Is this the right room for an argument?

The effects of an internet-based argumentation intervention on aspects of self-regulated learning and critical thinking in young adolescents.

Thesis

Coral Milburn-Curtis

Department of Education, University of Oxford; Green Templeton College

Supervisors

Dr Maria Evangelou
Dr Lars-Erik Malmberg

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Education

Michaelmas Term 2014

Abstract

The importance of developing learners' self-regulated learning and critical thinking has captured the attention of researchers in recent decades. But interventions that are effective in developing such skills are scarce in the English school context.

This study reports on the creation, design, implementation, piloting and evaluation of such an intervention - a four-week internet-based extra-curricular critical thinking course, *Is this the right room for an argument?* Designed by the researcher as a University of Oxford outreach course, it was provided, for the purposes of this study, to a group (N = 617) of eleven and twelve year old state-school students in England. Based upon self-regulated learning (SRL) and critical thinking (CT) research literature, the intervention comprised four elements which have been linked to progress in both SRL and CT: dialogic argumentation, written counter-argumentation, self-reflection in a learning journal and critical thinking tests.

The effectiveness of the intervention was evaluated by constructing three experimental conditions and one control condition, which isolated the effects of different combinations of the intervention activities.

Structural equation modelling, regression techniques and repeated measures analysis of variance were used to detect main and experimental effects on change in SRL and CT over time, whilst autoregressive cross-lagged path analysis models were specified to explore the process of SRL in action over the course of four learning events.

There was an anticipated overall main effect of the intervention, on change in aspects of SRL, with small and medium effect sizes, and on CT, with a medium effect size. It was hypothesised that the use of the learning journal would promote change in SRL and that participation in argumentation would promote progress in CT. Both hypotheses were supported. Participants who self-monitored their learning experience in a learning journal reported significantly more gains in aspects of their SRL, compared with those who did not, with small and medium effect sizes. The extent of participation in dialogic argumentation predicted change over time in CT, with a small effect size, in metacognitive knowledge, with a large effect size and in motivational beliefs and efficacy for learning, with medium effect sizes. Repeated measures analyses detected significant and positive trends over time in metacognitive activity (elaboration strategies, volitional strategies, critical thinking and metacognitive experiences), with effect sizes ranging from small (volitional strategies) to large (elaboration strategies). A process analysis tracked metacognitive activity over four learning events, detecting statistically significant reciprocal relationships between observed critical thinking and self-reported elaboration strategies, and between volitional strategies and metacognitive experiences.

Findings suggest that this intervention had a positive effect on change in both self-regulated learning and critical thinking. Specifically, the more participants engaged in dialogic argumentation the more the change in both outcomes. The effects of the intervention, on change in self-regulated learning, were significantly enhanced when the critical thinking activity (argumentation) was accompanied by the self-regulatory activity (the learning journal).

1 Chapter One: The Introduction

1.1 Thesis outline

This study will explore the development and implementation of an internet-based argumentation intervention, and will report on an evaluation of its capacity to effect change in aspects of self-regulated learning and critical thinking in groups of young adolescents in England.

The thesis will be structured in seven chapters. **Chapter One** will provide an introduction, setting the scene and introducing the rationale for the study. It will do this by locating the study in the context of existing work. Research questions will be related to the evaluation of the intervention, and will concern themselves with a) change in aspects of self-regulated learning over time, b) change in critical thinking over time and c) the relationship between aspects of metacognitive activity, as recorded in learning journals over the course of the intervention. A final research question will explore the extent to which the study data reflect the conceptual model initially proposed for the study.

Chapter Two will provide a literature review, drawing upon research on self-regulated learning and critical thinking, in order to evidence the rationale for the study, to construct a conceptual model for structuring an exploration of its effectiveness, and to provide a context for its design and methodology.

Chapter Three will detail the intervention: *Is this the right room for an argument?* This was a four-week internet-based course in critical thinking, designed for young

UK adolescents. The chapter will explain how the intervention was designed, how it was implemented on an internet-based learning platform, and how students accessed the principal learning activities: dialogic argumentation, written counter-argumentation (the Devil's Advocate), self-reflection in a learning journal and assessments in critical thinking.

Chapter Four, the Methodology, will detail the main study research questions, design, sampling process, measures, data collection, data analysis, methodological limitations and ethical considerations.

Chapter Five will report on the methodology and findings of the pilot study, which took place three months before the main data collection. The pilot was intended to provide a preliminary evaluation of the potential effectiveness of the intervention, to inform any necessary modifications, and to test the intended measurement instruments.

Chapter Six will report the findings of the main study. These findings detected that the intervention had a significant and positive influence on gains in aspects of learners' self-regulated learning and on change in their critical thinking achievement scores. The effect of intervention activities, on change over time in self-regulated learning, was doubled when students additionally made use of a learning journal. A process analysis revealed reciprocal links between critical thinking and the use of elaboration strategies, and between the use of volitional strategies (inhibitory control) and learners' self-reported metacognitive experiences. The conceptual model, proposed in Chapter Two was evaluated with reference to its relevance to the findings.

Chapter Seven will discuss the findings, addressing the implications of each research question in turn. It will also identify aspects of learning policy, pedagogy and curriculum design that may benefit from the findings of this study. It will discuss the potential limitations of the study, will propose questions that the study has generated, leading to aspects that require further study, and finally will discuss its potential contribution to knowledge.

1.2 Table of Contents

1.3 Table of Tables

1.4 Table of Figures

1.5 Definition of terms

The following terms, which have multiple meanings in common parlance and in the literature, are specifically defined for the purposes of this study:

Argumentation: the process of producing, evidencing and justifying reasons for asserting a position - leading to **argument,** the product of argumentation (Kuhn, 1992).

Cognition and Metacognition: *cognition* refers to the use of strategies employed in the learning process, whereas *metacognition* refers to the capacity to control, monitor and evaluate those strategies (Flavell, 1979; Forrest-Pressley and Waller, 1984).

Cognitive strategies: goal setting, summarising, rehearsal, self-testing, memorising, organising and time management (Weinstein and Mayer, 1983).

Counter-argument: the process whereby an arguer highlights a point, which could be seen to support an opposite point of view (Leitão, 2000).

Critical thinking: '… a metacognitive process that, through purposeful, reflective judgement, increases the chances of producing a logical conclusion to an argument or solution to a problem' (Dwyer, Hogan and Stewart, 2014, p.43).

Dialogic argumentation: a means of developing students' critical thinking through peer-feedback and tutor-led discussion of moral, ethical or social issues (Clark and Sampson, 2008).

Efficacy for learning (EfL): an aspect of self-efficacy which refers specifically to learners' belief in their capacity to self-direct their learning in a collaborative settting.

Inhibitory control: the volitional strategy of avoiding distraction and ignoring extraneous or irrelevant stimuli, controlling inappropriate attentional tendencies (Durston, Mulder, Casey, Ziermans and van Engeland, 2006). It is characterised as a metacognitive skill (Efklides, 2008, 2009).

Metacognitive knowledge: a self-regulated learning disposition: the learner's capacity to understand how they learn in a given situation (Sperling, Howard, Miller and Murphy, 2002).

Metacognitive strategies: control and evaluation strategies that monitor the learning process. These include elaboration (linking new learning to previous knowledge or experiences), self-evaluation, self-judgement and evaluation of metacognitive experiences (Zimmerman, 2008; Efklides, 2008).

Need for Cognition: 'disposition to engage in and enjoy effortful cognitive activity' (Cacioppo, Petty, Feinstein and Jarvis, 1996, p. 197).

Process: a series of learning events, which are linked together over time (Schmitz, 2006).

Self-efficacy: 'judgments about one's ability to accomplish a [learning] task as well as one's confidence in one's skills to perform that task' (Pintrich, Smith, Garcia and McKeachie, 1991, p. 13).

Self-regulated learning: commonly described in terms of the 'offline' person-level *attributes* or dispositions that learners bring to the process of learning (e.g., metacognitive knowledge, motivational beliefs and self-efficacy) (Winne and Perry, 2000), and/or the 'on-line' task-level cyclical self-regulatory *processes* that take place during a learning task (e.g., goal setting, task management, self-monitoring and self-evaluation) (Schmitz and Wiese, 2006).

Volition: a metacognitive skill, involving the control and regulation of effort, persistence, time management, prioritising goals, maintaining focus and inhibitory control required to implement cognitive strategies (Efklides, 2009).

1.6 Introducing the study

This section will introduce the intervention study, articulating its purpose, determining the principal constructs of interest, and attempting to justify why those constructs are important. It will thus identify problems that have the potential to be addressed in the research, and will outline how the study will address them. Finally it will propose a summary of potential contribution to knowledge.

This is a study of the design, implementation and evaluation of the effectiveness of an internet-based argumentation intervention. It is evaluated in terms of its potential to promote aspects of critical thinking and self-regulated learning in young adolescent learners from England. The first aim of the study is to report on the design of the intervention and how it was implemented. The second is to explore its effectiveness in terms of its capacity to promote change in self-regulated learning and critical thinking. The third aim is to use the intervention data to explore the process of self-regulation in real time, by evaluating relationships between aspects of metacognitive activity reported in student learning journals over the four learning events of the intervention.

The first major outcome construct is critical thinking. *Critical thinking* is defined as 'a metacognitive process that, through purposeful, reflective judgement, increases the chances of producing a logical conclusion to an argument or solution to a problem' (Dwyer, Hogan and Stewart, 2014, p. 43). It is widely accepted as a key higher order thinking skill that the learners of today require in order to meet the increasingly complex demands which new information, new technologies and new working methods are placing on them, in a rapidly changing world (Ku, 2009). Many educators and researchers agree that the development of critical thinkers should be a key objective for contemporary educational policy (Bailin and Siegel, 2003; Abrami, Bernard, Borokhovski, Wade, Surkes, Tamim and Zhang, 2008). Despite the considerable amount of research which has explored the development and assessment of critical thinking of college and university students (e.g., Garcia and Pintrich, 1992; Paul, 1993; Facione and Facione, 2001; Halpern, 2003; Abrami *et al.,* 2008), there remains a distinct lack of studies that focus upon the nature of critical thinking as it is taught, developed and assessed for school-aged students, especially in England. As it is suggested that the capacity for critical thinking emerges and develops rapidly in early adolescence (Inhelder and Piaget, 1958; Piaget, 1964, Flavell, 1979, Kuhn, 1991, 1999), it follows that there is merit in exploring ways of developing and assessing it, for this age group.

The second major outcome construct in this study is self-regulated learning. *Self-regulated learning* refers firstly to the person-level *attributes* or *dispositions* that learners bring to the process of learning (e.g., motivational beliefs, metacognitive knowledge and self-efficacy) (Winne and Perry, 2000), with metacognition, the control and regulation of the learning process, regarded as its *sine qua non* constituent (Efklides, 2008). Secondly, it refers to the task-level cyclical self-regulatory *processes* that take place during a learning task, such as the use of learning strategies and metacognitive reflection (Zimmerman, 2008). Links between aspects of self-regulated learning and academic achievement are by now well established (Boekaerts, 2010; Corno, 2001; Pekrun, Goetz, Titz and Perry, 2002; Schwinger, Steinmayr and Spinath, 2009; Schwinger and Stiensmeier-Pelster, 2012; Zimmerman 1990, 2009;

Zimmerman and Schunk, 2001; Caprara, Vecchione, Alessandri, Gerbino and Barbaranelli, 2011); thus for the past three decades, an increasing body of researchers has been defending the view that the promotion of self-regulated learning should become a major goal of the modern education system (Boekaerts, 1997; Zimmerman, Bonner and Kovach, 2002). They argue that in a post-modern age, characterised by shifting societal patterns, in which the workplace environment attempts to keep up with rapid technological advances, there should be an increasing emphasis on enabling the learners of today - the workers and leaders of tomorrow - to adopt flexible self-regulated learning strategies which can prepare them for lifelong learning. Thus they propose that the development and provision of effective interventions that have the potential to develop self-regulated learning should become a priority for today's researchers, policy makers and educators (Schmitz, Schmidt, Landmann and Spiel, 2007; Dinsmore, Alexander and Loughlin, 2008).

However, many students fail to employ such strategies in the learning process, and few programmes exist in England, which provide effective methods to support the development of such skills (Duckworth, Akerman, MacGregor, Salter and Vorhaus, 2009). This is despite contemporary meta-analysis, conducted for the Sutton Trust (Higgins, Katsipataki, Kokotsaki, Coleman, Major and Coe, 2013), which suggests that such programmes have the potential to impact[1] upon student learning outcomes

[1]Achievement is measured using United Kingdom National Test data (Higgins *et al.,* 2013).

(in this case secondary school national achievement tests) with an effect size[2] of $.62 < d < .69$. Studies have also revealed that teachers do not encourage the development of self-regulatory strategies in the classroom (Perels, Dignath and Schmitz, 2009). Even where such programmes do exist, as Schunk and Ertmer (2000, p.635) suggest, ‘teaching a strategy does not guarantee that students will continue to use it.’ The challenge, therefore, is to provide for the development of self-regulated learning strategies, which can be incorporated into existing programmes without disrupting them.

This study explores an intervention that includes two learning methods that have the potential to have an effect on change in critical thinking and self-regulated learning. The learning methods under scrutiny are, respectively, internet-mediated argumentation (the metacognitive process of producing, evidencing and justifying reasons for asserting a position) (Cavagnetto, 2010; Alagoz, 2013; Darabi, Liang, Suryavanshi and Yurekli, 2013; Sanders, Wiseman and Gass, 1994; Walton, 1989; Cross, Taasoobshirazi, Hendricks and Hickey, 2008) and the regular use of an electronic learning journal (a structure for reflection and self-evaluation of metacognitive activity) (Hübner, Nückles and Renkl, 2010; Stephens and Winterbottom, 2010).

The first learning method, *argumentation*, has been linked to the process of metacognition, since the activities involved in argumentation (reflection on one’s own

[2] Cut-off criteria for Cohen’s *d* effect sizes are given in Section 4.7.2.

thinking, reasoning, analysis, synthesis of ideas) are metacognitive processes which exercise critical thinking (Kuhn, Shaw and Felton, 1997; Pinto, Iliceto and Melogno, 2012: Kuhn, 2000a; Hidi, Berndorf and Ainley, 2002; Kuhn and Crowell, 2011; Cavagnetto, 2010). Research has established that young adolescents are unlikely to easily perceive an argument from more than one perspective, with the most common weakness in their argumentation technique being the lack of the critical thinking technique of counter-argument (Kuhn, 1991; Voss and Means, 1991; Kuhn, Shaw and Felton, 1997; Stein and Bernas, 1999; Kuhn and Udell, 2007). This process of counter-argument is felt to be the defining attribute of effective argumentation (Green and Klug, 1990; Andriessen, Baker and Suthers, 2003; Voss, Perkins and Segal, 1991). Green and Klug (1990) for instance suggest that any lack of opportunity for counter-argument presents a 'serious problem' for students, since critical thinking is best developed through the 'crucial components' of the defence of an individual's argument and the rebuttal of others' (p. 463). This notion reflects the socio-cognitive conflict perspective proposed by Piaget (1932) and Doise, Mugny, St. James and Emler (1984), which perceives conflict as a pivotal catalyst for changing otherwise stable beliefs and opinions. Therefore it is suggested that interventions that attempt to develop critical thinking skills are likely to be effective if they include counter-argument as a particular focus (Kuhn, 1991). This assertion will be explored in the current study, where the core activity in the intervention is a counter-argument exercise and where subsequent analysis explores the effectiveness of counter-argument in terms of raising critical thinking.

A specific prosecution of argumentation is *dialogic* argumentation, which in the context of this study is defined as a means of developing students' critical thinking through peer-feedback and tutor-led discussion of moral, ethical or social issues (Clark and Sampson, 2008). Whilst the notion of a relationship between dialogic argumentation and metacognition has inspired much theoretical debate with regard to the extent to which dialogic argumentation is a 'primary thinking skill from which other skills follow' (Fisher, 2007, p. 615), it is nevertheless suggested that there is a need for experimental research to support this perspective, especially across a range of educational settings (Cavagnetto, 2010). For instance, research has predominantly studied the influence of dialogic argumentation on the development of metacognition in the context of science education in the US (e.g., Kuhn, 2010; Kuhn and Crowell, 2011; Clark and Sampson, 2008; Cross, Tasoobshirazi, Hendricks and Hickey, 2008), and involving undergraduate students (e.g., Green and Klug, 1990), with less experimental research exploring its effects in other educational contexts such as the English secondary school setting. Previous studies have tended to explore group differences in order to estimate relationships between dialogic argumentation and the development of metacognition (e.g., Kuhn and Crowell, 2011; Kuhn, Zillmer, Crowell and Zavala (2013). However, it has also been established that learners participate in internet-based dialogue at varying rates, with some preferring to be 'in the thick of it,' and others preferring to monitor discussion from the sidelines (Beaudoin, 2002; Dennen, 2008). This study addresses this potential limitation by using count data of *actual participation* in dialogic argumentation, in order to detect whether aspects of self-regulated learning and critical thinking are affected by the *extent* to which learners engage in dialogic argumentation.

Whilst the literature focuses upon the predictive effect of argumentation and dialogue on the development of metacognition (e.g., Colbert, 1987; Green and Klug, 1990; Fisher, 2007; Kuhn, 2011), there is a dearth of research that explores the relationships between argumentation (both dialogical and written) and other aspects of self-regulated learning. This study addresses this gap.

The second intervention activity concerns the use of a *learning journal* to scaffold metacognitive reflection. Researchers have explored the capacity for journal-writing to serve as a learning tool for promoting academic achievement (e.g., Bangert-Drowns, Hurley and Wilkinson, 2004). Others have explored its use as a measurement tool, for instance to explore the process of self-evaluation (Schmitz and Wiese, 2006; Glogger, Schwonke, Holzapfel, Nückles and Renkl, 2012). Such studies have begun to employ process analyses of journal data to explore self-regulated learning in action over multiple time points, with reports hailing the effectiveness of the evaluation of learning trajectories over time (e.g., Schmitz and Wiese, 2006; Schmitz and Perels, 2011). To date, such promising methods have rarely been employed either in the English context, or using internet-mediated methodology. This study incorporates both operational perspectives, using the learning journal as both a learning tool and a research tool. Firstly, it concerns itself with the extent to which the use of a journal, as a learning tool for reflecting upon metacognitive skills, strategies and experiences, can affect change in aspects of participants' self-regulated learning. Secondly, it uses aspects of the learning journal as a measurement tool, with which to explore the process relationships between aspects of metacognitive activity over a series of learning events.

The study is important in four ways. Firstly, it proposes an innovative and cost-effective learning method that has the potential to improve aspects of self-regulated learning and critical thinking in young adolescents. The results of the study suggest that young learners who engage in internet-based dialogic argumentation, whilst maintaining a learning journal, can make significant gains in their self-regulated learning and critical thinking over time. Secondly, whilst a number of studies have explored the effectiveness of a learning journal in terms of its potential to effect change in metacognitive and motivational dispositions, this study goes further by proposing that the use of a learning journal can also effect change in efficacy for learning, a relationship that has so far received little attention. Thirdly, literature on the effects of argumentation activity focuses on the relationship between it and metacognition, whilst this study extends the field by proposing links between argumentation and self-regulated learning aspects other than metacognition. Finally, the results of the process analysis reveal new links between aspects of metacognitive skills and metacognitive experiences, in the context of this internet-based argumentation intervention. Specifically, it reveals firstly that, in this study, a young learner's metacognitive experiences (e.g., learning satisfaction, response to challenge and task-related self-confidence) are related to their volitional strategies (the capacity to avoid distraction), and secondly that the use of elaboration strategies (making connections between new learning and old) is reciprocally related to achievement in critical thinking.

This section has introduced the research, stating its aims and highlighting the some challenges that have the potential to be addressed by the study. It has summarised the principal constructs to be explored in the study (self-regulated learning and critical thinking) and has attempted to justify the choice of the main intervention activities. Finally the potential importance of the study has been briefly summarised.

Following a summary of the research questions, a literature review will explore relevant aspects of critical thinking and self-regulated learning, arguing a rationale for the study, and presenting a conceptual model for the evaluation of the intervention.

1.7 Research questions

Intervention effects on critical thinking outcomes

RQ1: What are the main and experimental condition effects of the intervention on change over time in participants' critical thinking?

RQ2: To what extent does participation in dialogic argumentation predict progress in critical thinking over time?

Intervention effects on aspects of self-regulated learning

RQ3: What is the main effect of the intervention on change in aspects of participants' self-regulated learning (metacognitive knowledge, motivational beliefs and efficacy for learning) from pre-test to post-test, on groups of young adolescents, after one month, compared to a group who had no intervention experience?

RQ4: What are the effects of the separate experimental conditions on change in self-regulated learning and its components over time?

RQ5: To what extent does participation in dialogic argumentation predict change in self-regulated learning and its components over time?

Relationships between aspects of metacognitive activity over time

RQ6: What are the relationships between aspects of metacognitive activity and critical thinking over time?

Relationships between SRL dispositions and SRL processes

RQ7: What is the relationship between self-regulated learning dispositions (metacognitive knowledge, motivational beliefs and efficacy for learning) and self-regulated learning processes (metacognitive activity) during this intervention?

2 Chapter Two: The Literature Review

2.1 Introduction

The aim of this chapter is to explore the theoretical background behind the concepts of self-regulated learning and critical thinking, in order to a) provide a context for the intervention and b) inform a conceptual framework for its evaluation. The literature review is divided broadly into two sections, exploring firstly critical thinking and then self-regulated learning.

The section on critical thinking (CT) begins with a clarification of the nature of the construct and how it is related to self-regulated learning. A critique of current English policy with regard to the importance of critical thinking in the National Curriculum underpins a rationale for the inclusion of critical thinking as a major construct in this study. The section goes on to explore ways of developing the critical thinking of young students, focusing in particular on the potential contribution that the development of *argumentation* can make to the fostering of the skill of critical thinking. It concludes with an attempt to identify gaps in research that have the potential to be addressed in this study.

The second major section in the literature review focuses upon self-regulated learning (SRL). It begins by placing the construct in its historical perspective, and considers the importance of SRL research. Focusing specifically on the development of SRL in early adolescence, the review attempts to situate the current study in the context of existing policy and provision for the development of SRL of young learners in

England. Following an examination of existing models of SRL, a conceptual model for this study, which locates the relationship between self-regulated learning and critical thinking, is proposed. The section continues with an exploration of methods for the development of SRL, focusing on the potential for the *learning journal* to provide both an SRL learning tool and an SRL research tool, before exploring the gaps in research, which have the potential to be addressed in this study.

2.2 Critical thinking (CT)

The development of critical thinking in young students is frequently cited as being one of the most important objectives of contemporary educational policy (e.g., Kuhn, 1992; Paul and Elder, 2009). Indeed, according to Marin and Halpern (2011), it is 'the most important reason for formal education' (p. 1), whilst Bailin and Siegel (2003) suggest that it is 'a fundamental and overriding ideal of education' (p.188). Specific inclusion of critical thinking as a curriculum learning objective is becoming increasingly important, since analytical and reasoning skills are becoming more and more necessary for rapidly evolving real-world applications (Dwyer, Hogan and Stewart, 2014). Yet critical thinking, as a specific learning objective, features only rarely in English National Curriculum literature. Furthermore, there are few studies which explore critical thinking interventions which relate to the English educational culture, or which offer guidance for school educators in how to develop this important skill. This study details an intervention, which has the potential to affect the critical thinking skills of young adolescents. In support of the rationale for including critical thinking as a learning objective of the intervention, which underpins this study, the following section explores the concept of critical thinking, how it is taught and assessed, its potential to be developed through the process of argumentation, and how this study may contribute to knowledge in the field. It begins with an important rationale for the inclusion of critical thinking as a key self-regulated learning skill.

2.2.1 The relationship between CT and SRL

Critical thinking is a complex construct with many definitions. Ennis (1962) for instance defines it in terms of the assessment of argument. Whilst some disagree, (such as Van den Brink-Budgen (2010, p.1), who claims that 'saying that critical thinking is all about arguments is something like saying that music is all about a series of notes'), it is generally agreed that it is a self-disciplined, self-reflective process in which the thinker demonstrates the capacity to identify the strength of an argument by the use of analysis, evaluation, inference, reflective judgement and reasoning (Kuhn, 1992; Bowell and Kemp, 2010; Dwyer, Hogan, Harney and O'Reilly, 2014). As such it is commonly characterised as a cognitive strategy within the construct of self-regulated learning (Schraw, Crippen and Hartley, 2006), and is widely recognised as an essential metacognitive process (Bailin and Siegel, 2003; Dwyer, Hogan, Harney and O'Reilly, 2014). Abrami *et al.,* (2008) relate critical thinking directly to self-regulated learning by describing it as 'the ability to engage in purposeful, self-regulatory judgement' (p. 1102). Phan (2010) also locates critical thinking as a self-regulatory process, defining it as, 'a complex process of reflection that helps individuals become more analytical in their thinking and professional development' (p. 284). He conceptualises that critical thinking and self-regulated learning operate dynamically such that critical thinking may relate to other self-regulatory processes. This is a notion which, it will be seen, resonates with the findings of the current study.

Until 1990 the interpretation of the construct was unclear. Early writers such as Ennis (1962) paved the way for later clarity by articulating the notion of an individual's engagement with reflective thinking, a process that should lead to 'the correct assessing of statements' (p.83). The same author (1993) later amended his definition in line with important conceptual developments that had taken place in the intervening period, suggesting that the term now described, 'reasonable reflective thinking focused on deciding what to believe or do' (p. 180). Those intervening important developments centred around the Delphi Project (Facione, 1990), a major effort involving leading philosophers and critical thinking writers, that was mandated by the American Philosophical Association to establish theoretical consensus around the concept of critical thinking. The product of their deliberations was a definition of critical thinking that has exerted evident influence on subsequent definitions and modes of assessment ever since. It is:

> *We understand critical thinking to be purposeful, self-regulatory judgement which results in interpretation, analysis, evaluation and inference as well as explanation of the evidential conceptual, methodological, criteriological or contextual considerations upon which that judgement was based* (Facione, *op cit.,* p. 2).

By 2014, the definition of critical thinking was being articulated in an even more succinct fashion. In their construction of a 'critical thinking framework for the 21st Century', Dwyer, Hogan and Stewart (2014, p. 43) now define critical thinking as:

> *... a metacognitive process that, through purposeful, reflective judgement, increases the chances of producing a logical conclusion to an argument or solution to a problem.*

They concur with the Delphi recommendation that the core elements of critical thinking are analysis (the capacity to recognise the strength of an argument), evaluation (the capacity to evaluate claims and biases), and inference (the capacity to draw a conclusion from reasoned evidence). For reasons of simplicity, and because it reflects most closely the constructs explored in this study, the Dwyer, Hogan and Stewart (*ibid.*) definition is the one considered most appropriate for adoption here. Therefore, for the purposes of this study, critical thinking is located as a metacognitive skill within the construct of self-regulated learning (also see the conceptual model presented in Figure 2-1, Section 2.3.6).

2.2.2 Current CT educational policy and practice (in England)

Educational reforms since 2010 have, according to British government statements, led to increased autonomy for teachers, more emphasis on ‘knowledge-base’, increased emphasis on school performance as measured by exam results, and less prescription of curriculum content (Gibb, 2014). The ideology which drives this agenda is exemplified by ministerial statements such as, ‘good government does not improve public services. It enables public services to improve themselves’ (*ibid.*, p.1). It will be argued in this study that the development of critical thinking is of benefit to students’ long term educational and career prospects. Furthermore, research has established that early adolescence is not only the time when the capacity for critical thinking, a metacognitive skill, emerges and develops rapidly (Inhelder and Piaget, 1958; Piaget, 1964, Flavell, 1979), but it is also the time when critical thinking interventions that have the potential to develop such skills are at their most effective (Kuhn, 1992; Abrami *et al*., 2008). It follows that provision for the development of such a skill should be a prime objective of educational policy and practice.

The provision of critical thinking teaching methods requires time for planning, organisation and implementation (Fisher, 2007). For instance, in order to facilitate effective argumentation, a key critical thinking skill, teachers must: ensure that learners are taught ground rules for effective debating; allocate sufficient time for open-ended reasoning, thinking and enquiry; and ensure that they themselves have the skills for effective and challenging questioning (Alexander, 2006). However, the common perception, whether accurate or not, and which is portrayed in the national press, is that teachers, despite the aforementioned notion of autonomy, are under extraordinary pressure to deliver a narrow curriculum within a limited timeframe. For instance, a 2014 survey of 944 teachers reported that:

> *Some 79 per cent said targets prevented them teaching a broad curriculum, while 60 per cent admitted they "teach to the test". More than a third said they spent "as much time as possible practising tests". In a further disclosure, 38 per cent of teachers said they do less practical work and 37 per cent admitted to skipping topics that they know will not be tested in an exam (Paton, 2014; p.1).*

National Curriculum guidance was revised in September 2014 (National Curriculum, 2014b) and now lays out some generalised requirements which include references to critical thinking:

- [Students should be taught to] articulate and justify answers, arguments and opinions (Spoken language requirements, p. 18)
- [Students should be taught to] support ideas and arguments with any necessary factual detail (Writing requirements, p. 82)
- Teaching should equip pupils with the skills and knowledge to explore political and social issues critically, to weigh evidence, debate and make reasoned arguments. (Citizenship requirements, p. 201)
- Students should be taught to think critically (Art requirements, p. 199)
- [Students should be] equipped with the skills to think critically and debate political questions (Citizenship requirements, p. 201)
- Teaching should equip pupils to ask perceptive questions, think critically, weigh evidence, sift arguments, and develop perspective and judgement. (History requirements, p. 219)

At the time of writing, however, this prescription does not provide statutory guidance or recommendations for schemes of work, for how to develop critical thinking skills. Neither is critical thinking assessed, unless students specifically opt to study it as an As Level course (for 17 year olds). Thus, there is little incentive for schools to devote time, energy or resources into its teaching, and it has been suggested that the teaching of critical thinking skills is very much left to the arguably hard-pressed teacher, who can be tempted to sacrifice the pursuit of critical thinking techniques to the ‘serious business of getting the students to pass the exams' (Morse, 2012, p.1). It is suggested furthermore, that secondary school teachers themselves lack the critical thinking skills, the vocabulary and even the training required to be able to effectively develop the critical thinking skills of their pupils (Marin and Halpern, 2011). Whilst the latter finding refers to the US secondary school setting, rather than England, it may go some way toward explaining why there is such little evidence of its effective delivery in the curriculum in England.

Eyre (2011) in a report for the Policy Exchange, suggested that extension activities which have the potential to develop critical thinking, are best provided as extra-curricular activities. This pragmatic approach seems to offer a compromise between the need to provide for the development of critical thinking and Morse's (2012) 'elephant in the room' (p.1) - the need to pursue exam targets.

This study explores the effectiveness of such an extension activity, and is evaluated in terms of its effect on change in critical thinking over time.

2.2.3 How is CT taught and assessed?

Critical thinking teaching approaches

Ennis (1989) separated critical thinking teaching approaches into four categories - general, infusion, immersion and mixed - and explored the merits of each. The *general* approach involves interventions that offer instruction in critical thinking separate from course content. For instance, Riesenmy, Mitchell, Hudgins, and Ebel (1991) used a quasi-experimental design, to evaluate the effect of specific instruction on the critical thinking skills of 66 primary age schoolchildren. They found that treatment groups performed better on critical thinking tests than control groups, with a large effect size (Cohen's $d = 1.31$).

Infusion interventions involve subject matter where students are required to think critically, and where critical thinking objectives are made clear, but not taught separately. For instance Zohar, Weinberger, and Tamir (1994) used an experimental design to evaluate the effect of the infusion method on 678 young adolescents' general critical thinking and their biological critical thinking, finding significant and positive large effects when comparing experimental and control groups (critical thinking skills effect size: $d = 2.0$ and biological critical thinking effect size: $d = 2.5$).

The *immersion* approach requires learners experiencing critical thinking activities, but where critical thinking objectives are not translated transparently into learning goals. Kamin, O'Sullivan, and Deterding (2002), working with medical students, employed a quasi-experimental design to evaluate differences between different levels of immersion (video + online discussion, video + face to face discussion, and text + face to face discussion). Whilst they did not report effect sizes, they did report that the online discussion group scored the highest in tests of critical thinking.

The *mixed* approach involves a combination of the *general* and either the *immersion* or the *infusion,* whereby students experience critical thinking activities at the same time as receiving specific critical thinking instruction. For instance, McCarthy-Tucker (1998), working with secondary age children, reported significant pre- to post-test differences between experimental and control groups, when applying the mixed approach to the teaching of critical thinking.

Abrami *et al.,* in their 2008 meta-analysis of 117 critical thinking interventions, reported effect sizes[3] for the four types of intervention, reporting that the instructional approach that most significantly outperformed all others was the mixed approach, with the immersion and infusion approaches being the least effective. The mean effect size for all 117 interventions was Hedge's $g+ = .34$. In comparison, the mixed approach resulted in an average effect size of $g+ = .94$, whilst the immersion approach produced an effect of $g+ = .09$. Abrami *et al. (op.cit.)* also compared different pedagogical grounding effects, reporting that CT interventions which involved trained instructors produced an average effect size of $g+ = 1.00$, compared with interventions where critical thinking was simply mentioned as a course objective, which produced an average effect size of $g+ = .13$. They also found greater intervention effects on secondary students ($g+ = .69$), compared with primary age children ($g+ = .52$) and VIth form students ($g+ = .10$)

Abrami *et al.*'s (*op cit.*) meta- analysis findings have influenced the design of the current intervention. It was designed for secondary age children and presented critical thinking materials using a mixed approach, whereby critical thinking was taught specifically through critical thinking exercises (Kurfiss, 1998; Ennis, 1993), with specific instructions being offered at different points in the course (see Appendix 13), and whilst students were also immersed in critical thinking activities through the

[3] Effect sizes were measured by Hedge's $g+$, an estimator related to Cohen's d, but adjusted for sample size (Hedges and Olkin, 1985).

experience of tutor-led argumentation (Kuhn and Crowell, 2011; Dwyer *et al.,* 2014). The argumentation approach is discussed more fully in Section 2.2.4.

Critical thinking assessment

Since attention to critical thinking has increased significantly since it emerged as a construct in the work of John Dewey (1910), and since it has now become an important goal of contemporary education (Dwyer, Hogan and Stewart, 2014), it follows that it should be assessed. Assessment is necessary for the purposes of examining the effectiveness of critical thinking interventions, diagnosing students' critical thinking learning competency, identifying instructional needs, providing students with feedback about their performance, motivating them to achieve more, informing teachers about the success of their teaching methods, and providing data regarding student achievement (Ennis, 1993). Commonly, critical thinking assessments evaluate the learner's capacity to make logical deductions and inferences, and evaluate the strength of an argument (Magno, 2010).

There are a number of critical thinking assessments that have been designed for use with college or university students (e.g., Watson and Glaser, 1980; Facione, 1990a; Educational Testing Service, 1990, 1994; American College Testing Program, 1992). For instance, the Watson Glaser Critical Thinking Test (Watson and Glaser, 1980), widely used in colleges, universities and the workplace, evaluates five aspects of critical thinking, all related to argumentation skills. These are inference, recognition of assumptions, deduction, interpretation and evaluation of argument. Assessment methods, also intended for older students, include self-report and peer-assessment (e.g., Spitzberg, 2011) and reflective essays (e.g., James, Hughes and Cappa, 2010). However, these methods have been found to be unsuitable for use with younger school students, both in terms of their language, effectiveness, execution and/or content. For instance, Spitzberg (2011) found that peer-assessment of student competence was unrelated to critical thinking, whilst James, Hughes and Cappa (2010), who attempted to assess critical thinking through the analysis of reflective essays, reported that assessors found the task so difficult that they ended up ‘ having recourse to their own internal or intuitive apprehension of the level of [critical thinking] demonstrated’ (p. 293).

Very few critical thinking tests have been constructed for use with children, and those that have tend to be either written for the American context, aimed at older ‘high-school’ students, are expensive to administer, and/or are subjective and time consuming to mark. For instance the Ennis-Weir Critical Thinking Essay Test (Ennis and Weir, 1985) requires each participant to write a US-related essay of nine paragraphs, which is then subject to teacher-grading of each paragraph. Similarly, the Cornell Critical Thinking Test X (Ennis, Millman and Tomko, 1985), which costs $30 per student, is designed for US high school students, and also uses US-related topics.

More recently, studies (e.g., Shehab and Nussbaum, 2015) have begun to use the Need for Cognition (NFC) scale (Cacioppo, Petty and Kao, 1984) as a measure of critical thinking, because of its high correlation to critical thinking dispositions (Stedman, Irani, Friedel, Rhoades and Ricketts, 2009). The latter authors, in a study of 317 undergraduates, found that critical thinking and NFC were significantly correlated ($r = .63$, $p < .05$). The attraction of this scale is its ease of administration (it is a self-report measure), but the disadvantage is that results vary according to the domain of interest (Shehab and Nussbaum, 2015). Nevertheless, studies such as those by Nussbaum (2005) have reported that students with high NFC tend to display higher argumentation skills, with a moderate effect (d = .57), compared with those with lower NFC. Wang, Pascarella, Nelson Laird and Ribera (2014), reporting on a study of 1914 undergraduates, reported that NFC correlated moderately (r = .17) with 'higher-order learning', a self-report measure of metacognition. However few studies so far have explored the use of the NFC scale with children.

The lack of critical thinking assessments, especially for use with children in England, is surprising in the face of evidence that the emergence of critical thinking is a crucial feature of cognitive development in early adolescence (Inhelder and Piaget, 1958; Piaget, 1964; Flavell, 1979; Kuhn, 1992), and in the face of the acknowledged value of critical thinking as an educational objective (Taylor, 2012). However, Ennis (1993), who was responsible for the creation of a number of the above mentioned critical thinking tests, suggests that it is appropriate for teachers to create such tests for themselves, proposing that such a test could be used for 'diagnosis, feedback, motivation, impact of teaching and research' (p. 184).

Offering useful guidelines for the construction of such tests (p. 180), Ennis (*op.cit.*) suggests that critical thinking assessment has been neglected because it is 'difficult to do well' (p. 179). Despite this, he confirms that it is possible. He reviews the purposes of critical thinking assessment, suggests a list of critical thinking topics which a critical thinking test should cover, gives advice regarding how to use the test for motivational and developmental purposes, and warns educators of the traps involved in critical thinking testing. Relevant to the creation of a test for the current study, he suggests that such a test should assess the learner's capacity to judge the quality of an argument, and that students should be able to see their scores, in order to be able to identify where their strengths and weaknesses lie. He warns that the use of identical tests (for pre-test to post-test purposes) may result in the students simply remembering the answers to the first test, especially where feedback has been given. However, he also warns that the use of 'alternate forms' may threaten reliability - therefore care must be taken to ensure the comparability of such parallel tests. An important limitation to the construction of critical thinking assessments involves the content of the material, which should be relevant to the cultural beliefs of the participants. This is an important consideration where the intended participants are from a multi-cultural background. So-called 'moral' issues, for instance, may be approached from very different viewpoints, according to cultural allegiances. Different beliefs about say, stealing to feed one's family, may lead to justifiably different responses to questions that explore the arguments involved in such a scenario.

An important consideration for this study is Ennis' suggestion that critical thinking tests are motivational tools in themselves. This can, however, potentially lead to a method-effect whereby taking the tests results in an improvement in subsequent tests.

Finally, Ennis offers basic challenges that the creator of a critical thinking assessment should address. In particular, he points out that such a test should be related directly to the operational definition of the construct, and that the test material should be suitable for the age group of the participants. In assessing commercially available critical thinking assessments, he rues the lack of objective, multiple choice tests, and concludes therefore that 'it is necessary to make your own test' (p. 184).

Ennis' considerations have been taken into account in the construction of the test employed in the current study. It was adapted from the Defining Issues Test (Rest, Narvaez, Thoma and Bebeau, 1999), a multiple choice test originally designed for older US students. In its adapted form, it assesses a younger learner's capacity to evaluate the strength of an argument. It is based upon scenarios relevant to young adolescents in England - scenarios that revolve around moral, social and ethical dilemmas. In the current intervention, following the test, students were able to see their scores, and were therefore able to learn from their mistakes (albeit potentially precipitating a method-effect). Alternate forms were used for each test (and were therefore evaluated for a*lternate forms reliability* testing). The application of the test is described more fully in Section 4.5.1 and the results of alternate form reliability analysis is reported in Section 5.6.

2.2.4 Developing CT through argumentation

> *'Together with inquiry, argumentation constitutes the best possible education we can offer young people for life' (Kuhn, 2005, p. 130).*

This section argues that critical thinking, in an educational setting, can be developed by providing for the practice of the socio-cognitive conflict of collaborative argumentation. In support of this argument, it explores related literature, and attempts to identify gaps in research that have the potential to be addressed in this study. It is presented as a series of questions that might commonly be asked about argumentation and how its associated skills might be developed.

2.2.4.1 Why argumentation?

Brookfield (1987) separates critical thinking into constituent components. He suggests that it includes the identification and challenge of assumptions, the exploration of alternative viewpoints and a sceptical approach to fixed belief systems - all elements of the process of argument. It has also been said, however, that critical thinking is 'substantially greater than the sum of its parts' (Simpson and Courtney, 2002, p. 93). This is because critical thinking is more than a skill - it is associated with a set of learner dispositions which influence their whole approach to learning - dispositions which include open-mindedness, inquisitiveness, the search for truth, an analytical approach, systematic organisation, perseverance, willingness to self-correct, cognitive maturity and self-confidence (Facione, Sánchez, Facione and Gainen, 1995). Whilst writers stress that critical thinking is more than just a skill, they nevertheless highlight *argumentation* as a key activity that engages critical thinking processes (Kuhn, 1992; Kurfiss, 1998; Watson and Glaser, 1980; McPeck, 1981; Paul and Elder, 2009; Brookfield, 1987; Facione, 1990).

There are frequent references to the value of developing critical thinking in educational literature (Felton and Kuhn, 2001) and a number of writers have concluded that argumentation is a key skill that develops it (Mayer, 1986; Kurfiss, 1988; Walton, 1989; Bell, 1991; Kuhn, 1992; Sanders, Wiseman and Gass, 1994; Kuhn, 2005; Jonassen and Kim, 2010; Cavagnetto, 2010; Kuhn and Crowell, 2011;

Preiss, Castillo, Flotts and San Martín, 2013; Dwyer *et al.,* 2014). They argue that the construction of an argument (making a claim, questioning and weighing evidence, organising information, challenging others' ideas, seeking clarification, identifying argument fallacies, challenging inferences, evaluating arguments, planning for rebuttal etc.) involves all of the steps that are essential to critical thinking (e.g., interpretation, analysis, evaluation and inference). Indeed Kuhn (1991, 2005) conceptualises critical thinking *as* argument. Walton (1989, p. 169) claims that argumentation is of 'key importance' to critical thinking, that reasoned dialogue is the 'root notion', and that the specific argumentation skills that should be developed are argument analysis, counter-argument and the detection of bias. Mayer (1986) suggests that the key to the development of critical thinking 'lies in creating conditions for participation rather than passivity, and in providing opportunities for emotional engagement with the materials' (p. 251).

Despite such theoretical argument, experimental studies that empirically evaluate the effect of argumentation on critical thinking however are rare. Sanders, Wiseman and Gass said as much in 1994, and it appears that the situation is still very similar today. Those papers that do report significant associations frequently do not report effect sizes (e.g., Bell, 1991; Sanders, Wiseman and Gass, 1994; Preiss, Castillo, Flotts and San Martín, 2013). Green and Klug (1990), for instance, in an analysis of various teaching methods comparing argumentation with essay writing, lectures, student presentations and reviews, conclude that critical thinking is best developed through debate. Effect sizes were not reported, but the authors concluded that debate is the appropriate arena for the development of critical thinking skills, since it gives protagonists an opportunity to challenge definitions, explore value premises, probe for misuse of evidence, question reliability of evidence and de-construct logical arguments. In view of the dearth of studies that report effect sizes, and in view of the increasing importance place upon the reporting of effect sizes (Coe, 2002), this study will attempt to contribute to the field by addressing this important gap in knowledge.

According to Kuhn (1992), argument is the most significant way in which critical thinking and advanced reasoning manifests itself in real-life situations. Indeed, she suggests that no other kind of thinking 'contributes more - to the quality and fulfilment of people's live, both individually and collectively' (p.156). Western society provides many examples of this notion in action - from school committees, to the jury system, to the House of Commons. It follows, that argumentation should feature as a key element of critical thinking intervention and that participation in argumentation should hypothetically predict critical thinking progress - both notions that are explored in the current study.

2.2.4.2 What is argumentation?

Argumentation is the process of producing, evidencing and justifying reasons for asserting a position - leading to argument, the product of argumentation (Kuhn, 1992). Since argumentation involves analysis, evaluation and reasoning, it is typically characterised in the literature, as in this study, as a process that develops critical thinking, a key metacognitive skill (Kuhn, 1991; Walton, 1989; Sanders, Wiseman and Gass, 1994; Voss and Means, 1991).

Argumentation is frequently described according to Aristotle's categories of *apodictic*, *dialogic* and *rhetorical* argument. *Apodictic argument* is based upon the presentation of facts in order to support a claim, the typical outcome of which is that the claim cannot be disputed (Van Eemeren and Grootendorst, 2003). Whilst teachers may employ apodictic argument in their presentations, students are rarely called upon to make use of such techniques (Jonassen and Kim, 2010).

Dialogic argument involves individuals engaging in discussion with each other, thinking 'on their feet', synchronously employing their critical thinking skills of evaluation and reasoning. Here, the aim is either to prove an opponent wrong, as in *debate* (Van Eemeren and Grootendorst, 2003), or to come to an agreed perspective, in order to reach a collaborative decision, as in *discussion* (Walton, 1989). The former model involves formal stages of confrontation, opening arguments, argumentation, counter-argument and conclusion. The latter requires protagonists to explore alternative arguments, attempting to prove or disprove them in order to reach agreement; its importance lies in the concept of cognitive co-construction, the notion of learners actively contributing to the learning process. Clearly, the subject matter of either type of discussion should provide for more than one solution - otherwise there is no argument. It follows that *counter-argument* features largely as a useful skill in this context.

Rhetorical (or monological) argument, whether presented as a speech or in a written piece, appears to be a less complex process, as the argument is generally constructed asynchronously, giving the arguer the opportunity to weigh evidence and provide balanced reasoning for a position. A rhetorical argument is generally characterised as a dialogue between the arguer and a non-participative audience, a 'missing other'

(Kuhn, 2005, p. 142), whereby the object is to attempt to convince the audience of a claim or position that the arguer holds, irrespective of the position that members of that audience hold (Toulmin, 1958). The principal skill in rhetorical argument is therefore persuasion. This skill is taught to a certain extent in the English National Curriculum, within the literacy syllabus, to children aged 7 to 11 (National Curriculum, 2014) and is used extensively in the context of essay writing, throughout the education system. Persuasive argument involves the arguer first making a claim, then qualifying it with data, and/or a supportive theory or evidence, before adding a warrant to support the claim. The argument may then back up that claim, add further qualifiers, and so on. However, according to Leitão (2003), this form of argument typically neglects an exploration of both sides of an issue, tends to be presented from just one perspective, and therefore neglects to make full use of counter-argument.

It is suggested that, in pursuit of a reasoned conclusion, learners make more progress in collaborative argument than on their own (Moshman and Geil, 1998) and that indeed a conclusion cannot be a reasoned and rational one unless it has stood up to the test of an opposing idea (Habermas, 1991). This suggests that engagement in dialogic argumentation is likely to be more effective, in terms of critical thinking outcomes, than participation in individual rhetorical argumentation (e.g., in expository writing) - a notion that has influenced the design of the intervention.

2.2.4.3 Why is argumentation important?

Authors, educators and researchers agree that an important goal of education is to develop self-regulated learning, and in particular, the metacognitive skill of critical thinking - the capacity to evaluate evidence, explore claims, theories and counter-theories and construct a reasoned argument, in order to reach a balanced conclusion (Phan, 2010; Hattie, 2009). Argumentation skills are seen as essential at all levels of education, especially if a learner is to succeed in higher education (Joiner and Jones, 2003; Nussbaum and Schraw, 2007). Weaknesses in argumentation skills can have negative implications for an individual's education progress; therefore studies have sought to develop interventions that are specifically targeted at the development of argumentation skills (e.g., Schwartz, Neuman, Gil and Ilya, 2003).

It has been theorised that metacognition, an element of self-regulated learning, should be amenable to development through the practice of argumentation, since reflection on one's own thinking, a prerequisite for the construction and development of an argument, is a metacognitive process (Kuhn, 2000a; Efklides, 2008). This notion has been tested empirically (e.g., Kuhn, Shaw and Felton, 1997; Cavagnetto, 2010; Kuhn and Crowell, 2011; Pinto Iliceto and Melogno, 2012), with studies reporting significant relationships between argumentation and metacognition. For instance, Pinto *et al*., (2012) in a study of 533 Italian undergraduates detected a strong correlation of $r = .511$ between metacognition and argumentation skills. The rationale for the existence of this relationship is that the process of argument raises a learner's thinking to a metacognitive level since the arguer must have a sophisticated capacity to explore their own thoughts, to self-regulate their reflection, and to be aware of their logical sequencing (Leitão, 2007). Gillies and Khan (2007) for instance, report that students whose lessons featured argumentation used more metacognitive techniques such as elaboration and justification for their ideas, compared to those whose lessons did not. Moshman (2004) adds the notion that argument is a means of serving one's purposes and involves a 'metacognitive commitment' to the coordination of ideas (p. 221).

In free-flowing '*dialogic argumentation'*, the discipline of formulating, developing and defending an argument exercises a range of metacognitive skills, requiring the learner to control and regulate their thinking process, to believe in their capacity to propound their argument and to remain motivated to maintain their stance in the face of counter-argument. Dialogic argumentation helps learners to rehearse, articulate, synthesise and expand upon their knowledge. As students are questioning each other and defending their own arguments, they engage in the metacognitive control and regulation of their critical thinking.

Studies have established positive associations between the development of critical thinking, a key metacognitive skill, and participation in argument (Fisher, 2007; Green and Klug, 1990; Kuhn and Crowell, 2011). Reporting on an experimental study of US undergraduates, Green and Klug (1990) noted statistically significant differences between an experimental (debate) group and a control group. However, they did not report effect sizes. In a quasi-experimental study of US secondary age school students, Kuhn and Crowell (2011) compared debate and non-debate groups, noting significant differences in reasoning outcomes, and reporting a medium effect size[4] of η^2 = .097 for the effect of the intervention condition. It is suggested that in

[4] Cut-off criteria for eta square effect sizes are given in Section 4.7.2.

the process of communicating ideas to other protagonists, the skills of the weighing of evidence, metacognitive evaluation and reasoning are practised, honed and enriched, and knowledge is advanced. At the same time the participant is exposed to others' ideas, viewpoints and reasoning strategies. As a participant responds to metacognitive questioning within argumentation ('How do you know that? Are you assuming that…?'), they are more likely to ask such metacognitive questions of themselves (Olson and Astington, 1993). Kuhn (1991) further suggests that the process of forming an argument and articulating that argument to someone else, aids the critical thinking process, whether there is a reaction or not. She suggests that, 'If I go into a room with an orangutan to share some new ideas, the orangutan just sits and eats its banana, but I come out of the room clearer about my ideas' (p. 115).

Jonassen and Kim (2010), who discuss the effects of argumentation on the development of scientific thinking, suggest that embedding and developing argumentation activities in learning programmes is a valuable method of promoting ways of thinking, conceptual change and problem solving. Firstly, they report that argumentation is central to the development of scientific thinking, since it encourages an inquisitive way of thinking and a search for knowledge. It assists in making conceptual change visible through exposure in dialogue, in contrast to more traditional methods of summarising and reporting, where the expectation is that the student will report 'facts'. Finally, they report associations between argumentation and problem solving, which they suggest is 'the most ubiquitous, engaging, and essential intellectual activity in everyday and professional lives' (p. 441).

2.2.4.4 *When is the optimum time to develop argumentation skills?*

The majority of researchers contest that most older students do not understand or engage in argumentative discourse. They tend to make overuse of generalisations; they place a greater emphasis on the analysis of their own side of the argument than on an examination and rebuttal of the other; they form assertions from insufficient evidence; and they tend to develop their argument from their own opinion rather than from an examination of the evidence (Jonassen and Kim, 2010; Reznitska, Anderson, McNurlin, Nguyen-Jahiel, Archodidou and Kim, 2001; Means and Voss, 1996).

Kuhn (1992) suggests that it is crucial that critical thinking is nurtured in early adolescence. Before this time, whilst children should be encouraged to develop the skills of discussion, she suggests that they tend to hold absolutist views about how evidence might be related to the evaluation of argument. This perspective reflects the Piagetian notion that the 'formal operations' phase (from around ages 11 - 16) is the one in which critical thinking, a metacognitive skill, emerges and grows rapidly (Inhelder and Piaget, 1958; Piaget, 1964; Flavell, 1979), and the neuroscientific perspective that adolescence is a time of major neurological change (Blakemore and Frith, 2005). Importantly, in Kuhn's (1992) study of 160 participants aged between early adolescence and their late sixties, she reported that if critical thinking, and in particular the related skill of argumentation, was not developed during early adolescence, and if those learners reached adulthood without such skills, then they were unlikely to further develop in adulthood. Thus it appears that the development of an intervention that has the potential to develop argumentation skills in this window of early adolescence may be an endeavour worth pursuing. This is especially poignant since the specific teaching of neither argumentation nor critical thinking skills figures prominently on national curricula in England until As Level (i.e., for 17 year olds), which according to Kuhn's research, is possibly too late.

