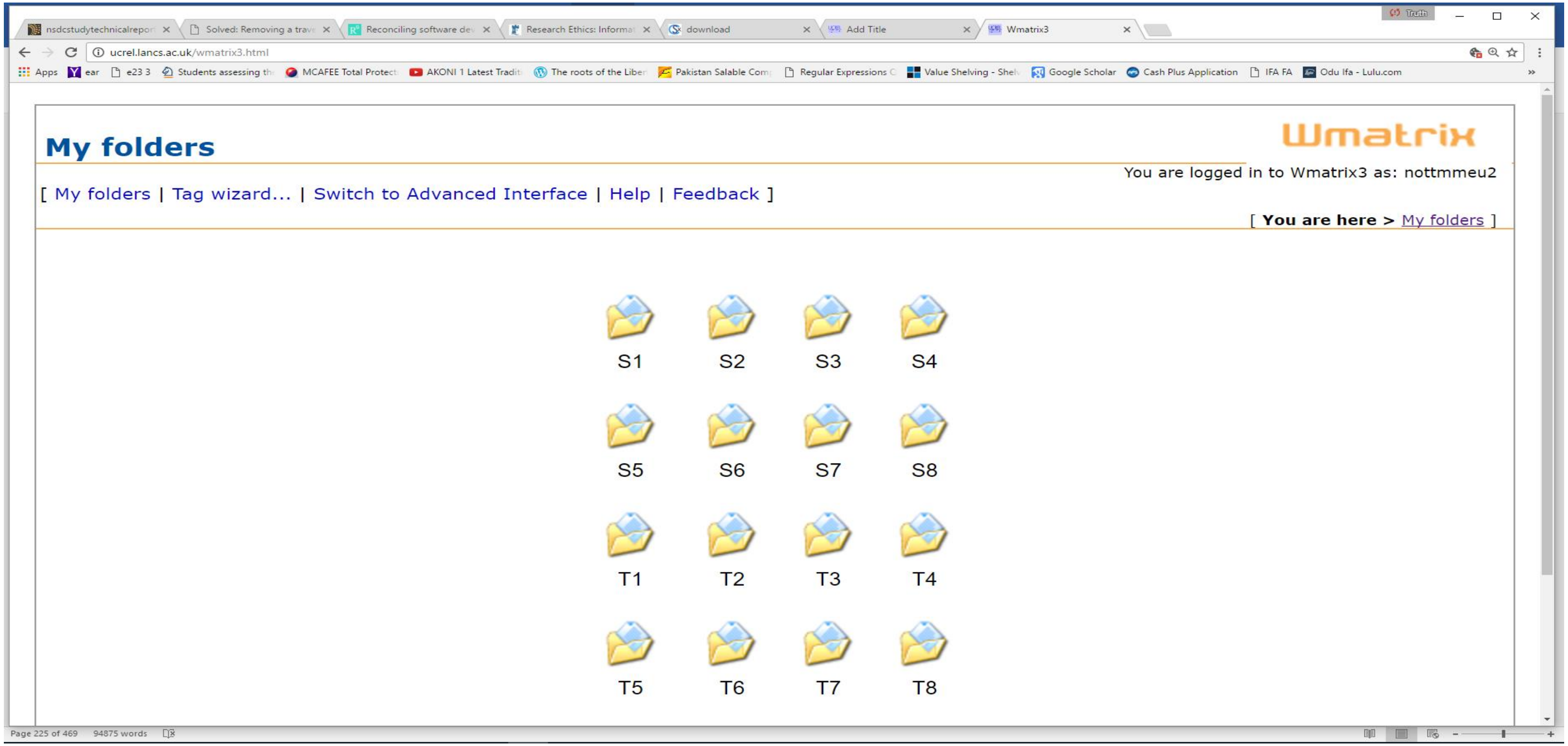
(iv) New information that may affect adversely the safety of the subjects or the conduct of the study.

(v) The attached End of Project Progress Report is completed and returned when the study has finished.

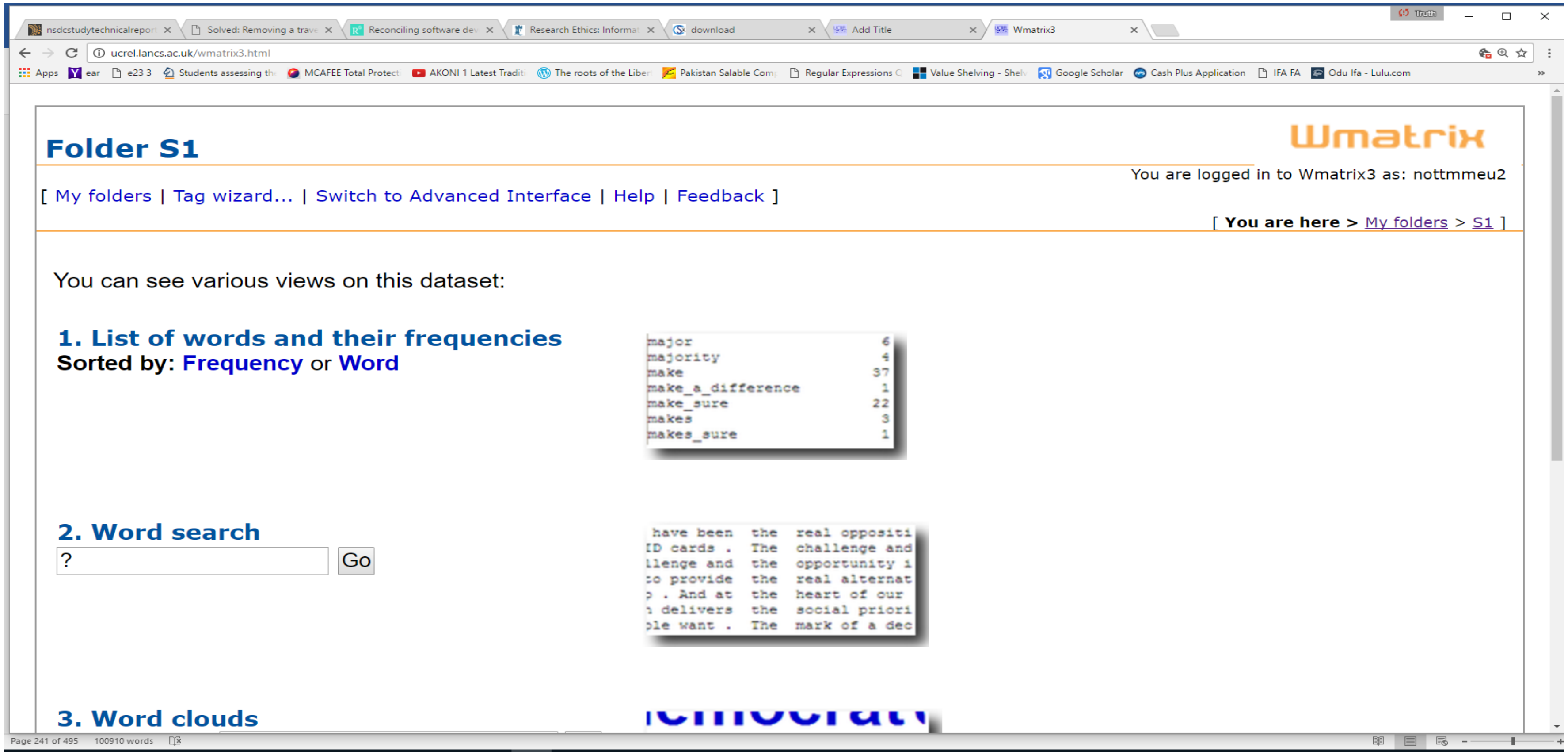
Yours sincerely

Professor R C Spiller
Chairman, Nottingham University Medical School Research Ethics Committee