2.2.4.5 *What prevents the natural development of argumentation skills in children?*

Kuhn (2005) suggests that children develop their argumentation skills in phases, but that they need assistance in moving from one phase to the next. Following her observation that a significant number of students never pass the critical thinking element of a US university 'high-stakes' entrance test, and are therefore barred from entry to university, she concludes that the reason that so many fail the test, is a basic failure to perceive an argument from a point of view other than their own. She theorises that this argumentation is 'not an intuitively given skill' (p. 147), but is developed through instruction and the specific practice of dialogic argumentation. Comparing the argumentation skills of three age groups (school, college and undergraduate), Kuhn and Felton (2000) found that, whilst critical thinking performance tended to increase overall with age, there was a common weakness across all age groups, in terms of the capacity to perceive the counter-argument. They concluded that critical thinking *progress*, and in particular the capacity to construct and de-construct a counter-argument, was less related to age, and more to instruction and experience.

Kuhn (2005) offers a three-stage framework for the categorisation of the phases of development of argumentation skills, suggesting that when a child first begins debating, they may adopt an *absolutist* point of view. The absolutist tends to be certain of their facts; they may try to close down a discussion point by saying something like, 'It's a fact,' or, 'They are experts, so they should know.' This can have the effect of rendering further discussion unnecessary, since a 'solution' is offered as a certainty. This is not a very helpful stance in dialogic argumentation, especially where the topic has typically been chosen specifically because there is likely to be more than one solution.

The *multiplist* viewpoint is a little more advanced. Here, the discussant proposes that nothing is certain - that everyone has an equally valid opinion, and that a 'solution' is therefore impossible to find. They may say, 'I don't think that we are going to find an answer to this question because everyone has a right to their own opinion,' or 'This is a very difficult question and all of the experts disagree, so how can we debate it?' However, the multiplist perspective does not advance the argument, since such individuals tend to believe that nothing is certain. In practice, multiplists often go round in circles. Within the boundaries of educational equality of opportunity and anti-discrimination, the challenge is to move children beyond the notion that all opinions are equally correct, to an evaluative position that is achieved through a reasoned examination of alternatives, evidence and counter-perspectives (Kuhn and Udell, 2007). However, despite the ability to reason being an implicit, if not explicit, aim of English education (Higgins, Hall, Baumfield and Moseley, 2005; Meyer, Haywood, Sachdev and Faraday, 2008), the notion of counter-argument rarely features in National Curriculum guidelines, therefore it is unsurprising that children are not practised in its use (Coffin and O'Halloran, 2009).

The most advanced debaters approach the process from an *evaluativist* perspective. The evaluative discussant sees the benefit of argument for its own sake. They demonstrate in debate that knowledge can be sought through an open-ended process of evaluation, evidence, reasoning and judgement. They may say, 'Experts disagree on this, but on balance it looks as if…' or 'James has offered a persuasive argument on this topic, but on balance I tend to agree with Margaret, because…' Thus the evaluativist will conclude that whilst everyone might have an equal right to hold an opinion, some opinions are better evidenced, better reasoned and better argued than others.

Moving a child from the multiplist phase, the prevailing mode in the early secondary classroom, to the evaluative phase is, according to Simon (2001) generally the most challenging. She suggests that whilst students are encouraged to share their views, it is commonplace for schools to promote the value that all opinions should afford equal prominence. This results in students willingly proposing their own viewpoints, but being unwilling to rebut or offer counter-arguments in response to those of others. A further obstacle to the development of balanced reasoning is the sense that the exposition of a counter-argument may weaken the individual's own argument (Baron, 1991, 1995; Nussbaum and Kardash, 2005; Kuhn and Udell, 2007). This has the effect of stifling any progress to the evaluativist phase. The challenge therefore is to provide the conditions for evaluative discussion, in such a way that students can progress incrementally.

2.2.4.6 How can argumentation skills be developed?

According to Kuhn (1992) argumentation skills are amenable to intervention particularly in early adolescence, with the 'most obvious method of strengthening it [being] practice.' (p.174). It is important, therefore, to capitalise upon this window of opportunity. If practice in argumentation is considered an effective method for its development, then the challenge for the educator is to provide opportunities, within that window, for all students who may potentially benefit. This study explores an intervention that may address that challenge.

It is suggested that argumentation skills flourish amid opportunities for open-ended and ill-structured dialogic social discussion on controversial moral, social or ethical issues, about which alternative perspectives could be argued to exist (Walton, 1989; Green and Klug, 1990; Arbaugh and Benbunan-Fich, 2005; Barton and McCully, 2007; Jonassen and Kim, 2010; Kuhn and Crowell, 2011).

As Walton (1995) puts it:

> *basically, in an argument, some key proposition is held to be in doubt, in contrast to an explanation, for example, where the proposition to be explained is generally taken for granted, or at least not subject to doubt or questioning as far as the purpose of the explanation is concerned (p. 49)*

For instance, students may discuss topics like wind-farms or the extension of the national rail network; dilemmas which offer the opportunity to explore the problem of improving conditions for some at the expense of others. It is suggested that engagement in debate on *controversial* issues is important in this context, since it exercises the skills of counter-argument (Kuhn and Udell, 2007).

One might assume that the school environment generates multiple opportunities for practice in dialogic argumentation. A review of the literature however, finds very few examples of open-ended dialogue happening in schools. Why should this be so? At the least the use of open-ended questions is considered good pedagogical practice in the classroom, and it is recommended as good teaching practice as a means of 'extending pupils' oral responses' (Higgins *et al.,* 2005, p. 1). Opportunities for students to discuss collaboratively are few however, made fewer by the lack of teacher skills, lack of National Curriculum direction, very tight timetabling, and curriculum and exam programme restrictions to which schools are currently subjected (Graff, 2003; Coffin and O'Halloran, 2009; Eyre, 2011; Morse, 2012). Furthermore, according to a number of writers including Marin and Halpern (2011), Driver, Newton and Osborne (2000) and Von Aufschnaiter, Erduran, Osborne and Simon (2008), teachers are not sufficiently skilled or confident in argumentation skills themselves and therefore shy away from using the approach in the classroom. Barton and McCully (2007) go further, and suggest that it is beyond the school's remit to provide either the opportunities for students to engage in controversial debate, or the training for teachers to become skilled in debate management - therefore teachers are unlikely to engage in this activity. Furthermore, they suggest that school is commonly perceived as a place of 'sanctuary' from controversy, where a 'culture of politeness' prevails (Barton and McCully, 2007, p. 6).

The urgent need to cover curriculum knowledge can lead to the teacher 'guiding' discussions towards a desired outcome, thus limiting the opportunity for open-ended, complex debate of unanswered questions. Kuhn (2005) observed middle-school[5] debate lessons that were, on the face of it, designed to promote argumentation skills. These, she argued, were typical examples of such lessons. Comparing a 'struggling' school with a 'best-practice' school, she found that even in the 'best-practice' school, students tended to read from prepared crib-sheets, presenting only their own argument, and failing to address or counter those of others. Students tended only to address the teacher, and rarely engaged in discourse with each other, lapsing into 'nothing but consecutive self-expression,' with students not even listening to each other (p. 122). The role of the teacher seemed to be restricted to organisation; indeed the presence of the teacher seems to constrict discussion, rather than encourage it. It appears that such practice served the purpose of, at best, knowledge construction, rather than the development of critical thinking.

The latter observation begs the question as to whether or not the provision of a discursive environment is sufficient for the generation of effective dialogic argumentation, or whether some sort of scaffolding is necessary. Perhaps the students should be left to their own devices and simply given appropriate topics to discuss amongst themselves? Milburn-Curtis (2011) reported significant differences in

[5] US 'middle-school' is for children aged 10 - 14 years.

exhibited levels of higher-order thinking, with a medium effect size[6] (r = .43) when comparing student-led and tutor-led discussions, with the quality of the questions being the factor that determined the level of discussion generated. Other authors have also articulated the necessity for expert questioning, suggesting that 'the [teacher] is only as good as the question [s/he] asks' (Lane Fox, 2008, p. 57) and the student 'will teach himself everything he needs to know - so long as he is asked the right questions' (Ryan, 2008, p. 73). It appears therefore, that the authors of argumentation interventions must pay attention to the necessity for teacher-questioning of the highest order, in order to specifically target critical thinking processes. Such questions, observed in the Milburn-Curtis (2011) study, would encourage learners to:

a) justify their assertions, 'Why do you say that?
b) explore metacognitive connections, 'How does this relate to our previous discussions?' 'Why have I asked this question?'
c) provide evidence: 'Can you give me an example of this?'
d) clarify conceptions: 'Are you saying … or …?' 'What do we already know about this?'
e) probe assumptions: 'You seem to be assuming…' 'How can you verify or disprove that assumption?'
f) probe rationale: 'How can you support your argument?'
g) probe reasoning: 'Are these reasons good enough?'
h) probe evidence: 'On what authority are you basing these assumptions?'

[6] Cut-off criteria for *r* effect sizes are given in Section 4.7.2.

i) reflect epistemologically: ‘Would it be possible for experts to find out for certain, if they studied this problem long and hard enough?’
j) question viewpoints: ‘Is there another way of looking at this?’
k) analyse the strength of an argument: ‘What are the strengths and weaknesses of this argument?’ ‘What argument would an opponent use - and how would you refute it?’
l) probe consequences: ‘What are the implications of…?’

Such observations therefore suggest that effective discussion should allow for free-flowing, open-ended expert questioning, that which does not necessarily direct itself towards a predetermined outcome, and that which allows for interaction between participants. Kuhn (2005) insists that this form of dialogic argumentation, on topics that explore a disputed claim, should become the norm for classroom-based discussion.

But how and why does this form of questioning develop critical thinking? The expert teacher will recognise the level at which the student is operating and will seek to use appropriate questioning to challenge them to the next level. The critical thinking of the absolutist discussant is limited to the presentation of facts and any required associated understanding. In order to move them into the multiplist phase, they will be encouraged to support their statement with evidence (‘Why do you say that? Can you provide evidence for that claim?’). The multiplist, who sits on the fence, cites equality as a desirable outcome of the discussion, and claims that a solution cannot be found because ‘everyone has a right to their own opinions,’ will be challenged to address alternative viewpoints (‘Is there another way of looking at this?’). The evaluativist will be challenged to hone their critical thinking skills by analysing the strength of argument, by challenging fallacies and by exploring alternative theories ('What are the strengths and weaknesses of this argument? Has anyone spotted the camel's nose in this argument? Can we unpick this circular / straw man argument?').

2.2.4.7 What is the importance of counter-argument?

Toulmin (1958) classified features of argument into six elements, the application of which has become widespread in argument studies (e.g., Voss and Means, 1991; Coffin, 2007; Cavagnetto, 2010; Kuhn and Crowell, 2011). The first four are *data, claims, warrants* and *backing*. A *claim* is supported by *data*; a *warrant* is a justification and a logical bridge between the claim and the data; and *backing* proves that the warrant is true, but does not necessarily relate to the main claim. For instance an argument might begin thus:

- 'You should join the Neighbourhood Watch (*claim*).
- Crime has been rising in this area by 8% over the last five years (*data*).
- Neighbourhood Watch is effective in detecting and preventing crime (*warrant*).
- There is a local Neighbourhood Watch community leader available for consultation (*backing*)'.

The final two elements of Toulmin's classification are *qualifiers* and *rebuttal*. A *qualifier* or *reservation* argues for the limitation of the original claim and can have the effect of offering the possibility that the claim is incorrect. The *rebuttal* acknowledges the possibility of counter-argument and attempts to pre-empt it. For instance the previous argument might progress thus:

- 'A Neighbourhood Watch sticker in your window is usually successful in deterring crime (*qualifier*).
- Most criticism of the Neighbourhood Watch scheme comes from people who have never been involved with one (*rebuttal*)'.

The final two are of interest to this study because they embody the process of *counter-argument,* a typical feature of debate, and one which is considered by many to be the most effective element of the debating process, in terms of its capacity to influence the development of critical thinking (Kuhn, 1991; Voss and Means, 1991; Kuhn *et al.*, 1997; Stein and Bernas, 1999; Anderson, Nguyen-Jahiel, McNurlin, Archodidou, Kim and Reznitskaya, 2001; Leitão, 2003). Counter-argument is the process whereby an arguer highlights a point, which could be seen to support an opposite point of view (Leitão, 2000). Kuhn (1991), who writes extensively on the topic of argumentation, suggests that the acknowledgement and handling of counter-argument is the mark of accomplished argumentation. Indeed, she suggests that failure to tackle an opposing viewpoint is the main obstacle, not only to effective argumentation, but also to critical thinking. The notion that the neglect of attention to alternative views is a serious weakness in argumentation techniques is a common thread in argumentation studies (e.g., O'Keefe, 1999; Nussbaum and Schraw, 2007; Stanovich and West, 2008; Wolfe and Britt, 2008). Nussbaum and Kardash (2005), in their experimental study of one hundred and seven US undergraduates, evaluated the effectiveness, in terms of the quality of essays, of direct instruction to use reasoning towards a single argumental outcome, compared with direct instructions to additionally explore the counter-argument. They reported that the most effective group condition was direct instruction to use counter-argument, noting medium effect sizes ($\eta^2 = .09$) when comparing experimental and control groups.

The tendency of students to propose and support only their own side of an argument is known as ‘myside bias’, and a number of argumentation and critical thinking studies have explored the disadvantage of this argument strategy, concluding that myside bias is a significant factor in reducing the persuasiveness of an argument (Perkins, 1985; Perkins, Farady and Bushey, 1991). Wolfe and Britt (2008) for instance suggest that the lack of counter-argument is a common factor of less-skilled argument, and manifests itself in reliance upon opinion rather than an examination of the evidence. A number of studies have explored the phenomenon of myside bias, and have concluded that it is prevalent amongst students of all ages. Toplak and Stanovich (2003) for instance, found that undergraduates tended to offer more myside arguments than ‘other-side’, and suggested that the balance improved with educational experience.

The process of developing the skill of counter-argument is made more difficult since some students believe that the articulation of the counter-argument can have the effect of weakening their own (Baron, 1991, 1995). The latter author reported that students themselves considered that arguments that were presented from only one perspective were qualitatively better than those which explored both sides of an argument. Such students also tended to employ myside bias in their own argumentation strategy.

Baron (1995) thus concluded that myside bias was influenced by students' perception of what constituted 'good' thinking (p. 221). It has also been suggested that the younger the student, the more likely it is that they will perceive an argument from a self-centred perspective, rendering the process of the development of argumentative skills one of 'sociocognitive decentering' (Leitão, 2003, p. 275). The latter perspective accords with the Piagetian notion that argumentation skills develop in the formal operations phase of child development. The assertion also resonates with that of Kuhn (1992) who suggests that argumentation skills emerge and develop rapidly in early adolescence, a notion that has influenced the design of the current intervention.

Despite the perceived difficulty of persuading students to explore counter-argument, its importance is explained by Wolfe, Britt and Butler (2009), who suggest that where an individual proposes an argument alongside its counter-argument, and where that individual subsequently rebuts the counter-argument, the result is a reduction of cognitive load on the part of the audience, thus making acceptance of the original argument more likely. Thus the inclusion of counter-argument strengthens the persuasiveness of the arguer. It follows that the use of counter-argument and its rebuttal is an important argumentation strategy.

Kuhn and Crowell (2011) compared experimental and comparison groups who had instruction in counter-argument for one year, for two years, or not at all. They found that those who did not receive specific intervention in counter-argument made 'little, if any' (p. 551) gains in that skill over time (one year and two years), whilst those who received instruction which specifically developed such a skill, made significant progress over time, with medium effect sizes (Cohen's d = .87 after one year of study and d = .94 after two years of study). They concluded not only that, 'counterfactual reasoning, dual-perspective reasoning, and integration of opposing arguments are essential building blocks of sophisticated, nuanced real-world argumentive reasoning' (p. 551), but also that the skill of counter-argument is only likely to develop if it is specifically taught. Interestingly, the notion of counter-argument is noticeably absent from the English National Curriculum, where references to argument are construed more in terms of conciliatory 'discussion'. English curriculum prescription tends to suggest that argument should involve, 'exploratory and tentative comments where ideas are being collected together, and reasoned, evaluative comments as discussion moves to conclusions or actions' (National Curriculum, 2014a, p. 15).

In view of the widespread acknowledgement that counter-argument is a key argumentation skill, which has a significant influence on the development of critical thinking, and in view of its noticeable absence in the National Curriculum, it follows that interventions that seek to develop argumentation skills should make specific provision for the development of counter-argument. Therefore, in view of the perceived importance of counter-argument, and the extent to which it is absent from the English National Curriculum, this factor forms the core of the current intervention. The study then proceeds to evaluate the capacity for counter-argument to raise individuals' critical thinking (see RQ2). Since studies which explore the relationship between counter-argument and critical thinking in young adolescents are rare, and even more rare in the English context, it is suggested that this element of the study has the potential to contribute to the body of knowledge on the subject.

2.2.4.8 How can argumentation be developed on a learning platform?

Arguing that discussion encourages the development of higher-order thinking, Bai (2009, p.157) contends that 'critical thinking happens in good discussions.' However, it has been suggested that dialogic argumentation rarely occurs in the school classroom, because of time constraints and lack of teacher skills (Coffin and O'Halloran, 2009). This is despite the classroom being 'the most promising arena for practicing and developing argumentive thinking skills' (Kuhn, 1992, p. 155). A number of recent studies have explored the notion that internet-based discussion forums can offer the opportunity for teachers to work directly and interactively with students, for the promotion of socio-cognitive processes and the facilitation of critical thinking through collaborative dialogic argumentation (Yang, *et al*., 2005; Newman, Webb and Cochrane, 1995; Hettinger, 1995; Bullen, 1998). Teachers can capitalise on the enthusiasm and fascination for internet-mediated forum discussion by using this medium to construct, deliver, manage and assess dialogic argumentation. It is only a short leap in conceptualisation from the sort of forum discussion in which young people already engage, to a tutor-led discussion that is presented on a medium already so familiar to them. This notion has been observed in empirical research. Reznitskaya, Anderson, McNurlin, Nguyen-Jahiel, Archodidou and Kim (2001) for instance found that students who discussed in collaborative groups produced more effective arguments in terms of the number of arguments, counter-arguments and rebuttals, than those who reasoned alone, with a small effect size (Cohen's d = .27).

It has been found however, that students rarely have the time, in a fast-flowing forum discussion, to construct a full argument and see it through to a conclusion. Rather they get caught up in the dynamic flow of the discussion, creating and following multiple threads, and multi-tasking to a degree that does not exist in a face-to-face discussion (Coffin, 2007). Researchers have concluded that the socio-cognitive scaffolding of discussions is therefore necessary, whereby teachers offer prompts and targeted support during the discussion (Clark, Sampson, Weinberger and Erkens, 2007; Jonassen and Kim, 2010; Darabi *et al.,* 2013). Through tutor-modelling and targeted questioning therefore, participants learn the conventions of effective dialogic argument whilst engaging in it. In practical terms, on an internet discussion board, the teacher leads the discussion, which is visible to the participants as a dynamic 'thread', prompting participants as necessary, with open-ended questions such as those exemplified previously in this section. The teacher can scaffold the development of argument by modelling the vocabulary of critical thinking, acknowledge effective reasoning, offer counter-arguments and summarise (Reznitskaya *et al.,* 2001).

In online forum debates it has been noted that the browsing of discussion posts can also be an effective argumentation activity (Dennen, 2008). Browsing, or 'pedagogical lurking' as it is sometimes known (*ibid.*), is the phenomenon whereby participants will watch a discussion, but will not actively make contributions. Several studies have suggested that it is has the potential to be an effective learning strategy (e.g., McKendree, Stenning, Mayes, Lee and Cox, 1998; Beaudoin, 2002; Dennen, 2008). The principle at hand is that replying to a post is not the only way that an individual can participate in a discussion - that there is much to be gained from following the discussions of others. It is suggested that the browsing of a forum discussion is equivalent to the notion of 'listening' to a discussion in a classroom

(Wise, Hausknecht and Zhao, 2014). In the face-to-face classroom, it is acknowledged that there is not enough time in a lesson for everyone to speak - yet listening is considered a valid learning activity (National Curriculum, 2014). Studies relying upon participant opinion have reported that students consider browsing a worthwhile activity (e.g., Beaudoin, 2002; Dennen, 2008), but few studies have empirically evaluated the *effect* of such activity on measured learning outcomes. This study offers an opportunity to do that. It could be argued that listening activity leaves little visible evidence of learning, whilst observing the forum discussions of others leaves even less of an obvious mark. Fortunately, the technology employed on a modern learning platform enables the teacher/researcher to observe exactly the number of times an individual accesses a discussion post, whether they reply to it or not. Thus with such technology, as in the present study, it has been possible to measure the extent to which browsing may have an effect on specific learning outcomes.

According to Joiner and Jones (2003) computer-mediated argumentation should, in theory, present a number of advantages over the classroom environment. Students should be able to ponder and construct their responses in their own time; they can see a record of the argument as it has developed so far; they can post their responses without waiting for another participant to finish; and they can participate equally. As a result, according to Hollingshead and McGrath (1995), participants are likely to be less inhibited than in face-to-face discussion. Whilst the pace of forum discussion can be frustrating, students in forum groups tend to explore a subject non-sequentially, resulting in broader discussions and greater information exchange (Benbunan-Fich, Hiltz and Turoff, 2002). Critics of the medium focus on the lack of visual cues (such as facial expressions, frowns etc.) and auditory cues (including silences and tone of voice) (Slagter van Tyron and Bishop, 2009). Others claim that the multi-tasking involved in following multiple threads exercises the capacity to deal with cognitive load (Alloway and Alloway, 2012). Comparisons between internet forum and face-to-face argumentation studies result in contradictory conclusions. Joiner and Jones (2003) in their study of seventy-three UK undergraduates, reported that the quality of forum discussion was lower than that of face-to-face, but that the difference was not statistically significant. However, they did not compare the two conditions within the same time frames; the face-to-face discussion took place over one hour and in a timetabled session, whilst the forum discussion took place over two weeks, in the students' own time. On the other hand, in a meta-analytic study, Darabi, Liang, Suryavanshi and Yurekli (2013) summarised that forum discussion was a particularly effective learning tool if there was evidence of structure, interaction, complexity, clear expectations, expert questioning and, crucially, a direct interaction between teacher and students. These considerations have informed the design of the current intervention (see Section 3.2.2.2).

Rhetorical argument can also be developed on a learning platform, using 'persuasion instructions' (Nussbaum and Kardash, 2005, p. 157). Here students are instructed to compose an argument on a specific issue, and are supplied with guidelines about how to develop their reasoning. The issue under exploration should again be one that lends itself to argument. Since persuasion instructions are reported to have the effect of reducing the quality of counter-argument (*ibid.*), it is important therefore to instruct students to specifically include counter-argument in their exposition (Kuhn and Udell, 2007). This consideration has also been addressed in the design of the current intervention, where the rhetorical argument element, the Devil's Advocate, was structured so that it focused upon the skill of counter-argument (see Section 3.2.2).

2.2.4.9 How can argumentation skills be assessed on an internet-mediated learning platform?

Computer-mediated arguments, whether dialogic or rhetorical, provide an effective means of data collection since they are automatically recorded and can therefore provide a useful resource for the subsequent analysis of argument, for learning purposes, and/or for assessment. According to Inch and Warnick (2002) and Keefer, Zeitz and Resnick, (2000), argument can be assessed by the coding of claims, counter-arguments and rebuttals, with the unit of coding being a sentence, an idea or a whole argument. For instance, in her experimental study of sixty-five young UK adolescents, Milburn-Curtis (2011) used content analysis to evaluate the quality of discussion posts in an online forum, in order to evaluate the effectiveness of teacher-participation in online discussion. Comparing teacher-led groups with student-led groups, she reported a significant group effect with medium effect size of r = .43. She reported that the process of analysing discussion posts was speeded up because they were already transcribed electronically. Associated learning outcomes can also be assessed on an online learning platform. For instance, an objective assessment such as the Defining Issues Test (Rest, Narvaez, Thoma and Bebeau, 1999), which features in the current intervention, can be administered, marked and graded automatically, in electronic format.

2.2.4.10 What are the gaps in argumentation research that have the potential to be addressed in this study?

Researchers of collaborative argumentation studies theorise that argumentation is a socio-cognitive process that promotes critical reflection and knowledge creation (Andriessen, Baker and Suthers, 2003). They argue that discussion written in the form of forum posts supports the development of argumentation skills by providing the time to track through, reflect upon and respond to complex argument. Other studies have suggested that the forum medium is particularly suited to the creation of counter-argument, since learners can ‘see’ the whole discussion; this promotes metacognitive activity by enabling learners to analyse the argument as it unfolds, adjusting their own argument, and thus creating new knowledge (Leitão, 2000). Given that counter-argument is a significant feature of forum debate (Coffin, 2007) and given the significant connection between counter-argument and the development of critical thinking (Leitão, 2000; 2001), it follows that there is merit in researching the relationship between participation in forum dialogic argumentation and measurable critical thinking. Whilst such relationships have been explored in the context of higher education outside England (e.g., Sanders, Wiseman and Gass, 1994; Preiss, Castillo, Flotts and San Martín, 2013), very little research has been conducted thus far in the context of secondary education in England.

According to Wecker and Fischer (2014), the field of research studying argumentation is fragmented and there is much still to do. They call for quantitative, rather than qualitative analysis of the effects of argumentation on measured learning outcomes. They insist, as do others (e.g., Cohen, 1995) that the reporting of effect sizes is essential in order to interpret outcome effects, even where findings are non-significant. They suggest that inter-rater reliability processes should be reported. Finally, they call for studies that evaluate the relationship between argumentation and domain-specific learning outcomes. The current study responds to these calls.

Whilst a particular field of study has focused upon the capacity for collaborative discussion to influence the development of the metacognitive element in the context of self-regulated learning (Kuhn, 1992; Kuhn and Crowell, 2011), there are however few studies that have explored the notion that collaborative discussion may also affect other aspects of self-regulated learning, such as efficacy for learning and/or motivation. Since a field of research has established positive relationships between these self-regulated learning variables, (Boekaerts, 1999; Zimmerman, 2008; Hadwin and Oshige, 2011), it follows that dialogic argumentation may have an effect on aspects of self-regulated learning other than metacognition. This study attempts to explore this notion by evaluating the effect of participation in dialogic argumentation on change over time in aspects of self-regulated learning other than metacognition - i.e., in motivational beliefs and efficacy for learning.

2.2.4.11 Argumentation Summary

Argumentation is recognised as a key metacognitive skill that supports problem solving and critical thinking, and is particularly amenable to intervention in early adolescence. Whilst there is evidence that schools make use of rhetorical argument throughout the curriculum, the use of dialogic argument is less evident. However, it is in dialogic argument that conceptual change takes place, an outcome supported specifically by the use of counter-argument and rebuttal. It has been suggested that dialogic argumentation intervention can be implemented, developed and assessed through the medium of the internet. Literature searches indicate that there is merit in exploring the relationship between critical thinking and argumentation, in the context of secondary education in England, since this phenomenon has not yet been adequately studied.

2.3 Self-regulated learning (SRL)

The literature defines self-regulated learning, which in England is often referred to as 'independent learning', in a number of ways. Indeed, it is becoming known as 'an elusive construct' (Zimmerman, 2011a, p.49). This is not surprising since the literature approaches self-regulated learning from a number of perspectives, according to setting, environmental culture, learner age and stage of development, subject domain and type of task (Zimmerman, 2011a).

However, there is core agreement that the term commonly describes a set of metacognitive dynamic 'on-line' *processes* that enable the learner to take control and to regulate their cognitive and metacognitive strategies, which processes in turn are a function of the learner's interrelated self-regulatory *traits, dispositions* or *competencies.* Those processes, which according to Zimmerman (2000, 2008) function in a cyclical pattern, include goal setting, task analysis, time management, adaptation, control of distraction, effort-management, self-monitoring, causal attribution and self-evaluation. They are driven by a set of traits, which include metacognitive awareness, motivational beliefs and self-efficacy (Schunk, 2008). Such processes involve the learners engaging in learning themselves, rather than being the recipient of something that is imposed upon them through the experience of being taught (Winne, 2005).

Thus the notion of self-regulated learning implies that students learn most effectively by taking responsibility, using their initiative, persevering through difficulties and adapting their learning skills to ever-changing contexts. Aware of their feelings and motivational beliefs, the independent learner governs their own learning process, sets goals, works autonomously, restrains disruptive or impulsive behaviours, maintains concentration, and monitors and evaluates the learning process.

2.3.1 The influence of the historical perspective

Over the past four decades, the study of self-regulated learning has developed into a major research field (Bandura, 1986; Boekaerts, 1999; Pintrich, 2004; Winne and Perry, 2000; Zimmerman and Schunk 2001; Zimmerman, 2008), with the assumed relationship between self-regulated learning and academic achievement being the principal focus. Interest in the study of this relationship persists to this day, with recent studies continuing to report positive relationships between achievement and aspects of self-regulated learning. Lee, Lee and Bong (2014) for instance, in their study of 500 middle school children, reported that self-efficacy was positively related to academic achievement with a large effect size (r = .61). Similarly, Zimmerman and Kitsantas (2014) found that latent measures of self-regulated learning which included a range of SRL components (such as motivation, learning processes and self-efficacy), predicted the academic achievement (Grade Point Average: a summary of grades for a range of academic subjects) of US high school students, with a large effect size (standardised regression weight β = .96). In another recent study of 5805 Italian undergraduates, Mega, Ronconi and De Beni (2014) compared the effects of metacognition and motivation on academic achievement, reporting that motivation was the greater predictor (β = .32, compared with β = .16 for metacognition).

However, the study of 'self-regulated learning' covers a broad field, and the terms metacognition, self-regulation (SR) and self-regulated learning (SRL) are used interchangeably in the literature. Indeed, authors such as Boekaerts (1999), Boekaerts and Corno (2005), Dinsmore *et al.* (2008) and Schunk (2008) suggest that these constructs remain poorly understood and ill-defined, and they universally rue the lack of standard definitions. As this study explores 'self-regulated learning' from a number of approaches (e.g., from a dispositional perspective as well as a process one), it is important therefore to attempt to establish clear conceptual boundaries.

The term *metacognition* originated with Flavell, who described the concept as 'thinking about thinking', a definition ubiquitously used by researchers (e.g., Miller, Kessel and Flavell, 1970, p.613) and education professionals (e.g., Gough, 1991). According to Flavell, the original concept situated metacognition in the mind of the individual, who employed cognitive structures to monitor, control and regulate abstract thinking. Subsequent researchers such as Baker and Brown (1984) developed the construct, focusing on Flavell's separation of metacognition into *knowledge* and *regulation* of learning. The focus on *regulation* of learning has been further developed over time to describe the metacognitive control of cognitive strategies such as planning and information management, and similar control over the use of metacognitive strategies such as the critical thinking and the monitoring and evaluation of the learning experience (e.g., Schraw and Dennison, 1994). Such metacognitive control has since become synonymous with the process of *self-regulation.*

The concept of *self-regulation* has been significantly influenced by the work of Bandura, who was developing the concept concurrently with Flavell's work on metacognition. While Flavell's (1979) work was emphasising the metacognitive aspect of learner development, Bandura was focusing upon its behavioural and emotional aspects, later to include self-efficacy as a prime regulatory influence of the process of self-regulation (Bandura, 1982). Thus, establishing the *social-cognitive* perspective, Bandura emphasised the interaction of learner dispositions, their self-regulatory behaviour and their environment (Bandura, 1986).

Towards the end of the 20th Century, Zimmerman and his colleagues were building upon the work of Bandura, developing the notion of *self-regulated learning* (Zimmerman, 1989; Zimmerman and Schunk, 2001) by focusing on the *process* of self-regulation. Here the concept of self-regulation was applied specifically to the academic domain, and was applied further still in subsets of academic learning such as writing (e.g., Graham, Harris, MacArthur and Schwartz, 1991) or mathematics (Cleary and Chen, 2009). Whilst the constructs of metacognition and self-regulation developed independently of each other, the evolution of self-regulated learning now incorporated aspects of both. This integration resulted in the development of the social cognitive theory of learning, now emphasising metacognition, self-efficacy and motivation as key learner *dispositions* that drive the metacognitive *process* of self-regulation (Zimmerman and Schunk, 2001). Interpretations of self-regulated learning as a *process* differ in their emphasis, but all share important similarities, focusing on the metacognitive control and regulation of a) cognitive learning strategies such as goal setting (Winne, 1996; Boekaerts and Nirmivirta, 2000), volition and task interest

(Corno, 2001; Vollmeyer and Rheinberg, 2006), planning and task management (Zimmerman, Bonner and Kovach, 2002; Sitzmann, 2010) and b) metacognitive strategies such as critical thinking, self-evaluation and self-monitoring (De Bruin and van Gog, 2012, Efklides, 2008; Zimmerman, 2008).

Self-regulated learning is related to *executive functioning*. The latter construct features in the field of cognitive science and describes *cognitive* functions such as working memory capacity, inhibition and task switching. Self-regulated learning on the other hand is widely recognised as a *metacognitive* process that controls and regulates the cognitive. Of particular relevance to this study is one particular element that inhabits the intersection of both executive functioning and self-regulation - *inhibitory control (*Aron, 2007*)*. This is described by cognitive scientists in terms of active inhibition ('Do not do X') and passive inhibition ('Do Y') (Hofmann, Schmeichel and Baddeley, 2012). In contrast psychologists situate inhibitory control as a volitional strategy ('Don't get distracted by Z'), within the construct of self-regulated learning. This volitional strategy is further discussed in Section 2.3.5.2.1.

2.3.2 Why study self-regulated learning?

The construct of self-regulated learning is important because, as Marsh, Trautwein, Lüdtke, Koller and Baumert (2006) report, there are consistent correlations between students' effective use of self-regulating strategies, their self-beliefs and their levels of academic success. This notion is echoed by a number of studies (e.g., Zimmerman and Kitsantas, 2005; Higgins, Hall, Baumfield and Moseley, 2005; Dignath, Buettner and Langfeldt, 2008). Indeed, in another Zimmerman study, the author went so far as to suggest that, ‘a student’s level of learning has been found to vary based on the presence or absence of key self-regulatory processes’ (Zimmerman, 2002, p.66). Furthermore, according to Allan, Cook and Lewis (1996), the self-regulated learner is more likely to meet course deadlines, enjoy their studies, reduce overall study time, ask for appropriate help, acknowledge gaps in their understanding, become lifelong learners and recognise their own achievements. Williams (2003) adds that the self-regulated learner can look forward to an increase in self-motivation and self-esteem, the capacity for independent judgement, a respect for peers, the excitement of challenge, active rather than passive learning and good social and communication skills.

As a number of authors have further explained, self-regulation is vital as a life-long skill, since once a student has left school or university they are likely to be presented with a variety of new learning situations, for which their formal education will not have prepared them (Boekaerts and Corno, 2005; Zimmerman, 2002). For instance, at the time of writing, the Guardian newspaper list of graduate jobs includes *app designer, consumer experience manager, digital campaign planner* and *bid coordinator* - jobs which require new sets of skills, increased self-regulation and a flexible approach to new knowledge - and jobs which rarely featured on the radar of the job-seeker a decade ago (Guardian, 2014). Clearly, in this environment, those who are adaptable and can regulate and refine their own skills are likely to succeed over those who cannot. Zimmerman (2002) also points out that the learner who undertakes an extended project like a work of art, a dissertation or an invention, needs to be even more self-regulated than most. According to the Office for National Statistics (ONS, 2014), the number of people self-employed in the UK has more than doubled since 1975 (4.6 million self-employed in 2014, compared with 1.9 million in 1975). In the self-employment environment, it is clear that the 'relationship of self-reliance to success in life' is a vital one (Zimmerman, 2002, p. 66).

However, as has previously been discussed, programmes that aim to develop self-regulated learning skills in young people from the UK are noticeable by their absence (Duckworth *et al.,* 2009). It is argued therefore, that despite at least four decades of active research in the field of self-regulated learning, and because of the rapidity of change in the learning environment, there remains an urgent need to develop interventions that assist young learners to develop their self-regulated learning skills. This study presents and explores the effectiveness of such an intervention.

2.3.3 The importance of the early-adolescent phase

Contemporary neurological research studies have identified that early adolescence is a crucial time during which the brain undergoes significant changes - changes which result in acceleration of cognitive development (Casey, Giedd and Thomas, 2000). Educational psychologists have explored this notion in tandem with the earlier work of Piaget (Inhelder and Piaget, 1958), who introduced 'formal operations' as a stage in child development, a stage which, according to more recent theorists such as Flavell (1979) and Kuhn (2006) is also characterised by an increased capacity for metacognition and the associated metacognitive skill of critical thinking. Despite widespread criticism which suggests that Piaget's is too simplistic an explanation of child development, there persists the idea that early adolescence is a period of rapid cognitive change, during which an individual's awareness of their thought processes develops at an increased rate (Kuhn, 2006). A contemporaneous explanation for this is encapsulated in the social cognitive approach, which situates the metacognitive development alongside experiential, behavioural, contextual, observational and environmental factors which also influence the learning process (Bandura, 1977, 1986; Zimmerman, 1989; Glaveneau, 2011). With the appreciation that the development of metacognitive capacity is driven not only by inherent ability but also by experiential factors, comes the reasoning that *exercising* such capacity is 'a sufficient condition for development' (Kuhn, 2006, p. 62). Thus it is also proposed that the 'use it or lose it' principal refers also to the development of metacognition and critical thinking in early adolescence, a notion that prompts Kuhn to suggest that if the opportunity to exploit that particular window of opportunity is not grasped, then that opportunity may not present itself again (Kuhn, 1991, 1992, Kuhn and Crowell, 2011).

Particular cognitive changes during this time refer to the development of metacognition, an executive function characterised by increased awareness of the learning process, a disposition towards reflective thinking and the perceived desire to use appropriate cognitive strategies in problem solving situations (Miller and Cohen, 2001; Banich, 2009; Hofmann, Schmeichel and Baddeley, 2012; Dwyer, Hogan and Stewart, 2014). This awareness is associated with the control, monitoring and regulation of basic information processing, affording the potential for an individual to become an increasingly effective learner. According to Kuhn (2001), this results in an increase of the use of effective learning strategies and a conscious decrease in the use of ineffective ones.

2.3.4 Current SRL educational policy and practice (in England)

An analysis of current (2105) Department of Education National Curriculum literature reveals very little specific direction with regard to the development of self-regulated learning in national curricula. This reflects the policy of removing the prescription of schemes of work from the National Curriculum, since September 2013. It is not surprising, therefore, that self-regulatory habits and strategies are still rarely taught in schools. Evidence from OFSTED, the English education inspectorate, verifies that there is some way to go before self-regulated learning strategies become embedded in educational practice in English schools. In their 2012 *OFSTED News* they highlight the necessity for, 'a focus on developing personal, learning and thinking skills, and skills necessary for working life such as team work, creative thinking and self-management.' In a survey of thirty-four schools, it was found that only four of them were helping students to develop these skills. The majority of schools studied 'usually had little or no knowledge of where the skills were being taught' (OFSTED, 2012, p. 1). The inspectorate's key finding was that schools needed support in developing students' personal, learning and thinking skills.

Operational challenges have, however, constrained the implementation of interventions that had had the potential to make a difference to such findings. The 2007 SEAL strategy (Social and Emotional Aspects of Learning) for instance, took a whole-school approach to the promotion of 'well-being.' Modules included those on 'aspects of learning' and focused on self-regulation, self-awareness, managing feelings and motivation (DCSF, 2007, p. 4). Findings of a 2010 Department of Education evaluation, however, noted that the SEAL implementation had been ineffective, having been held back by a lack of 'will and skill' on the part of participating staff. Furthermore, SEAL failed to have a positive impact upon school-level outcomes (Humphrey, Lendrum and Wigelsworth, 2010, p. 2). Other England initiatives have included the Personal, Learning and Thinking Skills (PLTS) framework, designed to address the need to develop life-long learning skills. Once part of the secondary curriculum, this quickly became restricted to the 14 - 19 vocational Diploma, a precarious project which according to the Institute of Education was doomed to, 'struggle along, with at best only limited or lukewarm recognition and support' (Waring, 2010, p. 7).

Meanwhile, England slips further and further down international league tables of achievement, with one marked decline being noted in recent findings from PISA, the Programme for International Student Assessment, which is supervised by the Organisation for Economic Development (OECD). From 2000 to 2012, England slipped from 8th to 26th in mathematics, 7th to 23rd in reading and 4th to 20th in science, with performance rated similar to Latvia and Portugal (OECD, 2012). Even critics of such performance tables acknowledge that, 'by any measure, […] schools are behind a group of top-performing countries and schools systems' (BBC, 2012). However, despite compelling PISA research which indicates that students who self-regulate their learning 'tend to perform better than those who do not' (OECD, 2010, p. 14), English educational policy continues to focus upon the acquisition of 'knowledge and understanding' as the major priority for teaching programmes, with direct prescription for the development of learning strategies having now been removed from National Curriculum guidelines (National Curriculum, 2014).

Duckworth *et al.,* (2009) in a literature review for the Institute of Education, recommended that education policy should make provision for high teacher expectations, teacher modelling of self-regulated learning strategies, providing challenging learning contexts, the specific teaching of thinking skills and enabling student choice and autonomy. However, features that involve learner autonomy, the pursuance of personal goals, taking initiative, freedom from external control and taking ownership of learning objectives are arguably not necessarily in line with the 'relatively prescribed National Curriculum that relies heavily on assessments' (Duckworth *ibid*, p.25).

It is argued therefore, that there is merit in exploring the effectiveness of interventions that have the potential to develop self-regulated learning skills in young people, and thus to address some of the above perceived shortcomings. This study explores the development, implementation and evaluation of such an intervention.

2.3.5 Models of self-regulated learning

There is no agreed universal model of self-regulated learning (SRL), but the majority of contemporary models derive from the social cognitive perspective, a model of behaviour that originated primarily with Bandura's work (1977b). Social cognitive theory incorporates and integrates cognitive and metacognitive influences with self-regulatory behaviour mechanisms that are driven by motivational factors (Bandura, 1986; Zimmerman, 1995). The essential premise, from an educational perspective is that a child can learn from the observation of others, in such a way that learning is not only self-directed but is guided by the behaviour of others. Emanating from this theoretical ideology, and fundamental to the most commonly cited self-regulated learning models (e.g., Boekaerts, 1999; Pintrich, 2000 and Zimmerman, 2008), is the notion that learners are active participants in a cyclical learning process, that learning is socially situated and that the regulation of learning is mediated and moderated by the learner's metacognitive knowledge, motivational beliefs and sense of self-efficacy.

These theoretical assumptions translate commonly to two strands of inquiry of self-regulated learning, both of which have informed the conceptual model for this study (Section 2.3.6, Figure 2-1). Thus self-regulated learning is commonly described in terms of the 'offline' person-level *attributes* or dispositions that learners bring to the process of learning (e.g., Winne and Perry, 2000) and the 'on-line' task-level cyclical self-regulatory *processes* that take place during a learning task (e.g., Schmitz and Wiese, 2006).

2.3.5.1 Offline self-regulated learning traits and competencies

Offline, person-level self-regulated learning dispositions refer to learner self-beliefs, traits and competencies that influence the process of self-regulated learning. Pintrich (1999) and Boekaerts (2010) focus extensively upon the importance of motivational beliefs as key influencers of self-regulated learning. Bandura (2006), Pajares (2002) and Caprara, Fida, Vecchione, Del Bove, Vecchio, Barbaranelli and Bandura (2008), on the other hand, emphasise the role of self-efficacy, asserting that a learner's degree of self-efficacy not only influences the extent to which they engage in self-regulation, but also mediates their academic achievements. Schunk (2008) and Efklides (2011) highlight the importance of metacognitive knowledge as a key driver of the self-reflection stage of self-regulation, whilst Zimmerman (2000, 2001) stresses the importance of all three. Thus a self-belief 'person-level' model of self-regulated learning would typically include three dynamic and composite learner characteristics: a) metacognitive knowledge: the learner's own knowledge of their cognition, motivation and learning behaviours, b) motivational beliefs, which regulate goals and effort management and c) self-efficacy: 'judgments about one's ability to accomplish a [learning] task as well as one's confidence in one's skills to perform that task' (Pintrich, Smith, Garcia and McKeachie, 1991, p. 13). Such models do, however, stress the interdependence of the separate elements (see Figure 2-1, Section 2.3.6). Thus a learner whose efficacy for learning was deficient would be less likely to be motivated to learn, or to use effective cognitive and metacognitive strategies. Similarly, according to Schunk and Ertmer (2000), a student with a high sense of self-efficacy is only likely to be an effective self-regulated learner if they also have the metacognitive knowledge and skills required to control and regulate the learning process.

2.3.5.1.1 Metacognitive knowledge

> *... if somebody knows something, then he knows that he knows it, and at the same time he knows that he knows that he knows. Spinosa (1632 - 1677)*
> (Nassar, 2012)

The notion of a learner's awareness of their cognitive processing is mentioned as far back as Plato, who attempted to explain the nature of human knowledge. Notable reflections on the construct include those of Locke (1689) who referred to the extent to which the mind was able to perceive its own operations, but the term 'metacognition' was first coined by Flavell in the 1970s. He defined metacognition in terms of the learner's knowledge and regulation of their cognitive strategy processing (Flavell, 1979).

Flavell's (*op cit.*) classification of metacognition into aspects of (static, trait-like) *knowledge* and (dynamic, process driven) *regulation* has resulted in a separation of the construct into *metacognitive knowledge* (sometimes referred to as metacognitive awareness) which refers to the learner's capacity to understand how they learn in a given situation (Sperling *et al.,* 2002), and *metacognitive regulation* which has become synonymous with the *process* of self-regulation itself. The latter confusion is not surprising since their definitions tend to be significantly similar. Georghiades (2004) for instance, describes metacognitive regulation as 'comparing, assessing and evaluating the content or the processes of one's learning' (p. 371), whilst Zimmerman (2008, p. 167) describes self-regulated learning in terms of learners being 'active participants in their own learning process'. It is not difficult therefore to understand how the two concepts have become confounded.

The conceptual model (Section 2.3.6) proposed for this study attempts to clarify the difference between metacognition and self-regulated learning by adopting Flavell's (1979) distinction between metacognitive *knowledge*, one of the 'offline' person-level competences that characterise self-regulated learners, and metacognitive *regulation*, the 'on-line' task-level process of executive control that the learner exerts over the use of cognitive and metacognitive skills and strategies during the learning process. Thus the process of self-regulation is characterised essentially as a metacognitive activity, whilst metacognitive knowledge is characterised as *one of* a number of self-regulated learning dispositions (Efklides, 2008). Self-regulated learning as a process therefore refers to the metacognitive control and regulation of a number of cognitive and metacognitive strategies and experiences that take place during the cycle of self-regulation (Zimmerman, 2008). This cycle, attributed to the work of Zimmerman (2000, 2008) separates the process of self-regulated learning into three phases: forethought, performance and self-reflection. During the forethought phase, the cognitive strategies thus controlled include task analysis, goal setting and planning, whilst the metacognitive strategies include self-motivation and synthesis of outcome expectations with goal orientation. During the performance phase, the cognitive strategies include task management, time management and self-recording, whilst the metacognitive strategies include critical thinking, control of volitional strategies, elaboration, self-monitoring and self-instructions. Finally, the self-reflection phase includes a number of metacognitive activities including self-evaluation, causal attributing, assessment of learning satisfaction and adaptation of future goal orientations.

To be self-regulated, learners require knowledge about their own cognition, an understanding of how, why and when to use appropriate cognitive strategies, and the working knowledge of how tasks are processed (Kuhn and Dean, 2004). This *metacognitive knowledge* resides in the long-term memory, and whilst it is described as a relatively stable trait that can be described, reflected upon and reported upon by the learner (Efklides, 2008), it is also theorised that it is amenable to enrichment through self-reflection (Zion, Michalsky and Mevarech, 2005). However, few studies are available which report effect sizes for the latter phenomenon. This omission is addressed in the current study, where self-reflection (operationalised in a learning journal) is tested in terms of its predictive effect on change in metacognitive knowledge over time.

Researchers argue that metacognitive knowledge is a key factor that impacts upon achievement (Garner and Alexander, 1989). It is thought, for instance, that learners with a high degree of metacognitive knowledge are able to regulate their learning processes to the extent that they directly improve their academic performance (Pressley and Ghatala, 1990). Other theorists have proposed the counter-intuitive notion that metacognitive knowledge operates independent of intellectual ability (Swanson, 1990; Schraw and Dennison, 1994). Despite a lack of effect sizes to support the latter suggestion, such claims are important because they suggest that if it is possible to increase a learner's metacognitive knowledge, then improvements in such can impact upon a wider population of learners, other than the most intellectually able. This study attempts to explore this notion further, evaluating sizes for the effects of intervention activities that may have the potential to effect change in metacognitive knowledge. Furthermore, analysis will attempt to detect whether gains in metacognitive knowledge are differentiated according to prior metacognitive disposition.

Measurement of metacognitive knowledge

The measure chosen to measure metacognitive knowledge in the current study is the Student Approaches to Learning (SAL) instrument, developed from the Boekaerts (1999) model and 'based on what has been judged to be the best of existing theoretical and empirical research' (Marsh, Hau, Artelt, Baumert and Peschar*, 2*006a, p.348). It has been important to employ a well-validated scale for the measurement of metacognitive knowledge since it is suggested that it is difficult to measure. This is because metacognitive knowledge reflects an inner awareness rather than a measurable behaviour. It follows that a learner needs to be metacognitively aware of their capacity for metacognitive knowledge, in order to be able to report on it. Fortunately, the well-validated SAL instrument has been consistently used in the Programme for International Student Assessment (PISA) study that takes place every three years across approximately twenty-five countries. The PISA study, commissioned by the Organisation for Economic Cooperation and Development (OECD), seeks to assist in the OECD's drive to raise educational standards by evaluating learning outcomes across a range of subject areas. Developed with relevance to the educational policies and the cross-cultural nature of participating countries in mind, the instrument has been subject to rigorous psychometric testing (reliability, item response theory, confirmatory factor analysis, uniformity of psychometric properties across countries and relationship to achievement), for more than a decade. The self-report SAL scale includes both self-regulated learning (cognitive, metacognitive and affective) and subject-specific (mathematics and verbal) items. As all of the constructs within the scale have been analysed as separate constructs in themselves, the measure is deemed adaptable and broadly applicable for a wide range of educational studies including correlational, large-scale survey and experimental (Marsh *et al., 2*006a). As in the PISA study itself, different combinations of the SAL items enable the researcher to focus on specific aspects of self-regulated learning and to separate general self-regulated learning items from subject specific ones. The current study has adopted the recommendation of Marsh *et al., (op cit.)* who suggest that it is not necessary to employ the entire instrument, and that the use of 'selected scales' is appropriate (p.348).

2.3.5.1.2 Motivational beliefs

Considerable research has gone into the exploration of self-regulated learning and yet the mystery remains as to why some students choose to self-regulate, whilst others do not. According to a number of authors including Boekaerts (2002) and Schunk and Zimmerman (2008), any search for the answer should include motivation as a key exploratory factor. Indeed, throughout the literature, there is an emphasis on the extent to which *motivational beliefs* and the related construct of *volition,* regulate the self-regulatory process (e.g., Pintrich, 2000; Boekaerts and Cascallar, 2006; Boekaerts, 2010; Wolters and Rosenthal, 2000; Wolters, 2003), with the essential notion being that the learner's mental exploration of their motivation beliefs affects their choices of self-regulatory activities in a contextualised setting (Boekaerts and Cascallar, 2006) thus predicting metacognitive strategy use (Sungur, 2007).

A number of studies have concluded that motivational beliefs are significantly related to other aspects of self-regulation, such as metacognitive knowledge and self-efficacy (Dembo and Eaton, 2000; Pintrich and De Groot, 1990; Valle, Cabanach, Nunez, Gonzalez-Pienda, Rodriguez and Pieniro, 2003). Dembo and Eaton (2000), for instance, theorise about the important cyclical link between self-regulation, motivation and self-efficacy, suggesting that when students take control over their learning they are more likely to employ strategies which lead to ever increasing motivation, and are thus more likely to maintain positive beliefs about their abilities. They suggest ways to improve student motivation, including highlighting the importance of managing distraction, and encouraging collaboration with peers. Pintrich and De Groot (1990), conducted a correlational study on 173 young adolescents in the US, finding that motivation was significantly correlated to aspects of self-regulation such as metacognition (r = .73) and strategy use (r =.63). They also reported that motivation was related to a range of academic outcomes such as essays, tests and grades ($.19 < r < .36$).

Motivational beliefs tend to be domain-specific, and refer to the internal values that learners hold about the subject they are studying and about the effectiveness of the particular learning methods and strategies at their disposal. Whilst it is held that motivational beliefs derive from previous learning experiences, once formed they are not necessarily easy to change, and unfavourable beliefs can impede the learning process (Niemivirta, 1999). Thus it is suggested that motivational beliefs influence the forethought phase of the cycle of self-regulation, thus playing a central role in the regulation of strategy choice, effort investment and volition, especially if either the task ahead, or the learning environment, is found to be challenging. It follows that there is merit in exploring methods for raising a learner's motivational beliefs.

The importance of examining motivation also lies in its relationship to learning processes and academic outcomes (Bandura, 1977a, Eccles and Wigfield, 2002). Motivational beliefs have, for instance, been found to be related to effort and persistence (Schunk, Pintrich and Meece, 2008), outcome expectancy (Schunk and Zimmerman, 2008), gender (Meece and Painter, 2008), cultural identity (McInerney, 2008), self-efficacy (Pajares, 2008), task value (Wigfield, Hoa and Klauda, 2008), volition (Corno, 1994, 2001) interest (Hidi and Ainley, 2008) and engagement (Martin, Papworth, Ginns and Liem, 2014). Most importantly, a range of studies from different research perspectives have come to the same conclusion - that motivation is a function of a learner's goal orientations (Boekaerts, Smit and Busing, 2012; Dweck, 2006; Zimmerman, 2011). Hence Boekaerts *et al*., (*op cit*) who approach the topic from an educational psychology perspective, categorise *ego* and *non-ego* goals, whilst Zimmerman reports on the analogous nomenclature of goal orientation theorists, and distinguishes *performance goals* from *learning goals* respectively. The distinction is that performance goal orientation involves an individual being motivated to seek confirmation of their personal competence; this leads to a motivational emphasis on avoiding negative judgements. Learning goal orientation, on the other hand, is concerned with an individual's attempt to self-improve by increasing their competence, and leads to the learner being motivated to seek self-improvement. Dweck theorises on the construct of a *mindset*, suggesting that the performance goal orientation emanates from a *fixed* entity mindset whereby the learner assumes (or is told) that their intelligence is innate and immutable. With this in mind, the learner's motivation is influenced by their seeking opportunities to prove themselves to others, and to measure their performance in comparison to that of others. A clear similarity can be seen between this notion and that of *ego goals* proposed by Boekaerts *et al.* (op cit). On the other hand, according to Dweck (*op cit*), the learning goal orientation reflects a *growth* mindset; here the notion is that intelligence is incremental, malleable and amenable to hard work and persistence. The relationship of learning goals to motivation is clear - the learner with the growth mindset is motivated to learn because self-improvement is seen as an attainable goal. It follows that the provision of opportunities for learners to discover that achievement is possible through effort and persistence (rather than it being a result of innate intelligence) has the potential to influence the development of motivation. According to Dweck (2006), this is an important consideration for educators and parents alike, especially with regard to more able children.

It is suggested that motivation is also amenable to improvement by direct instruction. For instance, Wolters, Benzon and Arroyo-Giner (2011) suggest that social interaction between peers, who are engaged in collaborative learning activities, has the potential to impact upon motivation. It follows that dialogic argumentation, a collaborative learning exercise, might be one such activity. The notion that participation in dialogic argumentation might influence a learner's motivational beliefs has not yet featured largely on the agenda of researchers in either the field of self-regulated learning or that of argumentation. This study contributes to the literature by exploring whether participation in a collaborative activity such as dialogic argumentation can influence change in a learner's motivational beliefs.

Measurement of motivational beliefs

Throughout the literature on self-regulated learning there is an emphasis on the role that motivational beliefs play in the self-regulated learning process (e.g., Boekaerts, 2010), with the idea that the learner's sense of motivation acts as a frame of reference that governs the self-regulating process. It follows that an assessment of a learner's self-regulated learning is incomplete without the inclusion of a measure of their level of motivation. Such measures, since they evaluate the strength of beliefs, can be verbalised by the learner, especially if the learners are offered prompts (Wolters, Benzon and Arroyo-Giner, 2011); therefore self-report measures feature as the most prominent means of measurement.

One of the most common such measures is the Motivated Strategies for Learning Questionnaire (MSLQ), which was developed by Pintrich, Smith, Garcia and McKeachie (1991) for use with college students. The MSLQ measures motivational beliefs and learning strategies and has been widely used to evaluate the effectiveness of university and college courses (e.g., Komarraju and Nadler, 2013; Phan, 2014), but has needed adaptation for use in some school studies (e.g., Nett, Götz, Hall and Frenzel, 2012; Ahmed, van der Werf, Kuyper, and Minnaert, 2013; Lee, Lee and Bong, 2014). The questionnaire items are presented using terminology which is more relevant to the US college system than to school children in other education systems (e.g., 'When I have the opportunity in this class, I choose course assignments that I can learn from even if they don't guarantee a good grade' or 'When reading for this course, I make up questions to help focus my reading.' (Duncan and McKeachie 2005, p.127)). For this reason, the language of the scale is frequently adapted for specific audiences (e.g., Schmitz and Wiese, 2006; Phan, 2014).

Another prominent measure is the Questionnaire on Current Motivation (Rheinberg, Vollmeyer and Burns, 2001), an adapted version of which has been chosen for this study (see Appendix 1). The questionnaire was chosen for its simplicity of language, its brevity, its level of accessibility for younger students and its ease of administration. It was considered suitable for this study also because it was designed to evaluate the motivational beliefs of participants once they had information about a task (in the case of this intervention - a course) and before they had actually made a start on that task (Vollmeyer and Rheinberg, 2006). The development of this measure is further detailed in Section 4.5.2.

2.3.5.1.3 Self-efficacy and Efficacy for learning

Contemporary research has detected that a learner's self-regulated learning processes are amenable to teaching, and that such teaching can influence learning outcomes (De Bruin and Van Gog, 2012). However, the teaching of cognitive and metacognitive learning strategies does not necessarily guarantee that those strategies will be used in the future. According to Schunk and Ertmer (2000) the key to the long-term effectiveness of the acquisition of learning strategies is the degree of *self-efficacy* that the learner brings to the process - that is the extent to which the student believes that academic success is under their control. Indeed it is felt that the cultivation and maintenance of positive self-beliefs is the principal motivational process in self-regulated learning, in terms of its capacity to influence the learner's selection of learning strategies, level of effort and persistence (Zimmerman, 1986, 1995; Pajares, 1996), application of cognitive and metacognitive strategies (Duckworth *et al.,* 2009) and academic goals and aspirations (Zimmerman and Bandura, 1994). In essence, those students who believe that effort, hard work and persistence are likely to lead to success, are the ones who tend to succeed the most (Schunk and Ertmer, 2000; Dweck, 2006). The firm belief that one can regulate one's learning provides the 'staying power' for doing so (Caprara *et al.,* 2008, p. 526). The construct of self-efficacy was chosen for this study because it is commonly referenced as a key driver of self-regulated learning (Bandura, 1977a; Zimmerman, 2000; Pajares, 2002), because of its 'growing primacy in contemporary life' (Caprara *et al.,* 2008, p. 525) and because of its important relationship to academic achievement (Bandura, Barbaranelli, Caprara and Pastorelli, 1996; Pastorelli, Caprara, Barbaranelli, Rola, Rozsa and Bandura, 2001; Marsh, *et al.,* 2006a; Caprara *et al.,* 2008; Zuffianò, Alessandri, Gerbino, Kanacri, Di Giunta, Milioni and Caprara, 2013). The original outreach intervention was designed to reflect a number of learning objectives, including self-esteem. However, self-efficacy (which according to Bong and Skaalvik (2003) is closely related to self-esteem) was chosen above self-esteem itself since self-efficacy is considered to be a forward-looking construct, rather than a backward-looking one (Bandura, 2006). This aspect of the 'self' measure was considered important since the construct was required, within the design, to measure the effect of an intervention.

Thus 'self-efficacy' describes an individual's 'beliefs in their capacity to produce given attainments' (Bandura, 2006, p. 307) and is a performance-based, domain-specific construct. Educators have long understood that an individual's self-beliefs affected their performance, but it was not until the late 1970s that studies began to explore self-beliefs from a domain-specific perspective. In particular, Bandura (1977a) proposed a concept of self-efficacy that has since become ubiquitously accepted in educational research as an important construct in the field of self-regulated learning. Self-efficacy refers to an individual's focus on their performance capability in a given situation. It is domain-specific, so in an educational setting, a student may for instance have a high mathematical efficacy but a low reading efficacy. Therefore efficacy beliefs are not global traits, but refer to specific areas of learning.

Studies have explored the role of self-efficacy beliefs in academic settings (e.g., Pajares, 1996; Wigfield, Eccles, Schiefele, Roeser and Davis-Kean, 2006) demonstrating that high self-efficacy is related to high self-regulation. Other studies have reported that the learner's sense of self-efficacy impacts upon the goals that they set for themselves, their persistence, effort investment and intrinsic interest in pursuing those goals, and indeed their performance (Zimmerman, 1989, 1990; Schunk, 1984, 1989). The argument is that self-efficacy regulates other self-regulatory beliefs and competencies, and in particular their motivation. If individuals have doubts about their capacity to succeed academically, they are unlikely to make use of their metacognitive knowledge and skills, and are unlikely to be motivated to invest effort in the learning task (Duckworth *et al.,* 2009).

Self-efficacy is related to *efficacy expectations*, but different from *outcome expectations*. The former refer to the individual's belief in their capacity to perform a task, whilst the latter refer to the expectation of the outcomes attached to that task (Zimmerman, 2000a). Another closely related concept is *self-esteem*. Here, the difference is that self-efficacy is a measure of performance capability, whereas self-esteem is a measure of self-worth. Historically, self-esteem can be traced back to the Ancient Greek maxim 'know thyself,' but modern interpretation is built upon the work of William James (1890) who laid the ground for the notion of self-concept, based upon what is important for the individual. He posited a difference between the I-self (the objective aspect of self-esteem) and the Me-self (the subjective aspect). The latter has become associated with the fields of self-esteem that have developed throughout the 20th Century and into the 21st. Until the 1970s, the concept of the self was seen as a unidimensional construct of self-esteem, reflecting the James (1890) model of the Me-self. However the Shavelson Model (Shavelson, Hübner and Stanton, 1976) heralded the arrival of a re-energised field of research that sought to identify the individual components of 'self-concept'. This model separated self-concept into a multidimensional and hierarchical construct which separated academic self-concept from social, emotional and physical self-concept, and which provided for the development of further sub-domains.

Self-concept (Marsh and Shavelson, 1985) is one of the most closely related constructs, but is nevertheless distinctly different (Ferla, Valcke and Cai, 2009; Scherer, 2013). Bong and Skaalvik (2003) provide a useful comparison between the two concepts, suggesting that self-concept is about perceived *competence*, is *past*-

oriented and *stable*, whereas self-efficacy is about perceived *confidence*, is *future*-oriented and is *malleable*.

Self-efficacy was chosen as a construct for this study because it is an important determinant of an individual’s capacity to self-regulate their learning, and a predictor of academic achievement (Zuffianò, Alessandri, Gerbino, Kanacri, Di Giunta, Milioni and Caprara, 2013), with its key role being its influence on motivation (Zimmerman, 2000a; McGeown, Putwain, Simpson, Boffey, Markham and Vince, 2014). It was chosen in preference over other ‘self’ concepts, since it is believed to be a superior predictor of learning outcomes; notably, it has been found to be a more accurate predictor, the more task-specific the measure (Pajares and Miller, 1994; Pajares, 1996). Self-efficacy is also predictive of a learner’s level of effort and their persistence (Komarraju and Nadler, 2013), their self-monitoring behaviour (Bouffard-Bouchard, Parent and Larivee, 1991) and their use of learning strategies (Zimmerman and Martinez-Pons, 1990). As this study was exploring an intervention which included these components, it followed that self-efficacy might provide a useful measure of the intervention’s effectiveness or otherwise (see also Hidi, Berndorff and Ainley, 2002).

Learners who are high in self-efficacy tend to believe that achievements can be made through effort, persistence and the pursuance of challenging mastery goals, whilst those with low self-efficacy tend to believe that their performance is based upon their innate and unchangeable level of intelligence (Dweck, 2006; Komarraju and Nadler, 2013). It follows that efforts to improve self-efficacy have the potential to influence learning outcomes.

According to Bandura (1977, 2006), powerful *mastery experiences* can have a transformational, profound and long-lasting effect on self-efficacy. Thus an individual's self-efficacy may change, for better or worse, as a result of a personal experience of success. This assertion is supported by Zimmerman (2000a) who further reports on two techniques that have the potential to develop an individual's self-efficacy. The first is teacher modelling of a cognitive process, whilst providing a commentary on its effectiveness. For example, in the current study, the discussion teacher modelled critical thinking processes by 'thinking aloud', exploring alternative viewpoints and proposing alternative ways of thinking. The second suggestion is to articulate and monitor goals. The latter strategy has also informed the design of the current study, where a group of students was required to report on how they set and monitored their goals in a learning journal. Caprara *et al.,* (2008) make similar suggestions, adding that the process of self-monitoring and self-evaluation also has the potential to increase self-efficacy, because it highlights to the learner the benefits of managing their academic activity.

Measurement of self-efficacy

According to Bandura (2006), there is no single global measure of self-efficacy. Rather a measure is likely to be more efficient, the more specific it is. This is because self-efficacy in any one individual varies widely according to context. This is not to say that there is no correlation between domains of self-efficacy; where similar subskills are involved in similar contexts, there is likely to be some covariation among respective self-efficacy beliefs. Zimmerman (2000a) adds that, as self-efficacy is a future-oriented concept, it is ideally assessed before students engage with a learning task.

Bandura (2006) offers guidelines for the construction of self-efficacy scales. He suggests most importantly, that because of domain-specificity, the scale items should accurately reflect the construct being measured. Responses to item questions should have enough gradations to avoid a ceiling effect. For instance, his items are scored on a 10-unit scale, where only the end points are anchored. Items should be worded in terms of self-efficacy *beliefs*, such that they ask the respondent how confident they are of their level of efficacy in a certain activity (rather than 'how well' they can do an activity). Zimmerman (2000a) adds that questions should be phrased in terms of confidence in performance, rather than competence. For instance, a self-efficacy item should be phrased, "How confident are you in your capacity to work independently on this task?' rather than, 'How good are you at working independently?' These recommendations were taken into account when constructing the efficacy scale used in the current study (see Section 4.5.2 and Appendix 1).

Whilst others such as Pintrich and De Groot, (1990), who created the extensively cited Motivated Strategies for Learning Questionnaire (MSLQ), have created 'self-efficacy for self-regulated learning' scales, such scales have been used primarily in the US college context (e.g., Lodewyk and Winne, 2005; Zimmerman and Kitsantas, 2007). The language of such scales is not appropriate to either the English context (e.g., references to 'end of chapter questions'), or the internet-based learning environment (e.g., references to 'chapters in my book').

In this study, the focus is on self-efficacy as it pertains to self-regulation in young English adolescents, who are studying via the internet. Research studies that have explored the effectiveness of internet-based interventions have concluded that *self-efficacy for self-regulated learning* is significantly related to internet self-efficacy (Debowski, Wood and Bandura, 2001; Joo, Bong and Choi, 2000). In order to differentiate such a construct from other measures of self-efficacy, an adapted scale, 'efficacy for learning' (EfL), created by the researcher, is introduced for the purposes of this study (see Appendix 1 and Section 4.5.2). Thus EfL is defined as 'a learner's belief in their capacity to self-direct their learning'. Inspired by the recommendations of Bandura (2006), and constructed from the perspective of social cognitive theory, it attempts to capture a learner's belief in their capacity to work independently, to communicate their ideas to others, to convert their hard work into academic success, and to manage effort and persistence.

2.3.5.2 On-line self-regulated learning processes

On-line process models describe self-regulated learning in terms of the metacognitive *regulation* of cognitive and metacognitive behaviour, skills and strategies that foster self-regulation. These include: the cognitive skills of task evaluation, goal setting, planning, time-management, organising ideas, self-checking, concentration, working collaboratively and working autonomously; the metacognitive regulation of critical thinking, volition, problem solving, self-evaluation and self-reflection; and metacognitive experiences such as learning satisfaction, response to challenge, task-related self-confidence and self-judgement (Flavell, 1979; Duckworth *et al.,* 2009; Efklides, 2008). This process of self-regulated learning is commonly represented by a cyclical series of states, involving three phases: *forethought* (e.g., task evaluation, goal setting, strategic planning), *on-task performance processes* (e.g., critical thinking, self-observation, self-instruction, self-monitoring, distraction avoidance, task strategies) and *reflection and reaction* (e.g., self-reflection, evaluation of metacognitive experiences, self-judgement, causal attribution, adaptation of future goal orientation) (Zimmerman, 2002, 2008). Both the three-stage model espoused by Zimmerman and a similar four-stage model (forethought, monitoring, control and reflection) proposed by Pintrich (Schunk, 2005) reflect the social cognitive perspective, which emphasises the interaction of cognitive, behavioural, motivational and environmental factors, over the course of a learning cycle. This cyclical feature suggests that self-regulation of learning is therefore amenable to practice, during which the learner develops a stock of learning strategies, which in turn enhance their self regulatory dispositions (Duckworth *et al.,* 2009).

Duckworth *et al.* in their 2009 literature review for the Institute of Education, noted that, despite more than three decades of self-regulated learning research, there were still questions to ask regarding the nature of self-regulated learning processes as they developed over time - and in particular how metacognitive skills develop alongside metacognitive experiences and behaviours, in terms of change and/or stability over time. A more recent sweep has revealed that such studies are now emerging. Bannert, Reimann and Sonnenberg (2014) for instance, explored the temporal patterns of student's self-regulated learning behaviours, but acknowledged that they had only uncovered 'the tip of the iceberg' methodologically (p. 181). Their process-mining models were able to provide only descriptive analysis, and the authors acknowledged that inferential statistical analysis such as significance testing, was not possible with their model. Other authors have begun to explore means of tracking self-regulated learning in real time (Schmitz, 2006; Schmitz and Wiese, 2006; Perels, Otto, Landmann, Hertel and Schmitz, 2007; Schmitz and Perels, 2011).

Process describes a series of relatively changeable *states* (Hertzog and Nesselroade, 1987). Schmitz and Wiese (2006) propose that self-regulated learning could be conceptualised as a three-stage process model (pre-action, action and post-action). Based upon the social cognitive model of Zimmerman (2000), it now describes self-regulated learning as a series of learning states only, rather than both personal dispositions and learning states. Thus the ordered sequence of those learning episodes represents a learning process that can be measured in terms of time-series data. Where those episodes are recorded as the cognitive and metacognitive activity that takes place during the learning process, it is thus possible to observe self-regulated learning trends over time (Schmitz and Wiese, 2006). Such a process is characterised by a design that contains multiple measurements (in contrast to the more common pre-to-post design), measured at intervals. Measures can either be manifest (i.e., observed variables such as test results) or latent (i.e., random variables such as measures of self-efficacy) (Schmitz, 2006). Process models are typically assessed dynamically at different times over a number of learning events, in order to capture a series of states of self-regulation. Rather than test the self-beliefs and competencies which the learner brings initially to the learning process, they attempt to capture self-regulatory behaviours such as self-monitoring and self-reflection, within the cyclical self-regulation process. This can be operationalised using trace methods such as 'think aloud' protocols, learning journals, or electronic tracking of self-regulatory behaviours such as self-monitoring, self-reflection and self-evaluation. Whilst the 'think aloud' method is deemed appropriate for face-to-face learning environments, it is according to Efklides (2006, p. 290) limited, because the evidence for metacognitive processing does not have a 'manifest form'. Rather she favours the efficacy of the *on-line* tracking of self-regulated learning behaviour, and argues that the validity of such measures is increased where it is measured in conjunction with other affective experiences and performance measures, within a situated task. She proposes that multiple, and repeated, measures not only enrich the understanding of the metacognitive process, but also the situational factors which accompany it - in particular where metacognitive skills and experiences are measured alongside measurements of performance. This recommendation has been incorporated into this study, where self-reported metacognitive skills (control of elaboration strategies and volitional strategies) are captured at the same time as self-reported metacognitive experiences and assessed critical thinking, across four learning events.

Analysis methods include trends over time (e.g., repeated measures), time-series analysis, or autoregressive crossed lagged panel analysis modelling. The latter method, described more fully in Section 4.7.5.2, is particularly effective for this type of data as it allows for the estimation of causal relationships between variables (Kenny, 1979). Despite being rare, process-analysis of self-regulated learning is, according to Schmitz (2006), a promising research tool. Perels *et al.* (2007) conducted time series analysis of a series of learning phases, in order to evaluate the process of self-regulation as it happened. They used learning diaries as a means of capturing these data. They were able to confirm the three-stage model of self-regulated learning, proposed by Schmitz (2006). Members of the same research team conducted a similar study (Schmitz and Perels, 2011), this time using standardised diaries as a means of *supporting* self-regulatory behaviour. They concluded that self-monitoring of the students' learning led to improved self-regulation and improved mathematics learning outcomes.

Thus, it appears that it is possible to attempt to capture self-regulatory activity across the different phases of self-regulation, differentiating between the various learning strategies employed. This notion has informed the design of the current study, where learners were required to report upon their use of a number of self-regulatory processes each week in their learning journals. The learning journal approach is described more fully in Section 2.3.7.

The following subsections refer to specific self-regulated learning processes that were captured in the current study, either through observed assessment or self-reported in learning journals.

2.3.5.2.1 Volitional strategies

Volition, or 'bucking down to work' despite distractions, is a function of the will and is considered one of the most significant aspects of self-regulated learning (Duckworth *et al.,* 2009, p. 24). It is defined as 'the tendency to maintain focus and effort toward goals despite potential distractions' (Corno, 1994, p. 229). The concept is often confused and confounded with that of motivation. Kuhl (1987) provided a useful differentiation by pointing out that motivation leads to the decision to act, whilst volition determines the extent to which the intention is fulfilled. According to Efklides (2009), and for the purposes of this study, volition is a metacognitive skill that involves the control and regulation of the effort, persistence, time management, prioritising goals, maintaining focus and distraction avoidance required to implement cognitive strategies. According to Kuhl (1987) and Corno (1993, 1994), the key indicator is the extent to which an individual can avoid distractions (inhibitory control). *Inhibitory control* is categorised as a metacognitive skill in education literature (Efklides, 2008), as a cognitive function related to executive functioning, in psychology (Fernandez-Duque, Baird and Posner, 2000; Kuhn, 2006), and as a frontal lobe executive function in neuroscience (Miller and Cohen, 2001; Aron, 2007). In an educational context, it refers to the learner's capacity to avoid distraction and ignore extraneous or irrelevant stimuli, controlling inappropriate attentional tendencies (Corno, 2001; Durston *et al*. 2006). Already perceived by researchers in the field of cognition as a major indicator of metacognitive processing in young adolescents (Harnishfeger, 1995; Luna, Garver, Urban, Lazar and Sweeney, 2004), it is nevertheless suggested that little is known about the situations, which promote the development of this important self-regulatory skill (Kuhn, 2006.) The presence of inhibitory control as a feature of metacognition may increase in prominence as educational studies explore the effect of internet distractions (such as social networking, text messaging etc.) on the study habits of young people (e.g., Alloway and Alloway, 2012; Calderwood, Ackerman and Conklin, 2014).

Hence volitional challenges tend to involve the management of distraction (Corno, 1993). This is especially poignant for the modern young adolescent who is surrounded by the lure of distraction in the form of email, phone texting, instant messaging and social media. It follows that an exploration of the influence of inhibitory control is important in a study that seeks to find effective ways to influence self-regulation in young individuals. An analysis of the relationship between volitional strategies and other self-regulatory processes was included in the design, since the control and regulation of volitional strategies is considered an essential metacognitive skill which is manifested in both the pre-action and the action phase of self-regulation, but which, according to Schmitz and Wiese (2006), is rarely included in learning strategy inventories.

A number of strategies can assist in maintaining volition in the face of potential distraction - strategies which are important and valuable, since managing volition is crucial to subsequent academic performance (Corno, 1993). These include cognitive strategies such as the use of a timer to limit concentration to a short, focused time period; switching off the distracting influences such as the TV; putting the mobile phone on 'silent'; studying in a separate room; and/or avoiding opportunities for multi-tasking. Potential distraction from emotional worries in children can be controlled by the use of strategies such as encouraging them to tell their worries (and hand over the responsibility for them) to a 'worry doll' before studying (Lieberman, 2011). It is further suggested that cognitive behavioural approaches, such as keeping a diary, can benefit students with weak volitional strategies (Boekaerts and Corno, 2005). This latter notion is explored in the current study, and is reported in Section 6.7.1.3.

2.3.5.2.2 Elaboration strategies

One of the key metacognitive skills in self-regulated learning is the learner's capacity to synthesise new material with that already existing in their memory. By employing this *elaboration strategy*, the learner can make meaning of new information. According to Marsh, Hau, Artelt, Baumert and Peschar (2006a) students who do not employ such higher-order learning strategies are unlikely to succeed in learning situations. Dignath, Buettner and Langfeldt, in their 2008 meta-analysis of self-regulated learning interventions, compared the training effects of elaboration, problem solving and organisational strategies. They concluded that the most effective interventions were those that focused upon learners' elaboration strategies.

Marsh *et al. (op cit.)* report that the use of elaboration strategies, as measured using the Student Approaches to Learning scale (see also Section 4.4.1.1), is positively linked to academic achievement in mathematics and reading. Whilst they estimate the relationship between elaboration strategies and aspects of self-regulated learning such as self-efficacy, control strategies and effort and persistence, they do not report upon its relation to critical thinking, to other metacognitive skills such as inhibitory control, or to metacognitive experiences. An opportunity is taken therefore, in this study, to contribute to a further understanding of the relationship between elaboration strategies and the latterly mentioned aspects of self-regulated learning.

2.3.5.2.3 Self-evaluation

Research has detected that self-evaluation, and in particular the monitoring and evaluation of learning performance, is an important metacognitive activity that fosters the development of students' ability to self-regulate their learning (Bandura, 1982; McCrindle and Christensen, 1995; Schunk and Ertmer, 2000; Winne and Perry, 2000; Zimmerman, 2001). Prominent recent meta-analyses have compared teaching approaches with one another, evaluating the effect, on general learning outcomes, of specific strategies, some of which can be identified as self-regulated learning approaches. For instance, Hattie's (2009) meta-analysis of the effectiveness of more than eight hundred teaching and learning strategies, reported effect sizes ranging from -0.34 to 1.44, noting that the most effective strategy of all was an element of self-regulated learning - the enhancement of the student's capacity to monitor and evaluate their own learning (d = 1.44).

Literature distinguishes two levels of self-evaluating activity. Zimmerman and Paulsen (1995) on the one hand refer to it in terms of the monitoring and evaluating of behavioural activity such as identifying mistakes, analysing time logs, or reading behaviours. They report the benefits of such activity in terms of its capacity to provoke a reaction - leading to a change in behaviour. Schmitz and Perels (2011) on the other hand, suggest that the self-evaluation process should involve 'a series of recordings of the whole cycle of self-regulation' (p. 256), thus potentially increasing the learner's metacognitive awareness of their learning experience. This study references the latter understanding of self-evaluation, which is presented throughout as a metacognitive activity, one in which the learner evaluates task preparation and performance, and reflects upon learning outcomes and how they can be improved.

In recent decades researchers have become aware of the importance of developing self-regulation of learning, and in particular metacognition (the control and regulation of cognitive learning strategies) (Flavell, 1979; Zimmerman, 2001; Efklides, 2011). A primary goal of many western education systems is to produce independent learners who can take control of the learning process. This is at the very least in order to be able to keep pace with the unprecedentedly rapid evolution of societal needs for diverse and newly-developing skills (Dinsmore *et al.,* 2008). A major focus in such studies has been the search for manageable and effective ways of developing such skills. As Zimmerman and Paulsen (1995) point out, 'teachers know self-regulated academic learners when they see them' (p.13). Identifying them is not the problem - the problem is knowing *how* to foster their self-regulation such that it becomes an habitual aspect of the learning process (Zimmerman, 2002).

Self-evaluation is a principal metacognitive skill in self-regulation (Zimmerman, 2008) and serves to inform the control and regulation of cognitive and metacognitive strategies such as goal setting, organisation and elaboration - strategies that are 'crucial for deep understanding and sustained retention' (Hübner, Nückles and Renkl, 2010, p. 20). Self-evaluation features as a key component of the most-cited self-regulated learning models (Puustinen and Pulkkinen, 2001) and is important for a number of reasons. It enhances learning by giving the student the opportunity to focus on, and take control of, the learning process (Bandura, 1986). It helps them to evaluate how performance is related to specific strategies, leading to the abandonment of less effective ones and the identification and implementation of new and more appropriate ones. It can promote systematic organisation of self-reflections, leading to more effective goal setting, task management and effort management. It can also assist with structuring time-management, an important self-regulated learning strategy

(Zimmerman, Greenberg and Weinstein, 1994), and can increase self-efficacy (Schunk, 1983) and motivation (Zimmerman and Paulsen, 1995; Schmidt, Maier and Nückles, 2012). Finally, it encourages reflective thought, a key critical thinking activity that has the potential to result in better organisation of the planning, goal-setting and evaluation process (Bandura, *op cit*).

The consequences of *not* self-evaluating are not to be ignored, and students who do not engage in metacognitive evaluation of their learning progress are at a disadvantage compared to those who do. Consequences stretch beyond the classroom, where those who self-evaluate can take control and direct their learning over and above the instructional framework provided by their teachers, thus giving themselves the opportunity to get ahead. Students who progress to college and university without self-evaluation skills will find themselves unprepared for the less structured regime of the tertiary environment. Further still, the modern workplace, where there are fewer and fewer 'jobs for life', offers an environment in which the person who can self-evaluate their performance has the advantage. Such a person can take pro-active control over their career-path, compared to the one who leaves the monitoring and evaluating to outside agencies such as the 'Human Resources' department. It is proposed therefore that the fostering of self-evaluation is a key self-regulatory skill that can enhance academic opportunities not only at school, but also beyond.

The effectiveness of self-evaluation is potentially moderated by a number of pre-requisite self-regulatory processes. Pressley and Ghatala (1990) suggest for instance that motivation to learn is a significant moderator to the desire to self-evaluate. Students must perceive the benefits of the effort of self-evaluation; therefore their self-efficacy perceptions are equally important - if they believe that self-evaluation will result in tangible benefits, then they are more likely to engage with the process. Their metacognitive awareness of why, how and when to self-evaluate is an important regulator in terms of their control over how much to self-monitor and self-evaluate (Paris and Byrnes, 1989). Thus three key self-regulatory characteristics (the learner's capacity for metacognition, the extent of their motivation and their level of self-efficacy) have the potential to influence, and to be influenced, by the process of self-evaluation. All three feature as key self-regulated learning characteristics under investigation in this study, in the context of their relationship to the use of a learning journal for the self-evaluation of learning experiences.

Zimmerman and Paulsen (1995) theorise that the effects of self-evaluation can be influenced by whether or not the process leads to a positive outcome. The student who observes a positive outcome is likely to continue self-evaluation, whilst the one who does not is less likely to do so. They suggest that one solution is to suggest to learners that they frame their observations from a positive rather than a negative perspective. For instance a student might be asked to reflect upon the extent to which they *avoided* distraction, rather than the extent to which they were *actually distracted*. They further suggest that it is important to support self-efficacy by focusing on strategy choice, rather than ability, when it comes to evaluating the evaluation of performance. Both suggestions have informed the design of the learning journal used in the current study.

Whilst it has been established that metacognitive self-evaluation is beneficial to the learner, students of all ages, including university students, do not typically self-evaluate their learning without specific support (Nückles, Schwonke, Berthold and

Renkl, 2004). This may be because (according to Honey and Mumford, 1989) the process of self-evaluation comes more easily to some learners than to others. The latter authors suggest that the recording of metacognitive activity in a *learning journal* not only develops the discipline of self-reflection, but it also hands the responsibility for the plenary process to the learner. A learning journal, which is a series of learning protocols written over time, can therefore provide a vehicle for self-regulation of learning by offering a framework for reflection and elaboration of the learning experience (Zimmerman and Paulsen, 1995; Berthold Nückles and Renkl, 2007; Schmitz and Perels, 2011; Schmitz, Klug and Schmidt, 2011).

Research studies, largely referring to work with university students, have reported that learners who have used learning journals have out-performed those who did not, in terms of both learning outcomes (Wong, Kuperis, Jamieson, Keller and Cull-hewitt, 2002; Berthold *et al*., 2007), and learning strategies (McCrindle and Christensen, 1995; Nückles, Hübner and Renkl, 2009; Clipa, Ignat and Stanciu, 2012; Glogger *et al*., 2012). This self-evaluation process, typically following a learning event, can be structured through a series of cognitive and metacognitive prompts, hints and questions that encourage the learner to formalise, structure, articulate and explore their metacognitive skills and experiences, their goals and learning strategies, monitor their understanding and evaluate their learning outcomes. It is suggested that such prompts can also activate cognitive strategies that the learner may already be capable of employing, but which are not necessarily activated without support (Hübner *et al*., 2010). Furthermore, the learner can proceed to make use of the self-evaluation information gathered, to explore issues of causality and to identify appropriate remedial strategies. The concept of a learning journal is further explored theoretically in Section 2.3.7, and experimentally in Section 6.5, RQ4.

2.3.5.2.4 Metacognitive experiences

Efklides (2008) suggests that metacognition has three facets: metacognitive knowledge, metacognitive skills and metacognitive experiences. The latter is a construct that, although it was defined in the 1970s (Flavell, 1979), has so far received less attention than other aspects of self-regulated learning. Indeed Efklides (*op cit.*) specifically calls for researchers to explore metacognitive experiences alongside the usual emphasis of metacognitive skills and metacognitive knowledge. *Metacognitive experiences* describes a dynamic, self-reflective and analytic process, whereby a learner expresses their awareness of feelings of self-efficacy with regard to the task in hand, their feelings of difficulty, challenge, task-related motivation, effort, success and their learning satisfaction (Efklides, 2006, 2009). De Bruin and van Gog (2012) suggest that 'to gain a full-fletched understanding of metacognitive processes in the classroom, more attention should be given to prior and current affective experiences and how these influence learning' (p. 251).

Metacognitive experiences describe the analytic awareness that the learner has of the learning experience: of the features of the task; the extent of their mastery of the cognitive process; the effort involved in completing the task. These metacognitive experiences in turn inform the control and regulation process of self-regulated learning (Efklides, 2008).

According to Flavell (1979) the importance of metacognitive experiences is threefold.

Firstly, they can influence the learner's goal orientations, leading them to generate new ones, and revise or abandon old ones. For instance the metacognitive experience of *feeling of difficulty* in a mathematics test may influence whether a learner sets themselves the goal of further revision or not. Secondly metacognitive experiences feed the metacognitive knowledge base, especially in adolescence; the metacognitive experience of learning satisfaction after a particular task, may inform the learner's metacognitive knowledge of the extent to which that particular learning strategy is an effective one. Thirdly metacognitive experiences activate cognitive and metacognitive strategies; the learner who enjoys the challenge of a self-evaluation task may be likely to choose that particular strategy another time; the one who experiences feelings of difficulty may be more cautious in their choice of strategies in the future.

Whilst the theoretical basis for the interaction of metacognitive skills and metacognitive experiences is well-established, with metacognitive experiences being described as the 'missing link in the self-regulated learning process' (Efklides, 2006, p. 287), few studies have conducted a process analysis of these variables over time. Efklides (2008) calls for further research into the characteristics of metacognitive experiences, suggesting that there is a need to explore them in specific populations, in collaborative situations and alongside both other aspects of self-regulated learning and performance measures. This study thus aims to contribute to the body of knowledge by exploring the relationship between aspects of metacognitive experiences and metacognitive skills, in a collaborative learning situation, in the particular setting of an internet-based learning course for young adolescents from England. The potential interaction over time, between metacognitive skills and metacognitive experiences is therefore the subject of a research question in this study (see RQ6, Section 6.7).

2.3.6 Conceptual model for the study of SRL and CT

It has been demonstrated that self-regulated learning can be characterised as a set of offline learner beliefs and/or a set of on-line processes, and that SRL is generally explored from these two perspectives. This dual perspective has informed the conceptual model for this study.

Firstly, self-regulated learning is described as a set of *traits*, *dispositions* or *attributes*, which the learner possesses, and which are considered pre-requisite characteristics of a self-regulated learner. Frequently these *person-level* traits are referred to as 'components' of self-regulated learning, and they include such aspects as self-efficacy, metacognitive knowledge or awareness, and motivational beliefs. As discussed in Section 2.3.5.1, various models place different emphasis on each of these components and are generally concerned with hypothesising the interactions between them, and their effects on learner achievement.

The second perspective views self-regulated learning as a *process* that directs learning at the *task level*. Thus the learner employs a set of self-regulatory skills in the pursuit of learning goals. Typical of process models is a description of activities involved in the control and regulation of cognitive and metacognitive skills, in the pursuit of a learning goal. Such activities include goal setting, task management, volitional strategies, motivation, maintenance of self-efficacy, critical thinking, self-monitoring and self-evaluation.

The model that best represents the approach adopted for the exploration of self-regulated learning in this study (Figure 2-1), integrates both of these approaches, reflecting predominantly the task/person level approach proposed by Efklides (2011) and the cyclical process model of Zimmerman (2000, 2008).

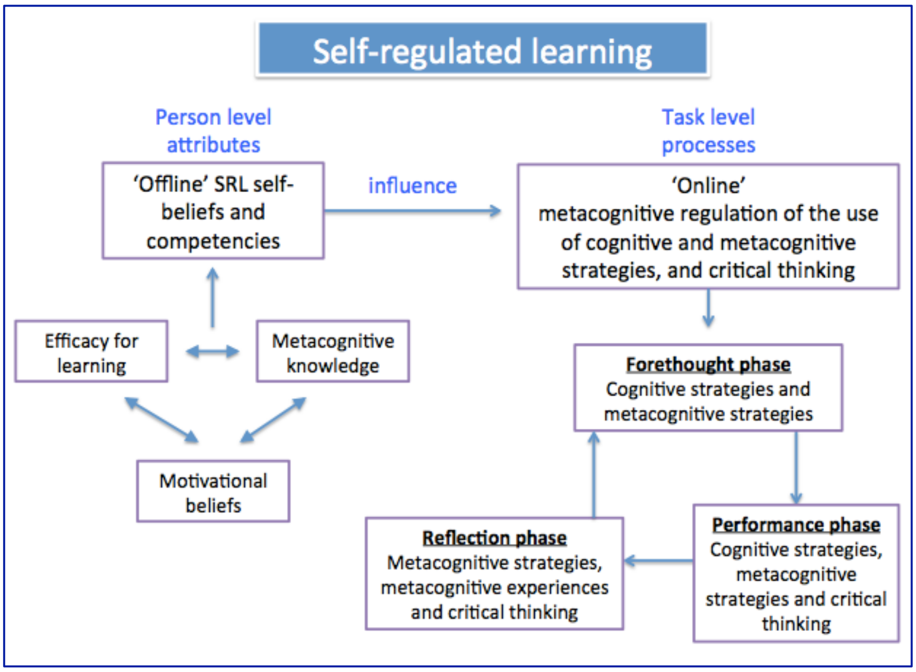


Figure 2-1 Conceptual model for the study of self-regulated learning

Thus this model conceptualises that the *person level* SRL competencies (metacognitive knowledge, efficacy for learning and motivational beliefs) inform and influence the *task level* processes (the cyclical process of forethought, performance and self-reflection).

The conceptual model influences both the structure of the study and the research questions. Self-regulated learning *dispositions* (metacognitive knowledge, motivational beliefs and efficacy for learning) are measured at pre-and post-test, whilst self-regulated learning *processes* (metacognitive skills, critical thinking and metacognitive experiences) are measured during the intervention. Person-level self-regulated learning attributes are captured in self-report measures, and provide pre- and post-test measures that contribute to the evaluation of the effectiveness of the intervention. Task-level processes are captured both in count measures of intervention activity, manifest critical thinking assessment measures and in self-report measures reported in the students' weekly learning journals. These provide process variables with which to analyse self-regulated learning in action over the course of the intervention.

Research questions RQ3 to RQ5 concern themselves with SRL dispositions, whilst RQ6 explores SRL processes. RQ7 tests the conceptual model itself, by exploring potential links between SRL dispositions and SRL processes as they are detected in this study. The latter question also seeks to explore whether post-intervention adjustments should be made to the conceptual model.

2.3.7 Developing self-regulated learning

UK and international literatures agree that students do not become self-regulated learners of their own accord; rather they need to be taught how to learn (e.g., Pressley, 1995; Winne, 1995). It is suggested that the development of self-regulated learning strategies is the key to successful learning (Zimmerman and Schunk, 2001). O'Malley, Chamot, Stewner-Mazanares, Kupper and Russo (1985), in an observational study of seventy US high school students and their twenty two teachers, reported that teachers were generally unaware of the learning strategies employed by their pupils, suggested that students who were not afforded the benefit of guidance in metacognitive development were 'learners without direction' (p. 24). Ellis and Sinclair (1989) echoed this conclusion and determined that without the explicit development of self-regulated learning strategies, learners were unlikely to be able to transfer learning skills from one situation to another.

Both Meyer *et al.,* (2008) in a review for the Department of Children Schools and Families, and Duckworth *et al.,* (2009) in their Institute of Education review, noted that the most prevalent and effective self-regulated learning teaching strategies in the UK involved a) scaffolding and b) the development of self-monitoring strategies.

Scaffolding refers to the teaching technique whereby learning is supported by a series of prompts, instructional aids, strategies or 'hooks' provided by teachers, which in the context of this study, structure the process of self-regulation (Meyer *et al.,* 2008). By providing a structured framework, teachers can encourage the development of strategies such as goal setting, monitoring of understanding, analysis of the effectiveness of particular learning strategies, and reflection of learning progress. Such a framework can be provided as prompts that are offered at certain stages during a task. At different times during a learning event, students can be encouraged to articulate and monitor relevant learning goals or use mind maps to develop metacognitive elaboration skills, linking new information with what they already know. Following the learning event they can use prompt sheets, which help them to evaluate their progress. Research has established that scaffolding can be provided in two ways: fixed scaffolding and adaptive scaffolding. *Fixed scaffolding* describes prompts that are unaffected by the learner's performance, whilst *adaptive scaffolding* is characterised by prompts which change according to the responses of the learner (Azevedo and Hadwin, 2005). Thus, a structure characterised by fixed scaffolding would offer students a generic set of prompts that could be used across a variety of learning situations, whilst adaptive scaffolding would typically involve dynamic interaction between the teacher and the learner. A number of studies have evaluated the comparative effectiveness of the two types of scaffolding, and in particular in the context of the hypermedia learning environment - with contradictory results. For instance, Chang, Sung and Chen (2001) studied 48 young adolescents, evaluating, in terms of pre- to post-test biology grades, the effectiveness of two computer-based fixed scaffolding conditions, in comparison with a control condition where students did not receive scaffolding. Both fixed scaffolding conditions provided students with test results accompanied by feedback which was automatically generated according to corresponding results, and one additionally provided a concept-mapping scaffold. Using analysis of covariance testing they found that both fixed scaffolding conditions were effective in raising students' post-test scores, compared with the control condition, with a medium to large effect size ($\eta^2 = .17$). Other studies, such as those by Jacobson and Archodidou (2000) have theorised that fixed scaffolding, using for instance statement banks for story generation, can be significantly effective as learning tools. On the other hand, studies have compared fixed scaffolding with adaptive scaffolding, finding that the latter approach is the more effective in terms of influencing learning outcomes. Azevedo, Cromley and Seibert (2004) for instance, in their experimental study of fifty-one US undergraduates, found that whilst both fixed and adaptive scaffolding techniques were effective in raising metacognitive activity, the latter technique effected the greater increase. Bangert-Drowns, Hurley and Wilkinson (2004), in a meta-analysis of forty-eight 'writing to learn' interventions for US school-aged children, explored the effectiveness of the scaffolding of self-reflection on academic achievement. Comparing writing interventions which employed metacognitive fixed-scaffolding prompts with those which provided no prompts, they reported an overall small effect size of Cohen's $d = .27$. However, those studies which reported that students specifically responded to scaffolded metacognitive prompts reported a greater effect size of $d = .44$. They concluded that fixed-scaffolding a learner's self-reflection process by offering prompts which guide a learner's self-reflection is an important aid to the development of self-regulated learning. Their assertion is consistent with other similar studies such as that by McCrindle and Christensen (1995), who compared two groups of Australian university students, both of whom wrote scientific reports, but one of which additionally completed a prompted (with fixed scaffolding) learning journal which drew their attention to the self-regulatory processes they had employed during the learning event. The journal group out-performed the non-journal group in overall achievement (d = .79), and used significantly more metacognitive strategies ($p < .01$). They attributed the effectiveness of the learning journal to its capacity to provide a structure for the reflection of the cognitive and metacognitive processes used during the learning process. They suggested that the students who were in the learning journal group 'saw learning less as a process of acquisition of knowledge and facts and more as a process of comprehension, analysis and interpretation' (p. 181).

The notion of *self-monitoring* is thought to be fundamental to the cyclical model of self-regulated learning (forethought, performance and reflection) (Zimmerman and Paulsen, 1995). Indeed such metacognitive reflection is believed to be the 'gateway to self-regulated learning' (Glogger *et al.,* 2012, p. 453). Metacognitive reflection involves the learner observing, monitoring and reflecting upon the effectiveness of their learning strategies (e.g., elaboration or organisation) in terms of their capacity to influence learning outcomes. Research has consistently shown that the process of reflection can enhance the self-regulatory process of metacognition - the control and regulation of the learning process (McCrindle and Christensen, 1995; Schmitz and Wiese, 2006; Nückles, Hübner and Renkl, 2009; Mehmet, 2010; Fabriz, Ewijk, Poarch and Büttner, 2014). For instance, in their study of 103 German undergraduates, Nückles, Hübner and Renkl (*op cit.*) explored the effects of combinations of different types of prompts (no prompts; cognitive prompts, metacognitive prompts and strategy prompts) on the use of metacognitive learning strategies and on comprehension test results. They detected a group effect, when comparing self-reflection groups with non-self-reflection groups, on the learning outcome of the use of metacognitive strategies, resulted in an effect size of $\eta^2 = .31$.

Similarly, they conducted a mixed analysis of variance, this time evaluating the effect in terms of the pre- to post-test comprehension test results, comparing self-reflection groups with those of non-self-reflection groups, finding an overall large effect for time ($\eta^2 = .25$) and a large group effect ($\eta^2 = .15$). They also reported that the most effective type of prompt was the mixed prompt ($\eta^2 = .10$), which included cognitive metacognitive and strategy prompts.

A number of related studies, including those by Berthold, Nückles and Renkl (2007) who worked with undergraduates, and Hübner *et al.* (2010) who worked with high-school children, have reported that, where a combination of cognitive and metacognitive prompts is operationalised in a *learning journal*, it can assist in the promotion of learning strategies. In their experimental study of 84 German undergraduates, Berthold, Nückles and Renkl (*op cit.*) found that students whose learning journals were scaffolded with a mixture of cognitive and metacognitive prompts outperformed those who did not receive prompts, with a large effect size of $\eta^2 = .68$. Similarly, Hübner *et al.* (*op cit*), in their experimental study of 70 German secondary school students, also compared the effectiveness of the use of prompts in learning journals. Interestingly they found that presenting students with examples of learning journals increased their use of metacognitive strategies, but did not have a significant effect on knowledge gain. However, where those learning journals also included cognitive and metacognitive prompts, then their effects on knowledge gains were significant and large ($\eta^2 = .15$). Furthermore, those gains remained significant when tested one week later. They concluded that both self-regulated learning strategies and knowledge gain could be positively affected by the use of learning journals that contain cognitive and metacognitive prompts. Furthermore, they theorise that the younger the student, the more necessary it is to provide structured instructional prompts.

This section has explored the two teaching strategies which, according to Meyer *et al.* (2008) and Duckworth *et al.* (2009), are not only prevalent in the UK, but also significantly effective in developing aspects of self-regulated learning, and in particular metacognition. A common theme has emerged - that these two strategies, *scaffolding* and *self-monitoring* can be effectively developed through the use of learning journals. The following section will explore this notion further.

2.3.7.1 Developing self-regulated learning with a learning journal

This section introduces one of the strategies that characterise the intervention - the potential for the *learning journal* to provide the vehicle for the development and analysis of self-regulated learning processes (e.g., Schmitz and Perels, 2011). It will further explore the argument that self-regulated learning can be developed through the regular use of a prompted learning journal. In support of this notion, it examines related literature, and attempts to identify gaps in research that have the potential to be addressed in this study. It is presented as a series of questions that might commonly be asked about the learning journal and how it might be used to develop aspects of self-regulated learning.

2.3.7.2 What is a learning journal?

In recent decades there has been an increasing awareness of the extent to which the learner can take responsibility for their own learning. This notion reflects the socio-cognitive perspective that the learner can play a significant role in knowledge development (Bandura, 1986). Empirical research (e.g. that by Hübner *et al.,* 2010) finds that, in maintaining a learning journal, the learner exercises self-regulatory skills by assessing the learning task, setting goals, using time efficiently, regulating the capacity to avoid distraction, monitoring the learning experience and revising goal orientations. This 'reactivity' process affects their metacognitive experience, which in turn affects the motivation, self-efficacy and metacognitive knowledge that they take forward to the next task (Zimmerman, 2002, p. 65). The latter author suggests that simply by asking a student to record their metacognitive experiences, the process acts as a trigger that increases the individual's metacognitive awareness, and thus feeds into the cycle of self-regulation.

It is suggested that the learning journal can be used to promote self-regulated learning both by providing the structure for self-reflection and also as a means of evaluating learning processes over a number of learning events (Schmitz, Klug and Schmidt, 2011). Thus learners self-regulate their learning by recording and monitoring their metacognitive knowledge, metacognitive skills and metacognitive experiences as they occur during the learning process (Hübner *et al.,* 2010).

A learning journal can be presented in different forms (e.g., audio, video, word processed or questionnaire) and in different levels of structure (e.g., unstructured free writing or structured with prompts) (Langer, 2002). At times described as a 'log' or a diary', it is a self-development tool which provides a framework for the self-regulated learning function of metacognitive reflection - the process of analysing, evaluating and reflecting upon the learning experience (Barclay, 1996; Cowan, 1992). In contrast to other types of writing tasks (e.g., summaries, précis, or scientific reports), the learning journal is specifically intended to target the development of cognitive and metacognitive self-regulated learning strategies. At the cognitive level, the learning journal should support the production of both organisational strategies (e.g., identification of main ideas and the relationships between them) and elaboration strategies (synchronising learning experiences by relating learned material to external or already learned examples or analogies). At the metacognitive level, learners use the learning journal to structure self-reflection of the learning process, leading to the control and regulation of the use of those same organisational and elaboration strategies (Hübner *et al.,* 2010). In other words, the metacognitive activity controls and regulates the cognitive activity.

The notion of the learning journal is considered appropriate for both university students (Schmitz and Wiese, 2006), and school students (Hübner *et al.,* 2010). Operationally, and as conceptualised in this study, the journal is a series of writing assignments, which are constructed to scaffold the planning, monitoring and evaluation of the learning process by presenting a set of motivational and positively phrased questions (see Appendix 2), which prompt the student to assess their own strengths and weaknesses, set goals, manage their time and monitor their learning, - emphasising, when the opportunity arises, the links between learning strategies and learning outcomes (Lee, Lim and Grabowski, 2010). It is typically administered as a follow-up activity following as soon as possible after a learning event such as a course, a lecture or a week of lessons (Schmitz, Klug and Schmidt, 2011).

As the learning journal is providing a structure for the self-regulation of learning, it is reasonable to construct it such that it scaffolds the phases of self-regulation as they occur. Hence, following Zimmerman's (2000) frequently cited cyclical model of self-regulated learning, a typical learning journal might prompt the learner to reflect upon how they might have prepared for a task, the goals they set, the metacognitive strategies they employed during the task, the extent to which they met their goals, their level of efficacy and their evaluation of the learning process. It may be structured as a series of Likert-style responses, where students are asked to consider a series of statements, reflecting upon how relevant they are to their recent learning experience. So, for instance, statements such as, 'I was able to put aside all of my negative thoughts and/or worries today while I was studying', would be designed to encourage them to reflect upon their distraction avoidance strategies, whilst, 'I tried to consider several ways of looking at this week's dilemma before forming my argument' would focus upon the reflection of elaboration skills. Depending upon the context of the learning journal, it may also be appropriate for students to contribute to their learning journal in text format, responding to questions such as, 'To what extent did you meet your learning goals today?' or 'Looking at your responses above, write down ONE learning strategy that you will focus on, as your goal for next week'. The journal would be kept for a set time, and would be presented at set intervals (e.g., daily, or weekly).

The delivery of learning journals has either been effected on paper (e.g., Cantrell, Fusaro and Dougherty, 2000; Langer, 2002; Schmitz and Wiese, 2006) or on a computer using scaffolded electronic prompts (e.g., Nückles *et al.,* 2004; Schwonke, Hauser, Nückles, and Renkl, 2006). Few learning journal studies, especially on young adolescents, have explored the opportunities that 'cloud technology' offers to prompt, record and measure metacognitive activity. This innovative methodology, which is employed in the current study, enables the researcher to track, monitor and record learning journal entries as they are made in real time. Such technology enables the journal to be accessed via the internet by the student, and saved in a cloud space that can be accessed immediately by the teacher. Journal items are recorded and saved electronically and can be downloaded into a spreadsheet programme. Responses can be recorded in either Likert-style statements, yes/no questions, or in full text form, thus enabling the teacher or researcher to interrogate the data quantitatively or qualitatively, as necessary. Thus the learning journal can serve as a teaching and learning tool, and also as a research instrument.

2.3.7.3 How effective is the learning journal as a learning tool?

It is suggested that the use of learning journals can have an effect, not only on self-regulatory outcomes, but also on academic learning outcomes. The meta-analysis of Bangert-Drowns, Hurley and Wilkinson (2004) identified that interventions that included metacognitive prompts had a large effect on achievement tests (Cohen's d = .44), compared with interventions that did not use metacognitive prompts. It seems reasonable to expect that self-reflection can afford an internal feedback loop that can help learners to move to the next step in their learning. Other studies however, have failed to detect such effects. Fabriz *et al.,* (2014), for instance, found no significant improvements on academic performance, comparing university students who used a learning journal with those who did not. Similarly, Hübner *et al.,* (*op.cit.*), who studied high school students, reported that journal writing only improved academic performance where a specific combination of metacognitive and cognitive (elaboration and organisation) strategies was targeted by the journal writing exercise. Indeed the latter authors suggest that such a combination was 'crucial for learning success' (p. 19). It should be said however, that 'learning outcomes' covers a wide range of academic achievement, which may account for the difference in research findings, with the outcomes ranging from knowledge of probability (Glogger *et al.,* 2012), to 'conventional measures of achievement' (Bangert-Drowns *et al.,* 2004, p. 50). Nevertheless, authors in general agree that the use of the learning journal has a positive effect on metacognitive control and regulation of the cognitive strategies employed in the learning process - and that this level of self-regulation has the potential to impact upon learning outcomes (Nückles *et al.,* 2009).