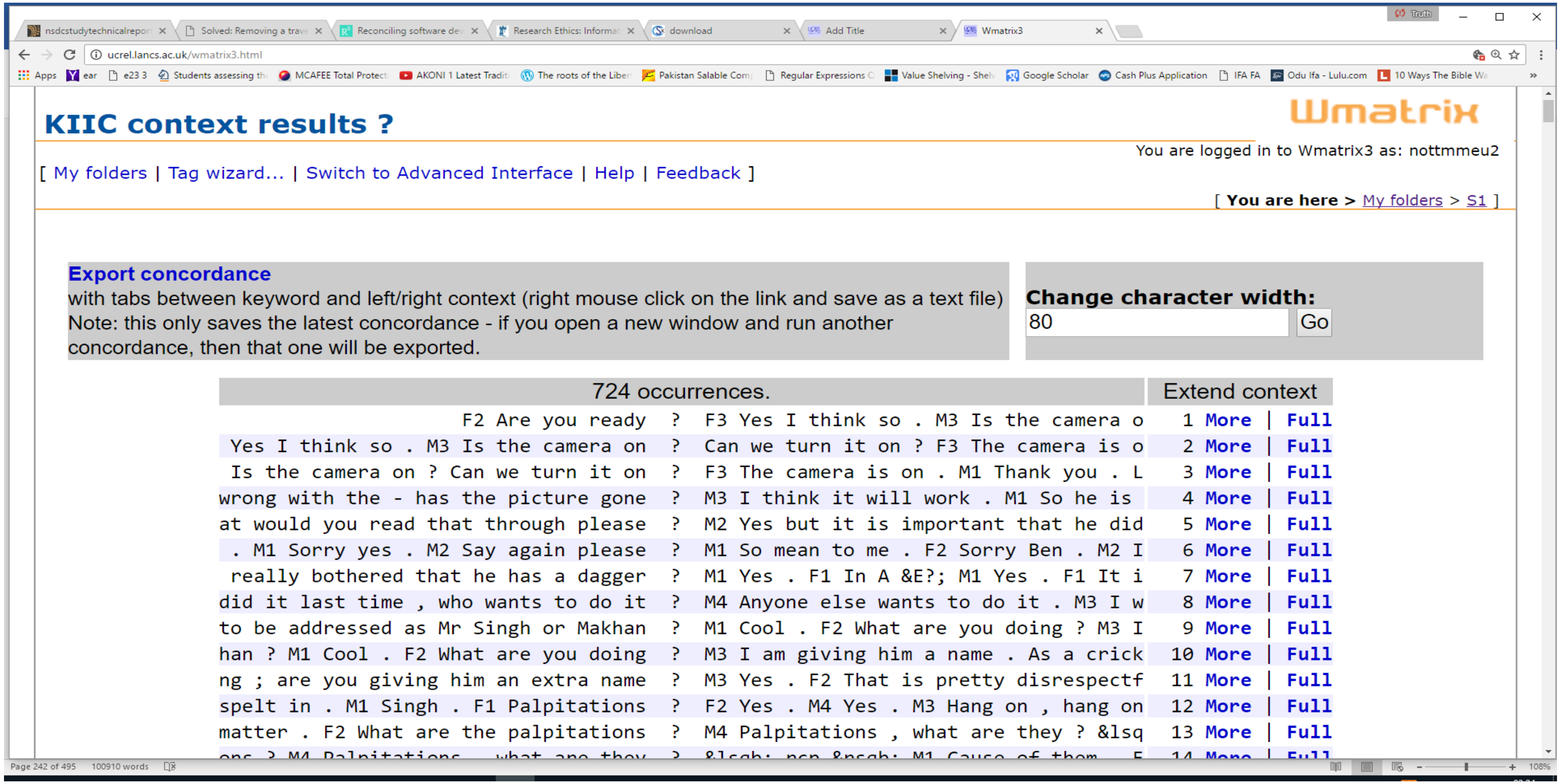
Appendix 5. Uploaded text files on Wmatrix 3 software



Appendix 6. Question mark search on Wmatrix 3 simple interface



Appendix 7. Concordance lines of question marks on Wmatrix 3



Wmatrix

KIIC context results ?

You are logged in to Wmatrix3 as: nottmmeu2

[My folders | Tag wizard... | Switch to Advanced Interface | Help | Feedback]

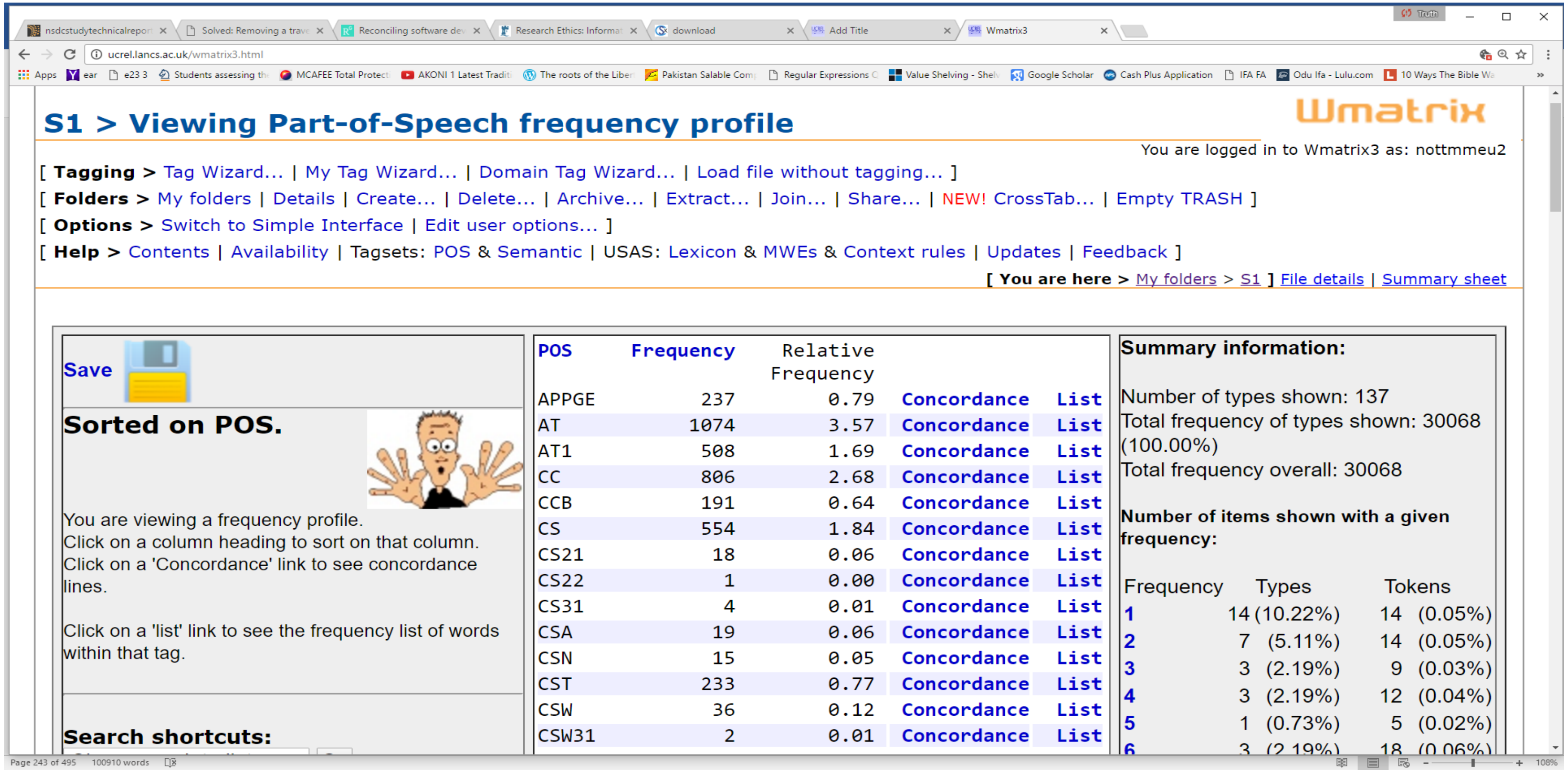
[You are here > My folders > S1]

Export concordance
with tabs between keyword and left/right context (right mouse click on the link and save as a text file)
Note: this only saves the latest concordance - if you open a new window and run another concordance, then that one will be exported.

Change character width: 80 Go

724 occurrences.			Extend context
F2 Are you ready	?	F3 Yes I think so . M3 Is the camera o	1 More | Full
Yes I think so . M3 Is the camera on	?	Can we turn it on ? F3 The camera is o	2 More | Full
Is the camera on ? Can we turn it on	?	F3 The camera is on . M1 Thank you . L	3 More | Full
wrong with the - has the picture gone	?	M3 I think it will work . M1 So he is	4 More | Full
at would you read that through please	?	M2 Yes but it is important that he did	5 More | Full
. M1 Sorry yes . M2 Say again please	?	M1 So mean to me . F2 Sorry Ben . M2 I	6 More | Full
really bothered that he has a dagger	?	M1 Yes . F1 In A &E?; M1 Yes . F1 It i	7 More | Full
did it last time , who wants to do it	?	M4 Anyone else wants to do it . M3 I w	8 More | Full
to be addressed as Mr Singh or Makhan	?	M1 Cool . F2 What are you doing ? M3 I	9 More | Full
han ? M1 Cool . F2 What are you doing	?	M3 I am giving him a name . As a crick	10 More | Full
ng ; are you giving him an extra name	?	M3 Yes . F2 That is pretty disrespectf	11 More | Full
spelt in . M1 Singh . F1 Palpitations	?	F2 Yes . M4 Yes . M3 Hang on , hang on	12 More | Full
matter . F2 What are the palpitations	?	M4 Palpitations , what are they ? &lsq	13 More | Full

Appendix 8. Parts of Speech frequency list



S1 > Viewing Part-of-Speech frequency profile

Wmatrix

You are logged in to Wmatrix3 as: nottmmeu2

[**Tagging >** Tag Wizard... | My Tag Wizard... | Domain Tag Wizard... | Load file without tagging...]
[**Folders >** My folders | Details | Create... | Delete... | Archive... | Extract... | Join... | Share... | NEW! CrossTab... | Empty TRASH]
[**Options >** Switch to Simple Interface | Edit user options...]
[**Help >** Contents | Availability | Tagsets: POS & Semantic | USAS: Lexicon & MWEs & Context rules | Updates | Feedback]
[You are here > My folders > S1 **]** File details | Summary sheet

Save

Sorted on POS.

You are viewing a frequency profile.
Click on a column heading to sort on that column.
Click on a 'Concordance' link to see concordance lines.

Click on a 'list' link to see the frequency list of words within that tag.