2.3.7.4 How effective is the learning journal as a research tool?

Self-regulated learning is generally understood to consist of situational factors as well as a set of more static dispositional characteristics, and thus lends itself to analysis (Winne and Perry, 2000). Researchers such as Schmitz and Wiese (2006) have found merit in using standardised learning journals as a means of capturing and evaluating such real-time situational self-regulation as it occurs, for research purposes. In a five-week study of forty university students, they used time-series analysis of standardised diaries to evaluate the effectiveness of an intervention to affect aspects self-regulated learning. They concluded that data collected in a standardised learning journal could provide an effective means of measuring the process of self-regulated learning over time. According to Schmitz, Klug and Schmidt (2011) such diary data can be as reliable a source as direct observation from an external observer. Separate elements of self-regulated learning (e.g., the use of metacognitive and cognitive learning strategies) can be tracked at different time points during the learning process, enabling the researcher to observe the trajectory of self-regulated learning as it happens.

Zimmerman (2008) found that the use of diaries as an event-measure of study habits, was more accurate that the use of questionnaires. However, others have detected difficulties surrounding the existence of a potential method effect, whereby the experience of recording reflections, the main function of the journal, potentially increases metacognitive capacity as a result (Barclay, 1996). It is important, therefore, to include a control element into the design of studies that seek to use the learning journal as both an assessment tool and as a learning tool. This consideration has been built into the design of the current study; whilst the learning journal provides data for the observation of self-regulated learning processes as they are reported during the intervention, the effect of completing the learning journal is also isolated by intervention condition (see the design in Section 4.3).

2.3.7.5 What are the limitations of the learning journal approach?

The learning journal is a personal record of learning experiences; therefore it relies upon the self-reporting of events, which, especially in children, may not accurately represent reality. However, this limitation is balanced against the notion that the journal is completed very soon after the learning event, and thus has the potential to be more accurate than say post-test self-reporting mechanisms.

A second potential limitation refers to the level of learner motivation required to sustain the continued completion of the journal over a number of learning events. However, very little research has gone into this phenomenon - as yet little is known about how many journal completions represent the optimum number of events that a learner can be expected to complete before tedium becomes a limiting factor. In considering these and other limitations (e.g., the need for high levels of compliance and the dependence upon a student's writing competence) Schmitz, Klug and Schmidt (2006) nevertheless conclude that the advantages of the learning journal as both a learning tool and a measurement tool outweigh its disadvantages.

2.3.7.6 What are the gaps in learning journal research that have the potential to be addressed in this study?

The use of the learning journal has been positively associated with gains in one particular aspect of self-regulated learning, namely metacognitive control of learning strategies (Berthold *et al.,* 2007; Nückles *et al.,* 2009). Thus when authors refer to the effects of the learning journal on self-regulated learning, they are usually referring to its metacognitive aspects (e.g., Nückles *et al.,* 2009; Hübner *et al.,* 2010). However, very little research has been conducted into potential links between the use of a learning journal and other trait aspects of self-regulated learning (e.g., motivational beliefs and self-efficacy) (Azevedo, 2014). This study explores the potential effects of learning journal use on three self-regulated learner characteristics: metacognitive knowledge, motivational beliefs and efficacy for learning - and thus has the potential to extend the field of learning journal study beyond the context of the metacognitive aspects of self-regulated learning.

Learning journal studies have yielded mixed results, in terms of links between the use of a learning journal and context-specific learning outcomes, with for instance small effects (η^2 = .05) being detected in mathematics learning (Schmitz and Perels, 2011), and large effects (*d* = .85) in story comprehension (Wong *et al.,* 2002), but not at all in biology learning (e.g., Stephens and Winterbottom, 2010). The effects of the use of a learning journal on the development of metacognitive skills are becoming well established (e.g., Nückles *et al.,* 2009), and critical thinking is generally categorised as a metacognitive skill (Phan, 2010), so it seems reasonable to explore whether or not the use of a learning journal influences critical thinking skills. English studies that include such a focus are rare (Higgins, Hall, Baumfield and Moseley, 2005); therefore this study aims to contribute to the field of learning journal study by exploring whether critical thinking achievement should be included in the list of potential benefits of implementing a learning journal into a learning programme.

A significant amount of research into the effectiveness of learning journals has emanated from Germany and has focused upon college or undergraduate students (Schmitz and Wiese, 2006; Schwonke *et al.,* 2006; Berthold *et* al., 2007; Nückles *et al.,* 2009) with very little research having been conducted on school-age children in England (Stephens and Winterbottom, 2010). This study aims to contribute to the field by exploring the effects of the use of a learning journal in an English school-age setting.

One strand of research has explored the use of a learning journal as a measurement tool for the assessment of self-regulated learning strategies themselves. Schmitz and Wiese, (2006) for instance, evaluated a self-regulated learning intervention by exploring a series of standardised diaries using time-series analysis. Using time-series analysis they were able to evaluate trends, relationships between learning variables and learning outcomes, and the effects of different intervention components. They concluded that learning diaries provided a useful means of evaluating the process of self-regulated learning, with the recommendation that future studies attempt to isolate the method effect whereby the recording of experiential learning itself may increase self-regulated learning. The exploration of the possible method effect of using learning journals has not yet featured prominently on the agenda of current researchers. This omission is addressed in the current study, where an attempt has been made to isolate the method effect of the learning journal from other intervention effects.

Learning journals have been assessed in a number of ways including single case studies, trend analysis and time-series analysis (Schmitz, Klug and Schmidt, 2011). Azevedo (2014) suggests that we are only on the verge of identifying which statistical methods can cope with the complex process analysis involved in event-based self-regulated learning data. Winne (2014) concurs, suggesting that this field of study is still at an early stage. This is an important consideration, and one that the current study attempts to address. Autoregressive cross-lagged panel analysis (ARCL: Kenny, 1979) is considered appropriate for exploring multiple variables captured over time. It follows that it should be suitable for the analysis of a diary that records multiple variables over multiple learning events. In the current study, by using ARCL to analyse learning journal process data over a series of learning events, self-regulatory behaviour will be analysed inferentially in terms of the relationships between variables over time, and the pattern of influence of self-regulatory behaviour on specific learning outcomes. Thus this study aims to make a methodological contribution to the field of learning journal research by demonstrating the effectiveness of the ARCL method, for evaluating the contents of learning journals over time.

Nett, Götz, Hall and Frenzel, (2012) issue challenges for researchers - challenges that have the potential to be addressed in this study. Firstly, they suggest that there is merit in investigating the interaction of cognitive, metacognitive and affective self-regulation strategies in the context of a learning situation, and over time, as a means to a) improve classroom practice and b) develop effective intervention programmes. This challenge is embraced in this study, which explores relationships between event-measures of cognitive and metacognitive activity as it is reported in a learning journal, over four learning events. Secondly, they echo the call of others such as Boekaerts and Corno (2005) and Hopfenbeck and Maul (2011), who suggest that self-report measures should be supplemented by observed measures where possible. In response, this study has been designed such that both observed and self-reported measures are included.

This section has explored the merits of the use of a learning journal to promote and measure aspects of self-regulated learning, attempting to justify its inclusion as one of two main components in the study.

2.4 Widening participation

The Milburn Report of 2009 explored the drivers of social mobility, citing stark statistics to support the notion that the prospect of career success of a child born in the 21st Century in England is determined more by 'birth than worth' (p.5). For instance, in 2009, over 50% of top professionals (e.g., barristers, Members of Parliament and civil servants) were educated in the private sector, compared to only 7% of the population as a whole, reflecting a trend in which professionals are increasingly hailing from wealthier backgrounds. The report concluded that attendance at English selective universities gives the most ready access to the professions, with such access being the key to opening new opportunities for 'a second great wave of social mobility in the years ahead' (p.5). However, applications to selective universities are themselves highly correlated with students' background and with the schools they attend: 'currently talented young women and men from disadvantaged backgrounds [...] who could apply to selective universities, are disproportionately not doing so, so reducing their chances of upward social mobility.' (Harris, 2010, p. 7).

A number of government reports have forcefully argued that the universities have a responsibility to support the drive for equality of opportunity (Dearing, 1997; Kennedy, 1997; Milburn, 2009; Jerrim, 2013). The 'Widening Participation' (WP) agenda is now enshrined in annual renewable agreements between universities and the Office for Fair Access (OFFA). Any university in England which chooses to charge above a basic annual fee[7] per undergraduate student must (since 2006) sign an agreement with OFFA which sets out how that university will improve access for students from a wide range of backgrounds. For instance, the University of Oxford, in its 2015/16 OFFA agreement (OFFA, 2015), sets out how it will aim to attract applicants from backgrounds which have been under-represented in the past. Whilst their *Widening Access* programme focuses upon older teenagers, aiming to convert those from under-represented (e.g., socio-economically disadvantaged) groups into applicants, their *Widening Participation* programme targets younger (not necessarily disadvantaged, but nevertheless from state schools) students, from age 10 upwards, promoting progression to a broader range of higher education institutions, including those other than Oxford itself (see Oxford, 2015). The objective of WP programmes is to inspire these students to situate the possibility of entry to higher education institutions, including other leading selective universities, on their educational radar. This objective is translated into a range of programmes that aim to introduce young students to intellectual activities that they may have not encountered before (such as debating), to introduce them to new subjects that they may have not considered studying, and to raise their academic achievement by enabling them to work on challenging intellectual projects. Learning outcomes typically articulated in university outreach literature include:

- To encourage awareness of the learning journey and the steps on it
- Inspire intellectual, personal curiosity

[7] The basic annual fee for English undergraduates, in 2015, was £6000.

- Enrich and enhance students self esteem, aspirations and hopes for the future. (Oxford, 2015, p.1)
- To encourage critical thinking. (Warwick University, 2015, p.1)

Thus a potentially effective outreach programme would include aspects of WP aims: intellectual activities which were not normally encountered in school, and activities which promoted aspects of self-regulated learning (such as metacognitive awareness, motivation to learn and self-esteem). Across the range of English universities, there are however, few WP programmes for these younger students. This is despite acknowledgement that the specific targeting of this age group plays, 'an important role in reducing socio-economic inequalities' (Jerrim, 2013). This study focuses upon the creation, implementation and evaluation of one such intervention, *Is this the right room for an argument?*, a critical thinking course for young adolescents, which aims to develop aspects of self-regulated learning.

2.5 Conclusion of literature review

The purpose of the literature review has been to explore principal constructs, establish the context and the rationale for the intervention, inform and construct a conceptual model to structure the evaluation of the intervention, and identify gaps in knowledge which this study has the potential to address.

The review began with an exploration of the relationship between critical thinking and self-regulated learning, as the constructs are operationalised in this study. It proposed metacognition as a common factor relating to both critical thinking and self-regulated learning and suggested that, for the purposes of this study, critical thinking should be conceptualised as a metacognitive skill within the construct of self-regulated learning.

A review of current educational policy regarding the importance of critical thinking in the National Curriculum in England provided a rationale for the inclusion of this construct in the intervention. It was argued that whilst critical thinking can be viewed as a key skill which is of particular intellectual benefit to learners in school and beyond, there are few national programmes in England which provide either for the development of critical thinking or its assessment. In acknowledgement that early adolescence is an age where critical thinking skills emerge and develop rapidly, and in view of the already constrained curriculum provision, it has been suggested that there is a need for interventions which can be delivered to this age group in such a way that they do not overload existing teaching provision. Opportunities for the teaching and assessment of critical thinking have been summarised, leading to the identification and justification for argumentation being an important critical thinking skill, and for including it as a key component of the study intervention.

A review of the literature on self-regulated learning (SRL) began by exploring the history of the construct, identifying key terms, and rationalising its working definition. It was noted that models of self-regulated learning present different emphases, but most agree that it is a complex and integrated process involving volition, cognitive and metacognitive skills and metacognitive experiences, and is a process which is driven by a learner's motivational self-beliefs, their sense of efficacy for learning and their metacognitive knowledge. An exploration of SRL studies revealed that the development of SRL can have a positive effect on academic achievement, and, most importantly, is amenable to intervention. A conceptual model was proposed, for the evaluation of self-regulated learning and critical thinking, as they are conceptualised for this study.

The rationale for the inclusion of self-regulated learning as a key component in the intervention was based upon the ubiquitous acknowledgement that it is a vital life-skill, and the simultaneous observation that there is a dearth of programmes in England which make specific provision for the development of such skills in young people. A review of potentially effective SRL interventions concluded that the learning journal had the potential to be an effective tool for both the delivery and assessment of self-regulated learning in action.

The review has identified problems that have the potential to be addressed in this study. The principal challenge is the need to develop an intervention which has the potential to be effective in developing the skills of self-regulated learning and critical thinking in young adolescents. More specifically, there is a need to provide for programmes which can be made available outside the classroom walls, and that make use of innovative technology.

This is a study of the development, implementation and evaluation of an intervention which is designed to address some of the problems identified in the literature review. Its design also reflects the Widening Participation aims which it was originally designed to address (a critical thinking programme which also had the potential to develop aspects if self-regulated learning). The main intervention components include a *learning journal*, which it is hypothesised will be an effective means of delivering and assessing self-regulated learning, and *argumentation,* which it is hypothesised will provide a vehicle for the development of both critical thinking and aspects of self-regulated learning. The following chapter details that intervention.

3 Chapter Three: The Intervention

This chapter will describe the intervention experienced by the participants of both the pilot and the main study. It will detail how they accessed both the course components and the pre- and post-test surveys, according to the design specified in Section 4.3. Intervention components were a) the weekly Devil's Advocate (argument and counter-argument), completed by all experimental groups IC2, IC3 and IC4, b) the weekly critical thinking assessments, also taken by all experimental groups IC2, IC3 and IC4, c) the daily debates, available only to experimental groups IC3 and IC4, and d) the weekly learning journal, available only to the experimental group IC4 (main study only). This was an internet-mediated intervention; therefore it has been possible to capture details of how intervention activities were implemented using screenshots.

3.1 Introduction

Adequate teaching is no longer regarded as a process of transferring knowledge from the teacher's into the student's memory, and interventions which focus on the process of learning, rather than the acquisition of exam-orientated knowledge, have the potential to constitute effective interventions (Boekaerts, 1996). In Chapter Two, it has been established that there is not only a demand for such interventions, but that they need not necessarily be restricted to the traditional classroom environment. A number of writers, including Duckworth *et al.,* (2009) and Eyre (2011) remind practitioners of the possibilities that present themselves outside the classroom walls, and call for programmes that deliver self-regulated learning and critical thinking interventions in innovative ways. Recently, studies have begun to explore the potential for internet-based learning platforms to provide a medium for the delivery of such programmes (Joo, Bong and Choi, 2000; Shea and Bidjerano, 2010). Internet-mediated programmes afford the benefits of accessible resources, flexible asynchronous and synchronous communication with a group of learners who model, support and aid each other's learning, with teachers who scaffold self-regulated learning skills, and with access to an innovative set of tools which facilitate 'anytime, anyplace' personalised learning (Mair, 2012, p. 149). Recent studies evaluating the benefits of online learning delivery have been encouraging (e.g., Azevedo, Cromley and Seibert, 2004), and articulate an 'implicit assumption' that computer-based hypermedia programmes can play an important role in the 'enabling environment' for self-regulated learning (Meyer *et al*., 2008, p. 21).

Is this the right room for an argument? was one such programme. This four-week internet-based course for young English adolescents provided the context for the intervention. The aim of the course was to develop a) critical thinking through argumentation activities and b) self-regulated learning through self-reflection in a learning journal. It is suggested that incorporating self-regulated learning strategies into a subject-based intervention in this way encourages the transfer of self-regulated learning strategies from one domain to another (Schunk and Ertmer, 2000; Perels, Dignath and Schmitz, 2009).

Background to the development of the intervention

The intervention had been originally created and delivered by the researcher as a critical thinking 'outreach' activity: from 2007 - 2009 for the UK Government

Department of Children, Schools and Families (DCSF) Young Gifted and Talented (YG&T) initiative; from 2009 - 2012 for the International Gateway for Gifted Youth (IGGY), based at the University of Warwick; and finally from 2010 – 2013, as a Widening Participation (WP) activity, for the University of Oxford. The timeframe for the current study falls within that of the University of Oxford WP outreach provision.

Prior to the commencement of this study, the course, known at the time as, *From Plato to Playstation* (a reference to some of the debating content) had been delivered approximately four times per year for four years (from 2007 - 2011) to more than 1000 UK state-school students aged between eleven and twelve (Year 7). At this stage the course consisted of a number of tutor-led daily debates only. At times the course was delivered as a one-day master-class, at others as a weekend activity and at others as a holiday course.

Following a literature review, and in preparation for the pilot, the intervention was reviewed and the additional components of a counter-argument activity (the Devil's Advocate) and the critical thinking tests were added. The new course, now renamed *Is this the right room for an argument?* was piloted, with 164 participants, in October 2012.

The original outreach course was intentionally created as a critical thinking intervention. In revising the intervention for the current study, and following the pilot, the focus was widened in order to situate critical thinking in the broader context of self-regulated learning. Of relevance here is that self-regulated learning shares considerable theoretical and practical overlap with the construct of critical thinking, since critical thinking is defined in the literature as a metacognitive skill (Phan, 2010, Efklides, 2009), and the intersection of the two constructs includes self-evaluation, metacognitive control of thought processes, capacity for analysis, deduction and inference etc. (Kuhn 1992, 1999, 2000; Olson and Astington, 1993; Heijltjes, van Gog, Leppink and Paas, 2014). Thus, following the pilot, a self-regulated learning component (the learning journal) was added. The final intervention now consisted of four components: daily debate, weekly Devil's Advocate, weekly critical thinking tests and the weekly learning journal. It was finally delivered, for the purposes of this study, to a further 617 participants, in January 2013.

Funding for tutor fees, database licenses, IT support and insurances (detailed in Appendix 3) was provided by the University of Oxford, Green Templeton College Oxford, The Francis Terry Foundation and the researcher's private funds respectively.

Intervention access

Intervention materials were accessed by the students in two ways: a) course activities and assessment materials were presented on a *Moodle* internet-based learning platform and b) the pre/post-test questionnaires and the learning journal were accessed via *Google Docs.*

Moodle is widely chosen by schools, university and government departments for the delivery of virtual learning materials (Lin, 2011). It is an open-source internet-based learning management system, which enables course managers to set up and deliver a set of activities and resources that can be accessed remotely, synchronously and asynchronously. Relevant *Moodle* learning tools include collaborative discussions, collaborative writing activities, quizzes (tests) and uploaded text and video resources. The relevant functions of a learning platform, which influenced the construction of this intervention, are a) that it facilitates and records discussion, b) student participation can be tracked electronically, c) students can be placed in separate groups such that students in one group are not aware of those in other groups, and d) students can create collaborative internet-based resources.

Google Docs is an open-source data collection and storage facility, residing on a *Google* 'cloud' server. The owner of a Google account, in this case the researcher, accesses their *Google Drive*, creates a questionnaire using the *Forms* tool, converts it into HTML code, and embeds the resulting code into a web-page on the Moodle platform. Participants access the questionnaire as an interactive form, complete the form by making choices or by entering text, and submit it when finished. The owner of the Google account is then able to access the resulting data, via a unique username and password. Data are downloaded directly into an Excel spreadsheet, which is formatted in such a way that it includes both questions and responses. Data can then be transferred from Excel into a statistics application.

Intervention components

The principal aim of the intervention was to develop participants' critical thinking skills, using computer supported collaborative learning instructional materials that also gave them the opportunity to self-regulate their learning (Yang, Newby and Bill, 2005; Stahl, Cress, Ludvigsen and Law, 2014; Wecker and Fischer, 2014). Based upon the critical thinking and self-regulated learning research literature explored in Chapter Two, the intervention comprised four elements which have been linked to progress in critical thinking and aspects of self-regulated learning: a) participation in dialogic argumentation (Kuhn and Udell, 2007; Fisher, 2007; Schellens and Valcke 2006), b) written argument and counter-argument (the Devil's Advocate) (Jonassen and Kim, 2010; Preiss, Castillo, Grigorenko and Manzi, 2013; Preiss, Castillo, Flotts and San Martín, 2013), c) self-reflection documented in a learning journal (McCrindle and Christensen, 1995; Barclay, 1996; Cantrell, Fusaro and Dougherty, 2000; Connor-Greene, 2000; Langer, 2002; Nückles, Schwonke, Berthold and Renkl, 2004; Schmitz and Wiese, 2006; Perels, Otto, Landmann, Hertel and Schmitz, 2007; Nückles *et al.,* 2009; Schmidt, Maier and Nückles, 2012; Glogger *et al.,* 2012; Clipa *et al.*, 2012) and d) critical thinking tests (Kurfiss, 1988; Ennis 1993).

More specifically, provision for the development of critical thinking took the form of a) daily weekday debates exploring a variety of moral, social and ethical dilemmas, b) weekly Devil's Advocate tasks that required the learner to form an argument, and a counter-argument, from multiple perspectives of a dilemma and c) weekly graded assessments in critical thinking. Self-regulation was fostered by encouraging participants to complete a weekly learning journal where they had the opportunity to set and monitor goals, plan their learning according to their personal timetable, monitor, evaluate and reflect upon their learning progress and self-report according to evaluation criteria. In engaging with the internet-based material, a student would also have the opportunity to further develop a range of self-regulation skills. For instance they would need to orientate themselves by negotiating through hyperlink instructions, plan their learning schedule around prescribed discussion timing windows, extract information from internet-based resources (not necessarily from the same link), complete tasks, and check and evaluate their learning record. According to Shapiro (2008) and Green, Moos and Azevedo (2011), these activities have the potential to improve self-regulated learning, since the navigation of hyperlinked materials requires an advanced degree of practice in metacognitive control. Course guidelines and messages from tutors (see Appendix 13) assisted the learners through this process.

Critical thinking was specifically modelled and developed by the course tutor, through open-ended questioning, alongside course activities (dialogic argumentation and the Devil's Advocate) which had the potential to have an effect on the development of critical thinking (Cho and Jonassen, 2003; Kuhn and Crowell, 2011). Meta-analysis of critical thinking interventions has determined that this 'mixed' approach can positively affect critical thinking outcomes (Abrami *et al.,* 2008, p. 1106).

Critical thinking skills were assessed at the end of each week, using an adapted version of the Defining Issues Test (Rest *et al.,* 1999). Students were presented with a decision-making scenario and then asked to assess the relevance of supporting arguments. For instance, one week, students were asked to envisage a scenario whereby students should imagine that they were about to vote for their school captain. They were then asked whether statements such as, 'one candidate is taller than the other' or 'one candidate was caught cheating in exams last month,' were 'very relevant', 'of some relevance' or 'not relevant' to the issue. Students completed one such assessment at the end of each week. The measure and mark scheme is discussed in more detail in Section 4.5.1.

Dialogic argumentation took the form of a weekday evening forum discussion on controversial moral issues about which alternative perspectives could be argued to exist (Jonassen and Kim, 2010; Kuhn and Crowell, 2011). This activity, which was not assessed, provided a vehicle for the development of critical thinking, since studies such as that by Kuhn and Crowell (2011) have concluded that argumentation intervention can affect critical thinking outcomes, and the capacity for metacognition (Fisher, 2007). For instance, Kuhn and Crowell (*op cit.*), in an experimental pre- to post-test study of 71 young adolescents, detected that a dialogic argumentation group performed significantly better on critical thinking outcomes, with a medium effect size (d = .68), compared with the non-debate group. Research by Colbert (1987) and Green and Klug (1990) reported that debaters made significantly more gains in their critical thinking scores than non-debaters. Similarly, Sanders, Wiseman and Gass (1994) determined that argumentation intervention enhanced students' capacity for determining the strength of an argument, a key critical thinking skill. In free-flowing dialogic argumentation, the discipline of formulating, developing and defending an argument exercises a range of metacognitive skills, requiring the learner to control and regulate their thinking process, to evaluate their argument and to re-evaluate in the face of counter-argument. According to Kuhn and Crowell (2011), it is 'the umbrella under which all reasoning lies,' with five skills being essential for skilful argumentation (p. 545). These are a) associating theories to causal claims, b) supporting claims with evidence, c) identifying and arguing alternative theories, d) anticipating counter-arguments to one's own theories and e) rebutting alternative theories. These five skills were designed into the current intervention, where, in order to develop their skills of argumentation, students were encouraged by tutors to present claims, identify roles and establish rules of engagement, defend, argue, counter-argue and rebut claims, before concluding an argument. Jonassen and Kim (*op cit.*) suggest that it is important to ensure that the subject for discussion is likely to stimulate strong views. Hence there is also a need to establish firm 'rules of engagement,' especially for novice debaters. Thus, students were given debating guidelines that encouraged them to acknowledge others' points of view, evidence their arguments and behave respectfully towards each other (see Section 3.2).

Students also completed a written argumentation activity, the Devil's Advocate, in which they were tasked to generate two rhetoric arguments per week, arguing and counter-arguing from opposing perspectives of a moral, social or ethical dilemma. Topics were taken from issues current at the time. The 2012 Olympics had recently taken place, so one of the debate issues explored whether or not Paralympians should be thought of as 'brave'. Students were asked to contribute to a collaborative glossary of arguments on the subject. They were asked to use no more than 50 words to put forward their arguments, and to include evidence and logical reasoning to support them. Firstly, they were encouraged to argue their opinion. However, they were then tasked to return to the activity and argue *against* their original perspective. So, for instance, one student argued:

> *Paralympians are showing determination to overcome their disabilities not bravery. The definition of bravery is showing courage, resolution and daring.*

> *Paralympians compensate for their disability by having super strength in the rest of their body, it could be argued they are stronger and fitter than able-bodied Olympians.*

Then the same student counter-argued:

> *I believe that Paralympians are brave to have done what they have done. They are not self - conscious on what they look like but we are, most people worry all the time on what they look like, Paralympians don't have time to worry about things. They don't know about being scared or worried or self - conscious. YES I DO THINK THEY ARE BRAVE.*

Students could return to this task as many times as they wished, in order to see and learn from others' contributions, and to adjust their own submissions if necessary. Thus they were not scored on this activity. Tutor-feedback, in the form of a notice to all students, was posted at the beginning of the week following each Devil's Advocate activity. The tutor thus used the opportunity to give generic formative feedback by highlighting instances of effective argument, sound reasoning, and evidenced points of view.

A learning journal, a standardised electronic document, provided the means for students to scaffold their self-reflection. It was implemented using *Google Docs* internet technology, using the link http://bit.ly/1hivddF. At the end of each week, learners were prompted with a series of statements that encouraged them to explore their metacognitive experiences and their use of metacognitive skills, and practise self-regulation whilst learning course content (Schunk and Ertmer, 2000; Perels, Dignath and Schmitz, 2009). Students were asked to respond, on a 1 - 10 Likert scale, to statements such as ‘I was able to relate this week’s topics to other areas of my interest’, or ‘I tried to consider several ways of looking at this week’s dilemma before forming my argument’. They were also encouraged to set learning goals and to write freely about their progress toward achieving them.

3.2 How the intervention course was delivered

This section offers a series of screenshots from the learning platform, demonstrating the intervention in action.

Firstly, students logged into the learning platform with their unique username and password:

You are logged in as Flossie Higginbottom (Logout)

… and were presented with the course page:

Is this the right room for an argument?

… a critical thinking course, developing your
learning strategies and your thinking skills

Tutor: Jen Oakton, MA
Date: January 2013

Effective communication opens doors. You are the
leaders of tomorrow (yes, you!) - and your words can change the world!

This course aims to do two things:

1. Improve your critical thinking skills
2. Improve your study skills

Periodically throughout the course, notices were posted at the head of the course page:

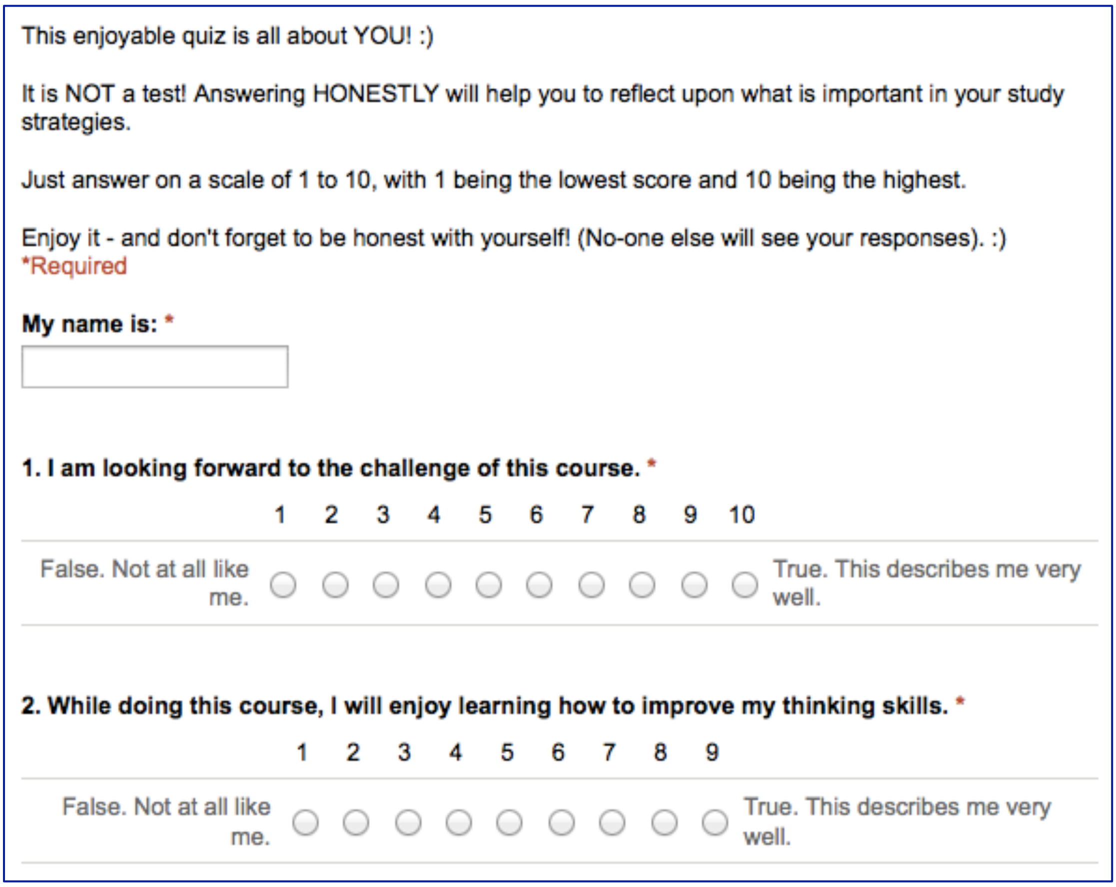
*** NOTICES ***
1. At 8am on the first day of the course, activity units will open up just below this unit. Please work through the course in the order in which it is presented. 2. The activities can be accessed by clicking on the blue links. 3. Red apple avatars are tutors (Jen) and moderators (people who keep you safe on the internet).

3.2.1 Induction activities

Induction activities were made available before the course started, so that, on enrolment, students could familiarise themselves with the course layout and complete the pre-test activities:

** Induction activities **
Please complete these induction activities before the course starts. 1. The Learning Quiz 2. News Forum 3. General guidelines 4. How to debate online 5. Guidelines for conducting research 6. Using quotations in your arguments 7. Contributing to glossaries

The 'Learning Quiz' was the pre-test self-regulated learning questionnaire. The link took the student to a *Google Docs* form, found at http://bit.ly/1nn2FPu:



This enjoyable quiz is all about YOU! :)

It is NOT a test! Answering HONESTLY will help you to reflect upon what is important in your study strategies.

Just answer on a scale of 1 to 10, with 1 being the lowest score and 10 being the highest.

Enjoy it - and don't forget to be honest with yourself! (No-one else will see your responses). :)
*Required

My name is: *

1. I am looking forward to the challenge of this course. *

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.											True. This describes me very well.

2. While doing this course, I will enjoy learning how to improve my thinking skills. *

	1	2	3	4	5	6	7	8	9	
False. Not at all like me.										True. This describes me very well.

A *News Forum* provided a link to one-way notification email, whereby the tutor could email specific groups of participants at once, giving general feedback and reminding students of forthcoming activities. Further examples of tutor feedback can be seen in Appendix 13.

Email to debating groups: Feedback after: This House would regulate the Press.

Today we had another spirited debate. There seemed to be an even split on either side of the House - always a good indicator!

We have developed a habit of starting with some researched evidence, which we then explore. This is great. (Thank you to the Georges, who have taken on the mantle of providing us with this invaluable evidence). **However,** we mustn't rely on that evidence to be 'true'. It is our job to question, probe and interrogate it! Just because someone has read it somewhere on the internet (or in the press), it doesn't mean that we have to believe it. Producing 'evidence' is fair enough - but how are you going to use that evidence to build your argument? That is the skill of argumentation. So, how do we interrogate such evidence?

1. Where did that information come from? Who wrote it? What was their motivation (Political? Emotional?)

2. What are the implications of this evidence? What does it mean? How does this evidence actually support the argument?

3. What is the counter-argument? What would someone say, if they were reasoning against your argument?

Above all - you don't need to have an opinion to start with! You can use the debate to explore both sides of the issue, before making your reasoned conclusion.

We are not necessarily here to reach an agreement - that's 'discussion'. This is 'debate'!

Tomorrow we will develop those skills of combat even further, when you will be able to conduct your own debate. If you would like to take up this challenge, then please create a title which begins, 'This House believes…' Make it controversial, but not offensive. For instance, a debate about 'animal rights' is likely to be controversial. One about 'what you want for Christmas' will produce fluffy bunnyness and plenty of exclamation marks and emoticons. !!!

Well done for today's debate. See you at 6pm tomorrow.

General guidelines gave information about organisational matters:

General Guidelines

General expectations

This is an intellectually challenging course, so we have high expectations of you! Please make sure that your contributions are presented with correct spelling, grammar and punctuation. Take every opportunity to improve your vocabulary and raise the academic level of your writing by making use of a thesaurus. This is a great opportunity to share your knowledge and experience in a supportive community of like-minded students and academics, so do take the opportunity to generate new knowledge, by engaging in discussion with each other. We particularly appreciate students who respond encouragingly to each other. This helps to build a collaborative community in which everyone can feel supported.

Avatars

If you would like to change your little picture (avatar) then you can do so by clicking on your name in the top right hand corner of the screen. However, you MUST ask your parents' permission if you choose to use your own photo, rather than just a picture. Red apple avatars indicate adults. The adults on this course are tutors and moderators.

Assessments

You have four learning quizzes (one in each unit), which will be graded. You can check the progress of your grades at any time by clicking on the menu on the left of the screen. Discussions and collaborative activities are never assessed, because some students prefer to ‘browse’ whilst others participate eagerly.

How to Debate Online gave specific guidelines for discussion:

How to debate online

In this course, we want you to take some responsibility for challenging and developing other students' thinking processes (and they will challenge yours!) So rather than just responding to an idea with your own response, why not try to 'unpack' the thread of an argument, by asking further questions, which challenge others' points of view?

The secret is to invite further comment, once you have expressed your own opinion. We'll be looking for this level of engagement with other students on this course! This keeps the interactions dynamic and demonstrates that you are interested in others' points of view as well as your own.

Have you noticed that, in the House of Commons, MPs refer to each other as 'My Honourable Friend', even if they are diametrically opposite in their opinions? Have you ever wondered why, when it's obvious that actually some of them don't even like each other very much? It's all to do with the conventions of civility - a sort of blanket of formality - which enables MPs to disagree (sometimes violently), without actually resorting to fisticuffs!

The difference between debating and just expressing your opinion is a matter of paying due attention to the opposing argument, whilst still respecting your own. The secret to this is to:

1. demonstrate that you have listened (or read in our case) carefully
2. reflect back your own understanding
3. state your own case, with evidence

So, if you find yourself disagreeing with another person's point of view (and this course encourages you, after all, to do exactly that!) one response might go something like this: 'I can see what you are trying to say' (some people say, 'I hear what you are saying') [demonstrating that you have listened] 'and that you think that global warming is a myth' [reflecting back your understanding of their case], 'but from the evidence I have read, I am led to believe the opposite' [stating your case] etc etc…

So in this course, we are watching out for respectful responses which demonstrate mature empathy. Before you respond from now onwards, check that you are showing us that you have paid attention and reflected back the other person's point of view, before stating your own.

This is a skill which you will find yourself using more and more as you grow older. The Leaders of Tomorrow (that's you, by the way) will find this one of the most useful skills to learn.

Students were expected to do some individual research before contributing to collaborative glossaries. The *Guidelines for Conducting Research* offered advice on this process:

Guidelines for conducting research

Sometimes we ask you to research a topic on your own,
so that you can contribute to a collaborative project.
If you asked to research a topic, where do you start?
It's very easy to type your search terms into Google or
Wikipedia and niftily copy and paste it straight into
your writing.

Well done! What have you proved? You have proved
that you know how to copy and paste - that's all!
What have you learned? Nothing!

Firstly it's very important that you do not believe everything that you read! Like history (which, it is often said, is rewritten by the victors), facts can usually be questioned. Most of us accept the written word as if it is the truth, but we encourage you to question everything, research different theories, contrast points of view and find alternative facts.

So let's say that we want to research the issue of the allocation of charity money in a famine crisis. That sounds quite straightforward, doesn't it? Your first thought might be that the money should be spent on food and given to hungry people. That sounds admirable doesn't it? But is there an alternative point of view? Research a little deeper and you will find sound arguments for using some of that charity money to build the infrastructure to take those starving people out of poverty and to enable them to help themselves, rather than use the money just to buy the food. 'Give a man fish and you feed him for a day. Teach a man to fish and you feed him for a lifetime.' Different charities will have formulated different policies on how they distribute charity money. This decision will have come about as a result of some significant debate.

So when you are researching topic, investigate opposing points of view, search, and build up a comprehensive picture so that you can make a balanced judgement.

So where do you start?

1. Search more than one source of information, e.g., Encyclopaedia, almanacs, newspapers, magazines, ezine articles.

2. Use Internet search terms such as ‘research on…’, ‘current theories on…’, ‘blogs on…’. Don’t just use Wikipedia. Follow links to deepen your understanding. People will have strong views on either side, so there will be plenty of material to investigate.

3. Don’t accept one point of view or theory. Look for alternatives and/or controversial theories. Are they credible? Do they merit discussion? Do all experts agree? Can you find any contrasting viewpoints? If so, which evidence is the most convincing? Search, and then make up your own mind.

Make sure that your work is always written in your own words, that you have demonstrated that you have considered all opinions, and that you have a balanced overview of the topic.

Do give a reference for each of your sources. If you quote directly from the source make sure that you only use a couple of sentences and cite your source of evidence. At the end of your writing you should include a reference list and list the sources of information you have used.

Remember we need to understand your writing in plain English, and (even apart from the issue of plagiarism) cutting and pasting from the internet just won’t to do!

Students were given tips on how to include quotations in their work:

How to add quotations into your arguments

A quotation is a small bit of text copied from the book. It can help you to prove that your ideas are correct.

For example, I might write:

I know that Bob bought sweets in Raj's shop regularly because it says 'Bob was one of his absolute best customers.' (p.68) This suggests that Bob spends lots of money in the shop and it might explain why he is overweight.

The paragraph starts with an idea and the quotation provides **support for this idea being correct.** Notice how the quotation is in quotation marks and that there's a page number in brackets straight after it. This is the number of the page where the quotation was found.

Notice also that there is another sentence after the quotation. This sentence **explains** how the quotation supports the writer's idea.

A practice collaborative activity prepared students for the Devil's Advocate (counter-argument) exercises, requiring them to explore an ethical dilemma by contributing to a collaborative glossary:

> We will be creating glossaries in this course. It is important that you learn how to do this, by contributing to this **practice** collaborative activity:
>
> Is the National Lottery good for society?

> ### Is the National Lottery good for society?
>
> This is an example of a collaborative activity which we will see later in the course. First, let's look at the argument.
>
> A while ago, a couple won £56 million in the Euromillions draw. News of one couple receiving such a staggering amount of money got us thinking and we'd like to hear your views on the topic too. Is the National Lottery valuable because of all the funding it gives to organisations and projects across the country, from sports facilities to museums and historic buildings? Or does it promote an unequal and money obsessed society?
>
> Your task is to write no more than 50 words, arguing either *yes* or *no* to the question, it is the National Lottery good for society?

3.2.2 Course activities

On the first day of the course, the units of activity opened:

> **Welcome!**
>
> We are looking forward to working with you on this course! Individual activities will be available all of the time, whilst the Debating Chamber will open up each weekday evening, from 6pm - 8pm.
>
> Have fun! ☺

3.2.2.1 The Core Activity: Devil's Advocate

As discussed in Section 2.2.4.7, the development of *counter-argument* is considered by many to be the most effective aspect of argumentation, in terms of its capacity to influence the development of critical thinking (Kuhn, 1991; Voss and Means, 1991; Kuhn *et al.*, 1997; Stein and Bernas, 1999; Anderson, Nguyen-Jahiel, McNurlin, Archodidou, Kim and Reznitskaya, 2001; Leitão, 2003). It follows that interventions that seek to develop critical thinking skills should include provision for the development of counter-argument. A number of studies have reported that when students are asked to generate reasons to support their own argument, they tend to provide more evidence and reasons *for*, than *against* their position (Stein and Bernas, 1999). Furthermore Jonassen and Kim (2010) found that the lack of counter-argument was the most common weakness in argumentation, especially for young adolescents. Thus the core intervention activity, *Devil's Advocate*, was designed to encourage participants to articulate *both* argument and counter-argument. Nussbaum and Kardash (2005) found that this was a more effective method of developing reasoning skills than instructions to simply articulate a one-sided argument alone.

At the beginning of each week, the Devil's Advocate exercise opened up, and stayed open for the rest of the course. The exercise was delivered using the *glossary* tool on the Moodle learning platform. In a Moodle glossary, participants are given the opportunity to add a component (a glossary entry) to a collaborative document. Thus the completed glossary document is a collection of short entries, available for all participants to see (and thus to learn from). In this case, each entry had a limit of 50 words, and students were asked to entitle their entry either *for* or *against*. Students were charged to make exactly two entries, one *for* and one *against*, attempting to make both arguments as strong as possible, regardless of their actual opinion.

It is suggested that unresolved problems of a social, moral or ethical nature - those that ultimately require a decision - provide robust topics for a counter-argument exercise (Jonassen and Kim, 2010). In Week One for instance, students were asked to consider the issue of badger culling, the controversial practice of reducing the badger population as a proposed means of controlling the spread of bovine tuberculosis:

> Developing your powers of argument
>
> The core learning activity this week (which is NOT graded) encourages you to develop your powers of argument by asking you to argue both FOR and AGAINST each side of an argument.
>
> You should write no more than 50 words, but those words need to make your argument clear, and they need to include EVIDENCE to back up your opinion. This is the issue under consideration:
>
> Badger culling: a necessary evil?

Badger culling - a necessary evil?

‘A target number of badgers to be killed in this year's pilot cull has been announced.

Authorisation letters for the cull have been issued by Natural England. A minimum of 615 in Gloucestershire and 316 in Somerset need to be killed.

In 2013, 921 badgers were killed in Gloucestershire and 940 in Somerset in a bid to tackle TB in cattle’.

Read the full article here: http://www.bbc.co.uk/news/uk-england-28942468

So what’s the evidence? Are they cute and harmless - or insidious killers? Can you argue both FOR and AGAINST?

It is likely that you will have strong opinions on the matter - one way or another - so please be respectful in this debate, no matter how strongly you feel.

Use EVIDENCE to back your opinions.

You have no more than 50 words.

Please click on ‘Add a new entry’ below.

Type FOR or AGAINST in ‘title’, and add your 50 words.

Don’t forget to click ‘SAVE’ each time you make an entry.

Students would add their own contribution for all other students to see:

FOR

I think it is necessary to introduce badger culls because their numbers have grown rapidly over the past few years. We have deer culls, so why not badger culls? Also, we would save lots more cows, as badgers are known to spread cattle tuberculosis.

Following the glossary activity, students were invited to vote:

Badger culling is a necessary evil?

I agree | I disagree | I don't have an opinion

Save my choice

3.2.2.2 Dialogic argumentation

As discussed in Chapter Two, studies have provided evidence of the potential for participation in dialogic argumentation to influence the development of metacognitive processes such as critical thinking (e.g., Colbert, 1987; Green and Klug, 1990; Kuhn, 2005; Fisher, 2007; Darabi, Liang, Suryavanshi and Yurekli, 2013). As a consequence, this intervention included a dialogic argument component. This activity took place in the 'Debating Chamber', which became visible from 6pm - 8pm each weekday evening. Students were first reminded of the previously seen debating guidelines:

> **The Debating Chamber**
>
> This is your opportunity to engage in intellectual
> discussion, with other course members.
>
> Debates will be open from 6pm to 8pm.
>
> Please read the induction activity How to debate online.
> There are also some general guidelines for you to read first:
>
> When you write a message it is important that you communicate as clearly as possible and avoid 'text-speak' (abbreviations like 'u' for 'you'). This will make it easier for us all to understand one another and write responses. Try to write conversational messages and create opportunities for other people to reply to you. For instance, when you post an idea, you could say 'what do you think?' By asking questions, you encourage others to respond to your post and get involved. Be as encouraging and supportive as possible. Remember! Some people may be nervous about sharing their ideas.

Particular attention, within the guidelines, was paid to the need for students to behave respectfully towards each other:

> **Safety Rules (Please read these carefully)**
>
> This is a closed learning platform and only those who have registered on the course have access to it. However, one of our aims is to help you to adopt safe internet practice whilst you are working with us. Therefore we have some rules which we would like you to take seriously, not only whilst you are with us, but when you venture onto the internet at other times too:
>
> Do not include any personal information when you are engaged in discussion, such as your surname, your home address, email address or phone number. If you do, your writing will be deleted or edited by the moderator.
>
> No bad language or bullying. It is fine to disagree with someone, but please do so respectfully. We will teach you how to debate effectively during this course. Please go NOW to this website and make sure that you understand how to stay safe online: www.thinkuknow.co.uk

Quality assurance of discussion texts, reported in Appendix 10, established that students had paid attention to the guidelines: there were few incidents of text speak, few incidents of off-topic discussion, and no unacceptable behaviour.

At 6pm each evening the discussions would become visible. Some discussions had already been set up by the tutor:

> **The Trolley Problem**
>
> Today we explore a 'thought experiment'.
>
> It asks you to consider a difficult situation, called
> the Trolley Problem, and is attributed to a philosopher
> called Phillippa Foot.
>
> A trolley (an unmanned railway carriage) is out of control and is hurtling down a track towards five people who have been tied to the rails and who will certainly die if it runs over them. You find yourself standing at a junction and realise that you can divert the train down another track. Unfortunately there is one person tied to the rails, down that track.
>
> What is your response to this dilemma? (Please note - we are not asking you, 'what is the solution?')

Other tutor-led debates included:

- This House believes that the internet is the greatest force for democracy in the world.
- This House would return cultural property, residing in museums (e.g., The Elgin Marbles) to its place of origin.
- This House would restrict advertising to children.
- This House would regulate the press.
- This House would ban cosmetic surgery.

The tutor who delivered the course (not the researcher) was an experienced teacher who had run internet-based learning courses for more than five years. She spread her time equally between discussions, stimulating argument by asking open-ended questions such as, 'there are some interesting ideas here, but where is the evidence? Are there any alternative perspectives here?' She would model critical thinking by challenging assumptions, highlighting the development of logical evidenced argument, offering counter-ideas, modelling thinking processes by ‘thinking aloud’, acknowledging good reasoning and demonstrating the vocabulary of reasoned argument (Kuhn, 1991). During the discussions, the researcher, who was not visible to the participants, observed all activities and moderated discussions by monitoring them for inappropriate behaviour (such as bullying), or for possible breaches of security, (such as students revealing their surnames or email addresses). This was important since learning via the internet brings with it a number of issues of child-protection (Salmon, 2000). No such breaches occurred however.

Students were also encouraged to set up and lead their own discussions. Before doing this they were encouraged to participate in an ungraded quiz designed to help them to identify the difference between intellectual discussion and chit-chat:

> This little quiz will help you to differentiate intellectual discussion from chit-chat. Please have a go at the quiz, BEFORE you enter the Debating Chamber for the first time.
>
> Ethical Dilemmas or General Topics for Discussion?

The quiz asked students to consider whether a particular issue was one of ethical importance or not:

True (this question/statement discusses an ethical issue)

OR

False (this question/statement does NOT discuss an ethical issue)

'Is breakfast the most important meal of the day?'

Select one:
- True
- False

Check

True (this question/statement discusses an ethical issue)

OR

False (this question/statement does NOT discuss an ethical issue)

'Does celebrity influence pose a danger for vulnerable teenagers?'

Select one:
- True
- False

Check

Having completed the quiz, they were encouraged to start their own discussions:

Now that you have participated in some of the tutor-led discussions, why not set up some ethical discussions of your own?

Your own discussions

Student-initiated discussions included topics such as:

- Does reading affect your popularity?
- Is 'Chance' a good thing?
- Do you believe that listening to Mozart helps your academic performance? If so how?
- Kindle or books?

- Is football a good or a bad thing?
- Can you be smart and cool?
- Is obsession for money bad?
- Should we have to do exams?
- What do you think about farmers killing animals that do not belong to them?
- Should we execute criminals?
- Prison Age?

Following each debate, the tutor sent an email to all debate participants, with feedback:

Feedback after_ 'This House would return cultural property residing in museums (e.g., the Elgin Marbles) to its place of origin.'

The debate was really effective today. Why? Because the debate fluctuated across opinions - because there was clearly no clear answer to the dilemma - because debaters supported their arguments with evidence (and sources) - because debaters asked as many questions as they answered - because you all did your research before the debate began and didn't just rely on your own opinion - because participants hardly went off topic (and when they did, they were swiftly brought back into line by other debaters).

Thank you to everyone who participated, for providing us with an interesting intellectual exercise.

How can we improve tomorrow? It would be good to see better attention to the conventions of punctuation. We all make mistakes in typing, but attention to capital letters and full stops is a way of demonstrating respect for other readers.

Looking forward to tomorrow!

PS: Just to remind you that the moderators will ruthlessly remove all instances of coded messages or posts which contain any hint of unpleasantness. This is a SAFE place for students to debate - without any potential for them to be made to feel uncomfortable.

3.2.2.3 Critical thinking assessments

Each Friday, all students in experimental groups (IC4, IC3 and IC2) were required to complete a critical thinking assessment:

> **Developing your critical thinking**
>
> There are four graded critical thinking quizzes, one of which will appear each Friday.
>
> Just click on the link:
>
> Reasoning 1

The relevant links became visible at the appropriate times. Clicking on the test link enabled the students to see the instructions for the assessment:

> **Reasoning (4)**
>
> Imagine that you are about to vote for the school captain.
>
> Which of these issues are relevant to your decision about whom to vote for?
>
> (Please click on Continue below):
>
> Continue

The test itself required students to consider whether ten statements were ‘very relevant’, ‘quite relevant’ or ‘not relevant’ to the issue under consideration. Two examples are pictured below:

One of the candidates is taller than the other.

Select one:

☐ This is VERY relevant to the issue.

☐ This has SOME relevance to the issue.

☐ This is NOT relevant to the issue.

One candidate has won the Bronze Duke of Edinburgh’s Award three years ago.

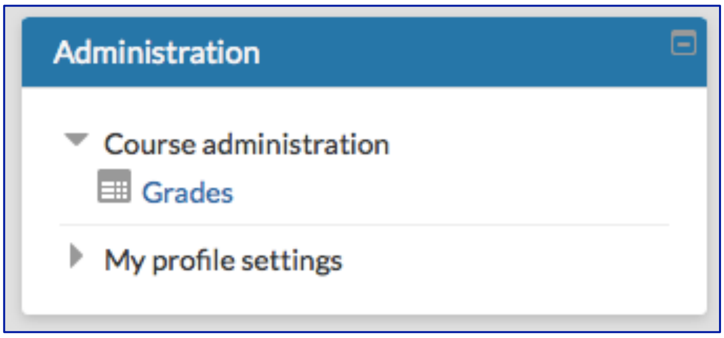
Select one:

☐ This is VERY relevant to the issue.

☐ This has SOME relevance to the issue.

☐ This is NOT relevant to the issue.

Following the assessment, students could check their grades by clicking on a link on the course page:



3.2.2.4 The learning journal

This activity was only available in the main study, and not in the pilot. At the end of each week, students in certain groups (IC4) were encouraged to make use of a learning journal:

Your learning journal
The learning journal gives you an opportunity to reflect upon your learning experience and to set, monitor and evaluate your learning goals. It is available from Friday until Sunday.
Just click on the link below.
My Learning Journal

This was accessed via a link that took them to a Google Docs form, the live version of which can be accessed at http://bit.ly/1lB9kG0 :

The Learning Journal

This journal is part of your coursework. It is a self-reflection exercise which will help you to develop effective learning strategies throughout this course and beyond.

It will be presented to you at the end of each unit of the course. It is not a test and you are not graded on it.

Answering honestly will help you to think about the way you learn.
*Required

What is your name? *

Please consider the following statements and decide how well they describe your learning experience.

5: I tried to consider several ways of looking at the dilemma before forming my argument. *

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.	○	○	○	○	○	○	○	○	○	○	True. This describes me very well.

8. I achieved my learning goals this week.

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.	○	○	○	○	○	○	○	○	○	○	True. This describes me very well.

The following is an example of a student response to the prompt, ‘Looking at your responses above, write down ONE learning strategy that you will focus on, as your goal for next week’.

> *‘Today I set myself a learning goal to participate in a discussion and put some points forward to contribute the discussion. I feel that I achieved that but would have liked to work for a little longer. Luckily I did not get distracted by email as I did it on a device that is not email enabled. I will be discussing topics today with my parents as I found some of the points raised very interesting and needing of more discussion I was able to relate to the topics today as I am an avid reader and like the challenge as a whole. I was not entirely satisfied with my learning today as I felt I should have made more time to do the tasks specified so I will change my timetable tomorrow to allow for that. I feel that I was a valuable contributed to the discussion as I asked questions to other students to get further answers and made some points that were then discussed further. The thing I most enjoyed today was the mental challenge involved, and the prospect of online debates was particularly exciting. I think I have tried my hardest but have a few things to improve on’.*

Finally, on the last day, students were given the opportunity to say goodbye to their course-mates and to complete the post-test questionnaire on their metacognitive knowledge, motivational beliefs and efficacy for learning:

Bye!

Time to go!

It's so sad to come to the end of a course.

But we hope that you have enjoyed yourself

and will look forward to the next challenge!

Please complete the course with the final self-evaluation activities:

Bye for now!

Learning Quiz (2)

In acknowledgement of their participation, students were able to download a Record of Participation, which detailed their course activities:

Is this the right room for an argument?

An introduction to critical thinking, reasoning and argument.

Record of participation for:

Course title: *Is this the right room for an argument?* A course in critical thinking, reasoning and argument.
Course dates: 1st - 28th January 2013
Course developer: Coral Milburn-Curtis MSc, Department of Education, University of Oxford
Course Tutor: Jen Oakton MA

Course Content

Critical thinking tasks:

1. Mustaq Singh and his family in India
2. The Oscar Pistorius 'blade runner' controversy
3. Mrs Bennett and the doctor
4. Voting for the school captain

Devils' Advocate tasks:

1. Badger culling - a necessary evil?
2. Are parents more important than schools?
3. Paralympians - a question of bravery?
4. Is homework a waste of your time?

Examples of debating topics in the Debating Chamber:

1. This House believes that the internet is the greatest force for democracy in the modern world.
2. The House believes that good will triumph over evil.
3. Students' own discussions

Self - reflection activity:
The weekly learning journals

3.3 Intervention treatment fidelity

Treatment fidelity (otherwise known as *treatment integrity*) is generally approached from three angles: treatment adherence, practitioner competence and treatment differentiation (Perepletchikova, Tresat and Kazdin, 2007). *Treatment adherence* refers to the extent to which the intervention was executed as planned; *practitioner competence* involves the level of skill, judgements and qualifications of the practitioner who is delivering the treatment; *treatment differentiation,* in the case of this study, refers to the extent to which group treatments were faithfully differentiated according to the design.

Treatment adherence

Treatment adherence is important firstly since if a treatment was not implemented as planned, then it would not be possible to identify accurately the various intervention effects. Such failure would result in a threat to internal validity. Secondly, it would be difficult to replicate the intervention, thus threatening external validity. For this reason, it is important to establish and describe the independent and dependent variables, in order to avoid the possibility of unknown and confounding factors influencing the results. Thirdly, if the treatment is not implemented as planned, there runs the risk of a threat to statistical conclusion validity, the potential for unsystematic error to creep into the data, thus reducing potential effect size and statistical power, with the possibility of a significant effect size being missed (Sanetti, Dobey and Gallucci, 2014).

Treatment adherence was optimised in this study by creating and delivering the intervention via an internet-mediated learning platform. This ensured that all participants were exposed to the same material in a standardised format. The electronic intervention, created and delivered using *Moodle* open-source learning software, could be exported and used for other learning programmes, replicating this one faithfully. All procedures, tasks, instructions and activities were included in electronic format, and therefore a manual was not necessary. Activities were timed to open and close at pre-set times. For instance, discussions always took place from 6pm until 8pm each evening. The discussions would only be visible to the participants during those times and would disappear at other times. Guidelines were also presented electronically and participants were reminded of their contents at critical points during the intervention. For instance, before discussions, they were reminded to read the discussion guidelines.

Dependent and independent variables were described clearly and unambiguously in each analysis in the study.

Practitioner competence

Practitioner competence is important in order to ensure adequate delivery of the intervention. This can be controlled and monitored by prescribing the training, procedures and evaluation of the tutors who deliver the intervention. In the case of this study, the practitioner was one single tutor, whose only input into the intervention was to lead the discussions. Therefore it was necessary to standardise the delivery of discussions in order to ensure that no participant was favoured or disadvantaged because of lack of tutor competence. All other aspects of the intervention were created by the researcher and delivered electronically (and therefore standardised for all participants) via the learning platform.

The tutor was an experienced and well-qualified teacher - a Masters graduate in English literature (University of Warwick), with a Post Graduate Certificate in Education (the UK teaching qualification). She had written learning and assessment programmes for the National Foundation for Educational Research and had created and delivered critical reading and critical thinking courses and workshops for at least ten years. She had delivered the course ‘Is this the right room for an argument’ five times prior to the delivery of the intervention featured in the current study.

The researcher (also the creator of the intervention) is a former school principal, Office for Standards in Education (OFSTED) inspector, Headmaster and Mistresses Conference (HMC) inspector and Department of Education panel leader. She has created and delivered critical thinking and critical reading programmes for the University of Oxford outreach programme, the International Gateway for Gifted Youth at the University of Warwick and the national Young Gifted and Talented initiative. She has also delivered the course ‘Is this the right room for an argument’, in collaboration with the tutor, five times previously.

The tutor responsible for the intervention was re-trained in the use open-ended and non-leading questions, in order to ensure that intervention questioning was a) standardised and b) of a high standard. This training had already taken place prior to a similar intervention (Milburn-Curtis, 2011), and then again before the October 2012 pilot intervention. Therefore the training was a refresher, and was focused on the use of open-ended questions. The tutor and the researcher worked together for half a day, exploring the nature of open-ended questioning by rating a set of discussion posts sampled from the learning platform. These transcripts (Appendix 11) were not part of this research study and were used solely for the purpose of coder training. It was agreed that higher-order questions were those that were likely to elicit higher-order thinking in the students. These would be typically phrased in the order of, 'How do we know that this is the cause?' (Kuhn, 1992). Lower order questions would be likely to elicit factual, or low-level understanding responses in the students. Bloom's Taxonomy (Figure 3-1) was used as a guide in the analysis of the posts (Bloom, Engelhart, Furst, Hill and Krathwohl, 1956).

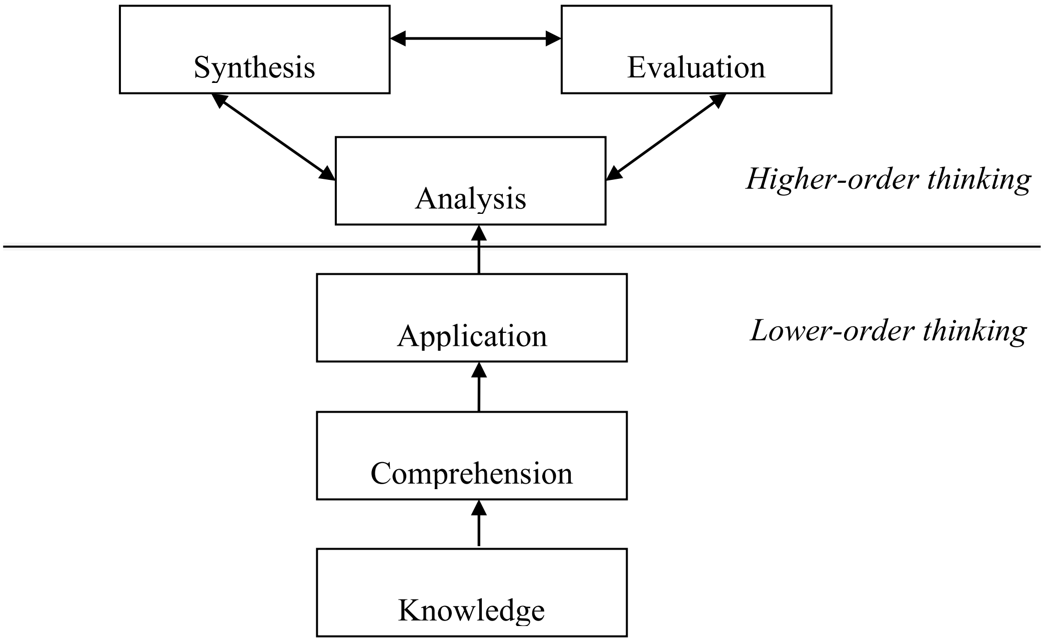


Figure 3-1 Bloom's Taxonomy of Educational Objectives

The two raters each independently analysed 70 sample posts and categorised them into higher-order (coded 1) or lower-order (coded 0) questions. At the end of 60 minutes, the results were compared for inter-rater reliability. Two tests that are commonly used to ensure the reliability of the coding activity are Holsti's (1969) coefficient of reliability and Cohen's *kappa* (1960). According to Wimmer and Dominick (2006), Holsti's coefficient of reliability represents an acceptable level of concurrence if it is greater than 90%, whilst the *kappa* is acceptable above 0.75. However, Rourke, Anderson, Garrison and Archer (2001) uncover some disagreement regarding the acceptable level of reliability of the *kappa*, reporting suggested *kappa* 'acceptability' scores ranging between 0.67 (McDonald, 1998) and 0.96 (Hillman, 1999). Thus, for the current study, Holsti's coefficient of reliability (CR)[8] was the formula of choice, and the minimum percentage of reliability was set, for the current purpose, at 90%. The inter-rater reliability tutor exercise produced an initial agreement rate of 94%, which rose to 100% after collaborative discussion between the two raters.

The tutor worked each evening from 6pm - 8pm, leading all discussions and sharing her time equally between the two debating groups. During this time, the researcher monitored discussions whilst they took place, but did not participate.

Tutor posts were analysed for quality assurance after the course by the researcher (Appendix 12). 5% of them were organisational (e.g., 'This discussion has branched in more than one direction. The discussion about favourite books might merit its own

---

[8] CR provides a percentage agreement between two coders and is calculated using the formula: $CR = 2m / (n_1 + n_2)$, such that m represents the number of coding decisions upon which the coders agree; n_1 represents the number of coding decisions made by one coder and n_2 the number of coding decisions made by the other.

discussion thread perhaps?'). 22% of them reinforced previous posts (e.g., 'Evidence, evidence, evidence please. Perhaps someone can research what Stephen Hawking has to say on the matter?'). 5% were lower order questions (e.g., 'Have national trends changed regarding the average age of first time parents? Can you find out?'). 72% were higher-order questions, that is they were likely to elicit higher-order thinking (e.g., 'Why do you think it was morally wrong? Is it not the moral thing to do to save the maximum number of lives? Or is there more to it than that?') Analysis revealed that the tutor asked significantly more higher-order questions than lower order questions ($\chi^2(1) = 56.823$, $p < .001$).

Treatment differentiation

The study design featured four participant groups, who received different combinations of the intervention activities. In order to ensure that the treatments were delivered to the right groups, in the intended format, and at the appropriate times, the intervention activities were set up electronically and delivered automatically according to a pre-set timetable. Importantly, it was possible to have different groups working on the same intervention activity at the same time, but without the possibility of interaction between groups. All activity was recorded electronically.

3.4 Summary

This chapter has detailed the intervention, explaining how the electronic material was accessed and has demonstrated how participants made use of the learning resources provided. Issues of treatment fidelity have also been addressed. The following chapter details the methodology.

4 Chapter Four: The Methodology

4.1 Introduction

This chapter defines the study operationally and details how the research questions have been addressed. The first objective of this study has been to report on the design and implementation of a critical thinking intervention. This has been reported in Chapter Three. A second objective is to evaluate the effectiveness of the intervention, in terms of its capacity to influence a) change in participants' critical thinking and b) change in their self-reported self-regulated learning characteristics - and to identify the predictors of any such change. The third objective is to use intervention data to explore the process of self-regulated learning (including critical thinking) in real time. This chapter describes the methodology for the analysis which attempts to fulfil the second and third of those objectives.

The methodology that guides analyses in this study is approached from a quantitative standpoint, whereby through the creation of experimental conditions, it is possible to test and verify underlying hypotheses within the field of investigation. In so doing, the researcher attempts to identify and relate variables of interest - those which can be measured, observed and analysed statistically. Hypotheses are specified and data are collected which can support or refute those hypotheses. According to Creswell (2007), the choice for this approach is influenced by the study aim, in the context of the field it aims to serve, and the audience for the research. Thus, the principal justification for this experimental approach is the current state of research studying self-regulated learning and critical thinking. For instance, research in self-regulated learning has burgeoned over recent decades and has matured from its origins as an exploration of a very broad and multi-faceted phenomenon, to a search for clarity of theory, and a thirst for explanation and meaning (Kaplan, 2008; Dinsmore, Alexander and Loughlin, 2008; Efklides, 2011). Within this research context, the study aims to contribute to the field by testing hypotheses that seek to better understand a) the extent to which the elements of this intervention can effect change in self-regulated learning and critical thinking, and b) the process relationship between metacognitive skills, including critical thinking, and metacognitive experiences.

Research questions attempt to evaluate the intervention from two main perspectives. Firstly, they explore main and differentiated effects of the components of the intervention (the learning journal, dialogic argumentation and the Devil's Advocate) on change in critical thinking and self-regulated learning. Secondly they use intervention data in order to explore the relationship between aspects of metacognitive activity and critical thinking over the course of the intervention. Thus, RQ1 and RQ2 investigate the effects of the intervention on change in critical thinking. RQ3, RQ4 and RQ5 explore intervention effects on change in aspects of participants' self-regulated learning. RQ6 is concerned with the exploration of aspects of metacognitive activity recorded in learning journals over the course of the intervention. RQ7 revisits the conceptual model, described in Section 2.3.6, and explores the extent to which the data are represented by the model.

4.2 Research questions and hypotheses

Intervention effects on critical thinking outcomes

RQ1*: What are the main and experimental condition effects of the intervention on change over time in participants' critical thinking?* It is hypothesised that there will be a main effect of the intervention in terms of students' progress in critical thinking across four equally spaced critical thinking assessments. It is further hypothesised that the main effect will be differentiated by intervention condition, such that the most effective experimental condition is predicted to be IC4 (critical thinking tests + Devil's Advocate + dialogic argument + learning journal), followed by IC3 (critical thinking tests + Devil's Advocate + dialogic argument), followed by IC2 (critical thinking tests + Devil's Advocate).

RQ2*: To what extent does participation in dialogic argumentation predict progress in critical thinking over time?* It is hypothesised that the extent to which participants engage in dialogic argumentation will predict change in their critical thinking over time.

Intervention effects on aspects of self-regulated learning

RQ3*: What is the main effect of the intervention on change in aspects of participants' self-regulated learning (metacognitive knowledge, motivational beliefs and efficacy for learning) from pre-test to post-test, on groups of young adolescents, after one month, compared to a group who had no intervention experience?* It is hypothesised that there will be a main effect of the intervention on change in a) overall self-regulated learning, b) metacognitive knowledge, c) motivational beliefs and d) efficacy for learning, when comparing the experimental groups (IC4, IC3 and IC2) with the control group.

RQ4: *What are the effects of the separate experimental conditions on change in self-regulated learning and its components over time?* It is hypothesised that there will be a differentiated effect of the intervention conditions when comparing each of the experimental conditions with the control condition, with the most effective experimental condition predicted to be IC4, followed by IC3, followed by IC2. The outcome will be measured in terms of change in a) overall self-regulated learning, b) metacognitive knowledge, c) motivational beliefs and d) efficacy for learning.

RQ5: *To what extent does participation in dialogic argumentation predict change in self-regulated learning and its components over time?* It is hypothesised that the extent to which a participant engaged in dialogic argumentation will predict gains in self-regulated learning over time, from pre-test to post-test. Outcomes will be measured in terms of change in a) overall self-regulated learning, b) metacognitive knowledge, c) motivational beliefs and d) efficacy for learning.

Relationships between aspects of metacognitive activity and critical thinking over time

RQ6: *What are the relationships between aspects of metacognitive activity and critical thinking over time?* It is hypothesised that critical thinking will be related to the use of elaboration strategies over time, whilst volitional strategies (inhibitory control) will be related to metacognitive experiences over time.

Relationships between self-regulated learning dispositions and self-regulated learning processes

***RQ**7: What is the relationship between self-regulated learning dispositions (metacognitive knowledge, motivational beliefs and efficacy for learning) and self-regulated learning processes (metacognitive activity) during this intervention?* It is hypothesised that pre-test SRL dispositions will predict SRL processes (metacognitive activity) during the intervention, and that SRL processes will predict post-test SRL dispositions.

4.3 Design

This was an experimental design aiming to evaluate the effect of an internet-based argumentation intervention in terms of its capacity to affect change in participants' self-regulated learning and critical thinking over time, and to estimate the predictors of any such change. The study also aimed to investigate the process-relationship between participants' self-reported metacognitive skills (control of elaboration strategies and inhibitory control), observed metacognitive skills (critical thinking) and their self-reported metacognitive experiences. Thus there were two design components: a) a pre-to-post test experimental design and b) a repeated measures time-series design, estimating relationships between aspects of metacognitive activity across four time points.

The pre-to-post design (Table 4-1) was a four (experimental condition) between-person factor x two (time) within-person factor design, such that participants were randomly assigned to one of four groups - three experimental and one control. There was one principal independent variable (the intervention) with four conditions (the groups).

All participants were invited to take the pre-test self-regulated learning questionnaire before the course started (Appendix 1). The purpose of the pre/post-test procedure was to isolate and quantify changes in self-regulated learning that might be associated with the intervention. A different treatment condition was attached to each of the groups such that each group experienced a different combination of the principal conditions IC4 (four components: learning journal + dialogic argumentation + Devil's

Advocate + critical thinking tests), IC3 (three components: dialogic argumentation + Devil’s Advocate + critical thinking tests), IC2 (two components: Devil’s Advocate + critical thinking tests) and CG (no intervention activity), and planned contrasts isolated the effects of the separate conditions (Gorard, 2001).

Table 4-1 Experimental design

Intervention conditions	Pre-test self-regulated learning	**Week 1**	**Week 2**	**Week 3**	**Week 4**	Post-test self-regulated learning
IC4		A, B, C, D	A, B, C, D	A, B, C, D	A, B, C, D	
IC3		B, C, D	B, C, D	B, C, D	B, C, D	
IC2		C, D	C, D	C, D	C, D	
CG		-	-	-	-	
Intervention components: A: Learning Journal; B: Dialogic argumentation; C: Devil’s Advocate; D: critical thinking assessment						

All groups, including the control group, undertook the pre-test and post-test self-regulated learning questionnaire, with negligible levels of item missingness in the data (see Section 4.7.1). All experimental groups experienced the core Devil’s Advocate activity (Section 3.2.1). Here, students were required to contribute two pieces of written argument every week to a collaborative document. One of their pieces was required to argue *for* a proposition and the other *against*. All experimental groups also took the weekly critical thinking tests (Section 3.2.2.3). Groups IC4 and IC3 were additionally given the opportunity to participate in dialogic argumentation (the evening debates) (Section 3.2.2.2). IC4 was additionally asked to complete a weekly learning journal (Section 3.2.2.4).

Various designs have been employed in the implementation of self-regulated learning intervention research, with the most common making use of a quasi-experimental approach, which assigns different conditions to different groups, with no control group (e.g., Magno. 2010). The dearth of randomised control-group studies in self-regulated learning research, highlighted by Winters, Greene and Costich (2008), reflects the ethical dilemma that emerges when one group has the benefit of an intervention, whilst another does not (Bryman, 2004). However, the 'cross-over' design, which features a control group experiencing the intervention after the main data have been collected, enables the researcher to implement an experimental design, without disadvantaging any group academically. This design is typically used in clinical experiments where it would be unethical for only one group to receive a potentially beneficial treatment (Piantadosi, 2005). Criticisms of the cross-over design tend to be confined to its use in medical research (Millar, 1983), and commonly cited issues such as placebo effects (*ibid*, p. 870) do not apply in this educational setting. This crossover aspect was chosen therefore in order to provide experimental conditions which enabled comparisons between experimental and control groups, without disadvantaging the control group. Thus those elements of the intervention that had been missing during the experimental period were made available after the post-test data had been gathered. Data from the crossover element were not used in the study, as the crossover element was not administered within a similar time frame.

The second design component was a series of weekly repeated measures of metacognitive activity, observing critical thinking achievement and metacognitive experiences and skills, captured in an electronically delivered learning journal, over four weeks. This is, according to Schmitz (2006), Schmitz and Wiese (2006) and Schmitz and Perels (2009), an effective method for the collection of repeated measures diary data. The students in one experimental group only (IC4) were enrolled into the learning journal, (thus also providing an opportunity for comparison between learning journal and non-learning journal groups). Each student in this group was asked to complete a standardised learning journal, using Likert-style items, at the end of each learning event (i.e., one week). The students completed each journal after all of the week's activities were over. This element of the design enabled two types of analysis: firstly it allowed for an analysis of trends over time in the different components of metacognitive activity recorded. Secondly it provided data for an autoregressive cross-lagged panel analysis of the relationships between the metacognitive activity variables over time (critical thinking, elaboration strategies, inhibitory control and metacognitive experiences).

Critical thinking tests were administered to all three experimental groups (IC4, IC3 and IC2), but not to the control group (CG). Whilst it is acknowledged that this may precipitate a limitation to the interpretation of the critical thinking results, it was designed as such because it was anticipated that there would be a method effect of taking the tests, such that the process of taking it would exercise critical thinking skills to the extent that the instrument itself would 'contribute variance to the scores beyond what [was] attributable to the construct of interest' (Sechrest, Davis, Stickle and McKnight, 2000, p. 64). This has previously been found to be the case, when students have taken critical thinking tests, which subsequently were found to be instrumental in raising critical thinking skills (e.g., James, Hughes and Cappa, 2010).

4.4 Population and Sample

The population comprised approximately the 10% of those 600,000+ Year 7 (11 and 12 year old) students who, in January 2013, were studying in state schools in England, and who had already reached Level 5 Literacy (DirectGov, 2013). This achievement benchmark was established in order to ensure that participants had a) the literacy ability to access the material used in the intervention and b) the capacity to comprehend questionnaire items, which, according to Hopfenbeck and Maul (2011), is dependent upon achievement levels.

This particular population was also chosen for a number of reasons. Higgins, Baumfield, Lin, Moseley, Butterworth, Downey, Gregson, Oberski, Rockett and Thacker (2004), in a report for the Institute of Education, recommended that, 'when introducing interventions that focus on improving specific cognitive strategies, it could be more efficient to target particular groups of pupils'. It was hence chosen because Year 7 students do not typically experience, as part of their National Curriculum provision, either direct instruction in critical thinking or programmes which develop argumentation skills (Coffin and O'Halloran, 2009; Duckworth *et al.,* 2009). Furthermore, it is believed that this is not only the age where metacognition (and in particular the metacognitive skill of critical thinking) begins to develop rapidly (Inhelder and Piaget, 1958; Piaget, 1964; Flavell, 1979; Kuhn, 1991, 1992), but this is also the 'optimal period' where argumentation interventions can impact upon achievement (Kuhn and Crowell, 2011, p. 545). It is also argued that it is worthwhile studying this English subgroup as it is rarely studied in the context of self-regulated learning research, and even more rarely in the context of internet-based dialogic argumentation. Furthermore, it is suggested that the relationship between learning strategies in the school setting is very different to that of the tertiary sector, where the majority of research in self-regulated learning, argumentation and critical thinking has taken place (Yip, 2013).

It is noted here that, as Gorard (2001) insists, an interpretation of findings should therefore be restricted to a specific population - in this case young English adolescents of above average literacy ability. The 'sampling frame' (*ibid,* p. 58) is further restricted to those individuals who attend schools who had volunteered to participate in the study. This is an important limitation of which to take account, for those students who did not attend such schools did not have the opportunity to be included in the sample.

The sampling process is described in Figure 4-1. The sample was taken from 339 state schools from different areas of England. Schools had responded to emails sent to the headteachers of 600 schools, comprising 150 randomly sampled from each of the North East, North West, Midlands and the South. Each school was asked to offer the intervention opportunity to two or (at the most) three students who met the achievement standard. Schools were asked to send out to parents of prospective participants the prepared letter provided for them (Appendix 4). In the letter, parents were directed to the Moodle web-based learning platform where they found a link where they could a) enrol their child on the course and b) sign an electronic consent form if they agreed to their child's course activity being included in the research study.

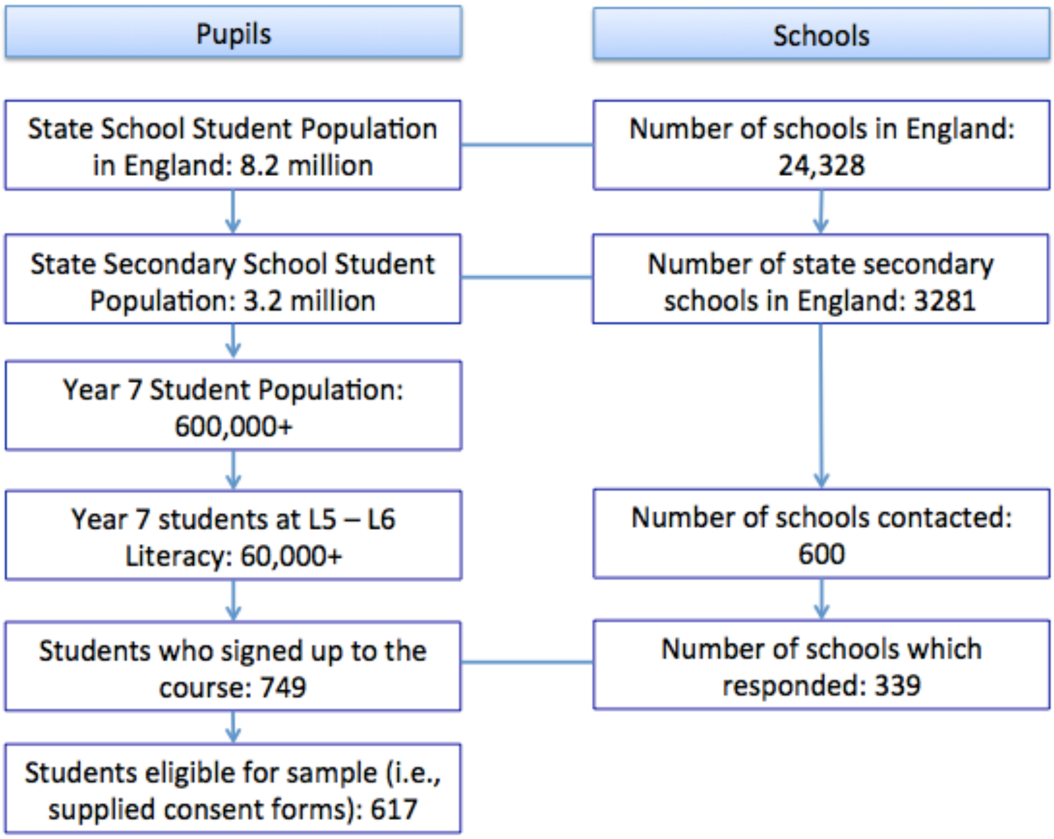


Figure 4-1 Sampling flow chart

Demographic figures correct at January 2013. Source: DirectGov 2013

Office for National Statistics (ONS) statistics report that 97% of families with children have home access to computers (ONS, 2013). Experience from the pilot study (summarised in Chapter Five) reflected this statistic, as all schools in the pilot study had either been able to offer loan laptops with internet accessibility, or they had found it easier to direct pupils to the local library where they could work on the internet. This was also found to be the case in the main study, where as far as is known, no child was disadvantaged by not having access to the internet.

Following the recommendation of Gorard (2001) computer-generated random numbers were used to initiate the process of random allocation (Appendix 5). As each student enrolled, they were allocated the next random number in a list of 1000 unique numbers. The day before the course started, those with numbers 1 - 250 were placed into the intervention group IC4; 251 - 500 into IC3; 501 - 750 into IC2; and 751 - 1000 into the control group. Once in their intervention groups, participants could only 'see' their own group and were not aware of the course structure in other groups. Thus four randomly assigned groups were created.

According to Field (2009), *a priori* estimation of sample size is informed partly by observing the findings of similar research. Findings of the pilot had identified medium and large effect sizes for the main effects of the intervention and therefore it was reasonable to expect a similar effect size for the main study. Thus, with a desired error probability of .05, a power calculation of at least .8 and an estimated effect size of Cohen's $d > .8$, calculations revealed that a minimum of 50 per condition was appropriate for a sample size (Cohen, 1992). However, as the intended analysis was to involve factor analysis, a larger sample was desirable. Field (2009) suggests that there should be 10 - 15 participants per variable, whilst others suggest that numbers need total no more than 300 (Kass and Tinsley, 1979). Tabachnick and Fidell (2007, p.613) agree with the latter suggestion and describe a sample of 300 as 'comforting'. Therefore, with the prospect of factor analysis as one of the tests, a sample size of at least 300 was identified as the minimum for this study.

Seven hundred and forty nine students initially signed up for the course, of whom 47 did not provide consent forms and a further 73 gave consents but did not participate. Twelve neither provided consent forms nor participated. The resulting 617 students, of whom 322 (52.2%) were female, were therefore those who a) participated and b) also provided consent forms.

4.5 Measures

The literature on self-regulated learning raises important questions about the nature of this composite construct, about the extent to which it can be measured, and about the validity and indeed limitations of instruments that try to capture the process of self-regulated learning. However, the conceptual model (Section 2.3.6) which underpins this study characterises self-regulated learning as a) a set of learner *characteristics* (metacognitive knowledge, efficacy for learning and motivational beliefs) which influence the process of self-regulation and b) the dynamic process *states* of self-regulatory activity such as metacognitive experiences, and metacognitive skills (critical thinking, volitional strategies and control of elaboration strategies) (Efklides, 2006). This study explores self-regulated learning through the lens of this conceptual perspective, by investigating a) the effect of the intervention on change in participants' critical thinking and self-regulated learning characteristics and b) the process of metacognitive activity over the course of the intervention.

The predominant method of collecting data on self-regulated learning (SRL) characteristics remains the self-report questionnaire (e.g., Sperling *et al.,* 2002), whilst measures such as think-aloud procedures and scaffolded learning journals are being championed as 'state of the art' assessment techniques for self-regulated learning processes such as metacognitive activity (Spörer and Brunstein, 2006; Schmitz and Wiese, 2006). In this study, SRL characteristics were measured by self-report questionnaire, whilst SRL processes were measured by a mixture of observed measures (critical thinking) and self-reported measures (inhibitory control, control of elaboration strategies and metacognitive experiences). Participation in course activity (e.g., dialogic argumentation) was measured by the counts of the number of times participants made course contributions. An overview of study measures is presented in Table 4-2:

Table 4-2 Measures

Parameter	Self report (SR) or observed (O)	Description	Representation	When
Learner Dispositions	SR	Metacognitive knowledge	Likert scale 1 - 10	Pre-test and post-test
Learner Dispositions	SR	Efficacy for learning	Likert scale 1 - 10	Pre-test and post-test
Learner Dispositions	SR	Motivational beliefs	Likert scale 1 - 10	Pre-test and post-test
Achievement	O	Critical thinking	% correct	End of each week
Strategy use	SR	Elaboration strategies	Likert scale 1 - 10	End of each week
Strategy use	SR	Inhibitory control	Likert scale 1 - 10	End of each week
Learner experiences	SR	Metacognitive experiences	Likert scale 1 - 10	End of each week
Participation	O	Dialogic argumentation	Count	End of each week
Participation	O	Browsing discussions	Count	End of each week

4.5.1 Critical thinking

For the purposes of critical thinking assessment, and according to the definitions adopted for this study (see Section 2.2), critical thinking is measured in terms of the learner's capacity to identify the strength of an argument (Paul, 1990). At the end of each week, students took a test designed to evaluate their progress in critical thinking. They were told that this was an assessment and that they would be able to see their grades for this task. As the course activity had focused upon the development of argument and counter-argument, the critical thinking assessment was designed to test students' capacity for identifying the strength of an argument in the context of scenarios that would be meaningful to the age group. As discussed in Section 2.2, the assessment that was selected for adaptation in this study was the Defining Issues Test (DIT2: Rest *et al.,* 1999). It was chosen for a number of reasons. It was important, in order to address potential issues of construct validity, to select a test that reflected as closely as possible the content of the material it was designed to assess. The content of both the Devil's Advocate task and the dialogic argumentation task was focused upon the exploration of moral, social or ethical dilemmas. Therefore one reason for choosing the DIT2 assessment, which had been developed from Kohlberg's (1984) works on moral development, was because it featured evaluation of arguments that focused upon such dilemmas. The DIT2 test was originally validated as a measure of cognitive ability and reasoning (*ibid*). Importantly for this study, Rest *et al.*, (1999, p.657) suggested that, 'it is possible to update, shorten, and revise the DIT without sacrificing validity. This should be encouraged for experimentation with new dilemmas and items.' It had initially been devised as a paper and pencil exercise. However Xu, Iran-Nejad and Thoma (2007) tested an internet-based version and found it to be comparable to the paper and pencil version in terms of ease of use. They reported that the test 'demonstrated robust validity and reliability in hundreds of studies' with a Cronbach's Alpha of 'in the upper .70s/ low .80s' (p. 13).

Four parallel assessments were adapted from the DIT2 test. In each test, participants were given one scenario that presented a dilemma that would require an ethical, moral or social decision. Thus the test content related specifically to the definition of dialogic argumentation[9] used in the context of this intervention. Students were given ten statements and were asked to identify the strength of a range of arguments by rating each of the statements as either *very relevant, of some relevance or not relevant* to each issue in question. Questions covered four topics: a) the election of a school captain, b) whether or not a doctor should give medication which might kill a terminally ill patient who has requested it, c) whether Oscar Pistorius, a 'blade runner' should have been allowed to compete against athletes who were not amputees in the London 2012 Olympics[10] and d) whether a father in India should steal food for his starving family. For example, on the subject of Oscar Pistorius, students were asked to consider whether the information 'one of his team-mates fell during practice' was relevant, of some relevance, or not relevant to their voting decision. Two marks were awarded to those students who considered that this information was not relevant to their decision. Thus each of the 10 questions had a maximum score of 2, with an overall score, for each of the four tests, of 20 marks. In each test, three of the 'NOT relevant' answers were correct, three of the 'SOME relevance' answers were correct and four of the 'VERY relevant' answers were correct. The order in which the NOT, SOME and VERY items were presented varied across the timepoints.

Following the recommendation of the DIT2 authors, Rest *et al.* (*op cit.*)*,* it had been necessary, in adapting the test for younger English adolescents, to make certain

[9] Dialogic argumentation: a means of developing students' critical thinking through peer-feedback and tutor-led discussion of moral, ethical or social issues.

[10] The intervention took place before Oscar Pistorius was involved in further controversy regarding the death of his girlfriend.

changes to the language and to the subject material. Specifically, 'the election of the President of the United States' was changed to 'the election of a school captain,' and the question about the London 2012 Olympics replaced a further US-specific question.

As some of the questions had been adapted from the original, the assessment was tested in the pilot study (Chapter Five). *Alternate forms reliability* testing (Erwin, 2000) was performed, with analysis of variance confirming that the tests were not significantly different to each other, in terms of difficulty, at each of the four time points (see Section 5.6). An example of the test can be found in Appendix 6.

The limitations of the critical thinking test

There are limitations to the potential effectiveness of this test that should be articulated here. Firstly, because of the dearth of available tests for young adolescents, it was adapted from an adult test; this poses a potential risk to validity. Secondly, item analysis was not available, since the test produced a sum total score only, rather than single item scores; therefore this poses a threat to internal validity of the test. Thirdly, it was originally written with an American audience in mind, and therefore the language was adapted for students from England. Fourthly, it was not validated across populations; therefore generalisations must be restricted to the population sampled here - above average young adolescents from English state schools. It is suggested therefore, as Simpson and Courteney (2002) have already pointed out, that there is a need to create and validate a critical thinking test that is suitable for use with children.

4.5.2 Self-regulated learning

The pre-test and post-test self-regulated learning questionnaires were identical. All three self-regulated learning components (metacognitive knowledge, efficacy for learning and motivational beliefs) were presented together, with the questions presented in random order. Following the advice of Bandura (2006) the questionnaire was given a nondescript title - the *Learning Quiz*. This was a deliberate attempt to ensure that learners did not try to second-guess the purpose of the questionnaire, and so bias their responses. However, again following Bandura's advice (*ibid*.) in order to encourage frank responses, the learners were told that the *Learning Quiz* was an important one because it was 'exploring the learning habits of young people'.

Measurement of metacognitive knowledge

Metacognitive knowledge items from the SAL instrument, described in Chapter Two, have been chosen to contribute to the set of measures for this study for a number of reasons. Firstly, the scale supports cross-cultural generalisability, in that it has been rigorously tested across cultures, many of which are likely to be represented in the current English[11] study. Secondly, its strong factor structure, its psychometric rigour and the independence of its sub-scales mean that it is possible to contextualise the scale by selecting certain components, without compromising the structure and reliability of the measurement. Finally, it recommends itself for use with younger

[11] According to DirectGov (2013), state school students in England are classified into 18 distinct ethnic groups.

students because of its brevity and its ease of use (Marsh *et al., 2*006a). Items focus on participants' metacognitive knowledge of their control strategies (e.g., When I study, I try to figure out which concepts I still haven't understood) and elaboration strategies (e.g., When I study, I try to relate new material to things I have learned in other subjects). Matell and Jacoby (1971) suggest that the decision about how many scale-points should be employed should relate to the context, whilst Preston and Colman (2000) report that the 10-point scale is particularly popular with respondents. In this case, and in the case of all of the self-report measures in this study, a 10-point scale was chosen, particularly in order to appeal to a child's familiarity with the concept of the language of, 'on a scale of one to ten'.

The scale was subject to exploratory factor analysis (EFA), using maximum likelihood estimation, at the pilot stage of the study. The analysis, with 164 participants, confirmed the *a priori* structure of the scale, with two clear factors: Elaboration Strategies and Control Strategies. The two factors correlated with each other (r = .483, p < .001), and together accounted for 61% of the variance. The Kaiser Meyer Olkin test (Kaiser, 1974) indicates whether the sample size is large enough for factor analysis. According to Field (2009), a KMO > .8 is good. Here, KMO was .811, and the Cronbach's α measure of reliability (Cronbach, 1951) was .810. Thus the scale was considered sufficiently robust to take forward for use in the main study. Confirmatory factor analysis of its capacity to model main study data is reported in Section 6.4.1.

Measurement of efficacy for learning

Bandura (2006) defines self-efficacy as 'people's beliefs in their capability to produce given attainments' (p.307). According to Schunk (1990), it is a characteristic which influences and is influenced by the process of self-regulation. Crucial to the design of this study is the notion that self-efficacy is a future-oriented concept - and therefore ideally measured before and after a learning task (Zimmerman, 2000a; Bong and Skaalvik, 2003). Based upon the definition of efficacy for learning (EfL) articulated in Section 2.3.5, an instrument was created consisting of seven items (Appendix 1) presented in the form of a Likert-scale, which, according to Maurer and Pierce (1998) is an acceptable method for measuring self-efficacy. Following the advice of Bandura (2006), the scale was created specifically for the current learning situation, and was presented on a 10-point scale, anchored from 1 (False - not at all like me) to 10 (True - this describes me very well). Questions focused upon the learner's self-belief in their capacity to self-regulate their learning, with questions such as, 'I am confident in my capacity to work independently' and 'I am confident that I can learn from my mistakes'. As with all of the scales in this study, it was constructed with the population in mind, with particular attention being paid to the accessibility of the language for young adolescents and the ease of use of the scale grading.

This instrument was piloted in October 2012, with 164 participants. Factor analysis was conducted using maximum likelihood estimation. This confirmed a unidimensional scale (α = .82) that explained 65% of the variance. The Kaiser Meyer Olkin test (KMO .80) verified the adequacy of the sample size and the scale was retained without further alteration. Analysis of its capacity to model the main study data is reported in Section 6.4.1.

Measurement of motivational beliefs

The motivational beliefs (MB) scale, adapted from Rheinberg, Vollmeyer and Burns (2001) consisted of nine items measuring the pre-course motivational beliefs of the participants. Students responded on a 10-point scale, anchored from 1 (False - not at all like me) to 10 (True - this describes me very well) to statements such as 'For activities like this course, I don't need a reward; they are lots of fun anyway,' and 'I will be happy to give up my free time to work on this course'.

The original scale (*ibid.*) included items which measure motivation with regard to the level of anxiety, probability of success, challenge and interest in the task. The latter two sets of items were chosen for adaption for this study, since research has demonstrated that these are the items which are most strongly correlated with educational performance (*ibid.*)

The original items are matched against their adapted versions in Table 4-3.

Also piloted in the October 2012 study, the motivational beliefs scale was subject to exploratory factor analysis (using maximum likelihood estimation). This confirmed a unidimensional scale (α = .82), explaining 66% of the variance. The KMO index of .89 confirmed the adequacy of the sample size and the scale was adopted for the main study without further modification. Confirmatory factor analysis of its capacity to model the main study data is reported in Section 6.4.1.

Table 4-3 Adaptation of Rheinberg *et al.* (2001) Motivation scale

Original item from Rheinberg, Vollmeyer and Burns, 2001	Adapted Motivational Beliefs scale (MB)
This task is a real challenge for me.	I am looking forward to the challenge of this course.
While doing this task I will enjoy playing the role of a scientist who is discovering relationships between things.	While doing this course, I will enjoy learning how to improve my thinking skills.
After having read the instruction, the task seems to be very interesting to me.	After having read the course guidelines, the course seems to be very interesting to me.
I am eager to see how I will perform in the task.	I am eager to see how I will perform on this course.
I'm really going to try as hard as I can on this task.	I'm really going to try as hard as I can on this course.
For tasks like this I don't need a reward, they are lots of fun anyhow.	For activities like this course, I don't need a reward; they are lots of fun anyway.
If I can do this task, I will feel proud of myself.	If I complete this course, I will feel proud of myself.
I feel petrified by the demands of this task.	When I think about this course, I feel excited.
I would work on this task even in my free time.	I will be happy to give up my free time to work on this course.

The aggregated self-regulated learning scale

The self-regulated learning dispositions scale was taken forward from the pilot to the main study data and consisted of three factors (8 metacognitive knowledge, 7 efficacy for learning and 9 motivational beliefs). Confirmatory factor analysis of main study data, of the compiled scale confirmed, as hypothesised, a three-factor measurement model, including the above three sub-scales (Pre-test: $\alpha = .90$; Post-test: $\alpha = .93$). All factor loadings were significant and greater than .43. The resulting CFA model of self-regulated learning dispositions provided a good fit to the data: Pre-test: χ^2 (227) = 492.647; $p < .001$; RMSEA= 0.047; CFI= 0.944; TLI= .932. Post-test: χ^2 (232) = 461.186; $p < .001$; RMSEA= 0.043; CFI= 0.962; TLI= .954. (The cut-offs for goodness of fit indices are explained in more detail in Section 4.7.5.) Overall these indices suggest that this measurement model fitted the data well.

4.5.3 Metacognitive activity

The use of self-report questionnaires is criticised in the literature, despite its ubiquitous use (Boekaerts and Corno, 2005; Hopfenbeck and Maul, 2011). Critics focus upon the possibility that participants (especially young students) may have a faulty recall of events, or indeed they may attempt to report according to what they perceive teachers or researchers might be expecting of them. Recently researchers have begun to employ event measures, such as those that can be recorded in a structured learning journal, as a means of recording self-regulated learning events in real time (Schmitz, 2006; Schmitz and Wiese, 2006). Thus, where such events represent episodes within the cyclical process of self-regulated learning, they can represent both predictor variables and outcome variables, and can contribute to both a synchronous and a cross-lagged analysis of the relationships between aspects of self-regulated learning over time.

A number of such learning events were captured during successive cycles of learning in the current intervention, at the end of each week. These plenary activities included critical thinking assessments and the completion of the learning journal, which recorded self-reports of metacognitive activity (elaboration strategies, volitional strategies and metacognitive experiences).

Self-reported metacognitive activity, which was measured in terms of metacognitive skills (control of elaboration strategies and volitional strategies) and metacognitive experiences, was captured in the learning journal and recorded either in text form or in Likert scale form (Appendix 2). Students answered on a scale of 1 to 10 to questions such as, 'I was able to put aside all of my negative thoughts and/or worries this week while I was studying' and 'I am satisfied with my learning progress this week'. Exploratory factor analysis, using maximum likelihood estimation, with Oblimin rotation, extracted a three-factor structure that explained 70% of the variance. The three factors were characterised as a) metacognitive experiences, b) the volitional strategy of inhibitory control (a metacognitive skill) and c) perceived use of task specific elaboration strategies (a metacognitive skill) (Efklides, 2009). The full three-factor scale had a reliability of Cronbach's $\alpha = .84$, whilst the three separate scales were: Inhibitory control: $\alpha = .70$; Elaboration strategies: $\alpha = .74$; Metacognitive Experiences: $\alpha = .86$. Elaboration strategies were weakly related to inhibitory control ($r = .19$, $p < .05$) but not significantly to metacognitive experiences. Inhibitory control, a metacognitive skill, was strongly related to metacognitive experiences ($r = .57$, $p < .001$). Confirmatory factor analysis is reported in Table 6-24, in Chapter Six.

4.5.4 Dialogic argumentation

In this study, the intervention was delivered electronically, via an internet-mediated learning platform, and all transactions were dated and recorded. Thus it was possible to make a weekly count of the number of times students contributed posts to forum discussions. The use of such count data (the number of occurrences of a behaviour within a fixed period of time) for measuring the frequency of learning events, is often used in research that attempts to explore learning processes as a predictor of learning outcomes (Greene, Costa and Dellinger, 2011) and is effectively gathered via electronic means. A quality assurance of this measure is presented in Appendix 10.

4.6 Data collection procedures

Data were collected electronically, a method which affords a number of advantages over others. Firstly, it is cost-effective, since the researcher does not have to travel to meet participants - there is no need for stationery or postage, and just one researcher can work with hundreds of participants. Secondly, a convenient distance is preserved between researcher and participant, such that the researcher is de-personalised to a certain extent, an advantage in experimental conditions, according to Gorard (2001). Thirdly, data are collected in digital and text form, saving time by removing the necessity for transcription, and by reducing human error. Fourthly, email communication to a large number of schools increases the chance of recruiting greater numbers of participants. Electronic transactions result in machine-readable, traceable digital data. Indeed, it is suggested that the unprecedented ease with which digital data can be traced and recorded represents ‘a potential shift in the nature of social science’ (Welser, Smith, Fisher and Gleave, 2008, p. 4).

Course activity took place on a *Moodle* learning platform (also described in Section 3.2), which was accessed by students, tutors and administrators via unique usernames and passwords. Students, identified to each other by first names only, had access only to their own course area and to their own course activity report. The researcher, with ‘administrator’ permissions on the learning platform, was able to create the course, allocate ‘teacher’ permissions to the course tutor, monitor course activity and download course statistics. The teacher was able to teach on the course, but was not able to change activities, or to access personal student data other than course progress results. At the end of the course, all activities became hidden from students and teacher, and the course data were made available to the researcher for downloading onto her own computer. The data, downloaded into Excel, were anonymised before being imported into statistical applications such as SPSS (2011) and MPlus (Muthén and Muthén, 2013), and the course was deleted once the data were collected from it.

The pre- and post-test self-regulated learning questionnaire and the learning journal, which captured metacognitive activity, were administered via *Google Docs* forms (described more fully in Section 3.2). Each questionnaire collected data about the participant's name and gender, alongside their question responses. Responses were given on a Likert-type rating scale, anchored on a scale of 1 to 10, where 1 represented 'False: not at all like me,' and 10 represented 'True: this describes me very well.' Participants were asked to complete self-regulated learning questionnaires just before the course started and then as the course closed. Emails were sent to the students two days before the course started and then again on the final day, to remind them to participate in these voluntary activities.

The Likert scale was devised such that it could be interpreted as an interval scale, rather than an ordinal one. It is understood that where a Likert scale consists of separate items that prompt participants to choose between specific statements, it is subject to treatment as an ordinal scale (Göb, McCollin and Ramalhoto, 2007) and is necessarily capable of generating little more than descriptive data. However, where a rating scale is devised, such that participants are asked to choose a point along a line 'akin to a tape measure' (Muijs, 2010, p. 42), and where such a scale has more than seven points, it is considered to approximate to a continuous scale (Muijs, *op cit.*). The choice of an even number of scale points in this study was informed by a desire to avoid the 'central tendency' problem whereby respondents may choose the central option, especially towards the end of a survey, as interest in the activity possibly declines. Whilst the range of the scale should be sufficient to encourage respondents to use all points of the scale (and thus produce a normally distributed set of responses), in practice, self-reported scales tend to result in a negative skew. This is particularly true where children are concerned, as there is the possibility that they see the question as a sort of test item, in which a positive response is better than a negative one. Van Laerhoven, Van der Zaag-Loonen and Derkx (2004), for instance, in a comparison of different types of questionnaires, found that whilst children preferred the Likert more than other types of questions, they tended to score higher on Likert scales than on other types of questionnaire. However, they concluded that the Likert scale questionnaire is nevertheless the most appropriate means of collecting self-report data from children.

Forms were submitted electronically, after which they became available to the researcher via a unique username and password. They were then downloaded from *Google Docs* into an *Excel* file, which was then stored on the researcher's computer and backed up appropriately and securely. As the original data were residing on a server not in the control of the researcher, they were deleted once they had been downloaded. Identification numbers were assigned to each participant, in order to ensure confidentiality, and only those who had submitted consent forms were included in the analysis.

Critical thinking assessments were constructed by the researcher, using the ‘quiz’ tool in *Moodle*. Students clicked onto a quiz link and their responses were automatically scored, according to the pre-determined scoring system (see Section 4.5.1). Students could view their own scores, but not those of others. Scores were downloaded by the researcher into *Excel* for export into statistical analysis packages.

Data on student course participation were captured and downloaded from the learning platform. For instance, each time a student contributed to a discussion, that action was electronically attributed to that student, with a timestamp detailing date, hour and second. Data were downloaded into *Excel* and thence into other appropriate statistical applications.

Exporting from SPSS into MPlus involved saving the SPSS file into a comma separated values (.csv) file, a format readable by MPlus. Examples of MPlus coding are given in Appendix 9.

4.7 Data analysis procedures

This section provides an overview of the data analysis methods employed in the study. Statistical techniques which are less commonly used in intervention studies (latent change score modelling and autoregressive cross lagged panel analysis) are outlined in this section, whilst the more common statistical techniques (ANOVA, ANCOVA and regression) are described only briefly here and then in further detail in Appendix 14.

The software programmes used for this study were SPSS v. 20 and MPlus v. 7. The SPSS package was used for all analyses except structural equation modelling, which used MPlus. SPSS analyses were recorded in a syntax file, and example of which can be seen in Appendix 8, whilst MPlus analyses were recorded in MPlus code, an example of which can be seen in Appendix 9.

This section details the procedures required for the analysis of data. It first presents an analysis of missing data, followed by an explanation of how statistical significance and effect sizes are reported. Data analysis procedures, summarised in Table 4-4 below, are explained in the order in which they appear in the Findings, i.e., ANCOVA, repeated measures ANOVA, latent change score modelling, regression and autoregressive cross-lagged panel analysis.

Table 4-4 Data analysis procedures

The pilot study	
Research Question	**Analytical procedure**
Pilot RQ1: What is the effect of the intervention on pre-test to post-test change in the SRL characteristics of a group of 11 and 12 year old students, after a period of one month, comparing experimental groups with a control group?	Analysis of Covariance (ANCOVA)
Pilot RQ2: What is the effect of the intervention on the repeated measures change in CT of a group of 11 and 12 year olds students, over four assessments?	Repeated measures analysis of variance (ANOVA)
The main study	
Research Question	**Analytical procedure**
RQ1: What are the main and experimental condition effects of the intervention on change over time in students' CT?	Repeated measures analysis of variance (ANOVA)
RQ2: To what extent does participation in dialogic argumentation predict progress in CT over time?	Regression
RQ3: What is the main effect of the intervention on change in aspects of participants' SRL (metacognitive knowledge, motivational beliefs and efficacy for learning) from pre-test to post-test, on groups of young adolescents, after one month, compared to a group who had no intervention experience?	Latent change score modelling
RQ4: What are the effects of the separate experimental conditions on change in SRL and its components over time?	Latent change score modelling
RQ5: To what extent does participation in dialogic argumentation predict change in SRL and its components over time?	Latent change score modelling
RQ6: What are the relationships between aspects of metacognitive activity and CT over time?	Autoregressive cross-lagged panel analysis
RQ7: What is the relationship between self-regulated learning dispositions (metacognitive knowledge, motivational beliefs and efficacy for learning) and self-regulated learning processes (metacognitive activity) during this intervention?	Path analysis

4.7.1 Missing Data

Previous research has determined that attrition rates for internet-based courses are significantly greater than for face-to-face courses (Sitzmann, 2010). Questionnaire participation rates in the current study corroborate this assertion, with person missingness ranging from 8.6% to 18.1% in the pre/post-test surveys. This was expected as students had been reminded that participation in the research element of the course (the questionnaires) was entirely voluntary (see Appendix 4: Ethical materials). However, students were enrolled on a 'course'; therefore it was expected that there would be a higher level of participation in the weekly critical thinking assessment, where person-missingness for the weekly assessments was much lower, and ranged from 1.2% to 5.2%.

Missing items were coded as -9 in the SPSS datafile. SPSS analyses then employed pairwise comparisons to estimate variables which included missing data. Missing data was handled in MPlus by the use of maximum likelihood estimation.

Missingness poses a problem for structural equation modelling if it is other than 'Missing Completely at Random' (MCAR) (Schafer and Graham, 2002). Missingness completely at random is important to establish in order to confirm that missing items are not related to other variables in the dataset. It is possible to establish whether or not missingness is MCAR or not by using Little's Test (Little and Rubin, 1987). This test, using SPSS v.20, was performed on the study variables of interest. Here, a non-significant result indicates MCAR. The syntax for Little's test is detailed in Appendix 8.