Search shortcuts:

POS	Frequency	Relative Frequency		
APPGE	237	0.79	Concordance	List
AT	1074	3.57	Concordance	List
AT1	508	1.69	Concordance	List
CC	806	2.68	Concordance	List
CCB	191	0.64	Concordance	List
CS	554	1.84	Concordance	List
CS21	18	0.06	Concordance	List
CS22	1	0.00	Concordance	List
CS31	4	0.01	Concordance	List
CSA	19	0.06	Concordance	List
CSN	15	0.05	Concordance	List
CST	233	0.77	Concordance	List
CSW	36	0.12	Concordance	List
CSW31	2	0.01	Concordance	List

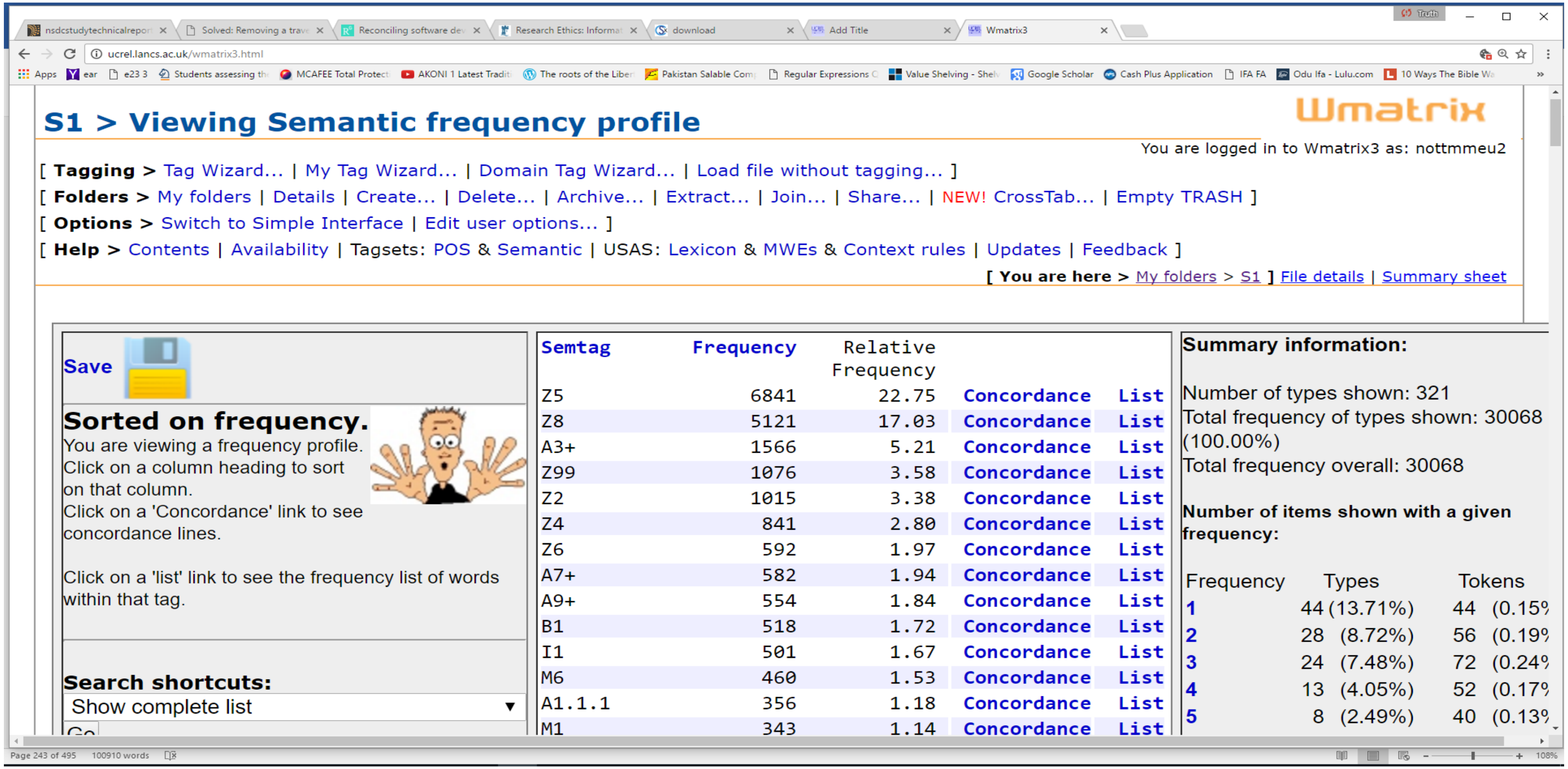
Summary information:

Number of types shown: 137
Total frequency of types shown: 30068 (100.00%)
Total frequency overall: 30068

Number of items shown with a given frequency:

Frequency	Types	Tokens
1	14 (10.22%)	14 (0.05%)
2	7 (5.11%)	14 (0.05%)
3	3 (2.19%)	9 (0.03%)
4	3 (2.19%)	12 (0.04%)
5	1 (0.73%)	5 (0.02%)
6	3 (2.19%)	18 (0.06%)

Appendix 9. Semantic tag frequency list



S1 > Viewing Semantic frequency profile

Wmatrix

You are logged in to Wmatrix3 as: nottmmeu2

[**Tagging >** Tag Wizard... | My Tag Wizard... | Domain Tag Wizard... | Load file without tagging...]
[**Folders >** My folders | Details | Create... | Delete... | Archive... | Extract... | Join... | Share... | NEW! CrossTab... | Empty TRASH]
[**Options >** Switch to Simple Interface | Edit user options...]
[**Help >** Contents | Availability | Tagsets: POS & Semantic | USAS: Lexicon & MWEs & Context rules | Updates | Feedback]
[You are here > My folders > S1 **]** File details | Summary sheet

Save

Sorted on frequency.
You are viewing a frequency profile.
Click on a column heading to sort on that column.
Click on a 'Concordance' link to see concordance lines.

Click on a 'list' link to see the frequency list of words within that tag.

Search shortcuts:
Show complete list

Semtag	Frequency	Relative Frequency		
Z5	6841	22.75	Concordance	List
Z8	5121	17.03	Concordance	List
A3+	1566	5.21	Concordance	List
Z99	1076	3.58	Concordance	List
Z2	1015	3.38	Concordance	List
Z4	841	2.80	Concordance	List
Z6	592	1.97	Concordance	List
A7+	582	1.94	Concordance	List
A9+	554	1.84	Concordance	List
B1	518	1.72	Concordance	List
I1	501	1.67	Concordance	List
M6	460	1.53	Concordance	List
A1.1.1	356	1.18	Concordance	List
M1	343	1.14	Concordance	List

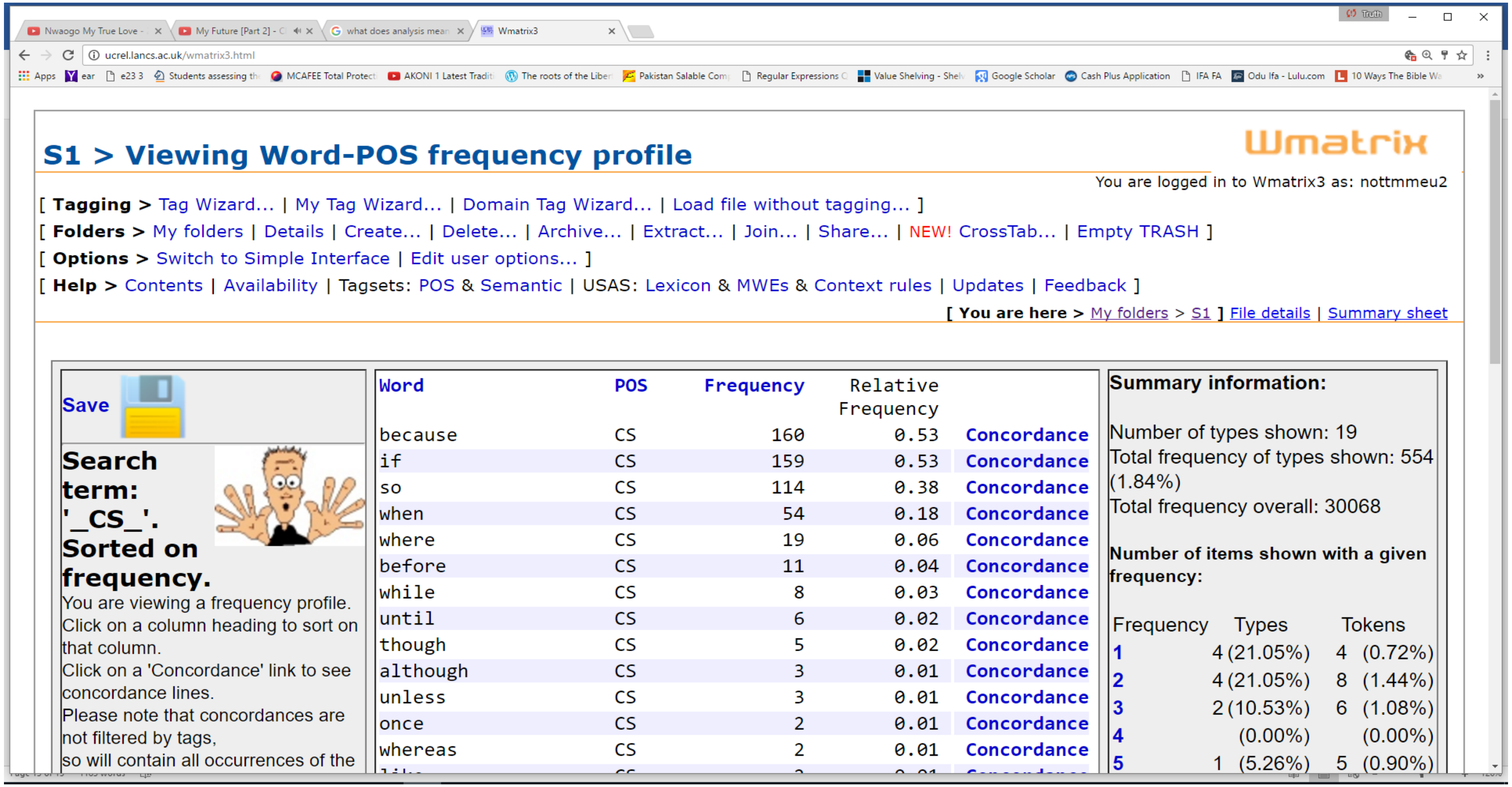
Summary information:

Number of types shown: 321
Total frequency of types shown: 30068 (100.00%)
Total frequency overall: 30068

Number of items shown with a given frequency:

Frequency	Types	Tokens
1	44 (13.71%)	44 (0.15%
2	28 (8.72%)	56 (0.19%
3	24 (7.48%)	72 (0.24%
4	13 (4.05%)	52 (0.17%
5	8 (2.49%)	40 (0.13%

Appendix 10. List of words and their frequencies on part-of-speech domain



S1 > Viewing Word-POS frequency profile

Wmatrix

You are logged in to Wmatrix3 as: nottmmeu2

[**Tagging >** Tag Wizard... | My Tag Wizard... | Domain Tag Wizard... | Load file without tagging...]
[**Folders >** My folders | Details | Create... | Delete... | Archive... | Extract... | Join... | Share... | NEW! CrossTab... | Empty TRASH]
[**Options >** Switch to Simple Interface | Edit user options...]
[**Help >** Contents | Availability | Tagsets: POS & Semantic | USAS: Lexicon & MWEs & Context rules | Updates | Feedback]
[**You are here >** My folders > S1] File details | Summary sheet

Save

Search term: '_CS_'. Sorted on frequency.