Little's Test estimated significance levels of $p = .189$ for IC4; $p = .246$ for IC3; $p = .154$ for IC2, and $p = .535$ for the control group. Levels of missingness, therefore did not reveal significant problems, and are reported in Table 4-5.

Table 4-5 Person missingness by group

				Missing	
IC4	N	M	SD	Count	%
CT1	149	12.13	3.284	2	1.3
CT2	147	12.97	2.472	4	2.6
CT3	147	14.05	2.315	4	2.6
CT4	146	14.87	2.430	5	3.3
PreMK	125	61.65	8.546	26	17.2
PoMK	124	69.01	7.577	27	17.9
PreMot	123	58.45	6.972	28	18.5
PoMot	124	65.42	4.621	27	17.9
PreEfL	110	58.25	6.495	41	27.2
PoEfL	124	64.06	5.090	27	17.9
Little's MCAR test: χ^2 = 101.644, *df* = 90, *p* = .189					
				Missing	
IC3	N	M	SD	Count	%
CT1	147	12.61	2.597	8	5.2
CT2	150	13.20	2.317	5	3.2
CT3	149	14.06	2.334	6	3.9
CT4	150	14.76	2.684	5	3.2
PreMK	132	62.93	8.309	23	14.8
PoMK	127	66.61	8.180	28	18.1
PreMot	129	58.33	7.071	26	16.8
PoMot	127	62.27	7.262	28	18.1
PreEfL	131	58.21	6.614	24	15.5
PoEfL	127	61.68	5.398	28	18.1
Little's MCAR test: χ^2 = 64.999, *df* = 58, *p* = .246					
				Missing	
IC2	N	M	SD	Count	%
CT1	158	12.67	2.560	2	1.3
CT2	158	12.95	2.591	2	1.3
CT3	158	14.16	1.989	2	1.3
CT4	157	14.44	2.427	3	1.9
PreMK	132	62.92	8.456	28	17.5
PoMK	140	66.69	8.715	20	12.5
PreMot	129	58.12	7.578	31	19.4
PoMot	140	60.96	8.380	20	12.5
PreEfL	134	58.29	6.328	26	16.3
PoEfL	140	61.02	6.725	20	12.5
Little's MCAR test: χ^2 = 77.690, *df* = 66, *p* = .154					

Table 4-5 (continued)

CG	N	M	SD	Missing Count	Missing %
CT1	n/a	n/a	n/a	n/a	n/a
CT2	n/a	n/a	n/a	n/a	n/a
CT3	n/a	n/a	n/a	n/a	n/a
CT4	n/a	n/a	n/a	n/a	n/a
PreMK	138	63.61	8.013	13	8.6
PoMK	133	64.32	8.774	18	11.9
PreMot	138	57.86	7.085	13	8.6
PoMot	134	59.21	7.908	17	11.3
PreEfL	140	57.69	6.199	11	7.3
PoEfL	135	58.87	7.066	16	10.6
Little's MCAR test: $\chi^2 = 15.842$, $df = 17$, $p = .535$					

Item missingness analysis revealed negligible levels of item missingness (less than 1%) in all pre- and post-test scales and the assessment grades (see Table 4-6).

Table 4-6 Item missingness

Scale	**Items missing**	**Total items**	**% missing**
Pre-test metacognitive knowledge	8	4264	.19%
Post-test metacognitive knowledge	15	4224	.36%
Pre-test motivational beliefs	27	4797	.56%
Post-test motivational beliefs	27	4761	.59%
Pre-test efficacy for learning	22	3731	.59%
Post-test efficacy for learning	7	3696	.19%
Critical thinking Assessment grades	5	1568	.3%

4.7.2 Statistical significance and effect size

Whilst statistical significance indicates the probability that a result can be achieved by chance, the effect size points to the importance of the result and is represented as a standardised measure of the magnitude of the observed effect. Increasingly, researchers report effect sizes in addition to significance tests. Therefore, in this study, effect sizes are calculated wherever possible, with cut-off criteria evaluated according to the metric detailed in Table 4-7.

Table 4-7 Effect size cut-off criteria

Effect size	Relevant for:	Small	Medium	Large
Cohen's *d*	Difference in terms of standard deviation units	.2	.5	.8
r	Strength of association	.1	.3	.5
η^2	% of variance explained	.01	.06	.14

It should be noted that whilst Cohen (1969, p. 23) suggests that an effect size of $d = .2$ is 'small', he does acknowledge that effect sizes should be evaluated in context (see also Cohen, 1994). Coe (2002) points out that very few educational interventions produce effect sizes that are anything other than small. Glass, McGaw and Smith (1981), in considering the use of effect sizes in education, are particularly critical of the use of arbitrary cut-offs, suggesting that effect sizes should be considered in the context of the practical importance of an intervention, suggesting that an effect size of as small as Cohen's $d = .1$ may be significant in educational terms. For the sake of comparison however, the analyses in this study will reference effect sizes according to the criteria detailed in Table 4-7.

4.7.3 ANCOVA

Analysis of Covariance (ANCOVA) was built into the design of the pilot study as it is a powerful technique for estimating the difference between groups at post-test (the dependent variable), whilst partialling out the effect of the pre-test (the covariate) (Field, 2013). The main question that ANCOVA poses is whether or not a difference between groups, once a covariate has been taken into account, could have occurred by chance (Tabachnick and Fidell, 2013). According to Field (*op cit.*) the advantage of doing this is twofold. Firstly it increases the power and sensitivity of the F test, by reducing within-group error variance, since some of the unexplained variance can be accounted for in the covariate. Secondly, it takes into account the most likely confounding variable - the pre-test score - since this is likely to be related to the post-test score.

According to Field (2013), general assumptions which apply to all analysis of variance (ANOVA) techniques, and which are addressed in the current study, assume that:

- Measures on the dependent variable are on a continuous scale
- Scores are obtained through random sampling
- Observations are independent of one another
- Scores are distributed according to a normal distribution
- Variances across groups are equal (homogeneity of variance).

Additional assumptions apply to ANCOVA, and are also addressed accordingly in the current study (Pilot study RQ3: Section 5.9.1):

- The covariate should be measured prior to the intervention treatment
- The covariate should have a good internal consistency (with Cronbach's α above .7)
- Covariates should not be strongly correlated (e.g., $r > .8$)
- The relationship between the covariate and the dependent variable should be linear
- The relationship between the covariate and the dependent variable should be the same for all groups (homogeneity of regression slopes).

Further details about the ANCOVA technique are described in Appendix 14.

4.7.4 Repeated Measures ANOVA

Repeated measures, or *within-subjects*, refers to situations where the same participants contribute to different means. This method of analysis is used in the pilot study (RQ2) and in the main study (RQ1). The repeated measures test evaluates systematic variation, i.e., differences in the same subjects in different conditions, or, as in the case of this study, differences in performance across time, that are due to an experimental effect. It is preferable to a *between-subjects* design in such a situation since the difference between the repeated measures can only be attributed to either the experimental manipulation, or a different factor that may influence how a participant performs from one time-point to another. However, in a between-subjects design, differences can be caused not only by the above two factors, but also by differences in the characteristics of the members of each independent group, a factor which may cause additional random variation between conditions. Field (*ibid.*) suggests that the repeated measures design is therefore capable of better identifying the effect of an experimental manipulation, where measures have been repeated over time.

Assumptions which apply to repeated measures ANOVA assume that:

- measures on the dependent variable are on a continuous scale
- scores are obtained through random sampling
- observations are independent of one another
- scores are distributed according to a normal distribution
- variances across groups are equal (homogeneity of variance)

- variances of the differences between pairs of groups (i.e., the levels of the independent variable) are equal (sphericity).

Criticism of the repeated measures designs focuses on the disadvantage of a possible carry-over effect from one assessment to another. This, according to Tabachnick and Fidell (2013) limits the generalisability of the result. This limitation is addressed in this study by using parallel critical thinking assessments, rather than identical ones. Parity of these assessments is tested and reported in Section 5.6.

The repeated measures ANOVA technique is further described in Appendix 14.

4.7.5 Structural equation modelling

Structural equations were estimated to test a number of the research questions (RQ3 – RQ7). Structural equation modelling (SEM) is a powerful confirmatory statistical technique, whose purpose is not only to analyse the covariance patterns between related manifest (observed) variables, but also to explain as much of their variance, determined by the model under analysis. It is particularly effective in modelling latent change in factors that represent theoretical constructs. Combining factor analysis, regression and path analysis, it allows for error correction, correlated disturbances, overall model fit assessment, categorical endogenous variables and missing data (Hox and Bechger, 1998; Blunch 2012). SEM is employed in the context of this study a) for the analysis of pre- to post-test latent change and its predictors, a process known as latent change score modelling (McArdle, 2009), and b) for path analysis of the relationships between manifest (observed) variables over time, a process known as autoregressive cross lagged panel analysis (Kenny, 1979).

The SEM approach requires that measures have good psychometric properties, i.e., that they have strong structural validity and reliability. Therefore the first analysis in the SEM process is to establish the structural validity of the proposed latent constructs through confirmatory factor analysis. This is done by estimating the extent to which the measurement model 'fits' the data, by evaluating factor loadings, unique variances and 'goodness of fit' indices. A latent path analysis is then specified (i.e., the structural part of the model), in which the relationships between latent constructs are estimated. Estimated relationships between latent constructs are net of measurement error (i.e., disattenuated), since residual errors of endogenous (dependent within the model) variables are also estimated (Kline, 2010). The resulting model can be represented graphically in a path diagram detailing the regression coefficients that connect the elements of the model.

Goodness of fit indices determine how well the hypothesised model describes the data. Evaluating goodness of fit involves establishing the extent to which the estimates generated by the model match the variances, covariances and means of the data (West, Taylor and Wu, 2012). Commonly the χ^2 test statistic (Jöreskog, 1969) is reported, as maximum likelihood estimates have a χ^2 distribution. Here the cut-off criterion is > .05. However, this is particularly sensitive to large sample sizes and, in view of its limitations, is frequently accompanied by other fit indices, such as the RMSEA Root Mean Square Error of Approximation (Steiger and Lind, 1980), the TLI Tucker-Lewis Index (Tucker and Lewis, 1973) and the CFI Comparative Fit Index (Bentler, 1980). Typical cut-off values are RMSEA < .06; TLI > .90; and CFI >

.90 (Hu and Bentler, 1999). The goodness of fit indices reported in this study therefore include RMSEA, CFI and TLI.

4.7.5.1 Latent Change Score Modelling

Latent change score modelling (LCSM), a structural equation modelling technique, is used to explore several hypotheses in this study. It is an emerging, but infrequently employed research technique for modelling change pattern predictors that have been measured with error (McArdle, 2009; McArdle and Prindle, 2008; Duncan, Duncan and Stryker, 2006). It is employed in this study to estimate the longitudinal effect of the intervention and the predictive effect of intervention activities. According to Mun, von Eye and White (2009), LCSM offers advantages in evaluating intervention effects of pre- and post-test designs. The technique is appropriate where an intervention has the effect of changing the characteristic behaviour of individual participants, whereby change can be modelled as a discrete event over time, from Time 1 to Time 2, and where change is captured in terms of interindividual differences in intraindividual change over time. In particular, it can be included in tests that estimate predictors of that change. Specific advantages, over methods such as analysis of variance (ANOVA), which influenced the choice of the LCSM technique for aspects of this study, include its function of taking account of measurement errors and individual differences in response to intervention, whilst also providing an interpretation of the appropriateness (‘fit’) of model to the data.

The latent change score modelling (LCSM) procedure is to create an unobserved variable that is the change in the score from Time 1 to Time 2, in this case pre-test to post-test self-regulated learning. Structural equation modelling (SEM) then treats that variable as latent and uses it in tests about changes from Time 1 to Time 2. The hypothesised models in this study, such as the model featured in Figure 4-2, specify a latent change (LC) phantom construct using metacognitive knowledge (a latent variable indicated by manifest scores on questionnaire items) at Time 2 as its indicator.

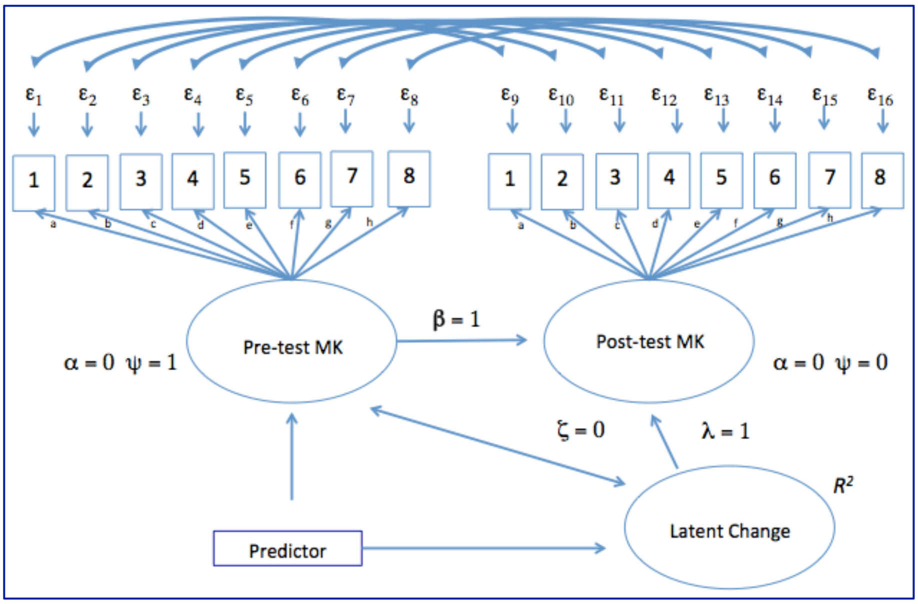


Figure 4-2 Hypothesised LCSM model for RQs 1, 2 and 4
MK: metacognitive knowledge; α Latent intercept; ζ disturbance (endogenous); ψ latent variance (endogenous variable); λ loading; ε error terms; R^2 variance explained; ellipses represent latent variables; rectangles represent manifest variables (which are simplified for clarity)

In this model, a latent change score is computed (Time 2 minus Time 1) so that each individual has an increase or a decrease, and then the change score becomes a dependent variable in the design. The LC model also extracts out two useful parameters: a) the correlation between the pre-test and the LC and b) the extent to which the experimental condition predicted how much each participant's scores changed over time. Several parameters are constrained in order to pull out the LC score and facilitate the identification of the change score model. Factor loadings are constrained to equality (e.g., MK at Time 1 and MK at Time 2). This is done in order to test invariance of the factor loadings, i.e., the two constructs have the same structural validity at both time points and hence are valid to compare. All loadings are estimated, (rather than, by default, fixing the first factor loading to one); latent means are set to zero; the latent regression path is fixed to 1, and intercepts are equated across time-points. This results in strong invariance of measurement, i.e., one is measuring the same construct at Times 1 and 2, and hence they are reasonable to compare. The residuals of parallel items at Time 1 and 2 are allowed to be correlated across time to control for the dependencies in the repeated measures. The pre-test variance is fixed at 1 and the post-test fixed at 0. The variance of the change score is estimated, as well as the latent mean of the change score. The post-test is then regressed on both the pre-test and the LC, and the LC is regressed on the predictor variable. Thus it is possible to estimate the effect of the predictor on change in the dependent variable, which, in the case of this example, is latent change in metacognitive knowledge. An example of the coding for this latent change score model is given in Appendix 9.

4.7.5.2 Autoregressive cross-lagged panel analysis

Autoregressive cross-lagged (ARCL) panel analysis is used to explore the penultimate research question of the main study (RQ6). This analysis is useful for exploring the relationships between two or more sets of latent or manifest variables over time (Kenny, 1979; Finkel 1995; Marsh and Yeung, 1997) and is often used for examining questions of reciprocal causality. This involves the sets of variables being measured at the same time points (e.g., x1, x2, x3, at the same time as y1, y2 and y3). Where the causal nature of the sets of variables is similar over time (i.e., x1 causes x2, which causes x3, and y1 causes y2 which causes y3), then correlations and cross-lagged correlations provide information about the effect of x on y over time. SEM is useful for exploring autoregressive models because a whole model can be estimated at once, and model fit indices of alternative models inspected. Furthermore, it enables the examination of direct and indirect effects within the same model, allowing any one variable to act as a dependent and an independent variable at the same time. ARCL analysis is useful in this study because four aspects of metacognitive activity were recorded simultaneously over four time points. The hypothesised model can be seen in Chapter Six, Figure 6-15. The four metacognitive activity variables which feature in the hypothesised model are a) critical thinking, a metacognitive skill, b) control of elaboration strategies, a metacognitive skill, c) inhibitory control, a metacognitive skill and d) metacognitive experiences.

Assumptions for SEM

Structural equation modelling is sensitive to a number of assumptions. These include temporal dependence of regression elements, normality, sensitivity to sample size and missing data, multicollinearity (independent variables should not be highly correlated) and singularity (one independent variable should not be included as a subset within another).

Normality is tested in this study through inspection of skewness and kurtosis values. In SEM, for normally distributed data, maximum likelihood (ML) estimates are appropriate. However, for skewed data robust maximum likelihood (MLR) estimates are requested (see Table 4-8). By specifying MLR, estimation ensures robust standard errors and compensates for non-normal data (Muthén and Muthén, 2013).

Table 4-8 Estimation procedures for research questions that employ SEM

RQ	Groups	Variables	Estimation
RQ3	1, 2, 3, 4	Pre- and Post-SRL	ML
RQ4	1, 2, 3, 4	Pre- and Post-SRL by group	MLR
RQ5	1,2, 3	Pre- and Post-SRL & Intervention activities	ML
RQ6	1	CT, ES, IC & ME	ML

ML: maximum likelihood; SRL: self-regulated learning; DA: dialogic argumentation; CT: critical thinking; ES: elaboration strategies; IC: control; ME: metacognitive experiences

Critics of SEM focus on two issues. Firstly, it is particularly sensitive to statistical assumptions, such as normality, and secondly to sample sizes - and as such, is a large-sample technique (Kline, 2010). In this study, normality issues are addressed by use of robust maximum likelihood estimation, and sample sizes are large.

4.7.6 Regression

The regression process fits a linear model to the data, exploring the effect of a predictor variable on an outcome variable, producing a linear equation analogous to the 'equation of a straight line', commonly understood as $y = mx + c$, where y is the outcome variable, m is the gradient of the line and c is the intercept. In the case of regression m and c are replaced by the 'regression coefficients' b_1 and b_0. Hence the regression equation for simple regression takes the form of:

$$\text{predicted value} = (b_0 + b_1 x) + \text{error}$$
$$y_i = (b_0 + b_1 x_i) + \varepsilon_i$$

The resulting equation is then used to predict the value of the dependent (outcome) variable.

The regression technique requires a number of assumptions to be met. These include:

- Sample size: Tabachnick and Fidell (2007, p. 123) provide the formula $N > 50 + 8m$, where m is the number of independent variables
- Multicollinearity: independent variables should not be highly correlated (above $r = .9$)
- Singularity: one independent variable should not be included as a subset within another
- Outliers: these, according to Tabachnick and Fidell (2007, p. 128) are those with standardised residuals above $|3.3|$
- Normality of residuals: these should be normally distributed about the predicted dependent variable scores

- Linearity of residuals: these should have a linear relationship with predicted dependent variable scores
- Homoscedasticity of residuals: the variance of residuals around the dependent variables should be the same for all predicted scores.

The regression technique is further described in Appendix 14.

4.8 Ethical considerations

This study involved working on the internet with children. Therefore ethical considerations were taken into account at various stages in the research, from the design stage right through to the reporting. The Central University Research Ethics Committee (CUREC) clearance, the University of Oxford's ethical requirement, was obtained and all data were anonymised before analysis. Acknowledging the recommendations of Eynon, Fry and Schroeder (2008), the informed consent of the parents of participants was obtained, not only because of the need to uphold institutional standards, but also for reasons of Child Protection and data protection. Stringent consent procedures were therefore followed, according to CUREC guidelines, including parental consent and the freedom of participants to remove themselves from the study, should they have chosen to do so. Participants were not paid, but did receive a 'Record of Participation' detailing their learning activities on the course (see Section 3.2).

The process of working with the students involved them participating in an internet-based course, which was active for one month, and then visible ('read-only') for a further month. 'Participation' involved them engaging in open discussion with tutors and with other students on the course, on a learning platform, accessed through the internet. According to ThinkUKnow (2011) the use of moderators is essential to safeguard children on the internet. Thus, a moderator, who was online for the duration of the course, scrutinised every post as it was uploaded to the learning platform, having been alerted by email as soon as a student made a written contribution.

Following the recommendations of the Child Exploitation and Online Protection Centre (CEOP, 2012), the students were directed to the *ThinkUKnow* website (www.thinkuknow.co.uk) for advice on how to stay safe online and were given clear 'Stay Safe' guidelines before the course started. These guidelines included instructions not to reveal personal details such as surname, email address, telephone number or social networking identity, and directives not to bully or intimidate other participants. Discussions were closed down as the course ended, so that students could not access the discussion space, once the tutor and moderators were no longer 'present'.

This study has been designed with close reference to the recommendations made in the professional guidelines for the Association of Internet Researchers (http://aoir.org/documents/ethics-guide/). In particular, it should be noted that the learning platform was not a public site and was accessed only through a username and password. The author of the learning material (the researcher), and the course tutor, both had current Enhanced CRB checks (Criminal Records Bureau). Participants did not have the opportunity to communicate with each other, other than through the course discussion channel, which was tutor-led and monitored. Students could only see each other's first names, and could not access any personal details of any other student. All participants were enrolled through contact with their schools.

Participants were informed that this research was exploring 'the way students study'. They were not informed specifically (although they may have guessed) that their pre-test survey was to be compared with the post-test, (as this may have precipitated a 'subject effect', whereby they responded according to perceived researcher expectations). Neither were they informed of the research questions. According to McMillan and Schumacher (2010), this procedure is appropriate where full disclosure would affect the validity of the results.

All participants in this study were volunteers. However, their school teachers recommended the course to them, so students may have felt some responsibility to their teachers to participate. Therefore it was made clear to participants that they were able to participate in the course but could opt out of the study (the pre- and post-test surveys) if they preferred - without academic penalty and at any time.

This study involved experimental groups and a control group and an intervention that may have benefited its participants. Therefore the 'cross-over design' (Piantadosi, 2005), which enabled a control group to access an intervention, once the comparative data were collected, ensured that students in the control group were not academically disadvantaged by having been randomly assigned to that group.

4.9 Methodological Limitations

This section will begin addressing issues of validity by considering four aspects, originally distinguished by Cook and Campbell (1979); these are internal validity, statistical conclusion validity, construct validity and external validity. It will then consider issues of objectivity and finally, challenges surrounding the use of self-report measures.

4.9.1 Internal Validity

Internal validity refers to the capacity to establish a causal relationship between variables within the experiment - variables such as treatment and outcome variables (Shadish, Cook and Campbell, 2002). According to Campbell and Stanley (1966), experimental design, with random assignment to groups, has the potential to control for a number of threats to internal validity. The threat of *selection bias* is addressed where participants are assigned into groups at random, as in this study. Potential interaction effects of *pre-testing* can pose a threat to internal validity if there is the possibility of a method effect. For instance, it could be argued that asking a participant about their efficacy beliefs could trigger self-reflections that could produce a reactivity effect. Bandura (2006) anticipated this threat and reported on ‘numerous’ studies which demonstrated that efficacy beliefs were not significantly amenable to change simply by the action of by self-report, and that this non-reactivity was corroborated across domains and diverse activities (p. 315). In this study, this threat was nevertheless controlled for, since both control group and experimental groups experienced identical surveys and at the same times (Bryman, 2004). *Maturation* of participants during the duration of the experiment was similarly controlled insofar as changes in individuals should have occurred equally across experimental and control groups. *Instrumentation* (Bryman, 2004), which refers to potential changes in the nature of the instrument (for instance if it is administered by different researchers) was controlled for by ensuring that the pre- and post-test surveys were electronically administered, fixed, unaltered and identically measured. *Programme leakage,* or diffusion of intervention (McMillan and Schumacher, 2010) was avoided by ensuring that the groups were not able to come into contact with each other.

4.9.2 Statistical conclusion validity

Statistical conclusion validity (Black, 2005) requires there to be a relationship among the proposed variables. This issue is explored in the literature review, where theoretical relationships have been established, which have in turn informed the research questions.

4.9.3 Construct Validity

Construct validity, at times subsumed into the concept of external validity (e.g., Campbell and Stanley, 1966) refers to the validity of the measurement procedures (such as the questionnaires in this study).

Care has been taken, in the literature review, to avoid an *inadequate explication of the constructs* (Bryman, 2004) by ensuring that the questionnaires were derived from established constructs. This has been important in this study as definitions of self-regulated learning vary considerably across the literature. So, for instance, the metacognition items were taken from the PISA questionnaire (Marsh *et al.,* 2006a), the motivational beliefs items from Rheinberg, Vollmeyer and Burns (2001), the efficacy for learning items were constructed according to the advice of Bandura (2006) and the critical thinking assessments were derived from Rest *et al.,* (1999).

The threat of *mono-method bias* (McMillan and Schumacher, 2010), which refers to the use of a single method of measurement for any one construct, is also relevant to this study as there is a need to address perceived criticism of the use of self-report measures for self-regulated learning measurement (e.g., Connor-Green, 2000; Hopfenbeck and Maul, 2011). Therefore the inclusion of multiple (self-report and observed) instruments of measurement has been built into the construction of the study. Analysis of the extent of the relationships between self-report and observed measures of self-regulation has therefore also been possible (See RQ6, Section 6.7).

4.9.4 External Validity

External validity refers to whether the results are generalisable. A number of writers (Campbell and Stanley, 1966; Robson, 2002; Figueredo, Olderback, Schlomer, Garcia and Wolf, 2011) point out that construct validity is closely related to the external validity issue of generalisability, since the strength of the conceptualisation of the construct and the validity of the measurement instrument will determine the extent to which generalisations can be made from the findings. Whilst this study design incorporated randomisation, the population from which the sample was drawn, was a specific one. Thus, in this study, any generalisations will be restricted to the setting of the intervention (in this case, web-based learning) and the population (young adolescents in England of above-average ability).

4.9.5 Objectivity

It is important for the current researcher to be particularly sensitive to the issue of objectivity in this study. She is the creator of this intervention and thus must avoid any assumptions that it might be ‘successful’. For this reason data are quantitative and are objectively evaluated. Furthermore, while the researcher developed the intervention, she did not teach on it, thus controlling for possible *experimenter effects* during the intervention. In order to further minimise bias which might have originated from instructor/student contact, the tutor kept her contact with participants to a minimum and standardised their interactions wherever possible, for instance by giving generic feedback, which all students could see. An example of such feedback can be seen in Appendix 13.

Campbell and Stanley (1966) counsel the researcher to consider potential and/or plausible rival research hypotheses and to be aware of pre-existing assumptions. For instance the participants were aware that they were being studied and may have responded, especially in the post-test set of surveys, with what they felt might have been expected of them. This was addressed within the design, which afforded a comparison between experimental groups and the control group.

4.9.6 Treatment integrity

Treatment integrity is addressed in the section that describes the intervention (Section 3.3).

4.9.7 Self-report scales

It should be acknowledged that some authors have cast doubt on the validity of self-report instruments (e.g., Boekaerts and Corno, 2005; Hopfenbeck and Maul, 2011). Such researchers conclude that the results of self-reported surveys, particularly with regard to those that survey students' use of learning strategies, should be treated with caution; they note that such measures should be relied upon only where they are triangulated with other types of measurement mechanisms. Care has been taken in the design of this study to incorporate the recommendations of this body of opinion and therefore the current design includes not only self-report measures (e.g., of self-regulated learning beliefs and competencies), and observed assessments (of critical thinking).

4.10 Summary of Chapter Four

This chapter has defined the study operationally, detailing how the research questions are to be addressed, specifying design, measures, data collection and analysis procedures, ethical considerations and potential methodological limitations. The following chapter presents the methodology and findings of the pilot study.

5 Chapter Five: The Pilot Study

The effects of a dialogic argumentation intervention on aspects of young adolescents' self-regulated learning over time.

5.1 Introduction

This section describes the pilot study. In order to avoid repetition, it should be noted that aspects of the methodology that were duplicated for both pilot and main study, have already been reported. These sections include Chapter Three and Section 4.4, Section 4.5.1, Section 4.5.2, Section 4.5.4, Section 4.6, Section 4.7, Section 4.8 and Section 4.9.

5.2 Pilot study aims

The pilot took place three months before the main data collection, in October 2012. The first aim was to test the structural validity and reliability of the research measures proposed for the main study. This involved factor analysis and reliability checking of the pre- and post-test self-regulated learning measures. The second aim was to explore the effectiveness of the pilot study, in order to inform the design of the main study. Research questions explored the main and experimental effects of the intervention on change in participants' a) self-regulated learning (metacognitive knowledge, efficacy for learning and motivational beliefs and b) critical thinking. A third research question explored the effect of participation in dialogic argumentation on change in critical thinking, and aspects of self-regulated learning.

5.3 Pilot research questions and hypotheses

Research Question One: *What is the effect of the intervention on pre-test to post-test change in the self-regulated learning characteristics (metacognitive knowledge, motivational beliefs and efficacy for learning) of a group of 11 and 12 year old students, after a period of one month, comparing experimental groups with a control group?* It was hypothesised that there would be a main effect of the intervention on change in participants' self-regulated learning characteristics from pre-test to post-test. It was further hypothesised that main effects would be differentiated by experimental condition, such that the most effective would be that experienced by Pilot Intervention Condition 3 (PIC3) (dialogic argumentation + Devil's Advocate + critical thinking tests), followed by PIC2 (Devil's Advocate + critical thinking tests), followed by the Pilot Control Group (PCG) (no intervention activity).

Research Question Two: *What is the effect of the intervention on the repeated measures change in the critical thinking of a group of 11 and 12 year old students, over four critical thinking assessments in a period of one month?* It was hypothesised that there would be a main effect of the intervention on repeated measures change in critical thinking over time. It was further hypothesised that the main effect would be differentiated by experimental condition, such that the most effective experimental condition would be that experienced by PIC3 followed by PIC2.

Research Question Three:

RQ3*a: To what extent does participation in dialogic argumentation predict progress in critical thinking over time?*

RQ3*b: To what extent does participation in dialogic argumentation predict progress in overall self-regulated learning over time?*

RQ3*c: To what extent does participation in dialogic argumentation predict progress in metacognitive knowledge over time?*

RQ3*d: To what extent does participation in dialogic argumentation predict progress in motivational beliefs over time?*

RQ3*e: To what extent does participation in dialogic argumentation predict progress in efficacy for learning over time?*

It was hypothesised that the number of times participants engaged in dialogic argumentation would predict change over time in a) critical thinking and b) aspects of self-regulated learning (metacognitive knowledge, motivational beliefs and efficacy for learning).

5.4 Pilot study design

Two hundred and forty students were initially enrolled into the intervention, a four-week internet-based course, described in Chapter Three. They were randomly assigned into three groups with 80 participants in each group. The final number of participants (N = 164) represents those students who a) supplied consent forms and b) completed the course. The dataset was therefore 'full set' with no missing data on questionnaire items. The final numbers were: PIC3: (dialogic argumentation + Devil's Advocate + critical thinking assessments): 71 students; PIC2 (Devil's Advocate + critical thinking assessments): 41 students and Pilot Control Group (PCG): 52 students. All participants completed the pre- and post-test self-regulated learning electronic questionnaires (described in Section 4.5). There were no missing data on the questionnaires, as all items were required to be completed before each questionnaire was submitted. Table 5-1 summarises the design.

Table 5-1 The pilot study experimental design

<table>
<tr><th>Intervention conditions</th><th rowspan="4">Pre-test SRL</th><th>Week 1</th><th>Week 2</th><th>Week 3</th><th>Week 4</th><th rowspan="4">Post-test SRL</th></tr>
<tr><td>PIC3</td><td>A, B, C</td><td>A, B, C</td><td>A, B, C</td><td>A, B, C</td></tr>
<tr><td>PIC2</td><td>B, C</td><td>B, C</td><td>B, C</td><td>B, C</td></tr>
<tr><td>PCG</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
</table>

Intervention components: A: Dialogic argumentation; B: Devil's Advocate; C: Critical thinking assessment

All students in the two intervention conditions (PIC3 and PIC2) participated in the Devil's Advocate and the critical thinking tests. Additionally, students in PIC3 also had the opportunity to engage in weekday evening debates (dialogic argumentation). Students in the PCG completed only the pre- and post-tests, after which they were given access to the intervention material. This element of the design, discussed further in Section 4.3 was chosen in order not to disadvantage those who were not placed in the experimental groups. Therefore, as a range of hypothesised predictors was to be associated with intervention group membership, a random-assignment, pre-test/post-test, cross-over experimental design was chosen, structured as 3 between-person (group) x 2 within-person (time). This explored the differential effect of two experimental conditions (PIC3 and PIC2), in comparison with a control group (PCG), over two time points. This could evaluate the differentiated intervention components (predictors) associated with group membership, whilst exploring the variables of interest under controlled conditions.

A key objective behind the choice of this design was to rigorously meet criteria for experimental conditions. Thus, individuals were randomly assigned to experimental and control groups; hypotheses informed the design of the experiment; the experiment was designed to be repeatable; since this was a pilot, inference was no more than provisional; a minimum statistical significance of $p < .05$ was set, and hypotheses were constructed having considered all possible outcomes (Fisher, 1935). As further recommended by Campbell and Stanley, (1966) the researcher-controlled experiment was applied uniformly in all experimental conditions; each dependent variable was measured according to scales derived from theoretical perspectives, and inferential statistics were applied. Scales were subject to confirmatory factor analysis and tested for reliability and structural validity (Nunnally and Bernstein, 1994).

5.5 Pilot population and sample

The population, described more fully in Section 4.4, consisted of 11 and 12 year old (Year 7) students. The sample of one hundred and sixty four participants provided consent forms, finished the course and completed the pre- and post-test questionnaires. The 80 participating boys and 84 girls between them attended 100 secondary state schools in England.

5.6 Pilot measures

5.6.1 Self-regulated learning

The self-regulated learning measures featured pre- and post-test self-reported metacognitive knowledge, efficacy for learning and motivational beliefs. The scales, which are described more fully in Section 4.5.1, were:

a) Motivational beliefs, a unidimensional scale adapted from Rheinberg, Vollmeyer and Burns (2001)
b) Metacognition, a composite of two subscales (Elaboration Strategies and Control Strategies) taken from the Student Approaches to Learning scale, used in the PISA studies, a global ranking of education performance (Marsh *et al.,* 2006a)
c) Efficacy for learning (a unidimensional scale created by the researcher for this study)
d) The aggregated self-regulated learning scale, created from the three scales above

e) A count of the number of posts each participant in PIC3 made to evening debates.

Reliability analysis and exploratory factor analysis (maximum likelihood estimation) was conducted on each of the separate self-regulated learning scales and on the aggregated scale. All factor loadings were significant and greater than .4. Table 5-2 reports the results of the exploratory factor analysis and details the Kaiser Meyer Olkin measure, which reports on the adequacy of the sample size for factor analysis (Kaiser, 1974), the amount of variance explained by the factors and the reliability measure in term's of Cronbach's α (Cronbach, 1951).

Table 5-2 Exploratory factor analysis and reliability of pilot SRL measures

Measure	KMO	Variance explained	Cronbach's α
Pre-test MK	.81	61%	.81
Pre-test EfL	.80	65%	.82
Pre-test MB	.89	66%	.89
Pre-test SRL (aggregated)	.90	64%	.93
Post-test MK	.84	66%	.87
Post-test EfL	.81	69%	.84
Post-test MB	.87	68%	.90
Post-test SRL (aggregated)	.89	64%	.94

SRL: self-regulated learning (aggregated scores); MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning

Confirmatory factor analysis, conducted using MPlus, confirmed the factor structure of the scales, with goodness of fit analysis (Table 5-3) confirming that the models fitted the data quite well.

Table 5-3 Confirmatory factor analysis of SRL measures (pilot)

Measure	$\chi 2$ (*df*)	RMSEA	CFI	TLI
Pre-test metacognitive knowledge	33.950 (14)	.093	,952	.903
Post-test metacognitive knowledge	22.656 (12)	.074	.983	.959
Pre-test efficacy for learning	18.736 (11)	.065	.980	.961
Post-test efficacy for learning	15.931 (11)	.052	.990	.981
Pre-test motivational beliefs	64.949 (25)	.099	.946	.922
Post-test motivational beliefs	25.611 (21)	.037	.995	.991
Pre-test SRL	387.925 (233)	.064	.920	.905
Post-test SRL	433.053 (223)	.076	.915	.895

5.6.2 Critical thinking

The critical thinking measure, described in Section 4.5.1, was a series of four parallel assessments, adapted from the DIT2 Defining Issues Test (Rest *et al.,* 1999). The pilot study provided an opportunity to perform reliability testing. Each test provided a sum total only, rather than separate item scores, so item analysis was not available. However, *Alternate forms reliability* (Erwin, 2000) was conducted to evaluate whether the four tests were parallel to each other. This was achieved by varying the order in which the tests were administered. To do this, all students who took the assessments (N = 112) were randomly allocated into four sets, with each set taking the tests in a different order. Thus Set One took the tests in the order of 1,2,3,4; Set Two in the order of 2,3,4,1; Set Three in the order of 3,4,1,2 and Set Four in the order of 4,1,2,3. Analysis of variance, reported in Table 5-4, confirmed that the tests were not significantly different to each other, at each of the four time points.

Table 5-4 Alternate forms reliability (critical thinking tests)

Critical thinking tests	Order of Test	M	SE
End of Week One *M:* 12.60; *SD:* 2.24 F (3, 108) = .702, p = .553	1,2,3,4	12.77	0.48
	2,3,4,1	12.42	0.39
	3,4,1,2	12.97	0.39
	4,1,2,3	12.17	0.46
End of Week Two *M:* 13.38; *SD:* 2.47 F (3, 108) = 1.71, p = .169	1,2,3,4	13.55	0.52
	2,3,4,1	13.85	0.43
	3,4,1,2	13.52	0.43
	4,1,2,3	12.42	0.50
End of Week Three *M:* 13.95; *SD:* 2.24 F (3, 108) = .46, p = .708	1,2,3,4	13.91	0.48
	2,3,4,1	14.06	0.39
	3,4,1,2	14.18	0.39
	4,1,2,3	13.50	0.46
End of Week Four *M:* 14.93; *SD:* 2.24 F (3, 108) = 2.38, p = .074	1,2,3,4	14.18	0.50
	2,3,4,1	14.58	0.41
	3,4,1,2	15.76	0.41
	4,1,2,3	14.96	0.48

5.7 Pilot data collection procedures

Data collection involved gathering data on a) pre- and post-test self-regulated learning measures (metacognitive knowledge, motivational beliefs, efficacy for learning), b) weekly critical thinking assessments and c) a count of the total number of times each participant engaged in dialogic argumentation. The procedures for data collection are the same as those for the main study and are thus described in Section 4.5.

5.8 Pilot data analysis procedures

The statistical software employed for the pilot study was SPSS v. 20.

Research Question One explored the effect of the intervention on pre-test to post-test change in the self-regulated learning characteristics of the participants. Analysis of Covariance (ANCOVA) was appropriate for the design of this study, as, according to Field (2009), it is a powerful technique for estimating the difference between groups at post-test (the dependent variable), whilst partialling out the effect of the pre-test (the covariate). The analysis is described theoretically in 4.7.3 and operationally in Section 5.9.1.

Research Question Two explored the effect of the intervention on change in critical thinking, over four equally spaced parallel assessments. Repeated measures analysis of variance (ANOVA) was chosen for this analysis, since the same students were tested on parallel assessments, at four equally spaced times. Repeated measures ANOVA detects systematic variance, as it significantly partitions unsystematic (error) variance. It is described theoretically in Section 4.7 and operationally in Section 5.9.2.

The independent variable was the group at two levels (IC3 and IC2), and the dependent variable was the repeated measures critical thinking scores.

Research Question Three used regression analysis to explore the predictive effect of the extent to which participants engaged in dialogic argumentation on change in a) critical thinking and b) aspects of self-regulated learning. Change in critical thinking was defined as progress from the first assessment to the last, and was calculated for each student by subtracting the first test score from the last test score, creating a new variable (CT Change). Change in aspects of self-regulated learning was defined as the difference between the pre-test and the post test for each participant, thus creating four new variables (SRL Change; MK Change; MB Change; EfL Change). This research question required data from PIC3 only, the group which engaged in dialogic argumentation.

5.9 Pilot Findings

5.9.1 Pilot RQ1

What is the effect of the intervention on pre-test to post-test change in the self-regulated learning characteristics (metacognitive knowledge, motivational beliefs and efficacy for learning) of a group of 11 and 12 year old students, after a period of one month, comparing experimental groups with a control group?

It was hypothesised that there would be a main effect of the intervention on change in participants' self-regulated learning characteristics from pre-test to post-test. It was further hypothesised that main effects would be differentiated by experimental condition, such that the most effective would be that experienced by Pilot Intervention Condition 3 (PIC3) (dialogic argumentation + Devil's Advocate + critical thinking tests), followed by PIC2 (Devil's Advocate + critical thinking tests), followed by the Pilot Control Group (PCG) (no intervention activity).

The first step in the analysis was to establish that the mean self-regulated learning (SRL) score of the three groups was not significantly different to each other at pre-test. A one-way between-groups analysis of variance (ANOVA) was conducted to explore the effect of the group on the pre-test self-regulated learning scores, finding that there was no significant difference in mean SRL scores between the three groups at pre-test: F (2, 161) = .866; p = .423: PIC3 group (M = 195.01, SD = 21.43, N = 71); PIC2 group (M = 195.85, SD = 26.08, N = 41); PCG (M = 200.48, SD = 23.35, N = 52). This was expected as the groups had been assigned at random.

Similar one-way ANOVAs were performed on each of the pre-test self-regulated learning components, with a summary of results being reported in Table 5-5.

Table 5-5 Summary of one way ANOVAs of pilot pre-test variables

Pre-test Efficacy for learning F (2, 161) = 1.41, p = .247			
Intervention condition	M	SD	N
PIC3	58.28	7.24	71
PIC2	59.41	8.24	41
PCG	60.56	7.05	52
Total	59.29	7.46	164
Pre-test motivational beliefs F (2, 161) = .18, p = .833			
Intervention condition	M	SD	N
PIC3	72.70	8.43	71
PIC2	73.41	10.83	41
PCG	73.79	11.48	52
Total	73.23	10.04	164
Pre-test metacognitive knowledge F (2, 161) = 1.40, p = .250			
Intervention condition	M	SD	N
PIC3	64.03	9.43	71
PIC2	63.02	10.04	41
PCG	66.13	8.61	52
Total	64.45	9.36	164
Pre-test self-regulated learning (aggregated) F (2, 161) = .87, p = .423			
Intervention condition	M	SD	N
PIC3	195.01	21.43	71
PIC2	195.85	26.08	41
PCG	200.48	24.35	52
Total	196.96	23.57	164

PIC3: Pilot intervention group 3; PIC2: Pilot intervention group 2; PCG: Pilot control group

Thus, it was concluded that the group means were not significantly different to each other at pre-test, for each of the self-regulated learning variables (SRL: p = .423; MK: p = .250; MB: p = .833; EfL: p = .247).

A one-way analysis of covariance (ANCOVA) was conducted to estimate the main effect and to compare the effect of membership of the three intervention conditions: PIC3 (dialogic argumentation + Devil's Advocate + critical thinking assessment); PIC2 (Devil's Advocate + critical thinking assessment) and PCG (no intervention activity). The independent variable was the group at three levels, and the dependent variable was the post-test self-regulated learning scores. Participants' scores on the pre-test self-regulated learning scale were used as the covariate in the analysis. The analysis involved data from all pilot intervention participants.

Assumptions which need to be addressed for ANCOVA include normality, outliers, multicollinearity, homogeneity of variance, homogeneity of regression slopes, reliability of the covariate and linearity (Field, 2013). Normality was verified by inspection of skewness and kurtosis values, summarised in Table 5-6, indicating, according to Tabachnick and Fidell (2007), no violation of normality:

Table 5-6 Descriptive statistics for RQ1 (pilot)

	N	M	SD	Skewness	Kurtosis
Pre-test SRL	164	196.96	23.573	-.760	.301
Post-test SRL	164	214.30	21.790	-1.230	1.693
Total	164				

SRL: self-regulated learning (aggregated scores)

Cook's Distance indicates the presence of outliers and influential cases in the data, which might bias the model. Distances less than 1 are acceptable (Larson-Hall, 2010). For these data, the maximum Cook's Distances indicated no violations of this assumption: SRL = .29, EfL = .08; MB = .39; MK = .14. Multicollinearity, which refers to the correlation between independent variables, was not relevant in this analysis, as there was only one independent variable (the group). Homogeneity of variance, an assumption that the variance of the dependent variable is equal across groups, is monitored by inspection of Levene's Test, which should be non-significant. In the current ANCOVA analysis, where the dependent variable was the post-test self-regulated learning and the independent variable was the Group, Levene's Test resulted in a significance of $p > .05$ in self-regulated learning and all of its components, indicating no violation of the equality of variances assumption: (SRL: $p = .982$; MK: $p = .143$; MB: $p = .451$; EfL: $p = .950$). Homogeneity of regression slopes was tested by checking for the significance level of the interaction term (e.g., Group * Pre-test SRL). All significance levels were > .05, indicating that this assumption had not been violated. Covariate reliabilities were all acceptable, with Cronbach's alphas of: Pre-test SRL: $\alpha = .93$; Pre-test MK: $\alpha = .81$; Pre-test EfL: $\alpha = .82$; Pre-test MB: $\alpha = .89$. Linearity between dependent variables and their covariates was confirmed by inspection of scatterplots illustrated in Figure 5-1.

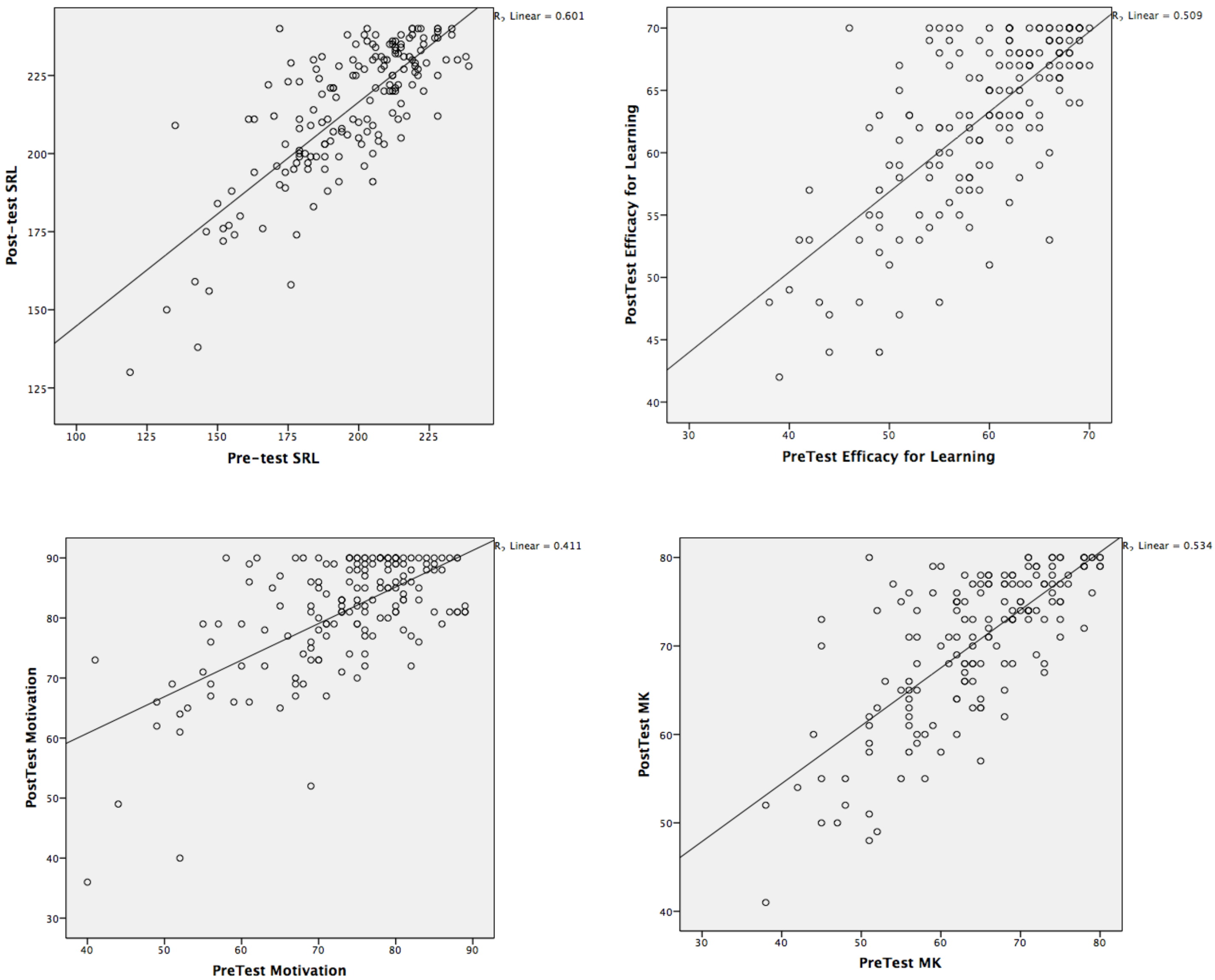


Figure 5-1 Linearity between dependent variables and their covariates for Pilot RQ1 (raw scores)

For the aggregated self-regulated learning variable, there was a main effect for time with a large effect size: F (1, 160) = 327.726; $p < .001$; partial $\eta^2 = .672$; $r = .82$. This effect was differentiated by group, such that 69% of the variance in post-test self-regulated learning, having controlled for pre-test self-regulated learning, could be explained by which experimental group a student was placed in, with a medium effect: F (2, 160) = 23.203; $p < .001$; partial $\eta^2 = .225$; $r = .47$.

Similar ANCOVAs were conducted on the separate elements of self-regulated learning (metacognitive knowledge, motivational beliefs and efficacy for learning), with the main effect results being summarised in Table 5-7 and the group effects in Table 5-8.

Table 5-7 Summary of the main effects of the pilot intervention

Measure	F	df	p	effect sizes
SRL	327.726	1, 160	< .001	$\eta^2 = .67$; $r = .82$
MK	207.316	1, 160	< .001	$\eta^2 = .56$; $r = .75$
MB	139.050	1, 160	< .001	$\eta^2 = .47$; $r = .68$
EfL	200.699	1, 160	< .001	$\eta^2 = .56$; $r = .75$

SRL: self-regulated learning (aggregated scores); MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning

Table 5-8 Summary of the group effects of the pilot intervention

Measure	F	df	p	effect sizes	variance explained
SRL	23.203	2, 160	< .001	$\eta^2 = .225$; $r = .47$	69%
MK	8.926	2, 160	< .001	$\eta^2 = .100$; $r = .32$	58%
MB	15.414	2, 160	< .001	$\eta^2 = .162$; $r = .40$	51%
EfL	10.794	2, 160	< .001	$\eta^2 = .119$; $r = .34$	57%

SRL: self-regulated learning (aggregated scores); MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning

Thus the intervention resulted in a main effect, with a large effect size, and group effects, with medium effect sizes, for all self-regulated learning components.

Post-hoc group comparisons on post-test scores, using Sheffé adjustment with a Bonferoni correction, detected significant differences between PIC3 and PIC2 for motivational beliefs ($p = .006$), efficacy for learning ($p = .001$) and overall SRL ($p <$ .001). Similarly, significant differences were detected between PIC3 and PCG for metacognitive knowledge ($p < .001$), motivational beliefs ($p < .001$), efficacy for learning ($p < .001$) and overall SRL ($p < .001$). However there were no significant differences between IC2 and PCG, for any of the post-test self-regulated learning variables. Thus it appears that membership of the PIC3 group (dialogic argumentation + Devil's Advocate + critical thinking assessments) was more effective than membership of the PIC2 group (Devil's Advocate + critical thinking assessments), when compared with the PCG (the control group).

In sum, analysis of covariance detected a main effect in the post-test, having controlled for the pre-test, with large effect sizes, for all self-regulated learning variables: metacognitive knowledge, motivational beliefs, efficacy for learning and the aggregated self-regulated learning scale. Significant group effects were also detected on all variables, with medium effect sizes in all instances. The group a participant was placed in (PIC3, PIC2 or PCG) accounted for more than 50% of the variance in post-test scores, across all variables. The most effective intervention condition was PIC3 (dialogic argumentation + Devil's Advocate + critical thinking assessments).

5.9.2 Pilot RQ2

What is the effect of the intervention on the repeated measures change in the critical thinking of a group of 11 and 12 year old students, over four critical thinking assessments in a period of one month? It was hypothesised that there would be a main effect of the intervention on repeated measures change in critical thinking over time. It was further hypothesised that the main effect would be differentiated by experimental condition, such that the most effective experimental condition would be that experienced by PIC3 followed by PIC2.

In order to establish parity between the groups at the first assessment, a one-way between-groups ANOVA was conducted, comparing group means on the first critical thinking assessment scores. Participants had been randomly divided into three groups for the whole intervention (PIC3, PIC2 and PCG), but only PIC3 and PIC2 participated in the critical thinking assessments. This was in order to control for a possible method effect whereby the critical thinking assessments themselves could be responsible for improving critical thinking scores (Kurfiss, 1988; Ennis 1993; Maul, 2013). Thus the scores on the first test assessment were as follows: PIC3 (M = 12.35, SD = 2.373, N = 71); PIC2 (M = 13.02, SD = 1.956, N = 41). There was no statistically significant difference between the groups at the first assessment F (1, 110) = 2.362, p = .127. This result was as expected since the groups had been allocated at random and there should have been no significant difference between them.

In order to address this research question a repeated measures analysis of variance was conducted, with critical thinking as the dependent variable, and time (at four time points) as the independent variable. Repeated measures designs must satisfy the assumptions of normality, homogeneity of variance, homogeneity of variance-covariance matrices and sphericity (Field, 2013). Normality of the dependent variable (critical thinking scores) was tested by inspection of skewness and kurtosis values (Table 5-9) and did not indicate any problems with normality.

Table 5-9 Summary statistics for RQ2 (pilot)

	N	M	SD	Skewness	Kurtosis
CT1	112	12.6	2.24	-.14	-.48
CT2	112	13.38	2.47	-.80	1.13
CT3	112	13.95	2.24	-.32	1.09
CT4	112	14.93	2.39	-.09	-.29

Homogeneity of variance, testing whether or not the error variance of the dependent variable (critical thinking scores) was equal across groups, was indicated by Levene's Test of Homogeneity. The significance of p =.117 revealed that there were no issues with homogeneity of variance. However, Levene's Test does not take into account equality of covariances between groups, so an additional test, Box's M Test of Equality of Covariance Matrices, was performed. Box's Test proved non-significant (p = .467), indicating that the assumption had not been violated.

Mauchly's Test of Sphericity (p = .575) indicated that the level of dependence between the four timed conditions was roughly equal, and therefore the assumption of sphericity had not been violated.

Repeated measures ANOVA detected a main effect for time in students' critical thinking scores, with a medium effect size, across the four assessments: F (3, 330) = 17.487, p < .001; partial η^2 = .137; r = 37. Thus the first hypothesis, that the intervention would have an effect on change over time in the critical thinking of the participants, was supported by the data. Contrary to the second hypothesis, this effect was not, however, differentiated by group, with the between-subjects effect being non-significant (p = .169).

An analysis of the separate repeated measures from Time 1 to Time 2, Time 1 to Time 3, and Time 1 to Time 4 revealed an increasing and significant effect for time (from a small, to a medium, to a large effect), but a non-significant effect for group:

CT1 to CT2:

> Repeated measures effect: F (1, 110) = 4.357, p < .05, η^2 = .038, r = .19
> Group effect: F (1, 110) = .141, p = .708

CT1 to CT3:

> Repeated measures effect: F (1, 110) = 18.901, p < .001, η^2 = .147, r = .38
> Group effect: F (1, 110) = 1.191, p = .277

CT1 to CT4:

> Repeated measures effect: F (1, 110) = 54.479, p < .001, η^2 = .331, r = .58
> Group effect: F (1, 110) = 2.944, p = .089.

5.9.3 Pilot RQ3

RQ3a: To what extent does participation in dialogic argumentation predict progress in critical thinking?

RQ3b: To what extent does participation in dialogic argumentation predict progress in self-regulated learning over time?

RQ3c: To what extent does participation in dialogic argumentation predict progress in metacognitive knowledge over time?

RQ3d: To what extent does participation in dialogic argumentation predict progress in motivational beliefs over time?

RQ3e: To what extent does participation in dialogic argumentation predict progress in efficacy for learning over time?

Pilot RQ1 and Pilot RQ2 had identified group differences when exploring change in both critical thinking and aspects of self-regulated learning. Significant group differences appeared to be related to membership of the group which engaged in dialogic argumentation (PIC3). However, not all participants of that group engaged in dialogic argumentation at the same rate, with the number of individual contributions over the course of the intervention ranging from 3 to 25 ($M = 13.06$; $SD = 4.329$). Thus this research question used regression analysis to explore the hypothesis that the *actual extent* to which participants engaged in dialogic argumentation would predict progress in a) critical thinking from the beginning to the end of the intervention and b) aspects of self-regulated learning from pre-test to post-test. Dependent and independent variables are set out in Table 5-10:

Table 5-10 Independent and dependent variables for RQ3 (pilot) regression analyses

	Dependent variable	Independent variable
RQ3a	Change in critical thinking	Contribution to dialogic argumentation
RQ3b	Change in SRL	Contribution to dialogic argumentation
RQ3c	Change in MK	Contribution to dialogic argumentation
RQ3d	Change in MB	Contribution to dialogic argumentation
RQ3e	Change in EfL	Contribution to dialogic argumentation

SRL: self-regulated learning (aggregated scores); MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning

The question made use of data from PIC3 only, the group that engaged in dialogic argumentation. Preliminary correlational analysis (Table 5-11) of the variables of interest revealed that contribution to dialogic argumentation was related change in critical thinking over time and to progress in motivational beliefs from pre-test to post-test. Assumption tests therefore proceeded using both critical thinking (CT) and self-regulated learning (SRL) variables.

Table 5-11 Correlational analysis of RQ3 (pilot) variables

Correlations	Contribution to debate	CT Change	MK change	MB change	EfL change	SRL change
Contribution to debate	1					
CT Change	**.299**	1				
MK Change	-0.095[a]	0.092[a]	1			
MB Change	**.370**	0.146[a]	0.184[a]	1		
EfL Change	0.01[a]	-0.008[a]	**.407**	0.167[a]	1	
SRL Change	0.15[a]	0.127[a]	**.781**	**.694**	**.623**	1

[a] Non-significant values; Significant correlations in bold ($p < .05$).
MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning; SRL: self-regulated learning

Regression analysis requires a number of assumptions to be addressed. These include: normality, sample size, multicollinearity, singularity, linearity, homoscedasticity and independence of residuals. Normality was established by inspection of skewness and kurtosis values, revealing no violation of this assumption (see Table 5-12).

Table 5-12 Descriptive statistics for RQ3 (pilot)

	Predictor variable	Outcome variables				
		RQ3a	RQ3b	RQ3c	RQ3d	RQ3e
	Dialogic argumentation	CT change	SRL change	MK change	MB change	EfL change
N	71	71	71	71	71	71
M	13.06	3.25	23.62	8.17	11.27	4.18
SD	4.33	2.87	12.50	6.96	6.75	3.83
Skewness	0.38	0.56	0.67	0.76	0.56	0.42
Kurtosis	0.56	0.17	0.54	0.37	0.94	-0.32

MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning; SRL: self-regulated learning

Tabachnick and Fidell (2007) suggest that the sample size for regression should be at least 50 + 8*m*, where *m* is the number of predictor variables. In this case, that would involve a sample size of at least 58, since there is only one predictor variable. As the sample size for the predictor variable was 71, it was concluded that it did not present a problem for regression, in this particular analysis.

Field (2009) makes recommendations about the analytics necessary for regression. Cook's distance indicates the presence of influential cases in the data, which might bias the model. Distances less than 1 are acceptable. Inspection of Cook's Distances, reported in Table 5-13 indicates no violation of this assumption on any of the outcome variables.

Table 5-13 Cook's Distances for regression analysis for RQ3 (pilot)

	Outcome variable	Max Cook's Distance	Mean Cook's Distance
RQ3a	Critical thinking change	.202	.019
RQ3b	SRL change	.079	.011
RQ3c	MK change	.074	.013
RQ3d	MB change	.194	.016
RQ3e	EfL change	.058	.011

SRL: self-regulated learning (aggregated scores); MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning

The Normal P-P plot of standardised residuals for each dependent variable approximated a straight line, indicating no major deviations from normality and the scatterplot of each of the standardised residuals was roughly rectangular, which according to Tabachnick and Fidell (*op.cit.*), indicates homoscedasticity (examples are presented in Figure 5-2).

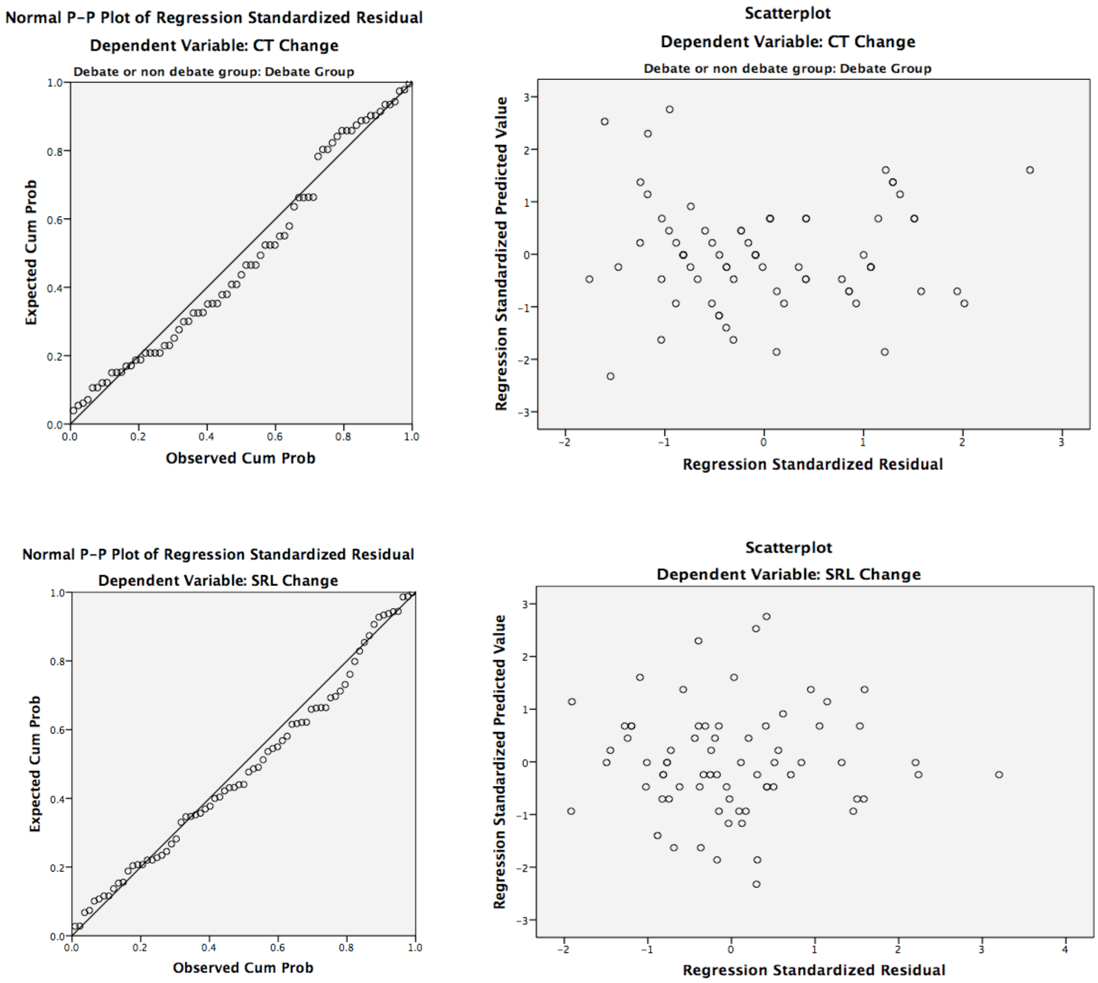


Figure 5-2 Analysis of P-P plots in preparation for regression analysis (pilot)

Standardised residuals, which indicate the level of error in the model, should be no more than |3.29|. Inspection of Table 5-14 reveals no violation of this assumption:

Table 5-14 Standardised residuals for analysis for RQ3 (pilot)

Outcome variable	Minimum	Maximum	Mean
Critical thinking change	-1.759	2.675	.000
SRL change	-1.918	3.2	.000
MK change	-1.647	2.964	.000
MB change	-2.685	2.739	.000
EfL change	-1.866	2.295	.000

SRL: self-regulated learning (aggregated scores); MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning

Tests for independence of residuals and the collinearity diagnostics are reported in Table 5-15. The Durbin-Watson test looks for violation of independence between adjacent residuals. Values close to 2 are acceptable, with those less than 1 or greater than 3 giving cause for concern. None of the Durbin-Watson results gave cause for concern. Collinearity diagnostics did not reveal any problems, with all tolerance figures greater than .10 and VIF figures less than 10:

Table 5-15 Durbin Watson and collinearity diagnostics for analysis for RQ3 (pilot)

Outcome variable	Durbin Watson	Tolerance	VIF
Critical thinking change	2.027	1.000	1.000
SRL change	1.971	1.000	1.000
MK change	1.879	1.000	1.000
MB change	2.290	1.000	1.000
EfL change	1.874	1.000	1.000

SRL: self-regulated learning (aggregated scores); MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning

The regression analyses are summarised in Table 5-16:

Table 5-16 Parameters from regression functions for RQ3 (pilot)

					Constant (intercept)			*Slope*		
Outcome variable	*F (1, 69)*	R^2	*p*	*SE*	*b0*	*SE*	*t val*	*b1*	*SE*	*t val*
CT change	6.786	.09	.011	2.76	.67	1.05	.64	.20	.08	2.61
SRL change	1.579[a]	n/a	.213	n/a	n/a	n/a	n/a	n/a	n/a	n/a
MK change	.631 [a]	n/a	.430	n/a	n/a	n/a	n/a	n/a	n/a	n/a
MB change	10.913	.14	.002	6.32	3.74	2.40	1.56	.58	.174	3.30
EfL change	.007 [a]	n/a	.936	n/a	n/a	n/a	n/a	n/a	n/a	n/a

[a] Non-significant results; n/a: not applicable

In sum, the regression models detected that the count of the number of contributed posts to the evening debates (dialogic argumentation) was a significant predictor of change in critical thinking (with a medium effect size) and motivational beliefs (also with a medium effect size), but not to change in metacognitive knowledge, efficacy for learning or overall self-regulated learning. Thus the hypothesis that critical thinking and self-regulated learning would be influenced by the extent of students' participation in dialogic argumentation was only partially supported by these data.

5.9.4 Summary of Pilot Findings

Findings of Pilot RQ1 suggest that the intervention had a main effect for change over time in students' self-reported self-regulated learning and all of its components, with large effect sizes ($.68 < r < .82$). Furthermore, participants who had the opportunity to engage in dialogic argumentation (i.e., were members of the PIC3 group) made significantly greater gains in their self-reported self-regulated learning than those who did not (PIC2). Again, this pattern was repeated for all self-regulated learning variables, with medium effect sizes ($.32 < r < .47$), with the largest effect being detected on participants' motivational beliefs.

Pilot RQ2 findings suggest that there was a main effect for change over time, with a medium effect size ($r = .37$), in students' assessed critical thinking scores. However, that effect was not differentiated by group; there was no significant difference in gains in critical thinking comparing students who had the opportunity to engage in dialogic argumentation (the PIC3 group) with those who did not (PIC2). This suggests that those who completed the Devil's Advocate activity made as much progress in critical thinking as those who additionally had the opportunity to engage in evening debates, and that membership of the group which engaged in those debates did not have a significant effect on critical thinking. It is also possible that the practice of taking the critical thinking tests was responsible for the improvement in critical thinking. This phenomenon has been reported elsewhere (Ennis, 1993, Kurfiss, 1998) and will be explored further in the main study.

Pilot RQ3 explored the effect of actual participation in dialogic argumentation, finding that the *extent* of students' participation (as opposed to membership of the group which had the *opportunity* to engage in dialogic argumentation) predicted change in both critical thinking (r = .30) and motivational beliefs (r = .37), but not in other aspects of self-regulated learning.

In sum, the pilot intervention had a significant main effect on self-regulated learning with a large effect size, and on critical thinking with a medium effect size. Self-regulated learning gains were further differentiated by group, with a medium effect size, such that participants who engaged in dialogic argumentation made significantly greater gains in their self-regulated learning compared with those who had not experienced the dialogic argumentation activity. The extent of participation in dialogic argumentation appeared to influence change in critical thinking and motivational beliefs. It was concluded that these findings merited further investigation in the main study.

6 Chapter Six: The Main Study Findings

6.1 Introduction

This chapter details the findings of the main study. These relate to the exploration of:

- The main and group effect of the intervention on change over time in critical thinking.
- The effect of specific participation in dialogic argumentation on change over time in critical thinking.
- The main effect of the intervention on change over time in aspects of self-regulated learning (metacognitive knowledge, motivational beliefs and efficacy for learning).
- The experimental (group) effect of the intervention on change over time in aspects of self-regulated learning.
- The effect of specific participation in dialogic argumentation on change over time in aspects of self-regulated learning.
- A process analysis of the relationships between metacognitive skills (including critical thinking) and metacognitive experiences, over four learning events.
- An analysis of the relationship between self-regulated learning dispositions (at both pre- and post-test) and SRL processes (critical thinking and metacognitive activity reported in intervention learning journals).

6.1.1 Summary of research questions

Intervention effects on critical thinking outcomes

RQ1: *What are the main and experimental condition effects of the intervention on change over time in participants' critical thinking?*

RQ2: *To what extent does participation in dialogic argumentation predict progress in critical thinking over time?*

Intervention effects on aspects of self-regulated learning

RQ3: *What is the main effect of the intervention on change in aspects of participants' self-regulated learning (metacognitive knowledge, motivational beliefs and efficacy for learning) from pre-test to post-test, on groups of young adolescents, after one month, compared to a group who had no intervention experience?*

RQ4: *What are the effects of the separate experimental conditions on change in self-regulated learning and its components over time?*

RQ5: *To what extent does participation in dialogic argumentation predict change in self-regulated learning and its components over time?*

Relationships between aspects of metacognitive activity over time

RQ6: *What are the relationships between aspects of metacognitive activity and critical thinking over time?*

Relationships between self-regulated learning dispositions and self-regulated learning processes

RQ7*: What is the relationship between self-regulated learning dispositions (metacognitive knowledge, motivational beliefs and efficacy for learning) and self-regulated learning processes (metacognitive activity) during this intervention?*

A range of statistical analyses are employed, and have been described in Chapter Four. Further analytical detail appears within each particular research question reported in this chapter. Analyses include one-way analysis of variance (ANOVA), latent change score modelling (a structural equation modelling technique), repeated measures ANOVA, simple regression and autoregressive cross-lagged panel analysis (a structural equation modelling technique).

6.1.2 Initial description of the sample

Six hundred and seventeen participants provided the data for this study. The sampling process, which is described in Section 4.4, resulted in students being randomly assigned into four groups (three intervention conditions and one control condition). Those four conditions are labelled throughout this study as:

- IC4 (four intervention components): learning journal + dialogic argumentation + Devil's Advocate + critical thinking assessments
- IC3 (three intervention components): dialogic argumentation + Devil's Advocate + critical thinking assessments
- IC2 (two intervention components): Devil's Advocate + critical thinking assessments
- CG: control group (no intervention activity).

Table 6-1 details the distribution of males and females amongst the four conditions:

Table 6-1 Group participation

Intervention condition	Male	Female	Total
IC4	70	81	151
IC3	78	77	155
IC2	82	78	160
CG	65	86	151
Totals	295	322	617

6.1.3 Verification of the randomisation process

In order to verify the randomisation process, a one-way between-groups analysis of variance was performed on the pre-test self-regulated learning scores. Descriptive statistics are reported in Table 6-2.

Table 6-2 Descriptive statistics for SRL pre-test

Pre-test	Metacognitive knowledge			Motivational Beliefs			Efficacy for learning		
Condition	Mean	SD	N	Mean	SD	N	Mean	SD	N
IC4	61.65	8.55	125	58.45	6.97	123	58.25	6.50	110
IC3	62.93	8.31	132	58.33	7.07	129	58.21	6.61	131
IC2	62.92	8.46	132	58.12	7.58	129	58.29	6.33	134
CG	63.61	8.01	138	57.86	7.09	138	57.69	6.20	140
Total			527			519			515

Levene's Tests for equality of variance did not report any violations, as reported in Table 6-3.

Table 6-3 Homogeneity of variance tests for pre-test SRL variables

	Levene's Statistic	*df*	*p*
Pre-test MK	1.19	3, 523	.31
Pre-test MB	0.88	3, 515	.45
Pre-test EfL	0.69	3, 511	.56
MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning			

On average, there was no significant difference between the four groups across any of the pre-test self-regulated learning variables, as reported in Table 6-4.

Table 6-4 One-way between groups ANOVA of pre-test SRL variables

	F	*df*	*p*
Pre-test MK	1.25	3, 523	0.29
Pre-test MB	0.17	3, 515	0.91
Pre-test EfL	0.26	3, 511	0.85
MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning			

This lack of significant difference between the groups at pre-test provided supporting evidence of the random allocation to groups. Post-tests are reported in Table 6-5.

Table 6-5 Descriptive statistics for post-test SRL

Post-test	Metacognitive knowledge			Motivational beliefs			Efficacy for learning		
Condition	M	SD	N	M	SD	N	M	SD	N
IC4	69.01	7.58	124	65.42	4.62	124	64.06	5.09	124
IC3	66.61	8.18	127	62.27	7.26	127	61.68	5.40	127
IC2	66.69	8.72	140	60.96	8.38	140	61.02	6.73	140
Control	64.32	8.77	133	59.21	7.91	134	58.87	7.07	135
Total			524			525			526

A one-way analysis of variance (Table 6-6) established that there were significant differences in all aspects of self-regulated learning, between intervention conditions, at post-test.

Table 6-6 One-way between groups ANOVA of post-test SRL variables

	F	*df*	*p*
Post-test MK	6.75	3, 520	0.001
Post-test MB	16.80	3, 521	0.001
Post-test EfL	15.48	3, 522	0.001

MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning

This potential intervention effect is explored further in RQ3 and RQ4.

6.2 Research Question One

What are the main and experimental condition effects of the intervention on change over time in participants' critical thinking?

This research question explored the effects of the intervention, on progress in critical thinking, across four assessments. It employed data from Experimental Conditions IC4, IC3 and IC2 only, the groups that took the weekly critical thinking tests, and explored the following hypotheses:

a) It was hypothesised that there would be a repeated measures effect in terms of students' progress in critical thinking across four equally spaced critical thinking assessments.
b) It was further hypothesised that the main effect would be differentiated by intervention condition, such that the most effective experimental condition was predicted to be IC4 (critical thinking tests + Devil's Advocate + dialogic argument + learning journal), followed by IC3 (critical thinking tests + Devil's Advocate + dialogic argument), followed by IC2 (critical thinking tests + Devil's Advocate).

The analysis used data from the experimental groups, the students who took the critical thinking tests:

IC4: Learning journal + dialogic argumentation + Devil's Advocate + critical thinking test
IC3: Dialogic argumentation + Devil's Advocate + critical thinking test
IC2: Devil's Advocate + critical thinking test

Critical thinking measures are described in 4.5.1. Students took one critical thinking test at the end of each week of the intervention. Thus each student had four critical thinking scores. In this analysis, critical thinking was the dependent variable, whilst time (at four time points) was the independent variable. Repeated measures analysis of variance was used for this analysis, since the same students were tested on parallel assessments, at four equally spaced times. This, according to Field (2013), is a powerful technique for detecting systematic variance, as it significantly reduces the unsystematic (error) variance, (since the same participants are used across conditions). The theoretical procedure is described in 4.7.4.

Before conducting formal analyses, and to satisfy the assumptions for repeated measures analysis of variance, summary statistics means, standard deviations and normality tests were performed (Table 6-7 and Table 6-8). No serious violations of missingness were noted (as reported in Section 4.7.1). Sphericity tests were conducted as part of the repeated measures analysis.

Normality of data on the dependent variables was tested by inspection of skewness and kurtosis values (Table 6-7), neither of which indicated violation of normality.

In order to confirm random assignment, a one-way analysis of variance was conducted on the first set of scores. Here it was established that there was no significant difference between the three intervention groups at the first critical thinking assessment $F\ (2, 451) = 1.675$, $p = .188$.

Table 6-7 Summary statistics for RQ1

	CT1	CT2	CT3	CT4
N	454	455	454	453
M	12.47	13.04	14.09	14.68
SD	2.834	2.462	2.209	2.517
Skewness	-0.821	-0.627	-0.529	-0.54
Kurtosis	1.628	1.241	0.383	0.895

Table 6-8 Mean assessment scores for critical thinking, by intervention group

Test	Group	M	SD	N
CT test	IC4	12.13	3.284	149
Week One	IC3	12.61	2.599	147
	IC2	12.67	2.560	158
CT test	IC4	12.97	2.472	147
Week Two	IC3	13.20	2.317	150
	IC2	12.95	2.591	158
CT test	IC4	14.05	2.315	147
Week Three	IC3	14.06	2.334	149
	IC2	14.16	1.989	158
CT test	IC4	14.87	2.430	146
Week Four	IC3	14.76	2.684	150
	IC2	14.44	2.429	157

In the repeated measures analysis, Mauchly's test of Sphericity was violated, $\chi^2(5) = 18.579$, $p < .01$; therefore degrees of freedom were corrected using Greenhouse Geisser estimates of sphericity ($\varepsilon = .97$). There was a significant effect for time, with a medium effect size in critical thinking: F (2.922, 1271.231) = 83.324; $p < .001$; partial $\eta^2 = .161$; $r = .40$. However, there was no significant effect for group: F (2, 435) = .699, $p = .498$. This suggested that whilst critical thinking improved over time for all students, that increase was not differentiated by group. This latter finding is described in Figure 6-1, which shows groups progressing at a similar rate:

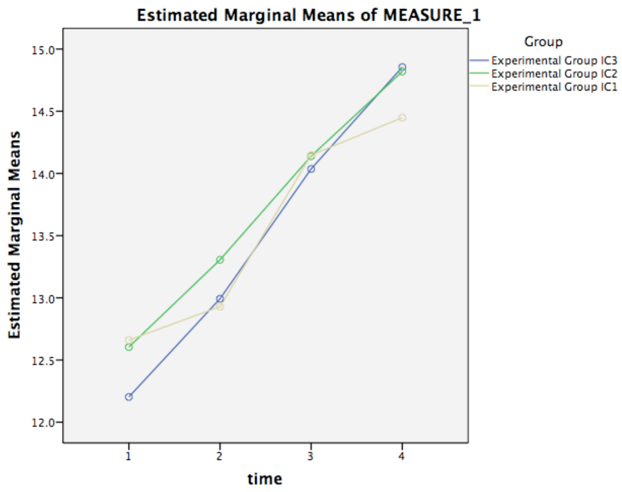


Figure 6-1 Repeated measures analysis of intervention group CT scores over time

The hypothesis that there would be a repeated measures effect of the intervention, on critical thinking, was therefore supported by these data, with a medium effect size (r = .40, $p < .001$). However, the hypothesis that there would also be a group effect was not upheld.

An analysis of the separate repeated measures between each time point, revealed an increasing effect for time, but a decreasing and non-significant effect for group:

CT1 to CT2:

Repeated measures effect: F (1, 444) = 13.451, $p < .001$ $r = .17$
Group effect: F (2, 444) = 1.420, $p = .243$

CT1 to CT3:

Repeated measures effect: F (1, 444) = 112.564, $p < .001$, $r = .45$
Group effect: F (2, 444) = .975, $p = .378$

CT1 to CT4:

Repeated measures effect: F (1, 443) = 195.234, $p < .001$, $r = .55$
Group effect: F (2, 443) = .392, $p = .676$

In sum, repeated measures analysis of variance detected a repeated measures effect of the intervention, across four assessment events, in participants' critical thinking scores. The result was statistically significant and represented a medium effect size. This effect was however, not differentiated by group.

All experimental groups (IC4, IC3 and IC2) completed the critical thinking tests. The above analysis estimated the effect of intervention condition on change in critical thinking over time. However, participants in IC4 and IC3 engaged with the dialogic argumentation activity at different individual rates. Therefore, the following research question (RQ2) explores the effect of actual participation in dialogic argumentation, (rather than group membership) on change in critical thinking.

6.3 Research Question Two

To what extent does participation in dialogic argumentation predict progress in critical thinking over time?

Regression analysis was used to explore the hypothesis that the extent to which participants engaged in dialogic argumentation would predict progress in their critical thinking, with progress being defined as the increase from the baseline test scores to the final assessment scores. The question required data from IC4 and IC3, the two groups that engaged in dialogic argumentation.

The dependent variable was the progress from the first critical thinking assignment (CT1) to the final assignment (CT4). *Progress* is defined in English educational policy as the difference in attainment between two fixed points (DfE, 2007). Therefore the critical thinking progress variable was created by subtracting CT1 from CT4 for each participant. The independent variable was the count of the number of times participants a) actively and b) passively participated in the dialogic argumentation activity. Active participation was measured by the total count of the number of times a participant contributed a post to a discussion. Passive participation was measured by the total number of times a participant browsed a discussion without actually contributing to it.

Preliminary correlational analysis (Table 6-9) revealed that dialogic argumentation, but not passive browsing of dialogic argumentation, was likely to predict progress in critical thinking. Therefore assumption tests and regression analysis proceeded using the single potential predictor variable of active participation in dialogic argumentation.

Table 6-9 Correlations between progress in critical thinking and intervention activity

	Correlation with progress in critical thinking
Total active contribution to dialogic argumentation	.22**
Total browsing discussions	0.08[a]
[a]Non-significant correlations; Significant values: ** $p < .01$	

Regression is sensitive to a number of assumptions including normality, sample size, multicollinearity, singularity, linearity, homoscedasticity and independence of residuals. Descriptive analysis of both dependent and independent variables (Table 6-10) revealed no violations of the assumption of normality:

Table 6-10 Descriptive statistics for RQ2

	Active dialogic argumentation	Progress in critical thinking
N	130	290
Mean	16.38	2.33
Std. Deviation	8.168	3.156
Skewness	.353	0.003
Kurtosis	-.232	-.499

Tabachnick and Fidell (2007) suggest that the sample size should be at least 50 + 8*m*, where *m* is the number of predictor variables. In this case, that would involve a sample size of at least 58, since there is only one predictor variable. As the sample size for the predictor variable was 130, it was concluded that it did not present a problem for regression, in this particular analysis.

The Normal P-P plots of standardised residuals approximated a straight line, indicating no major deviations from normality, and the scatterplot of the standardised residuals was roughly rectangular, which according to Tabachnick and Fidell (*op.cit.*), indicates homoscedasticity (Figure 6-2).

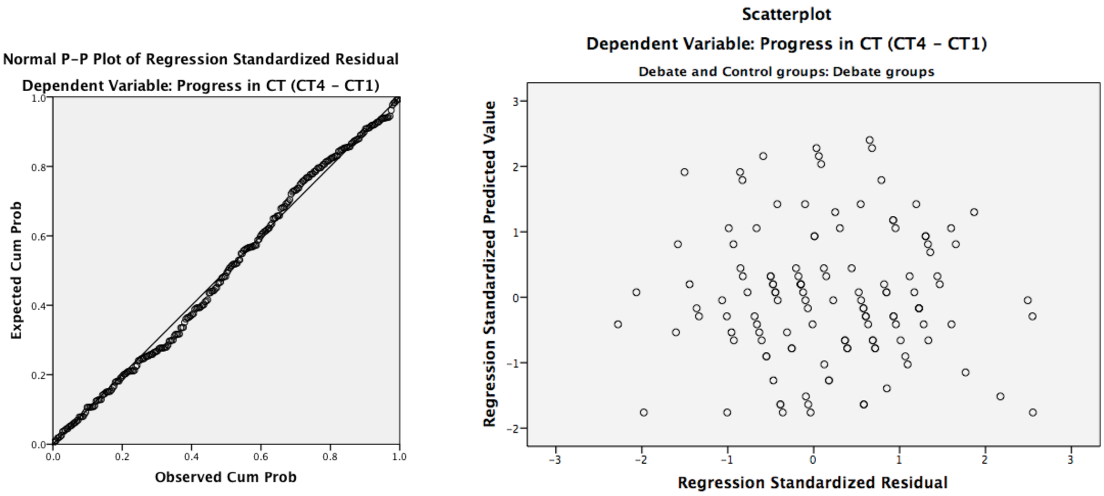


Figure 6-2 Analysis of residuals in preparation for regression analysis

Field (2009) makes recommendations about the analytics necessary for regression. Cook's distance indicates the presence of influential cases in the data, which might bias the model. Distances less than 1 are acceptable. In the case of the analysis for this research question, the maximum Cook's distance was .111, with the mean being .007, indicating no violation of this assumption. Standardised residuals, which indicate the level of error in the model, should be no more than |3.29|. Here, standardised residuals ranged from -2.278 to 2.551, with a mean of .186, indicating no violation of this assumption. The Durbin-Watson test looks for violation of independence between adjacent residuals. Values close to 2 are acceptable, with those less than 1 or greater than 3 giving cause for concern. Here, the value of 2 suggested no violation of independence of error terms. Collinearity diagnostics did not reveal any problems of collinearity, as tolerance figures were greater than .10 and VIF figures were less than 10.

The simple regression model detected that 4.8% of the variance in critical thinking progress could be accounted for by participation in dialogic argumentation, $F(1, 127) = 6.382$, $p < .05$, $r = .22$, a small effect size. Critical thinking (CT) progress could be predicted from participation in dialogic argumentation by the following equation: Predicted CT progress = .944 + (.085 x participation in dialogic argumentation), $R^2 = .048$.

In sum, the hypothesis that critical thinking progress scores would be influenced by the extent of students' participation in dialogic argumentation was supported by these data. The extent to which a student engaged in dialogic argumentation (i.e., the total number of posts to debates during the course of the intervention) predicted progress (change) in their critical thinking scores, with a small effect size ($r = .22$).

6.4 Research Question Three

What is the main effect of the intervention on change in participants' self-regulated learning dispositions (metacognitive knowledge, motivational beliefs and efficacy for learning) from pre-test to post-test, on groups of young adolescents, after one month, compared to a group who had no intervention experience?

This research question explored the main effect of the intervention, in terms of change from pre-test to post-test, on the self-reported self-regulated learning of the participants, comparing the experimental groups with the control group. The effect was estimated on change in a) aggregated self-regulated learning scores, b) metacognitive knowledge, c) motivational beliefs and d) efficacy for learning.

It was hypothesised that there would be a main effect of the intervention on change in self-regulated learning and its separate components, when comparing the experimental groups (IC4, IC3 and IC2 together) with the control group, in accordance with the hypothesised model described in Figure 6-3.

This research question used data from all four intervention groups. Groups IC4, IC3 and IC2 were exposed to the intervention activities; whilst CG (the control group) experienced the intervention activities only after the main data collection was complete. The measures are described in detail in Section 4.5.2, and consisted of pre- and post-test metacognitive knowledge, motivational beliefs, efficacy for learning and a self-regulated learning variable which was the aggregate of these three components. The dependent variable for this research question was the latent change in participants' self-reported self-regulated learning scores, whilst the independent variable was the group at two levels (experimental and control).

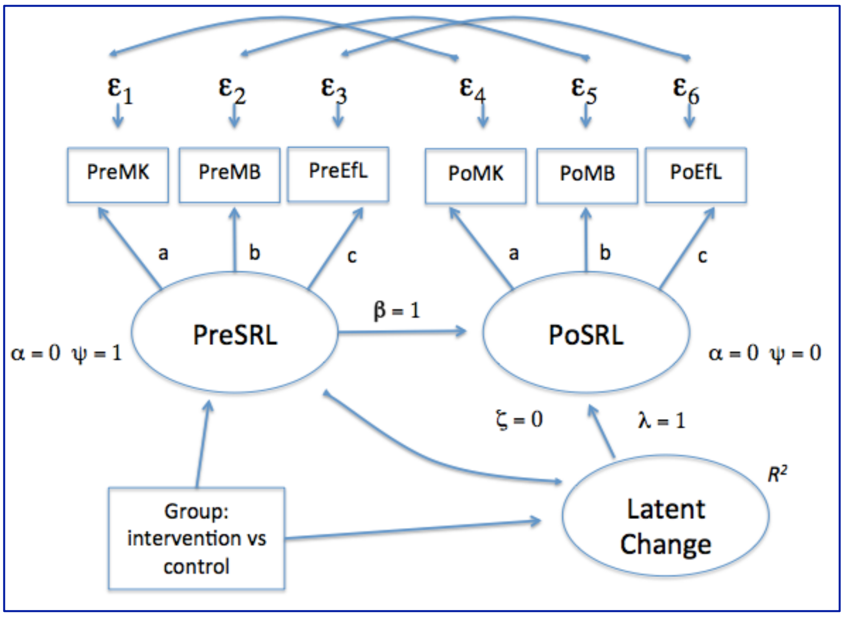


Figure 6-3 Hypothesised LCSM model: main intervention effect on change in SRL
α Latent intercept; ζ disturbance (endogenous); ψ latent variance (dependent variable); λ loading; ε error terms

The method of analysis used for this research question is latent change score modelling (McArdle, 2009; McArdle and Prindle, 2008; Duncan, Duncan and Stryker, 2006; Mun, Von Eye and White, 2009). The latent change score hypothesised model is specified according to the model described in Section 4.7.5.1, with corresponding parameters constrained in the same way. The intervention effect is explored by regressing the self-regulated learning change scores on the predictor variable, which in this case is the dichotomous variable Group (experimental v control).

Before conducting formal analyses, and to satisfy necessary assumptions for SEM, summary statistics, means, standard deviations and normality tests were performed (Table 6-11).

Table 6-11 Correlations and summary statistics for RQ3

	PreMK	PoMK	PreMB	PoMB	PreEfL	PoEfL
PreMK	1					
PoMK	**0.46**	1				
PreMB	**0.50**	**0.31**	1			
PoMB	**0.31**	**0.60**	**0.48**	1		
PreEfL	**0.65**	**0.35**	**0.48**	**0.28**	1	
PoEfL	**0.39**	**0.68**	**0.35**	**0.62**	**0.53**	1
N	527	524	519	525	515	526
M	62.8	66.62	58.18	61.88	58.1	61.34
SD	8.333	8.481	7.164	7.566	6.39	6.414
Skewness	-0.646	-0.865	-0.900	-1.261	-0.772	-0.867
Kurtosis	0.179	0.908	0.822	1.390	0.650	0.419

MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning
correlations in bold are significant at $p < .001$

Correlational relationships between all SRL variables were found to be positive and statistically significant ($p < .001$).

Normality was tested by inspection of skewness and kurtosis values, which indicated no violation of normality. Temporal dependence was confirmed (group allocation preceded pre-test, which preceded post-test) and no serious violations in terms of missing data were noted (as reported in Section 4.7.1) so the latent change score model proceeded using maximum likelihood (ML) estimates.

6.4.1 The measurement model

Confirmatory factor analysis (CFA), the first (measurement) stage of structural equation modelling (SEM), was first carried out on pre- and post-test self-regulated learning scales. It was conducted to a) explore hypothesised structures, b) reduce scales down to their most parsimonious, c) overcome possible collinearity problems which might affect SEM analyses and d) establish robust and stable constructs for structural equation modelling. Each manifest indicator was constrained to load onto the hypothesised factor, and the first factor loading of each factor was constrained to 1, leaving other parameters free to be estimated. Errors were allowed to correlate (Kenny, 1976; Marsh, 1989) allowing dependencies between measurement errors of corresponding responses over time.

CFA was used to estimate the measurement models for the pre-to post-test analyses of metacognitive knowledge, motivational beliefs, efficacy for learning and aggregated SRL measures. Goodness of fit indices, the cut-offs for which are explained in Section 4.7.5, report the extent to which the measurement model represents the data. All measurement models provided a good fit to the data (Table 6-12).

Table 6-12 CFA of SRL measures for RQ3

Scale	$\chi2$ (*df*)	RMSEA	CFI	TLI
Pre-MK	26.964 (12)	.048	.989	.974
Pre-MB	35.329 (17)	.045	.990	.979
Pre-EfL	13.214 (8)	.035	.993	.982
Pre-test SRL	492.647 (227)	.047	.944	.932
Post-MK	10.575 (8)	.025	.999	.995
Post-MB	52.303 (25)	.045	.988	.982
Post-EfL	15.154 (11)	.027	.996	.992
Post-test SRL	461.186 (232)	.043	.962	.954

MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning SRL: self-regulated learning (MK + MB + EfL)

6.4.2 Latent change in self-regulated learning

The latent change score structural equation model (previously described in Figure 6-3) was employed to test the hypothesised model, such that the self-regulated learning latent change scores were regressed on the group (experimental or control). The estimated model for the aggregated SRL scale (Figure 6-4) provided a good fit to the data: χ^2 (13) = 27.259; RMSEA= 0.043; CFI= 0.992; TLI= 0.987.

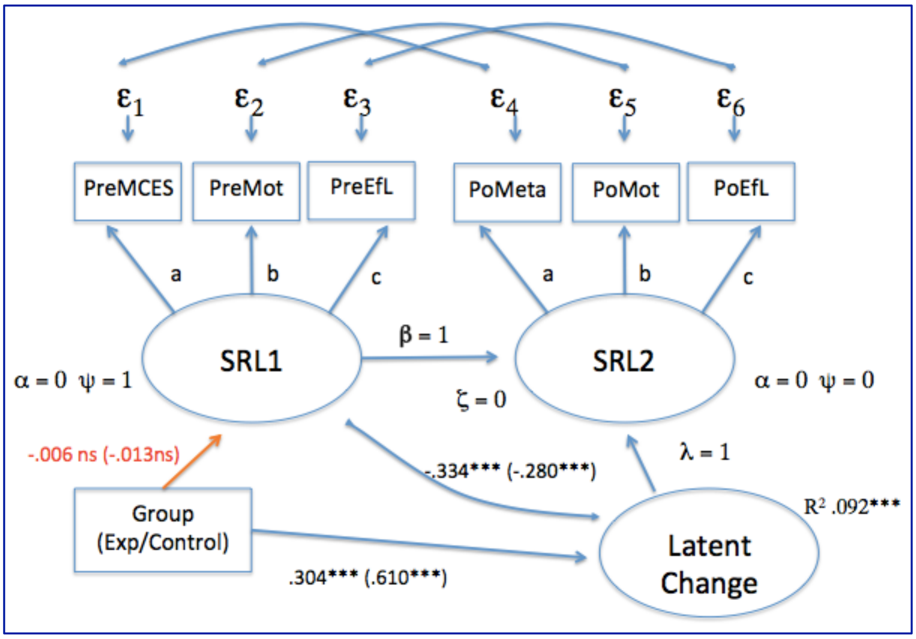


Figure 6-4 Estimated LCSM model: main intervention effect on SRL

Standardised estimates (unstandardised in parentheses); α Latent intercept; ζ disturbance (endogenous); ψ latent variance (dependent variable); λ loading; red connectors denote non-significant paths

There was no significant difference between experimental and control groups at pre-test, denoted by the non-significant path from Group to SRL at Time 1. This was expected, as the groups had been assigned at random.

The main effect of the intervention, on change from SRL1 to SRL2, was indicated by a mean latent change score of .22, which was statistically significant at $p = .009$. The residual variance of the latent change score indicated that 70% of the latent change from SRL1 to SRL2 could be accounted for by the intervention.

There was a significant group effect (experimental vs control) on change scores ($\beta = .304$). The group a participant was placed in (either experimental or control) accounted for 9.2% of the variance in SRL change scores, with a medium effect size ($r = .30$). It was also noted that the higher the students' initial SRL score, the less the change from pre-test to post-test. This was represented by a negative correlation of ($r = -.334$, $p < .001$).

Similar latent change score models were constructed and estimated for latent change in the separate components of self-regulated learning: metacognitive knowledge, motivational beliefs and efficacy for learning. These are reported in the following three sub-sections.

6.4.3 Latent change in metacognitive knowledge

The latent change score self-regulated learning model (previously seen in Figure 6-3) was adapted to estimate the intervention effect on latent change in metacognitive knowledge (MK). The estimated model (Figure 6-5) provided a good fit to the data: χ^2 (100) = 225.782; RMSEA= 0.047; CFI= 0.964; TLI= 0.951.

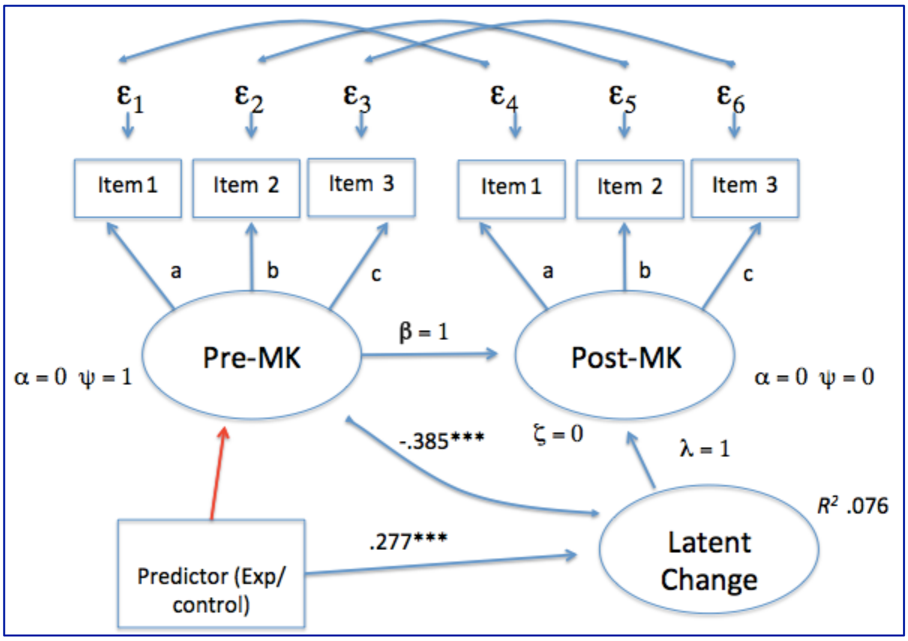


Figure 6-5 Estimated LCSM model: main intervention effect on metacognitive knowledge Standardised estimates; α Latent intercept; ζ disturbance (endogenous); ψ latent variance (dependent variable); λ loading; error terms omitted for clarity; red connectors denote non-significant paths

There was no significant difference between experimental and control groups at pre-test, as the path from Exp/Control Group to pre-test MK at Time 1 was non-significant. This was expected, as the groups had been assigned at random.

The main effect of the intervention, on change from MK1 to MK2, was indicated by a mean latent change score of .14, which was statistically significant at p = .013. The residual variance of the latent change score indicated that 65% of the latent change from MK1 to MK2 could be accounted for by the intervention.

There was also a significant group effect on change scores ($\beta = .277$, $p < .001$), with a small effect size ($r = .28$). The group a participant was placed in (either experimental or control) accounted for 7.6% of the variance in metacognitive knowledge change scores.

It was also noted that the higher the students' initial metacognitive knowledge score, the less the change from pre-test to post-test. This is represented by a negative correlation of $r = -.385$.

6.4.4 Latent change in motivational beliefs

The latent change score self-regulated learning model was further adapted to estimate the intervention effect on latent change in motivational beliefs (MB). The estimated model (Figure 6-6) provided a good fit to the data: χ^2 (141) = 420.983; RMSEA= 0.058; CFI= 0.941; TLI= 0. 929.

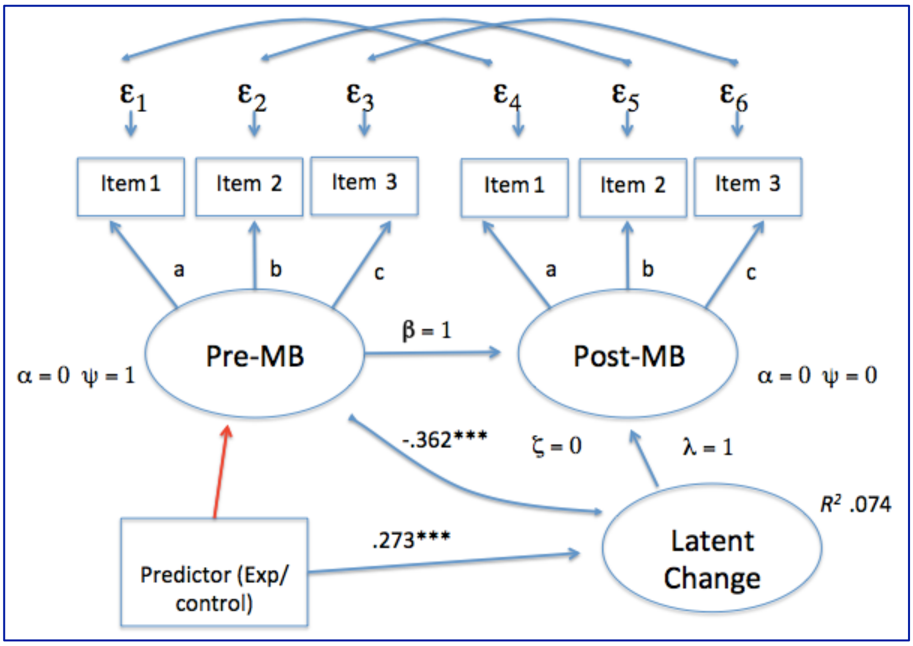


Figure 6-6 Estimated LCSM model: main intervention effect on MB
Standardised estimates; α Latent intercept; ζ disturbance (endogenous); ψ latent variance (dependent variable); λ loading; error terms omitted for clarity; red connectors denote non-significant paths

There was no significant difference between experimental and control groups at pre-test, as the path from Exp/Control Group to pre-test MB at Time 1 was non-significant. This was expected, as the groups had been assigned at random.

The main effect of the intervention, on change from MB1 to MB2, was indicated by a mean latent change score of .26, which was statistically significant at p = .001. The residual variance of the latent change score indicated that 68% of the latent change from MB1 to MB2 could be accounted for by the intervention.

However, there was a significant group effect on change scores, with a small effect size (β = .273). This was also expected since meta-analysis has revealed that most self-regulated learning interventions result in an improvement in motivational beliefs (Dignath, Buettner and Langfeldt, 2008). Here, the group a participant was placed in (either experimental or control) accounted for 7.4% of the variance in motivational beliefs change scores.   It was also noted that the higher the students' initial motivational beliefs score, the less the change from pre-test to post-test. This is represented by a negative correlation of (r = −.362).

6.4.5 Latent change in efficacy for learning

The latent change score self-regulated learning model was further adapted to estimate the intervention effect on latent change in efficacy for learning (EfL). The model (Figure 6-7) provided a good fit to the data: χ^2 (84) = 162.698; RMSEA= 0.040; CFI= 0.969; TLI= 0.961.

There was no significant difference between experimental and control groups at pre-test, as the path from Exp/Control Group to pre-test EfL at Time 1 was non-significant. This was expected, as the groups had been assigned at random.

The main effect of the intervention, on change from EfL1 to EfL2, was indicated by a mean latent change score of .25, which was statistically significant at p = .006. The residual variance of the latent change score indicated that 62% of the latent change from EfL1 to EfL2 could be accounted for by the intervention.

However, there was a significant group effect on change scores, with a small effect size (β = .257). The group a participant was placed in (either experimental or control) accounted for 6.6% of the variance in efficacy for learning change scores.

It was also noted that the higher the students' initial EfL score, the less the change from pre-test to post-test. This is represented by a negative correlation of (r = −.431).

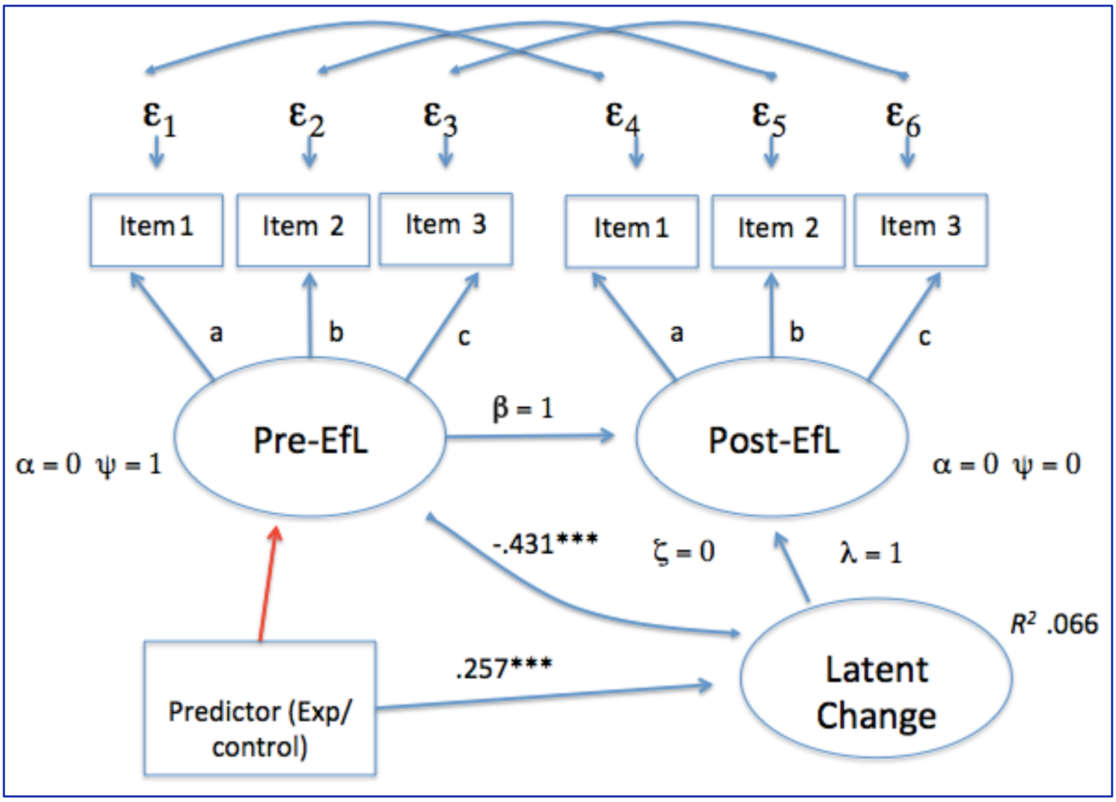


Figure 6-7 Estimated LCSM model: main intervention effect on efficacy for learning

Standardised estimates; α Latent intercept; ζ disturbance (endogenous); ψ latent variance (dependent variable); λ loading; error terms omitted for clarity; red connectors denote non-significant paths

Summary of Research Question Three

This research question explored the main effect of the intervention on change over time in aspects of participants' self-regulated learning. Small and medium effects were detected: aggregated SRL: r = .30; metacognitive knowledge: r = .28; motivational beliefs: r = .27; efficacy for learning: r = .26.