You are viewing a frequency profile.
Click on a column heading to sort on that column.
Click on a 'Concordance' link to see concordance lines.
Please note that concordances are not filtered by tags,
so will contain all occurrences of the

Word	POS	Frequency	Relative Frequency	
because	CS	160	0.53	Concordance
if	CS	159	0.53	Concordance
so	CS	114	0.38	Concordance
when	CS	54	0.18	Concordance
where	CS	19	0.06	Concordance
before	CS	11	0.04	Concordance
while	CS	8	0.03	Concordance
until	CS	6	0.02	Concordance
though	CS	5	0.02	Concordance
although	CS	3	0.01	Concordance
unless	CS	3	0.01	Concordance
once	CS	2	0.01	Concordance
whereas	CS	2	0.01	Concordance

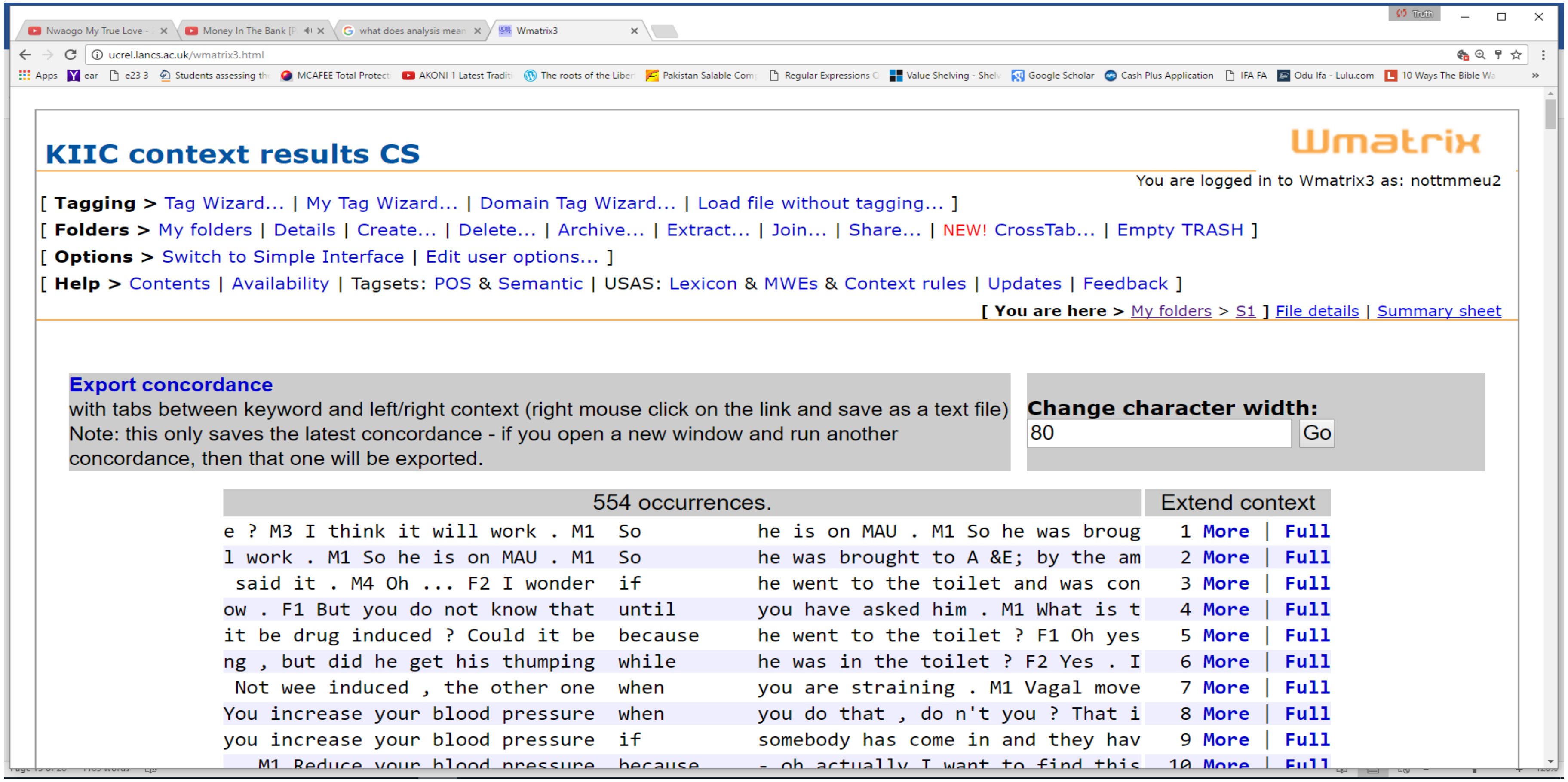
Summary information:

Number of types shown: 19
Total frequency of types shown: 554 (1.84%)
Total frequency overall: 30068

Number of items shown with a given frequency:

Frequency	Types	Tokens
1	4 (21.05%)	4 (0.72%)
2	4 (21.05%)	8 (1.44%)
3	2 (10.53%)	6 (1.08%)
4	(0.00%)	(0.00%)
5	1 (5.26%)	5 (0.90%)

Appendix 11. Keywords in the context result of CS



KIIC context results CS

Wmatrix

You are logged in to Wmatrix3 as: nottmmeu2

[**Tagging >** Tag Wizard... | My Tag Wizard... | Domain Tag Wizard... | Load file without tagging...]
[**Folders >** My folders | Details | Create... | Delete... | Archive... | Extract... | Join... | Share... | NEW! CrossTab... | Empty TRASH]
[**Options >** Switch to Simple Interface | Edit user options...]
[**Help >** Contents | Availability | Tagsets: POS & Semantic | USAS: Lexicon & MWEs & Context rules | Updates | Feedback]
[**You are here >** My folders > S1] File details | Summary sheet

Export concordance
with tabs between keyword and left/right context (right mouse click on the link and save as a text file)
Note: this only saves the latest concordance - if you open a new window and run another concordance, then that one will be exported.

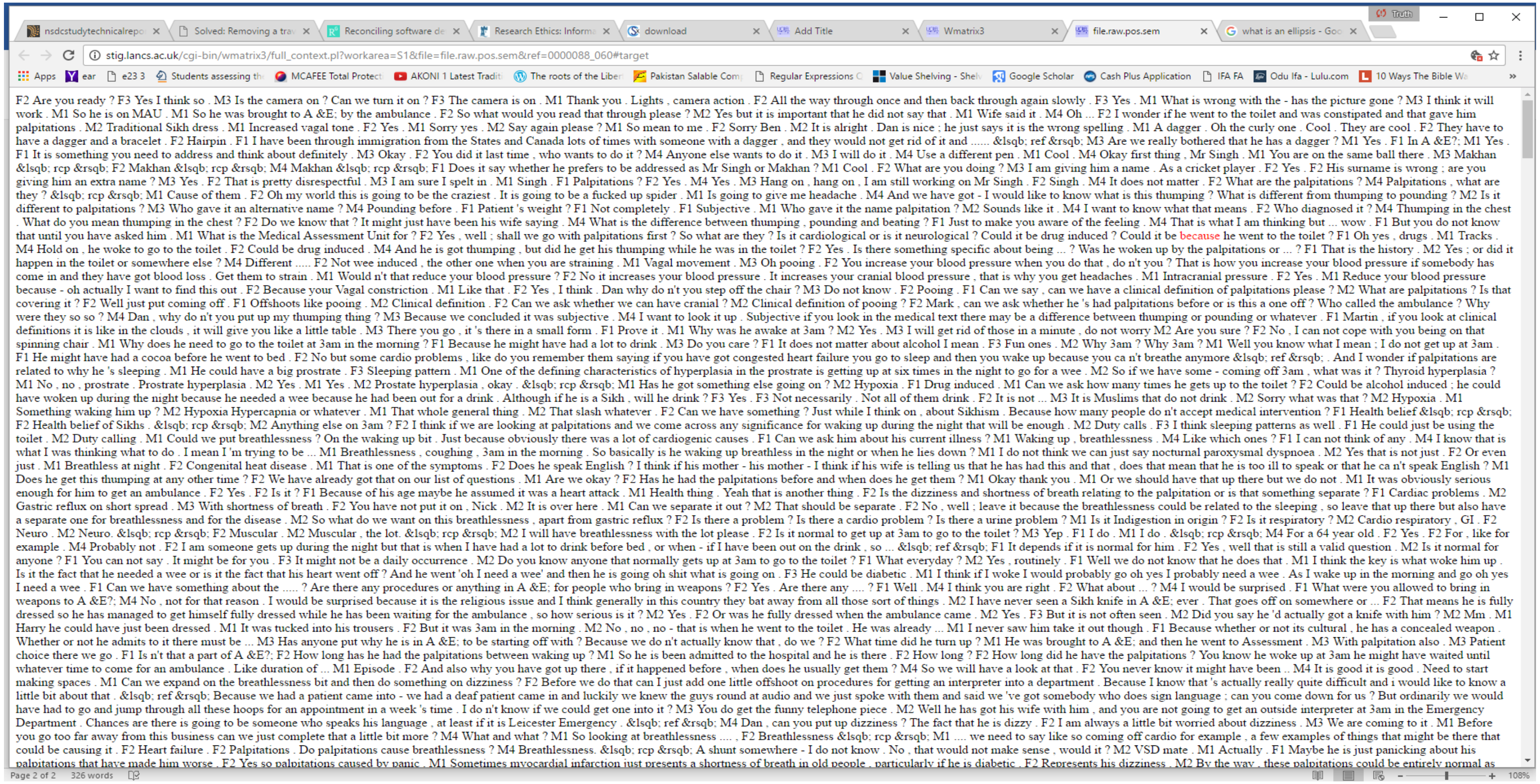
Change character width:
80 Go

554 occurrences.			Extend context
e ? M3 I think it will work . M1	So	he is on MAU . M1 So he was broug	1 More | Full
l work . M1 So he is on MAU . M1	So	he was brought to A &E; by the am	2 More | Full
said it . M4 Oh ... F2 I wonder	if	he went to the toilet and was con	3 More | Full
ow . F1 But you do not know that	until	you have asked him . M1 What is t	4 More | Full
it be drug induced ? Could it be	because	he went to the toilet ? F1 Oh yes	5 More | Full
ng , but did he get his thumping	while	he was in the toilet ? F2 Yes . I	6 More | Full
Not wee induced , the other one	when	you are straining . M1 Vagal move	7 More | Full
You increase your blood pressure	when	you do that , do n't you ? That i	8 More | Full
you increase your blood pressure	if	somebody has come in and they hav	9 More | Full
M1 Reduce your blood pressure	because	- oh actually I want to find this	10 More | Full

Appendix 12. KWIC results of CS exported to excel

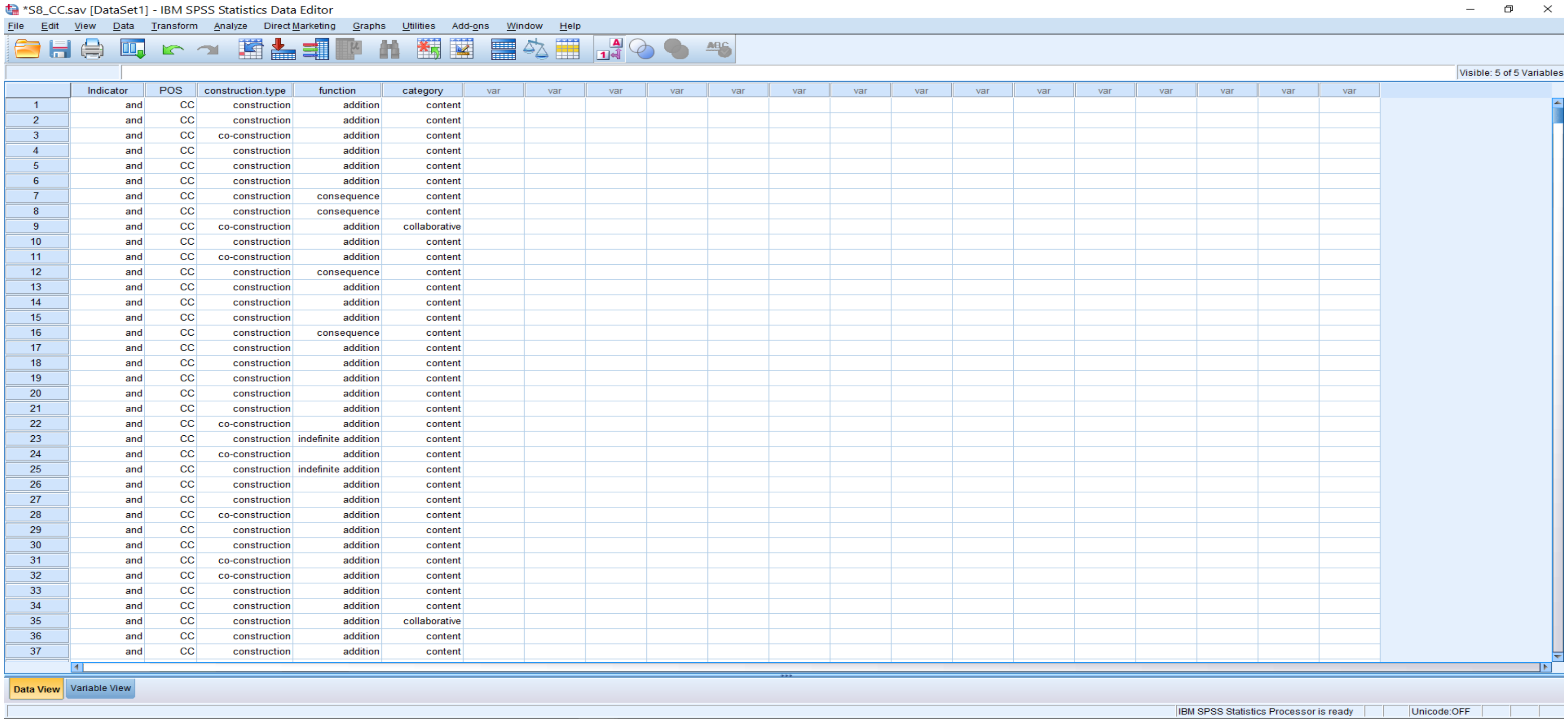
	Left co-text	keyword	Right co-text	keyword function	consruction type
2	said it. M4 Oh ... F2 I wonder	if	he went to the toilet and was con	possible situation	autonomous
3	it be drug induced? Could it be	because	he went to the toilet? F1 Oh yes	reason	autonomous
4	you increase your blood pressure	if	somebody has come in and they have	conditional	autonomous
5	. M1 Reduce your blood pressure	because	- oh actually I want to find this	incomplete	autonomous
6	lly I want to find this out. F2	Because	your Vagal construction. M1 Like	reason	co-construction
7	ou put up my thumping thing? M3	Because	we concluded it was subjective.	reason	co-construction
8	want to look it up. Subjective	if	you look in the medical text ther	conditional	autonomous
9	unding or whatever. F1 Martin,	if	you look at clinical definitions	conditional	co-construction
10	oilet at 3am in the morning? F1	Because	he might have had a lot to drink	reason	co-construction
11	like do you remember them saying	if	you have got congested heart fail	conditional	autonomous
12	go to sleep and then you wake up	because	you ca n't breathe anymore [	reason	autonomous
13	he night to go for a wee. M2 So	if	we have some - coming off 3am, w	possible situation	co-construction
14	d have woken up during the night	because	he needed a wee because he had be	reason	autonomous
15	he night because he needed a wee	because	he had been out for a drink. Alt	reason	autonomous
16	been out for a drink. Although	if	he is a Sikh, will he drink? F3	conditional	autonomous
17	ile I think on, about Sikhism.	Because	how many people do n't accept med	reason	autonomous
18	nything else on 3am? F2 I think	if	we are looking at palpitations an	possible situation	autonomous

Appendix 13. Full extension of the concordance lines in Wmatrix 3 software



F2 Are you ready ? F3 Yes I think so . M3 Is the camera on ? Can we turn it on ? F3 The camera is on . M1 Thank you . Lights , camera action . F2 All the way through once and then back through again slowly . F3 Yes . M1 What is wrong with the - has the picture gone ? M3 I think it will work . M1 So he is on MAU . M1 So he was brought to A &E; by the ambulance . F2 So what would you read that through please ? M2 Yes but it is important that he did not say that . M1 Wife said it . M4 Oh ... F2 I wonder if he went to the toilet and was constipated and that gave him palpitations . M2 Traditional Sikh dress . M1 Increased vagal tone . F2 Yes . M1 Sorry yes . M2 Say again please ? M1 So mean to me . F2 Sorry Ben . M2 It is alright . Dan is nice ; he just says it is the wrong spelling . M1 A dagger . Oh the curly one . Cool . They are cool . F2 They have to have a dagger and a bracelet . F2 Hairpin . F1 I have been through immigration from the States and Canada lots of times with someone with a dagger , and they would not get rid of it and ...... [ref] M3 Are we really bothered that he has a dagger ? M1 Yes . F1 In A &E?; M1 Yes . F1 It is something you need to address and think about definitely . M3 Okay . F2 You did it last time , who wants to do it ? M4 Anyone else wants to do it . M3 I will do it . M4 Use a different pen . M1 Cool . M4 Okay first thing , Mr Singh . M1 You are on the same ball there . M3 Makhan [rcp] F2 Makhan [rcp] M4 Makhan [rcp] F1 Does it say whether he prefers to be addressed as Mr Singh or Makhan ? M1 Cool . F2 What are you doing ? M3 I am giving him a name . As a cricket player . F2 Yes . F2 His surname is wrong ; are you giving him an extra name ? M3 Yes . F2 That is pretty disrespectful . M3 I am sure I spelt in . M1 Singh . F1 Palpitations ? F2 Yes . M4 Yes . M3 Hang on , hang on , I am still working on Mr Singh . F2 Singh . M4 It does not matter . F2 What are the palpitations ? M4 Palpitations , what are they ? [rcp] M1 Cause of them . F2 Oh my world this is going to be the craziest . It is going to be a fucked up spider . M1 Is going to give me headache . M4 And we have got - I would like to know what is this thumping ? What is different from thumping to pounding ? M2 Is it different to palpitations ? M3 Who gave it an alternative name ? M4 Pounding before . F1 Patient 's weight ? F1 Not completely . F1 Subjective . M1 Who gave it the name palpitation ? M2 Sounds like it . M4 I want to know what that means . F2 Who diagnosed it ? M4 Thumping in the chest . What do you mean thumping in the chest ? F2 Do we know that ? It might just have been his wife saying . M4 What is the difference between thumping , pounding and beating ? F1 Just to make you aware of the feeling . M4 That is what I am thinking but ... wow . F1 But you do not know that until you have asked him . M1 What is the Medical Assessment Unit for ? F2 Yes , well ; shall we go with palpitations first ? So what are they ? Is it cardiological or is it neurological ? Could it be drug induced ? Could it be because he went to the toilet ? F1 Oh yes , drugs . M1 Tracks . M4 Hold on , he woke to go to the toilet . F2 Could be drug induced . M4 And he is got thumping , but did he get his thumping while he was in the toilet ? F2 Yes . Is there something specific about being ... ? Was he woken up by the palpitations or ... ? F1 That is the history . M2 Yes ; or did it happen in the toilet or somewhere else ? M4 Different ..... F2 Not wee induced , the other one when you are straining . M1 Vagal movement . M3 Oh pooing . F2 You increase your blood pressure when you do that , do n't you ? That is how you increase your blood pressure if somebody has come in and they have got blood loss . Get them to strain . M1 Would n't that reduce your blood pressure ? F2 No it increases your blood pressure . It increases your cranial blood pressure , that is why you get headaches . M1 Intracranial pressure . F2 Yes . M1 Reduce your blood pressure because - oh actually I want to find this out . F2 Because your Vagal constriction . M1 Like that . F2 Yes , I think . Dan why do n't you step off the chair ? M3 Do not know . F2 Pooing . F1 Can we say , can we have a clinical definition of palpitations please ? M2 What are palpitations ? Is that covering it ? F2 Well just put coming off . F1 Offshoots like pooing . M2 Clinical definition . F2 Can we ask whether we can have cranial ? M2 Clinical definition of pooing ? F2 Mark , can we ask whether he 's had palpitations before or is this a one off ? Who called the ambulance ? Why were they so so ? M4 Dan , why do n't you put up my thumping thing ? M3 Because we concluded it was subjective . M4 I want to look it up . Subjective if you look in the medical text there may be a difference between thumping or pounding or whatever . F1 Martin , if you look at clinical definitions it is like in the clouds , it will give you like a little table . M3 There you go , it 's there in a small form . F1 Prove it . M1 Why was he awake at 3am ? M2 Yes . M3 I will get rid of those in a minute , do not worry M2 Are you sure ? F2 No , I can not cope with you being on that spinning chair . M1 Why does he need to go to the toilet at 3am in the morning ? F1 Because he might have had a lot to drink . M3 Do you care ? F1 It does not matter about alcohol I mean . F3 Fun ones . M2 Why 3am ? Why 3am ? M1 Well you know what I mean ; I do not get up at 3am . F1 He might have had a cocoa before he went to bed . F2 No but some cardio problems , like do you remember them saying if you have got congested heart failure you go to sleep and then you wake up because you ca n't breathe anymore [ref] . And I wonder if palpitations are related to why he 's sleeping . M1 He could have a big prostrate . F3 Sleeping pattern . M1 One of the defining characteristics of hyperplasia in the prostrate is getting up at six times in the night to go for a wee . M2 So if we have some - coming off 3am , what was it ? Thyroid hyperplasia ? M1 No , no , prostrate . Prostrate hyperplasia . M2 Yes . M1 Yes . M2 Prostate hyperplasia , okay . [rcp] M1 Has he got something else going on ? M2 Hypoxia . F1 Drug induced . M1 Can we ask how many times he gets up to the toilet ? F2 Could be alcohol induced ; he could have woken up during the night because he needed a wee because he had been out for a drink . Although if he is a Sikh , will he drink ? F3 Yes . F3 Not necessarily . Not all of them drink . F2 It is not ... M3 It is Muslims that do not drink . M2 Sorry what was that ? M2 Hypoxia . M1 Something waking him up ? M2 Hypoxia Hypercapnia or whatever . M1 That whole general thing . M2 That slash whatever . F2 Can we have something ? Just while I think on , about Sikhism . Because how many people do n't accept medical intervention ? F1 Health belief [rcp] F2 Health belief of Sikhs . [rcp] M2 Anything else on 3am ? F2 I think if we are looking at palpitations and we come across any significance for waking up during the night that will be enough . M2 Duty calls . F3 I think sleeping patterns as well . F1 He could just be using the toilet . M2 Duty calling . M1 Could we put breathlessness ? On the waking up bit . Just because obviously there was a lot of cardiogenic causes . F1 Can we ask him about his current illness ? M1 Waking up , breathlessness . M4 Like which ones ? F1 I can not think of any . M4 I know that is what I was thinking what to do . I mean I 'm trying to be ... M1 Breathlessness , coughing , 3am in the morning . So basically is he waking up breathless in the night or when he lies down ? M1 I do not think we can just say nocturnal paroxysmal dyspnoea . M2 Yes that is not just . F2 Or even just . M1 Breathless at night . F2 Congenital heat disease . M1 That is one of the symptoms . F2 Does he speak English ? I think if his mother - his mother - I think if his wife is telling us that he has had this and that , does that mean that he is too ill to speak or that he ca n't speak English ? M1 Does he get this thumping at any other time ? F2 We have already got that on our list of questions . M1 Are we okay ? F2 Has he had the palpitations before and when does he get them ? M1 Okay thank you . M1 Or we should have that up there but we do not . M1 It was obviously serious enough for him to get an ambulance . F2 Yes . F2 Is it ? F1 Because of his age maybe he assumed it was a heart attack . M1 Health thing . Yeah that is another thing . F2 Is the dizziness and shortness of breath relating to the palpitation or is that something separate ? F1 Cardiac problems . M2 Gastric reflux on short spread . M3 With shortness of breath . F2 You have not put it on , Nick . M2 It is over here . M1 Can we separate it out ? M2 That should be separate . F2 No , well ; leave it because the breathlessness could be related to the sleeping , so leave that up there but also have a separate one for breathlessness and for the disease . M2 So what do we want on this breathlessness , apart from gastric reflux ? F2 Is there a problem ? Is there a cardio problem ? Is there a urine problem ? M1 Is it Indigestion in origin ? F2 Is it respiratory ? M2 Cardio respiratory , GI . F2 Neuro . M2 Neuro. [rcp] F2 Muscular . M2 Muscular , the lot. [rcp] M2 I will have breathlessness with the lot please . F2 Is it normal to get up at 3am to go to the toilet ? M3 Yep . F1 I do . M1 I do . [rcp] M4 For a 64 year old . F2 Yes . F2 For , like for example . M4 Probably not . F2 I am someone gets up during the night but that is when I have had a lot to drink before bed , or when - if I have been out on the drink , so ... [ref] F1 It depends if it is normal for him . F2 Yes , well that is still a valid question . M2 Is it normal for anyone ? F1 You can not say . It might be for you . F3 It might not be a daily occurrence . M2 Do you know anyone that normally gets up at 3am to go to the toilet ? F1 What everyday ? M2 Yes , routinely . F1 Well we do not know that he does that . M1 I think the key is what woke him up . Is it the fact that he needed a wee or is it the fact that his heart went off ? And he went 'oh I need a wee' and then he is going oh shit what is going on . F3 He could be diabetic . M1 I think if I woke I would probably go oh yes I probably need a wee . As I wake up in the morning and go oh yes I need a wee . F1 Can we have something about the ..... ? Are there any procedures or anything in A &E; for people who bring in weapons ? F2 Yes . Are there any .... ? F1 Well . M4 I think you are right . F2 What about ... ? M4 I would be surprised . F1 What were you allowed to bring in weapons to A &E?; M4 No , not for that reason . I would be surprised because it is the religious issue and I think generally in this country they bat away from all those sort of things . M2 I have never seen a Sikh knife in A &E; ever . That goes off on somewhere or ... F2 That means he is fully dressed so he has managed to get himself fully dressed while he has been waiting for the ambulance , so how serious is it ? M2 Yes . F2 Or was he fully dressed when the ambulance came . M2 Yes . F3 But it is not often seen . M2 Did you say he 'd actually got a knife with him ? M2 Mm . M1 Harry he could have just been dressed . M1 It was tucked into his trousers . F2 But it was 3am in the morning . M2 No , no , no - that is when he went to the toilet . He was already ... M1 I never saw him take it out though . F1 Because whether or not its cultural , he has a concealed weapon . Whether or not he admits to it there must be ... M3 Has anyone put why he is in A &E; to be starting off with ? Because we do n't actually know that , do we ? F2 What time did he turn up ? M1 He was brought to A &E; and then he went to Assessment . M3 With palpitation also . M3 Patient choice there we go . F1 Is n't that a part of A &E?; F2 How long has he had the palpitations between waking up ? M1 So he is been admitted to the hospital and he is there . F2 How long ? F2 How long did he have the palpitations ? You know he woke up at 3am he might have waited until whatever time to come for an ambulance . Like duration of ... M1 Episode . F2 And also why you have got up there , if it happened before , when does he usually get them ? M4 So we will have a look at that . F2 You never know it might have been .. M4 It is good it is good . Need to start making spaces . M1 Can we expand on the breathlessness bit and then do something on dizziness ? F2 Before we do that can I just add one little offshoot on procedures for getting an interpreter into a department . Because I know that 's actually really quite difficult and i would like to know a little bit about that . [ref] Because we had a patient came into - we had a deaf patient came in and luckily we knew the guys round at audio and we just spoke with them and said we 've got somebody who does sign language ; can you come down for us ? But ordinarily we would have had to go and jump through all these hoops for an appointment in a week 's time . I do n't know if we could get one into it ? M3 You do get the funny telephone piece . M2 Well he has got his wife with him , and you are not going to get an outside interpreter at 3am in the Emergency Department . Chances are there is going to be someone who speaks his language , at least if it is Leicester Emergency . [ref] M4 Dan , can you put up dizziness ? The fact that he is dizzy . F2 I am always a little bit worried about dizziness . M3 We are coming to it . M1 Before you go too far away from this business can we just complete that a little bit more ? M4 What and what ? M1 So looking at breathlessness .... , F2 Breathlessness [rcp] M1 .... we need to say like so coming off cardio for example , a few examples of things that might be there that could be causing it . F2 Heart failure . F2 Palpitations . Do palpitations cause breathlessness ? M4 Breathlessness. [rcp] A shunt somewhere - I do not know . No , that would not make sense , would it ? M2 VSD mate . M1 Actually . F1 Maybe he is just panicking about his palpitations that have made him worse . F2 Yes so palpitations caused by panic . M1 Sometimes myocardial infarction just presents a shortness of breath in old people , particularly if he is diabetic . F2 Represents his dizziness . M2 By the way , these palpitations could be entirely normal as

Appendix 14. Analysed data exported to SPSS version 22



*S8_CC.sav [DataSet1] - IBM SPSS Statistics Data Editor

	Indicator	POS	construction.type	function	category
1	and	CC	construction	addition	content
2	and	CC	construction	addition	content
3	and	CC	co-construction	addition	content
4	and	CC	construction	addition	content
5	and	CC	construction	addition	content
6	and	CC	construction	addition	content
7	and	CC	construction	consequence	content
8	and	CC	construction	consequence	content
9	and	CC	co-construction	addition	collaborative
10	and	CC	construction	addition	content
11	and	CC	co-construction	addition	content
12	and	CC	construction	consequence	content
13	and	CC	construction	addition	content
14	and	CC	construction	addition	content
15	and	CC	construction	addition	content
16	and	CC	construction	consequence	content
17	and	CC	construction	addition	content
18	and	CC	construction	addition	content
19	and	CC	construction	addition	content
20	and	CC	construction	addition	content
21	and	CC	construction	addition	content
22	and	CC	co-construction	addition	content
23	and	CC	construction	indefinite addition	content
24	and	CC	co-construction	addition	content
25	and	CC	construction	indefinite addition	content
26	and	CC	construction	addition	content
27	and	CC	construction	addition	content
28	and	CC	co-construction	addition	content
29	and	CC	construction	addition	content
30	and	CC	construction	addition	content
31	and	CC	co-construction	addition	content
32	and	CC	co-construction	addition	content
33	and	CC	construction	addition	content
34	and	CC	construction	addition	content
35	and	CC	construction	addition	collaborative
36	and	CC	construction	addition	content
37	and	CC	construction	addition	content

Appendix 15. Description and examples of indicator function

Shared knowledge indicator

Functional category	Description	Example
Simple elaboration	Shared knowledge indicator with expanded content of reformulation such as repetition, paraphrase and clarification, mutual completion, etc.	M1 Slightly superior right atrium yeah. M3 So it is on the top, quite high up yeah.
Elaborative elaboration	Shared knowledge indicator with expanded content of non-causal new information such exemplification, addition, condition, etc.	M2 Pulse 150. M1 Yeah, but it is also atrial fibrillation as well.
Causal elaboration	Shared knowledge indicator with expanded content of cause-effect information inference, consequence, cause, etc.	M1 Yeah, so basically that is the reason why you get such a quick depolarisation…
Integration-oriented shared knowledge	Shared knowledge indicator with expansive content of elaboration	M1 Yeah, the reason it looks like that is because he is got fast conduction.

Conflict-oriented shared knowledge	Negation shared knowledge indicator with expansive content of elaboration	F2 No, because you need to know your cholesterol levels.
Agreement	Affirmative shared knowledge indicator accompanied by minimal expansive content of idea acceptance or affirmative shared knowledge indicator alone in response to a verify question.	F2 So normal heart rhythms and arrhythmias. M1 Yes that is a very good point. M1 Do you want to do that? F3 Yes.
Acknowledgement	Affirmative shared knowledge indicator occupying the whole response move.	M1 If I lie down after I have done too much exercise I can hear it in my rear. F2 Yes.
Simple negation	Negation shared knowledge indicator occupying the whole response move.	F1 But can you actually tell when it is going to contract from looking at the trace? M1 No.
Disagreement	Negation shared knowledge indicator accompanied by minimal expansive content indicating conflict of ideas.	F2 Do you get why the membrane is negative inside in relation to the outside to start with? M2 No, I don't.
Preface question	Shared knowledge indicator preceding a question.	F1 Yeah, is that what you mean?
Mutual completion	Affirmative shared knowledge is used to preface completion of prior speaker's talk.	It is called …….. F2 yeah yeah sick sinus syndrome.

Orientate to new information	Utterance that indicates that a speaker's frame of reference changes due to prior speaker's talk.	F1 Oh the beta-blockers
Self-recall	Utterance that indicates that prior speaker's talk stimulates recall of previous information.	M1 I remember he mentioned it in the lecture.
Incomplete	The statement is incomplete for interpretation.	So caffeine……………………

Knowledge extension indicators

Functional category	Description	Example
Addition	Addition is made to the first clause.	M3 They put me on the machine and they took some blood. I was a little bit anxious and sweaty.
consequence	Second clause results from the 1st clause.	F1 There is going to be blood stuck in different areas and not moving through properly.
Temporal sequence	Second clause comes after the 1st clause.	F2 … if you have got congested heart failure you go to sleep and then you wake up because you can't breathe anymore. M1 …… you get this indiscernible base line and then a QRS and then another indiscernible baseline….
Use of additive indicator for non-additive function	The use of AND for a subordinating function.	M1 A lot of the time if you try and cardiovert your way out of it they go into atrial flutter.
Concession	One clause concedes something while other clause gives the actuality or truth.	F2 We argue about the nitty gritty about whether that word is right for that particular… and it doesn't matter.

Indefinite addition	There is an addition of indefinite, nonspecific, or vague element to the first clause.	So I think with regard to specific sort of surgery and stuff like that….
Explanation	Second clause comments or explains the first clause.	F3 That occurs when the irregular rhythm appears indefinitely and cannot be corrected with treatment. M1 But the atrioventricular node is here down and filtering out some of those.
Similarity	Second clause makes a point like that of the first.	F2 ….it is a lot worse than if they are ready for it and they know that it is coming.
Contrast	Second clause contrasts with the 1st clause.	F2 ... you get some deep signal moving away and some moving toward…..that's why you get and up and down.
Conditional	First clause is a condition for the second.	F2 … some deep signal moving away and some moving toward and depends on which one is greater …
Continuative	And is used to initiate continuing talk or re-enter after interruption.	M4 Obviously exercise induced because he did not do any other exercise. And they said well what do you do that provokes this and he said it is when I am having sex.

Incomplete	The aspect of the talk prefaced by AND is missing.	F1 I have been through immigration from the States and Canada lots of times with someone with a dagger, and they would not get rid of it and……

Knowledge extension indicators (cont.)

Functional category	Description	Example
Concept options	Or is used to link two alternative ideas.	F1 It is not a name you inherit; it is a name you are given or give yourself.
Question options	Or is used to provide two alternative questions.	F2 Do you want to do that now or do you want to wait until we have done the history?
Definition options	Or is used to provide two or more alternative meanings to a concept.	F2 Lion means a jewel, or precious stone.
Clarification options	Or is used to provide two or more alternative clarifications.	M1 ... this is the problem because normally people with normal rhythms do have palpations or people with fairly benign rhythms, but not necessarily normal,
Coordination tag option	Or is used to provide an alternative indefiniteness.	….you get the actual potential forms go to 70 millivolts and in the nerve fibres or whatever it is.

Explanation options	The speaker provides options in argument to mark different pieces of support and evidence for a position.	M1 This is why you get Q wave or what looks like a downward deflection
Disjunctive continuative	A speaker provides an alternative to own statement after an interruption	F2 Age, sex and race. M1 Male. F2 Or ethnicity rather.
Question option turn	Or is used to ask an alternative question by a new speaker.	M4 Or is it misuse?

Knowledge extension indicators (cont.)

Functional category	Description	Example
Semantic opposition	But is employed to link contradiction to a statement.	M3 My mum she did, she has a high blood pressure and diabetes but my Dad's okay he just has a bit of arthritis.
Denial of expectation	But is used to link denial to an expectation.	M2 And we gave him 500 mcg of digoxin-4 to slow the ventricular rate but the atrial rate the AF rate was 110 bpm after that. F2 It is interesting you can get yourself a sword in the hospice but you cannot get yourself a sandwich.
Restriction	But is used to link constraint to an assertion.	M1 Well not in a nasty way but you are very critical of me. M1 Yeah, not in depth but you get the gist.
Concession	But is used to link acknowledgement or yieldedness to an assertion.	F2 There are allowances for situations where you cannot leave your bed but it is still going to be quite distressing.
Knowledge-reality	But is used to associate what is known with what exists.	F3 ……even if they do speak English it does not mean they will understand. Even though they can function in society, they can listen, but they would not actually understand.

Conditional	But is used to link an assertion to a condition.	M2 They can but that would require a very special effort by the staff sometimes.
Correction	But is used to link correction to a statement.	M2 It is not quite zero but there is much more, it is much more polarised at the…
Extension	But is used to link second clause, which expands the scope of the first clause.	F2 Your modifiable risk factors like the weight and those sorts of things but also your age, and your ethnicity, your family history and then multiplied by one.
Prediction	But connects a prediction to a statement.	F2 I have got this on the cardiac cycle which is kind of not really relevant for today but it will be useful later. F3 It can be like a few minutes, hours or days but it will usually go back by itself.
verification	But is used to link verification question to an assertion.	M4 And he is got thumping, but did he get his thumping while he was in the toilet?
Qualification	But is used to link feature specification to an assertion.	M2 More channels open but they are all operating at the same speed.
Analogy	But is used to connect similar assertions.	F2 …..well bradycardia is not arrhythmia but like arrhythmia you have palpitations and dizziness.
Reassurance	But is used to link encouragement to assertion.	F2 It is a little bit dirty on the front but never mind.

Debate	But is used to link a debate to an assertion.	M1 ….. some physiologists believe that there is actually a communication of cells between here kind of like this bundle that makes it go across faster but it is debatable.
uncertainty	But is used to link uncertainty to an assertion.	M1 Actually the outcome from the whole thing was that disease prevalence but I do not know what you would call it,

Knowledge enhancement indicators

Functional category	Description	Example
Mechanistic	Because is used to connect underlying mechanism or multistage process to an assertion.	M1 Yeah, the reason it looks like that is because he is got fast conduction.
Mechanical	Because is used to link physical reason to an assertion.	F2 …. your blood pressure went up because you jumped up …..
Conditional	Because is employed to relate a condition to an assertion.	M1 I was just going to do it in order because if you look at the classification you can then apply those classifications to the makes a bit more sense that way.
Functional	Because is used to link bodily function to an assertion.	M1 You pee out the glucose because your pancreas cannot remove it from the blood. F2 ...you go to sleep and then you wake up because you can't breathe anymore.
Genetic	Because is used to link sequence of events to an assertion.	M3 I used to start with Wikipedia when I wrote essays because you would look at it, read it and then it is got all the links so that they had references at the bottom.

Psychological	Because is used to link psychological reasons to an assertion.	F3 …. you came in at 3 o'clock because your wife was concerned about you having palpitations… F1 …and ran up the stairs or he ran somewhere because he panicked.
Fact-based	But is used to link concrete life event reason to a statement.	F2 I reckon Asha should do it because Asha does not go for the simulators.
Metaphysical	Because is used to link supernatural agent or religious reason to an assertion.	F3 Yes because in our religion Krishna is called Kanha which because he really used to eat lots and lots of butter, so it is Kanha, Krishna is one of the gods.
Epistemic	Because is used to relate state of knowledge to an assertion.	F1 We cannot do race and ethnicity because we do not know what it is.
Teleological	So and because are used to link consequence or characteristics to an assertion.	F2 …normally you record on lead 2 because that is the clearest one. M3 Are you drinking your pee because it is sweet?
Hypothetical	Because and so are used to link hypothetical reason to an assertion.	M1 He is got so many causes it is hard to work out which one it is because it could be hypertension it could he heart failure it could be the coronary artery ischemia.

Subsumption	Because is used to link categorical basis to an assertion.	We have to multiply it by 1.4 because he is from south Asia.
Tautological	Because is used to link mere reformulation to a statement.	M1 Obviously exercise induced because he did not do any other exercise. F2 The food is a separate learning area because it is kind of separate to that.
Rationale	Because is used to link warrant to a statement.	F2 It is in the patient's best interest just to be able to get the meals from the canteen even if they are having one slice of toast because we are going to have to pay for it.
Anthropological	Because is used to link animated concept to an assertion.	F2 Atrial wave takes time to propagate because it is obviously going from cell to cell. M1 Yeah, so basically the reason why you get such a quick depolarisation is because you have got these guys travelling really very quickly and let in a big flood of positive ions into the cell getting rid of the polarity. M1 It takes a long time to propagate because it is obviously going from cell to cell.
Analogical	Linguistic feature is used to link two assertions considered similar.	F2 ….she said that thing like we have got some deep signal moving away and some moving toward…

Order	A connective is used to link order of a phenomenon to an assertion.	F1 He might have had a cocoa before he went to bed. M1 If I lie down after I have done too much exercise I can hear it in my ear.
Simultaneity	A connective is used to link a simultaneous occurrence of a phenomenon to an assertion.	F2 Isn't it harder for your ventricles to pump while you are lying down though?
Durative	A connective is used to link a durative phenomenon occurrence to an assertion.	F1 But you do not know that until you have asked him. M3 I think that is more since this week.
Spatial	A connective is used to link a spatial dimension of a phenomenon to an assertion.	M3 Yes I am trying to think of where to put it. M3 Up where?
Consequential	So is used to link a statement to its consequence.	F2 ….it is the biggest one of the three so you get a better look at the P wave, which is up here. F2 … if it is 20% or over then that is a significant risk, so they should be started on something like Simvastatin.

Inferential	So is used to relate inference to a statement.	M3 I suppose they should have said he did that at lunchtime so there is no way that it could have been….. M2 He is on a Atenolol which we just mentioned, he is on Ramipril which is the ACE inhibitor, so he is on the ACE inhibitor which is…
Possible situation	If is used to link possible situation to an assertion.	M1 If you were to measure the atrial rate that is what you would get.

Facilitation techniques

Function	Description	Example
Facilitating	Facilitation occurs when the tutor guides the students in a certain direction, suggests what to do next, or attends to group dynamics e.g. suggestion, advice.	T1 The suggestion then would be you guys want to look at mechanisms under the line of pathology, perhaps you actually want to explore the normal situation as well.
Prompting	This technique is used by the tutor to gather more information by giving a hint e.g. concept completion.	T1 Transient ischaemic attacks and he is been having increasing bouts of palpitations so a lot of this sounds like he is potentially having what going on?
Elicitation	This involves the tutor posing a question to the individual or to the group as a whole.	T1 How did you do an EKG? T1 What would account for kind of sporadic symptoms that seem to be associated with lack of oxygen to places?
Informing	Informing occurs when the tutor passes on facts, information, explanations, opinions, or ideas to the students.	T1 It is usually thyroid storm which we are not talking about and you guys will talk about storm when you get to endocrine. But that is typically one of the things you will see is you will see thyroid problems that acutely precipitate atrial fibrillation attacks.

Give feedback	Feedback is given when the tutor confirms that he has seen or heard an appropriate response.	T1 The gap between the Q wave and the PR interval as it is called. Yes good, I think you guys did most of the basics, T1 I think that is actually a good point from a clinical perspective…
Interpersonal behaviour	Interpersonal behaviour occurs when the tutor has an awareness of his/her strength and weakness and relates to the students in an amiable manner.	T1 Sorry, I think that about all my lectures. I am still not entirely sure to be fair why some of them are where they are.
Stimulate elaboration	This occurs when the tutor asks questions that get the students to expand on something e.g. explanatory, judgemental questions.	T1 And it can be quite difficult and that is when you start doing things that you guys have discussed about the idea, maybe trying to do some sort of intervention as a therapeutic, so carotid sinus massage what else?
Summarising	The tutor recapitulates the salient points of discussion.	T1 You guys have discussed the idea of doing carotid sinus massage as a potential diagnostic here between atrial fibrillation and atrial flutter if this guy has atrial fibrillation which we have said that he does.

Reflection	The facilitator reflects on his own and the group's performance.	T1 ….I was also trying to push to wrap up most of the case last time if you had not noticed. I was not sure how we were going to make up in terms of the work.

Appendix 16. Workshop for hands-on training in the use of Wmatrix 3

Metaphor in end-of-life care (MELC)

- Home
- Team
- Background
- Dissemination
- Conference presentations
- For health professionals and the public
- In the media
- Publications

Using Corpus Methods to Analyse Metaphor in Discourse: One-day workshop at Lancaster University

This event has now finished. To see presentation slides from the event please follow these links:

- Welcome and introduction
- Using Wmatrix
- Annotating texts with eMargin
- Linking eMargin tags with USAS tags in Wmatrix
- Using Wmatrix for metaphor identification and analysis

The Metaphor in End of Life Care ('MELC') Project at Lancaster University is using a corpus-based approach to investigate the role of metaphor in language used to talk about end-of-life care by patients, carers, and healthcare professionals. As part of our project activities, we will be holding a one-day workshop on 'Using Corpus Methods to Analyse Metaphor in Discourse' at Lancaster University on 10th January 2014. This event is aimed at those with an interest in using corpus linguistic methods to investigate metaphor, particularly in the area of health communication. The workshop will include presentations and training sessions, and the programme is available here.

There is no fee to attend this event, which includes lunch and refreshments. Places are limited, however. We are able to offer funding for UK economy-class travel and one night's accommodation on campus, up to a combined maximum of £100, for 15 participants, on a first-come, first-served basis. A further 9 places are available without funding for travel or accommodation.

To register for a place at the workshop, please e-mail Dr Jane Demmen (j.demmen@lancaster.ac.uk) by 31st October 2013, indicating the name

Appendix 17. Wmatrix 3 statistical advice and support