Since a significant difference was detected between experimental groups and control groups in terms of time-ordered change in their self-regulated learning scores and in all of the self-regulated learning components, it followed that an exploration of the separate experimental conditions would lead to a finer grained analysis of the intervention effect. That effect is explored in the following research question.

6.5 Research Question Four

What are the effects of the separate experimental conditions on change in self-regulated learning and its components over time?

This research question explored the effects of the three experimental conditions (IC4, IC3 and IC2), in terms of change from pre-test to post-test, on the self-reported self-regulated learning of the participants, comparing each of the experimental conditions (groups) with the control group. The question specifically explored these effects on a) aggregated self-regulated learning, b) metacognitive knowledge, c) motivational beliefs and d) efficacy for learning, as illustrated in the hypothesised model featured in Figure 6-8.

It was hypothesised that there would be a differentiated effect of the intervention conditions when comparing each of the experimental conditions with the control condition, with the most effective experimental condition predicted to be that experienced by the IC4 group, followed by IC3 and followed by IC2.

This research question used data from all four group conditions, which were:

> **IC4:** Learning journal + dialogic argumentation + Devil's Advocate + critical thinking test
>
> **IC3:** Dialogic argumentation + Devil's Advocate + critical thinking test
>
> **IC2:** Devil's Advocate + critical thinking test
>
> **CG:** Control group (no intervention activity)

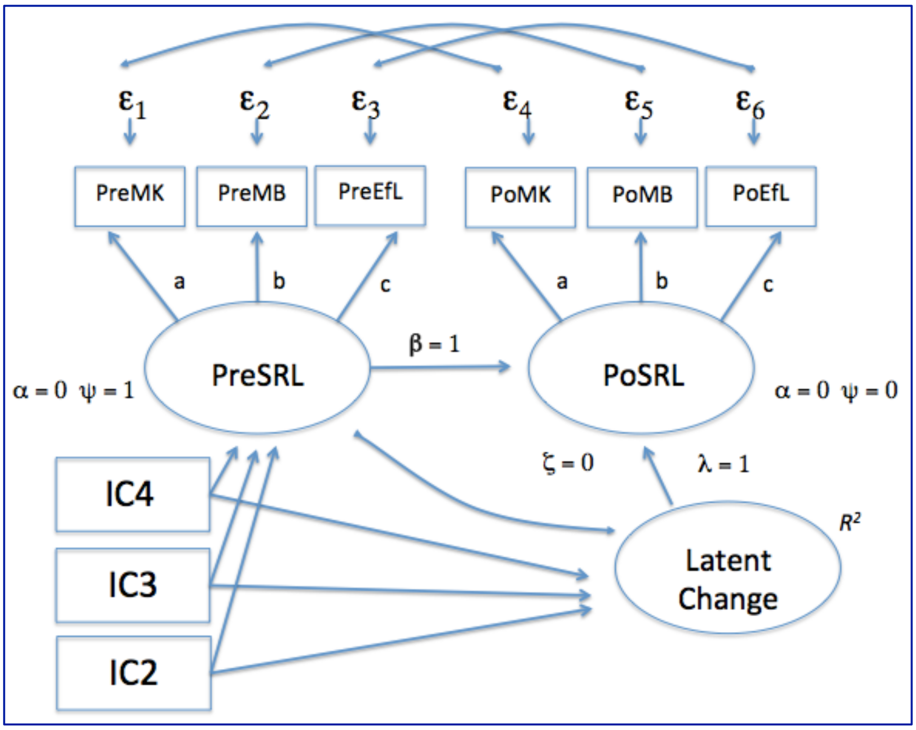


Figure 6-8 Hypothesised LCSM model: effect of experimental condition on change in SRL
α Latent intercept; ζ disturbance (endogenous); ψ latent variance (dependent variable); λ loading; R^2 variance explained; PreMK: Pre-test metacognitive knowledge; PoMK: Post-test metacognitive knowledge; PreMB: Pre-test motivational beliefs; PoMB: Post-test motivational beliefs; PreEfL: Pre-test efficacy for learning; PoEfL: Post-test efficacy for learning

The measures are described in detail in Chapter Four, and consisted of pre- and post-test measures of a) aggregated self-regulated learning scores, b) metacognitive knowledge, c) motivational beliefs and d) efficacy for learning. The dependent variable was the latent change in participants' self-reported self-regulated learning scores, whilst the independent variable was the experimental condition at four levels (three experimental and one control).

The principal method of analysis used for this research question was latent change score modelling (McArdle, 2009; McArdle and Prindle, 2008; Duncan, Duncan and Stryker, 2006; Mun, Von Eye and White, 2009).

The latent change score hypothesised model was constructed according to the model described in Section 4.7.5.1, with corresponding parameters constrained in the same way. The intervention effect was explored by dummy coding each of the three experimental conditions (groups) against the control group then regressing the self-regulated learning change scores on the dummy coded groups in the latent change score model. Before conducting formal analyses, and to satisfy necessary assumptions for SEM, summary statistics, means, standard deviations and normality tests were performed. Details are reported in Table 6-13.

Table 6-13 Summary statistics for RQ4

	IC4		IC3		IC2		CG	
	PreSRL	PoSRL	PreSRL	PoSRL	PreSRL	PoSRL	PreSRL	PoSRL
N	109	123	129	127	127	140	138	133
M	179.47	198.59	179.56	190.55	180.35	188.67	179.27	182.62
SD	18.059	15.215	18.136	17.935	17.218	20.676	18.604	20.767
Skewness	-0.405	-1.285	-0.875	-0.52	-0.429	-0.862	-1.47	-0.571
Kurtosis	0.061	2.112	1.004	-0.13	-0.052	0.164	2.479	-0.229
	PreMK	PoMK	PreMK	PoMK	PreMK	PoMK	PreMK	PoMK
N	125	124	132	127	132	140	138	133
M	61.65	69.01	62.93	66.61	62.92	66.69	63.61	64.32
SD	8.546	7.577	8.309	8.18	8.456	8.715	8.013	8.774
Skewness	-0.242	-1.262	-0.724	-0.709	-0.455	-1.018	-1.224	-0.578
Kurtosis	-0.274	2.213	0.634	0.93	-0.379	1.395	1.456	0.28
	PreMB	PoMB	PreMB	PoMB	PreMB	PoMB	PreMB	PoMB
N	123	124	129	127	129	140	138	134
M	58.45	65.42	58.33	62.27	58.12	60.96	57.86	59.21
SD	6.972	4.621	7.071	7.262	7.578	8.38	7.085	7.908
Skewness	-1.022	-1.577	-0.738	-1.236	-0.708	-1.149	-1.191	-0.921
Kurtosis	2.156	3.393	0.218	1.774	-0.012	0.808	1.411	0.378
	PreEfL	PoEfL	PreEfL	PoEfL	PreEfL	PoEfL	PreEfL	PoEfL
N	110	124	131	127	134	140	140	135
M	58.25	64.06	58.21	61.68	58.29	61.02	57.69	58.87
SD	6.495	5.09	6.614	5.398	6.328	6.725	6.199	7.066
Skewness	-0.513	-0.883	-0.864	-0.562	-0.566	-0.945	-1.158	-0.636
Kurtosis	-0.179	0.15	1.241	-0.284	-0.029	0.294	1.496	0.019

MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning
IC: intervention condition; CG: control group

Temporal dependence was confirmed (group allocation preceded pre-test, which preceded post-test) and no serious violations in terms of missing data were noted (as reported in Section 4.7.1). Normality was tested by inspection of skewness and kurtosis values, which indicated a violation of normality in some of the kurtosis values. Fortunately the statistical SEM software MPlus v.7 compensates for non-normality if the Robust Maximum Likelihood (MLR) method of estimation is chosen. Therefore the latent change score models for this analysis proceeded using MLR estimates.

Confirmatory factor analysis (the measurement model) for this analysis is reported in Section 6.4.1.

A latent change score structural equation model, employing robust maximum likelihood estimation, was therefore employed to test the hypothesised model. The estimated model (Figure 6-9) provided a good fit to the data, with fit indices of: χ^2 (21) = 37.784; RMSEA= 0.037; CFI= 0.989; TLI= 0.982.

A comparison of path estimates between each intervention condition and the latent change score provided a comparison of the effects of the interventions conditions, relative to the control condition. The intervention condition a participant was placed in accounted for 16.7% of the variance in the latent change scores. Membership of all experimental groups predicted significantly greater gains in their self-regulated learning scores than membership of the control group, with overall a medium effect size ($r = .41$, $p < .001$). This effect size also represents the correlation between predictor and outcome, indicating that those with higher self-regulated learning change scores tend to come from the most complex experimental conditions, (with Experimental Condition IC4 being the most complex).

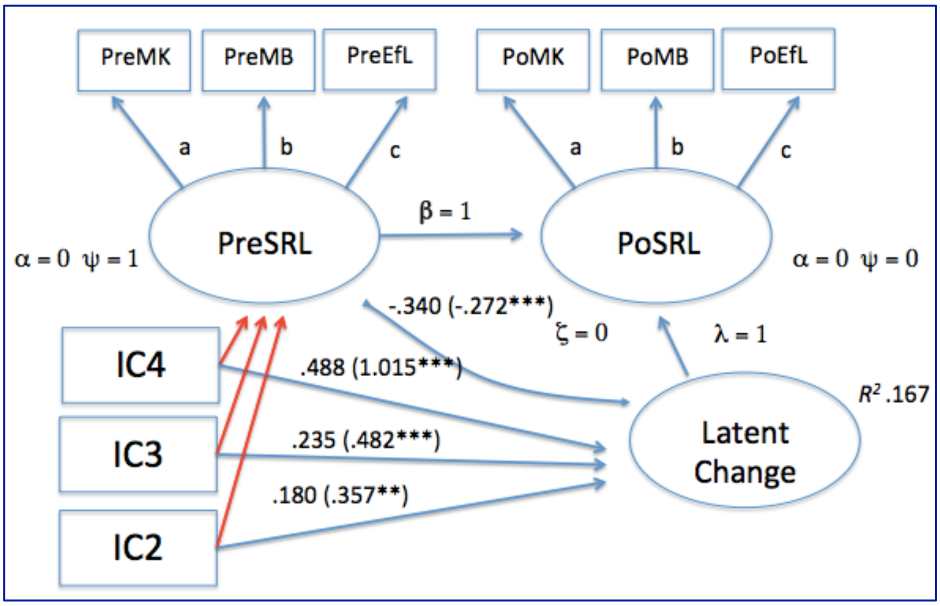


Figure 6-9 Estimated LCSM model: effect of experimental condition on change in SRL

Standardised estimates (unstandardised in parentheses); α Latent intercept; ζ disturbance (endogenous); ψ latent variance (dependent variable); λ loading; R^2 variance explained; error terms omitted for clarity; red connectors denote non-sig paths

PreMK: Pre-test metacognitive knowledge; PoMK: Post-test metacognitive knowledge; PreMB: Pre-test motivational beliefs; PoMB: Post-test motivational beliefs; PreEfL: Pre-test efficacy for learning; PoEfL: Post-test efficacy for learning

Significant intervention condition effects, on change in self-regulated learning, with small and medium effect sizes, were detected as follows: IC4: $\beta = .49$, $p < .001$; IC3: $\beta = .24$, $p < .001$; IC2: $\beta = .18$, $p < .01$. The β values represent the standardised differences in the group means, comparing each group with the control group. Thus, for instance, the effect of IC4 membership (which included the students who additionally completed a learning journal), on change in self-regulated learning over time, was more than twice that of IC3 (who did not complete the journal).

Un-standardised Beta estimates revealed that students in experimental condition IC4 (those who experienced all four intervention conditions: the learning journal, dialogic argumentation, Devil's Advocate and the critical thinking tests) were predicted to gain 1.015 points more in their self-regulated learning change scores, than those who were in the control condition. Those in IC3 (who experienced three intervention conditions: dialogic argumentation + Devil's Advocate + critical thinking tests) were predicted to gain .482 points more. Those in IC2 (those who experienced two conditions: Devil's Advocate + critical thinking tests) were predicted to gain .357 points more. All estimates were statistically significant at the $p < .01$ level.

Similar models were constructed for the separate components of self-regulated learning (metacognitive knowledge, motivational beliefs and efficacy for learning), with the estimates and model fits reported in Table 6-14.

Table 6-14 Standardised *Beta* estimates and model fit for LCSM models of the separate experimental conditions, compared with the control condition

Model	Experimental condition			Variance explained	Model fit			
	IC4	IC3	IC2	R^2	χ^2 *(df)*	RMSEA	CFI	TLI
SRL	.488	.235	.180	.167	37.784 (21)	.037	.989	.982
MK	.449	.223	.223	.138	241.964 (133)	.038	.969	.961
MB	.419	.215	.163	.123	361.430 (163)	.046	.943	.927
EfL	.455	.193	.091[a]	.158	184.983 (102)	.037	.958	.945

SRL: aggregated SRL scale (MK + MB + EfL); MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning; [a]Non-significant estimates; all significant estimates p < .01

The intervention condition a participant was placed in accounted for between 12% and 17% of the variance in latent change scores. In each model, IC4 was the most effective condition, with medium effect sizes, and with β estimates more than twice those of the other experimental conditions. This finding suggests that students who completed the intervention activities whilst also having access to a learning journal, made more progress in all aspects of self-regulated learning, compared with those who experienced the intervention but did not have access to the learning journal activity. The aspect of self-regulated learning for which membership of an intervention group made the most difference was efficacy for learning.

The above estimates for IC2 provide an estimation of the effect of the Devil's Advocate condition, since the analysis reported in Table 6-14 compared each condition with the control condition. Here it can be seen that the effect of condition IC2 (the Devil's Advocate plus the critical thinking tests) was positive and significant for change in overall self-regulated learning ($\beta = .18$, $p < .01$), change in metacognitive knowledge ($\beta = .22$, $p < .01$), and change in motivational beliefs ($\beta = .16$, $p < .01$), but not in efficacy for learning. Thus the principal critical thinking activities (the Devil's Advocate and the critical thinking tests) had the greatest effects on change in participants' metacognitive knowledge, with a small effect size.

In order to specifically isolate the learning journal condition, the latent change score analyses were repeated comparing IC4 (the group which experienced the learning journal) with IC3 (the group which did not). Results are reported in Table 6-15.

Table 6-15 The effect of the learning journal on aspects of self-regulated learning

Model	Estimates		Model fit			
	β	R^2	χ^2 *(df)*	RMSEA	CFI	TLI
SRL	.302***	.10	15.664 (13)	.026	.995	.993
MK	.319***	.10	144.344 (96)	.043	.971	.958
MB	.235**	.06[a]	206.560 (123)	.050	.958	.942
EfL	.294***	.09	121.032 (84)	.040	.962	.953

SRL: aggregated SRL scale (MK + MB + EfL); MK: metacognitive knowledge; MB: motivational beliefs; EfL: efficacy for learning; [a]Non-significant estimates; *** p < .001; ** p < .01

Thus a comparison of the effect, on change in SRL and its components, of the learning journal group (IC4) with the non-learning journal group (IC3) detected significant effects of the learning journal condition on change in overall self-regulated learning ($\beta = .30$, $p < .001$); change in metacognitive knowledge ($\beta = .32$, $p < .001$); change in motivational beliefs ($\beta = .24$, $p < .01$) and change in efficacy for learning ($\beta = .29$, $p < .001$). Thus the learning journal condition had the greatest effect on change in metacognitive knowledge, with a medium effect size.

Summary of Research Question Four

The hypothesis that there would be a differentiated effect of the intervention conditions on change in self-regulated learning and its components, was supported by these data. Results of this group comparison suggest that the IC4 (learning journal + dialogic argumentation + Devil's Advocate + critical thinking tests) condition was more than twice as effective as either the IC3 condition (dialogic argumentation + Devil's Advocate + critical thinking tests) or the IC2 conditions (Devil's Advocate + critical thinking tests). Members of the learning journal group reported significantly more change in SRL and all of its components, compared with those who did not, with small and medium effect sizes ($.24 < \beta < .32$).

6.6 Research Question Five

To what extent does participation in dialogic argumentation predict change in self-regulated learning and its components over time?

The analysis of RQ4 detected group effects between *membership* of the three experimental conditions, the three experimental conditions being:

- IC4: Learning journal + dialogic argumentation + Devil's Advocate + CT tests
- IC3: Dialogic argumentation + Devil's Advocate + CT tests
- IC2: Devil's Advocate + CT tests

Thus the 'core' of the intervention, which was provided for students in all experimental groups, consisted of the Devil's Advocate exercise and the weekly critical thinking test. As students were participating in a 'course', there was an expectation that participants would complete these core activities. Participation figures and missingness analysis confirmed this assumption, with mean missingness figures of: critical thinking: 2.6%; Devil's Advocate: 3%. However, the debating activities, which were provided for IC4 and IC3 at the end of each day, attracted only 42% to make an active contribution. Thus membership of intervention groups did not necessarily imply full participant activity. Therefore the current research question explores the effect of *actual participation* in dialogic argumentation activity (both active and passive), rather than simply group membership.

Active participation in dialogic argumentation

This research question explored the predictive effect of active participation in dialogic argumentation (rather than group membership) on change (from pre-test to post-test) in participants' self-reported self-regulated learning and its components (metacognitive knowledge, motivational beliefs and efficacy for learning). The analysis was conducted on groups IC4 and IC3, the two intervention groups whose members had the opportunity to engage in the evening debates. It was hypothesised that the total extent to which a participant engaged in dialogic argumentation would predict gains in self-regulated learning and its components over time, from pre-test to post-test, according to the hypothesis exemplified in Figure 6-10.

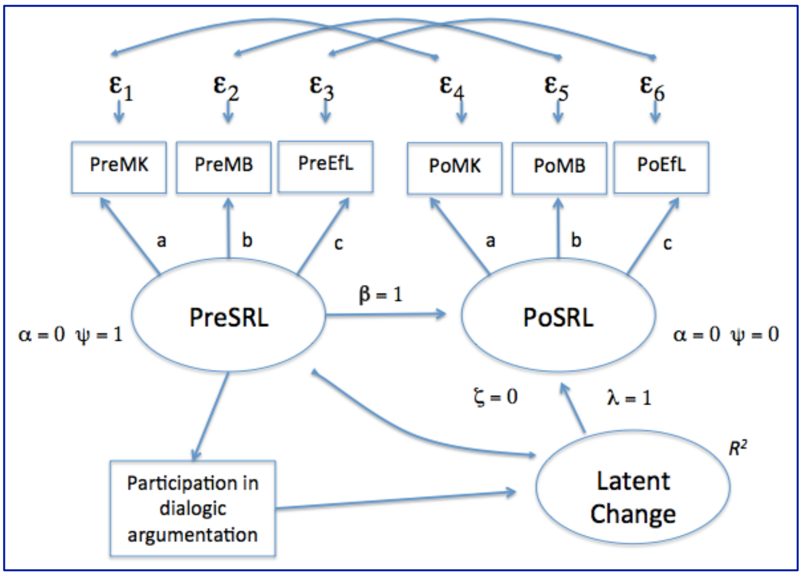


Figure 6-10 Hypothesised LCSM model: effect of dialogic argumentation on SRL change

α Latent intercept; ζ disturbance (endogenous); ψ latent variance (dependent variable); λ loading; R^2 variance explained. PreSRL: pre-test self-regulated learning; PoSRL: post-test self-regulated learning

Four separate analyses evaluated the effect of the independent (manifest) variable (the total count of a participant's posts to all the evening debates over the course of the intervention) on the dependent variables (latent change in self-reported overall self-regulated learning, metacognitive knowledge, motivational beliefs and efficacy for learning).

The principal method of analysis used for this research question was latent change score modelling (McArdle, 2009; McArdle and Prindle, 2008; Duncan, Duncan and Stryker, 2006; Mun, Von Eye and White, 2009). The analytical process is described in Section 4.7.5.1. The latent change score hypothesised model was constructed according to the model described in Section 4.7.5.1 with corresponding parameters constrained in the same way. Here the model specified a latent change phantom construct using self-regulated learning (a latent variable indicated by its manifest items) at Time 2 as its indicator, and with participation in dialogic argumentation activity replacing the predictor variable.

Before conducting formal analyses, and to satisfy necessary assumptions for structural equation modelling (SEM), summary statistics, means, standard deviations and normality tests were performed. Initial descriptive analysis of the dialogic argumentation variable detected the presence of zero inflated scores and significant outliers, resulting in a distribution unacceptable for structural equation modelling, as can be seen in Figure 6-11.

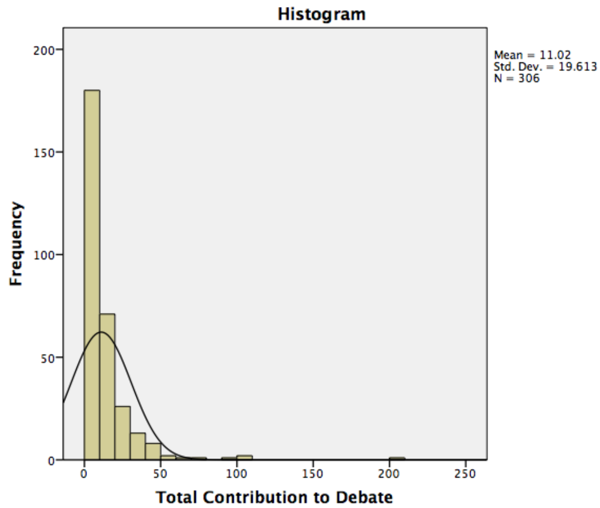


Figure 6-11 Untransformed distribution of overall dialogic argumentation counts

This distribution was to be expected for count data which, according to Greene, Costa and Dellinger (2011), are often non-normally distributed, and often positively skewed. The latter authors suggest various methods for transforming non-normally distributed count data, including dichotomising them (although Cohen (1983) would advise otherwise) and transforming them by square-rooting or by logarithmising. Whilst it is possible to adjust for non-normality in terms of skewness and kurtosis, by choosing the Robust Maximum Likelihood method of estimation in MPlus (Muthén and Muthén, 2013), the issue of outliers remains a problem for structural equation modelling (Kline, 2010). Therefore the treatment of such is an important preparation for this analysis. Zero inflation dialogic argumentation scores and upper outliers with *z* scores above 3.3 (an appropriate threshold according to Tabachnick and Fidell, 2007) were therefore excluded from further analysis, resulting in a variable (Table 6-16), which was then considered acceptable for structural equation modelling:

Table 6-16 Total dialogic argumentation

N	130
Mean	16.38
Std. Deviation	8.17
Skewness	.35
Kurtosis	-.23

Structural equation modelling requires there to be an estimation of the measurement model before that of the structural model. This measurement model (confirmatory factor analysis) for self-regulated learning and its components is reported in Section 6.4.1.

Correlational analysis confirmed the presence of significant relationships between dialogic argumentation and self-regulated learning at post-test (Table 6-17), so the latent change score model proceeded as specified in the hypothesised model.

Table 6-17 Correlational analysis for RQ5b

	DArg	PreMK	PoMK	PreMB	PoMB	PreEfL	PoEfL	PreSRL	PoSRL
Dialogic Argumentation	1								
Pre-test MK	0.057	1							
Post-test MK	**.341**	**.433**	1						
Pre-test MB	0.125	**.380**	**.267**	1					
Post-test MB	**.330**	**.263**	**.599**	**.476**	1				
Pre-test EfL	0.097	**.532**	**.268**	**.423**	**.227**	1			
Post-test EfL	**.429**	**.304**	**.681**	**.300**	**.622**	**.492**	1		
Pre-test SRL	0.019	**.809**	**.403**	**.713**	**.374**	**.755**	**.431**	1	
Post-test SRL	**.429**	**.361**	**.878**	**.408**	**.834**	**.374**	**.841**	**.473**	1

Correlations in bold: $p < .01$ two tailed

A latent change score structural equation model, employing maximum likelihood estimation, was thus employed to test the hypothesised model. Goodness of fit analytics suggested that the model provided a good fit to the data: $\chi^2(13) = 15.535$; RMSEA = .031; CFI = .994; TLI = .990.

The number of posts made to group discussions over the course of the intervention was a significant predictor of latent change in overall self-regulated learning from pre-test to post-test ($\beta = .56$, $p < .001$). Furthermore, the total extent to which a participant engaged in dialogic argumentation over the course of the intervention accounted for 23.9% of the variance in self-regulated learning change scores.

A similarly constructed latent change score model was used to estimate the effect of dialogic argumentation on change in each of the separate components of self-regulated learning (metacognitive knowledge, motivational beliefs and efficacy for learning). A summary of estimates and model fits is reported in Table 6-18.

Table 6-18 Latent change score estimates and model fits for the effect of overall participation in dialogic argumentation on change in SRL and its components

	Standardised estimates			Model fit			
	Change ON predictor	Predictor ON Pre-test	R^2	χ^2 *(df)*	RMSEA	CFI	TLI
SRL	.558***	.147[a]	.239	16.535 (13)	.031	.994	.990
MK	.524***	.093[a]	.223	136.189 (96)	.038	.976	.965
MB	.423***	.247**	.046[a]	180.405 (123)	.041	.971	.960
EfL	.434***	.165[a]	.087[a]	112.524 (84)	.035	.971	.964

[a]Non-significant estimates *** p < .001; ** p < .01

The extent to which participants engaged in dialogic argumentation was predicted by their pre-test motivational beliefs, with a small effect size ($\beta = .25$, $p < .01$). It appears that the greater their motivational beliefs at the outset, the more they were inclined to participate in discussions. However other pre-test self-regulated learning dispositions had no such effect. The number of posts which individuals made to the evening discussions was a significant predictor of change in overall self-regulated learning and all of its components, with medium and large effect sizes: self-regulated learning: ($\beta = .56$, $p < .001$); metacognitive knowledge ($\beta = .52$, $p < .001$); motivational beliefs (β

= .42, $p < .001$) and efficacy for learning ($\beta = .43$, $p < .001$). Thus the more students engaged in dialogic argumentation, the greater the increase over time in their metacognitive knowledge, their motivational beliefs and also their efficacy for learning.

The extent to which a participant engaged in dialogic argumentation accounted for 23.9% of the variance in self-regulated learning change scores, and 22.3% of the variance in metacognitive knowledge scores, whilst the variances in motivational beliefs scores (4.6%) and efficacy for learning (8.7%) were non-significant. Thus, in terms of individual self-regulated learning components, participation in dialogic argumentation had the greatest effect on change in their metacognitive awareness.

In sum, participants' pre-test motivational beliefs predicted the extent to which participants engaged in dialogic argumentation, with a small effect size, whilst other aspects of self-regulated learning had no such effect. The hypothesis that the total extent to which a participant engaged in dialogic argumentation predicted gains in self-regulated learning and its components over time, was supported by these data. Participants' contribution to dialogic argumentation was a significant predictor of change in self-regulated learning and all of its components, with medium and large effect sizes. The findings of this research question are summarised in Figure 6-12:

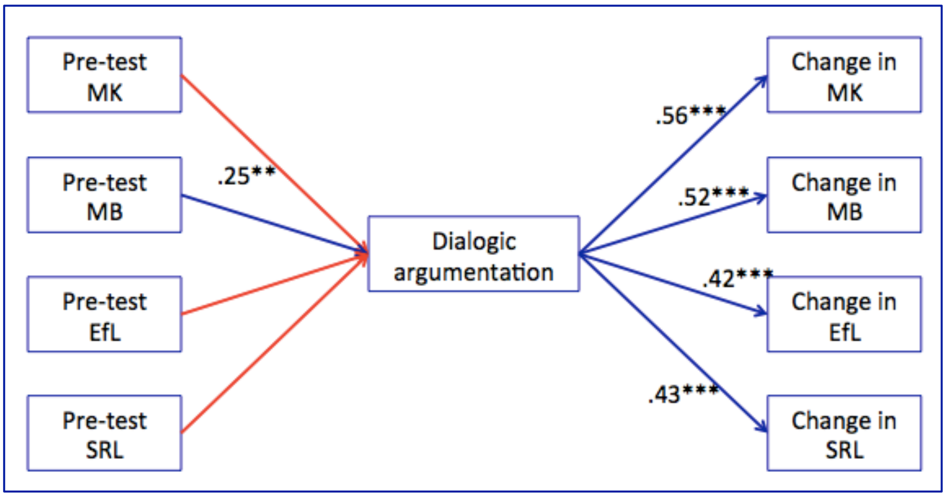


Figure 6-12 Overview of effects of dialogic argumentation on change in SRL (standardised estimates)

Passive participation in dialogic argumentation (browsing)

Active participation in dialogic argumentation (i.e., actively submitting posts into a discussion) accounted for 12% of overall dialogic argumentation activity, with the browsing of posts (without actual posting of replies) accounting for the remaining 88%. Thus, the variable 'browsing discussions' is the total count of the number of times a participant accessed discussions, without making a contribution. It refers to data from groups IC4 and IC3, the groups who had the opportunity to engage in debate. Thus this research question explored the predictive effect of *passive* participation in dialogic argumentation on change (from pre-test to post-test) in participants' self-reported self-regulated learning and its components. It was hypothesised that the total extent to which a participant browsed the debates would predict gains in self-regulated learning and its components over time, according to the hypothesis exemplified in Figure 6-13.

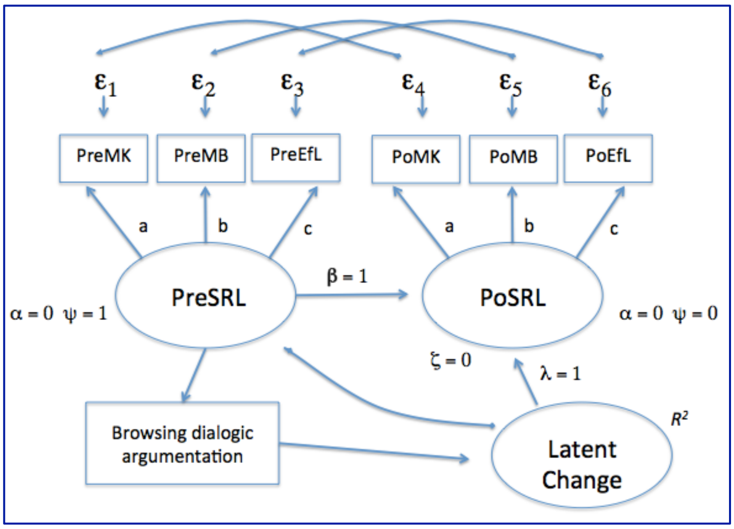


Figure 6-13 Hypothesised LCSM model: effect of browsing dialogues on SRL change

α Latent intercept; ζ disturbance (endogenous); ψ latent variance (dependent variable); λ loading; R^2 variance explained. PreSRL: pre-test self-regulated learning; PoSRL: post-test self-regulated learning

Four separate analyses evaluated the effect of the independent (manifest) variable (the total count of the total number of times a participant observed a debate without actually posting into it, over the course of the intervention) on the dependent variables (latent change in self-reported overall self-regulated learning, metacognitive knowledge, motivational beliefs and efficacy for learning).

The principal method of analysis used for this research question was latent change score modelling (McArdle, 2009; McArdle and Prindle, 2008; Duncan, Duncan and Stryker, 2006; Mun, Von Eye and White, 2009). The latent change score hypothesised model was constructed according to the model described in Section 4.7.5.1 with corresponding parameters constrained in the same way. Here the model specified a latent change phantom construct using self-regulated learning (a latent variable indicated by its manifest items) at Time 2 as its indicator, and with the browsing of dialogic argumentation activity replacing the predictor variable.

Before conducting formal analyses, and to satisfy necessary assumptions for SEM, summary statistics, means, standard deviations and normality tests were performed (Table 6-19), revealing no violations of normality.

Table 6-19 Descriptives: Browsing of dialogic argumentation

N	175
Mean	80.94
Std. Deviation	43.904
Skewness	.732
Kurtosis	.738

This measurement model (confirmatory factor analysis) for self-regulated learning and its components is reported in Section 6.4.1.

Correlational analysis confirmed the presence of significant relationships between browsing activity and self-regulated learning at both pre- and post-test (Table 6-20), so the latent change score model proceeded as specified in the hypothesised model.

Table 6-20 Correlations between browsing activity and aspects of self-regulated learning

	Browsing	PreMK	PoMK	PreMB	PoMB	PreEfL	PoEfL	PreSRL	PoSRL
Browsing	1								
Pre-test MK	**.246****	1							
Post-test MK	**.170***	**.433****	1						
Pre-test MB	-0.016	**.380****	**.267****	1					
Post-test MB	0.163	**.263****	**.599****	**.476****	1				
Pre-test EfL	0.075	**.532****	**.268****	**.423****	**.227****	1			
Post-test EfL	**.228****	**.304****	**.681****	**.300****	**.622****	**.492****	1		
Pre-test SRL	0.173	**.809****	**.403****	**.713****	**.374****	**.755****	**.431****	1	
Post-test SRL	0.145	**.361****	**.878****	**.408****	**.834****	**.374****	**.841****	**.473****	1

* $p < .05$; ** $p < .01$ two tailed

Latent change score structural equation models, employing maximum likelihood estimation, were thus employed to test the hypothesised model. Estimates and goodness of fit analytics are reported in Table 6-21.

Table 6-21 LCSM standardised estimates and model fits for the effect of overall browsing of dialogic argumentation on change in SRL and its components

	Change ON predictor	Predictor ON Pre-test	χ^2 (*df*)	RMSEA	CFI	TLI
SRL	.206*	.142[a]	21.386(13)	.047	.986	.977
Metacognitive knowledge	.075 [a]	.207**	155.526 (96)	.046	.964	.949
Motivational beliefs	.103 [a]	.067 [a]	184.246 (123)	.041	.969	.956
Efficacy for learning	.175*	.137[a]	103.877 (84)	.029	.979	.974

[a]Non-significant estimates; * $p < .05$; ** $p < .01$

Summary of Research Question Five

In summary, the greater a participant's self-reported initial metacognitive knowledge, the more they tended to browse discussions, with a small effect size ($\beta = .21$, $p < .01$). Other pre-test SRL characteristics did not influence browsing behaviour. Browsing activity itself had a statistically significant effect on change in participants' efficacy for learning ($\beta = .18$, $p < .05$) and their overall self-regulated learning ($\beta = .21$, $p < .05$), but not on either metacognitive knowledge or motivational beliefs. The findings of this aspect of the research question are summarised in Figure 6-14.

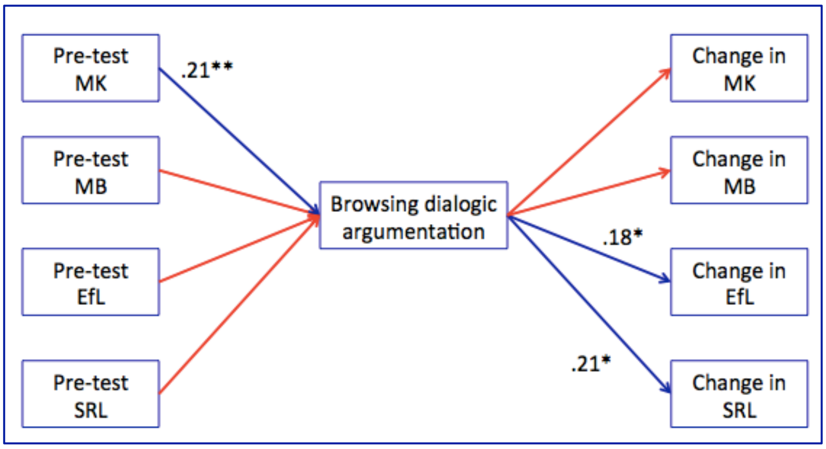


Figure 6-14 Overview of effects of browsing dialogues on change in SRL dispositions (standardised estimates)

6.7 Research Question Six

What are the relationships between aspects of metacognitive activity and critical thinking over time?

A process analysis, estimating an autoregressive panel model, evaluated the relationship between metacognitive skills (MS) and metacognitive experiences (ME), observed and reported in learning journals on four occasions. MS were measured in terms of observed assessments in critical thinking (CT), perceived control of elaboration strategies (ES) and inhibitory control (IC). ME were measured in terms of the extent to which learners were a) satisfied that they had made progress, b) were happy to give up their free time in order to study, c) enjoyed the intellectual challenge, d) made a valuable contribution to the course, and e) felt confident in themselves as learners (Efklides and Petkaki, 2005). These event-measures took place on the same day, at the end of each week of the intervention, and were captured in the learning journal (Appendix 2) and in CT (Appendix 6). It was hypothesised that CT, ES, IC and ME would be reciprocally related over time.

This analysis refers to data from IC4 only (the group that used the learning journal). CT scores were computed from the end of week assessments. The other three process variables represented the cognitive and metacognitive activity scores described in Section 4.5.3. Variables acted as both dependent and independent variables within the model. They are labelled:

- Critical thinking (Metacognitive Skill): CT1, CT2, CT3 and CT4
- Use of Elaboration Strategies (Metacognitive Skill): ES1, ES2, ES3 and ES4
- Inhibitory Control (Metacognitive Skill): IC1, IC2, IC3 and IC4
- Metacognitive experiences: ME1, ME2, ME3 and ME4.

Table 6-22 Correlations and summary statistics for RQ6

	CT1	CT2	CT3	CT4	ES1	ES2	ES3	ES4	IC1	IC2	IC3	IC4	ME1	ME2	ME3	ME4
CT1	1															
CT2	.26**	1														
CT3	.33**	.20*	1													
CT4	.18*	.19*	.27**	1												
ES1	.22*	.23**	.24**	.25**	1											
ES2	.17*	0.05	.27**	.19*	.36**	1										
ES3	0.12	.22*	0.15	.26**	.38**	.22*	1									
ES4	.28**	0.09	.39**	.28**	.27**	.22*	0.14	1								
IC1	-0.12	-0.03	-0.03	-0.18	0.10	.21*	-0.07	-0.11	1							
IC2	-0.08	-0.13	-0.09	-0.14	0.05	-0.01	-0.09	-0.12	.26*	1						
IC3	-0.03	0.03	-0.04	-0.09	0.13	-0.07	0.078	-0.08	.25*	.53**	1					
IC4	-.21*	-0.10	-0.10	-0.07	-0.03	0.00	0.01	-0.05	0.05	.32**	.38**	1				
ME1	0.01	-0.01	0.03	-.19*	0.02	-0.10	-.23*	-0.10	.39**	.41**	.33**	0.21	1			
ME2	-0.05	-0.04	0.07	0.02	0.10	-0.03	-0.05	0.02	.28**	.52**	.45**	.37**	.65**	1		
ME3	-0.07	-0.02	0.10	-0.10	0.11	-0.03	-0.04	-0.09	.39**	.62**	.52**	.42**	.64**	.59**	1	
ME4	-0.10	-0.13	-0.11	-0.09	0.01	0.03	-.22*	-0.08	.31**	.47**	.44**	.50**	.51**	.57**	.70**	1
N	149	147	147	146	136	136	131	135	117	121	113	109	117	121	113	109
M	12.13	12.97	14.05	14.87	12.54	12.80	14.35	15.73	17.24	18.13	18.29	18.41	41.82	44.53	43.95	44.61
SD	3.28	2.47	2.32	2.43	2.97	3.05	2.45	2.30	2.83	2.24	2.27	2.46	7.46	5.48	5.45	5.54
Skewness	-1.10	-0.64	-0.55	-0.56	-0.20	-0.39	-0.67	-0.40	-0.46	-0.78	-0.964	-1.07	-0.60	-0.89	-0.91	-0.85
Kurtosis	1.56	1.07	0.20	-0.03	-0.32	-0.36	0.95	-0.49	-0.63	-0.62	-0.0	0.16	-0.25	0.00	0.05	-0.01

CT: critical thinking scores in weekly assessments; ES: elaboration strategies;
IC: Inhibitory control; ME: Metacognitive experiences; significant correlations in bold * p < .05; ** p < .01

Correlational analysis detected relationships only between critical thinking and elaboration strategies, and between inhibitory control and metacognitive experiences. These relationships informed the hypothesised model featured in Figure 6-15. Thus it was hypothesised that there would be a relationship between critical thinking and elaboration strategies, and between inhibitory control and metacognitive experiences.

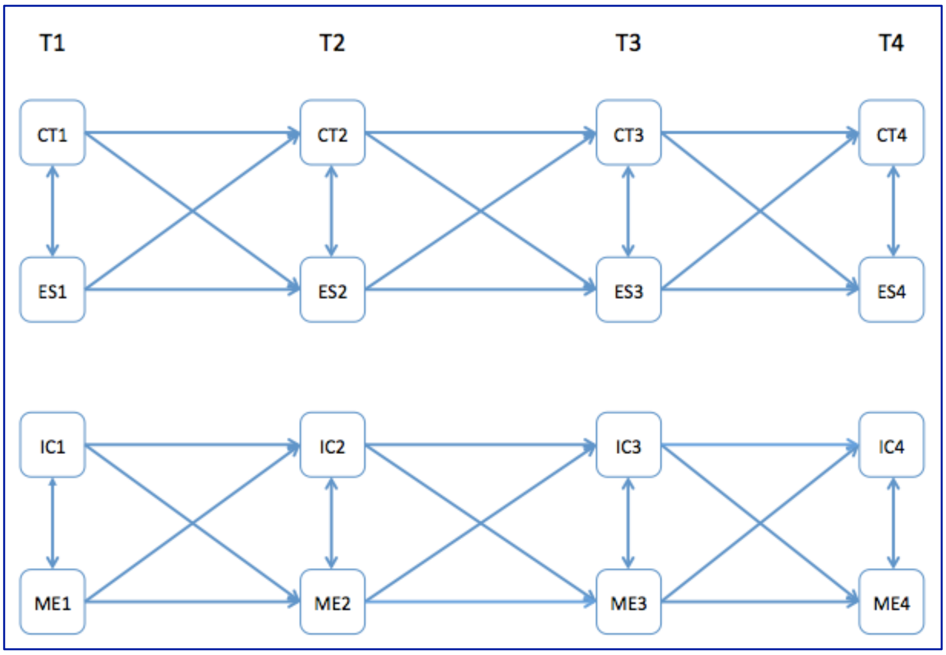
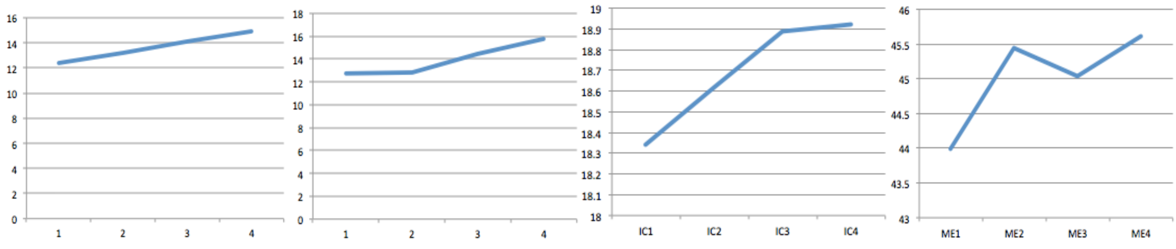


Figure 6-15 Hypothesised model for relationships between aspects of cognitive and metacognitive activity
CT: critical thinking; ES: elaboration strategies; IC: inhibitory control; ME: metacognitive experiences

Summary statistics are reported previously in Table 6-22 and linear trends are described in Figure 6-16. Analysis of the trends for all four variables revealed a general positive trend in each case.



Critical thinking | Elaboration strategies | Inhibitory control | Metacognitive experiences

Figure 6-16 Trends for cognitive and metacognitive activity in IC4, over four learning events

6.7.1 Repeated measures analyses

In order to establish the overall pattern of change for each of the four equally spaced measures of cognitive and metacognitive activity, one-way repeated measures analyses of variance were performed on each variable.

6.7.1.1 Critical thinking

Mauchly's test of Sphericity was violated, $\chi^2(5) = 13.620$, $p < .05$; therefore degrees of freedom were corrected using Greenhouse Geisser estimates of sphericity ($\varepsilon = .94$). There was a significant main effect for time, with a medium effect size in critical thinking: $F\ (2.808, 384.730) = 34.389$; $p < .001$; partial $\eta^2 = .201$; $r = .45$.

6.7.1.2 Elaboration strategies

Mauchly's test of Sphericity was not violated, $\chi^2(5) = 3.284$, $p = .656$; therefore it was not necessary to adjust degrees of freedom for the main effect of the repeated measures analysis for this variable. The repeated measures ANOVA for elaboration strategies estimated a significant main effect for time, with a large effect size in perceived control of elaboration strategies: F (3, 390) = 51.883; $p < .001$; partial η^2 = .285; r = .53.

6.7.1.3 Inhibitory control

Mauchly's test of Sphericity was not violated, $\chi^2(5) = 6.059$, $p = .301$; therefore it was not necessary to adjust degrees of freedom for the main effect of the repeated measures analysis for this variable. There was a significant main effect for time, with a small effect size in inhibitory control; F (3, 180) = 3.551; $p < .05$; partial η^2 = .056; r = .24.

6.7.1.4 Metacognitive experiences

Mauchly's test of Sphericity was violated, $\chi^2(5) = 22.417$, $p < .001$; therefore degrees of freedom were corrected using Greenhouse Geisser estimates of sphericity (ε = .83). There was a significant main effect for time, with a small effect size in metacognitive experiences; F (2.477, 173.368) = 5.922; $p < .01$; partial η^2 = .078; r = .28.

6.7.2 Auto-regressive cross-lagged panel analysis

Before conducting formal analyses, and to satisfy necessary assumptions for structural equation modelling (SEM), summary statistics, means, standard deviations and normality tests were performed. Temporal dependence was confirmed as measures of CT, ES, IC and ME were all taken at the same. Little's MCAR test of missingness was conducted on each of the process variables (Table 6-23), with non-significant p values indicating that missingness was distributed completely at random.

Table 6-23 Missingness analysis of process variables for RQ6

Process variable	Little's MCAR test of missingness
CT1, CT2, CT3 and CT4	χ^2 (13) = 8.687; p = .796
ES1, ES2. ES3 and ES4	χ^2 (5) = 6.905; p = .228
IC1, IC2, IC3 and IC4	χ^2 (28) = 37.504; p = .108
ME1, ME2, ME3 and ME4	χ^2 (28) = 25.346; p = .609

Normality was tested by inspection of skewness and kurtosis values (Table 6-22), and indicated no deviations from normality. Therefore the autoregressive panel model proceeded using maximum likelihood (ML) estimates.

Confirmatory factor analysis (CFA) of the nine-item cognitive and metacognitive activity scale confirmed a three-factor structure with a good level of reliability (Cronbach's α = .85). One factor described metacognitive experiences and two described metacognitive skills (inhibitory control and elaboration strategies). Goodness of fit indices suggested that the confirmatory model represented the data well: χ^2(24) = 45.442; RMSEA= .081; CFI= .942; TLI= .913. Parameter estimates are reported in Table 6-24.

It was hypothesised that critical thinking, elaboration strategies, inhibitory control and metacognitive experiences, reported weekly in the learning journal, would be related over time.

Table 6-24 CFA estimates for metacognitive activity recorded in the learning journal

Questionnaire item	Aspect of metacognitive activity	B	β	R^2
I was able to relate this week's topics to other areas of my interest.	elaboration strategies	1.000	.611	.37
I tried to consider several ways of looking at this week's dilemma before forming my argument.	elaboration strategies	1.177	.678	.46
I was able to put aside all of my negative thoughts and/or worries this week while I was studying.	inhibitory control	1.000	.760	.58
I coped well with distraction this week.	inhibitory control	.816	.702	.49
I am satisfied with my learning progress today.	metacognitive experiences	1.000	.747	.56
I was happy to give up my free time for this course this week.	metacognitive experiences	1.195	.737	.54
I enjoyed the challenge of intellectual activity this week.	metacognitive experiences	1.205	.806	.65
I felt that I made a valuable contribution to this week's activities.	metacognitive experiences	1.012	.699	.49
I feel confident this week.	metacognitive experiences	1.069	.726	.53

Correlational analysis (Table 6-22) informed the hypothesised autoregressive 4 x 4 cross-lagged panel model featured in Figure 6-15 above. An autoregressive panel model, relating the four variables, was built in two stages. First-order analysis explores relationships over time, such that each value of a time point (t) is explained by the value of the previous time point (t - 1), and whereby past values of a measure are used to predict future values of that measure (Bollen and Curran, 2004; Kline, 2010). Thus the autoregressive coefficients are the parameter values given by regressing each time point measure on its previous measure. For the current data, the final autoregressive model (Figure 6-19) represented four constructs, with each construct having four distinct time-ordered variables and with cross-lagged paths being specified to represent reciprocal influences over time. Residuals were correlated at time 2, time 3 and time 4.

Each of the sets of variables was first subject to a freely estimated first order analysis, in order to establish time ordered relationships from one week to the next. For example it was necessary to establish that critical thinking at one time was influenced by critical thinking at the previous measurement. According to Bollen (1989) it is important to establish this invariance, since if the data do not support this basic model, then they will not support subsequent, more constrained models. Freely estimated first order analysis of the four sets of variables (Figure 6-17) estimated significant correlations between the time-ordered events, suggesting stability over time (Kline, 2010). At this stage, however, the model resulted in a poor fit: χ^2 (96) = 235.905; p < .001; RMSEA= .122; CFI= .630; TLI= .560.

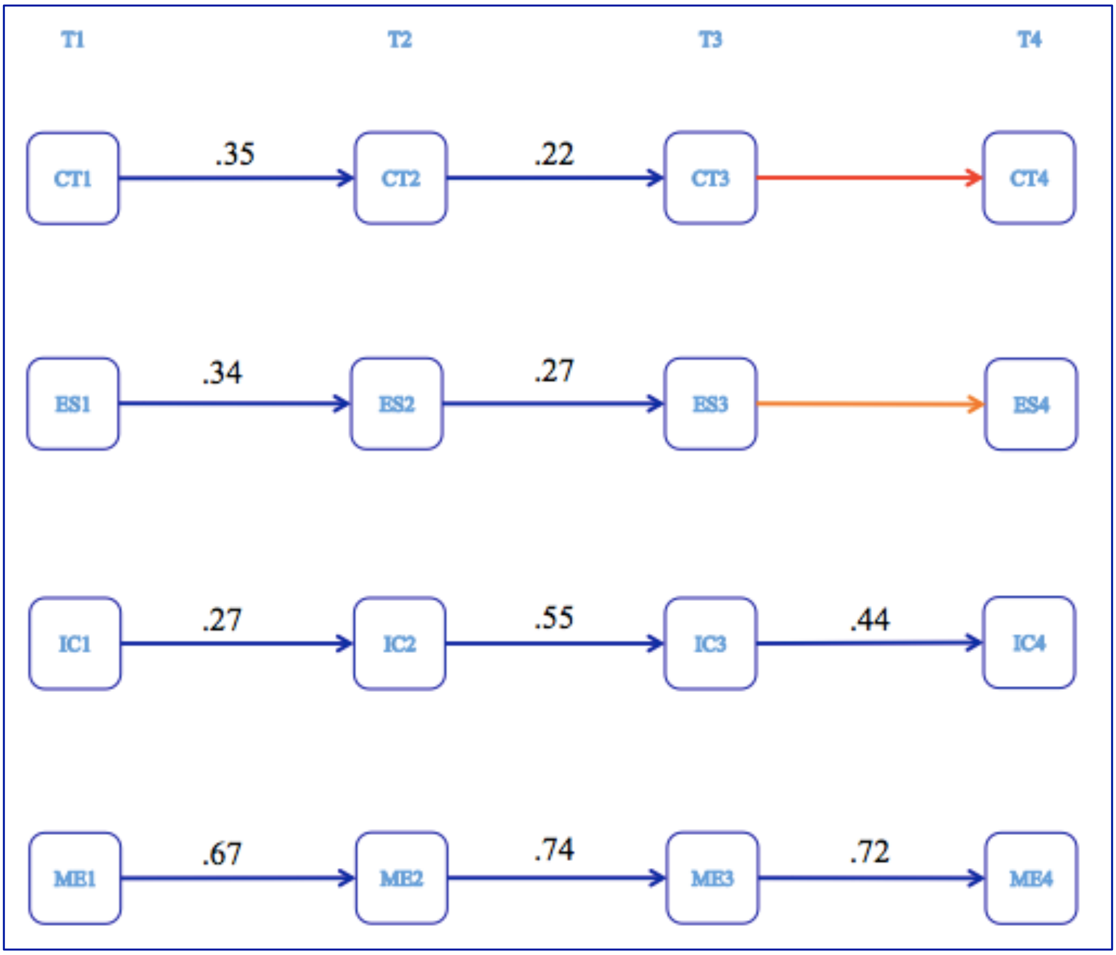


Figure 6-17 First order panel analysis freely estimated: standardised estimates
CT: critical thinking; ES: elaboration strategies; IC: inhibitory control; ME: metacognitive experiences; red lines indicate non-significant paths

In order to make meaningful comparisons when subsequently estimating cross-lagged coefficients, it was then necessary to constrain temporal path coefficients to equality (e.g., CT1 to CT2, CT2 to CT3, CT3 to CT4 etc). Thus by equating temporal pathways it would later be possible to isolate cross-lagged effects, for instance, estimating the effect of critical thinking at one time point on ES at another, and vice versa. Controlling temporal pathways would therefore result in a more accurate estimation of reciprocal cross-lagged effects (see Figure 6-18). Estimating temporally constrained first order analyses, resulted in a slightly better (but still not acceptable) goodness of fit (χ^2 (104) = 246.270; p < .001; RMSEA= .118; CFI= .623; TLI= .587). This improved fit suggests that metric invariance was satisfied and the analysis of the hypothesised model could proceed.

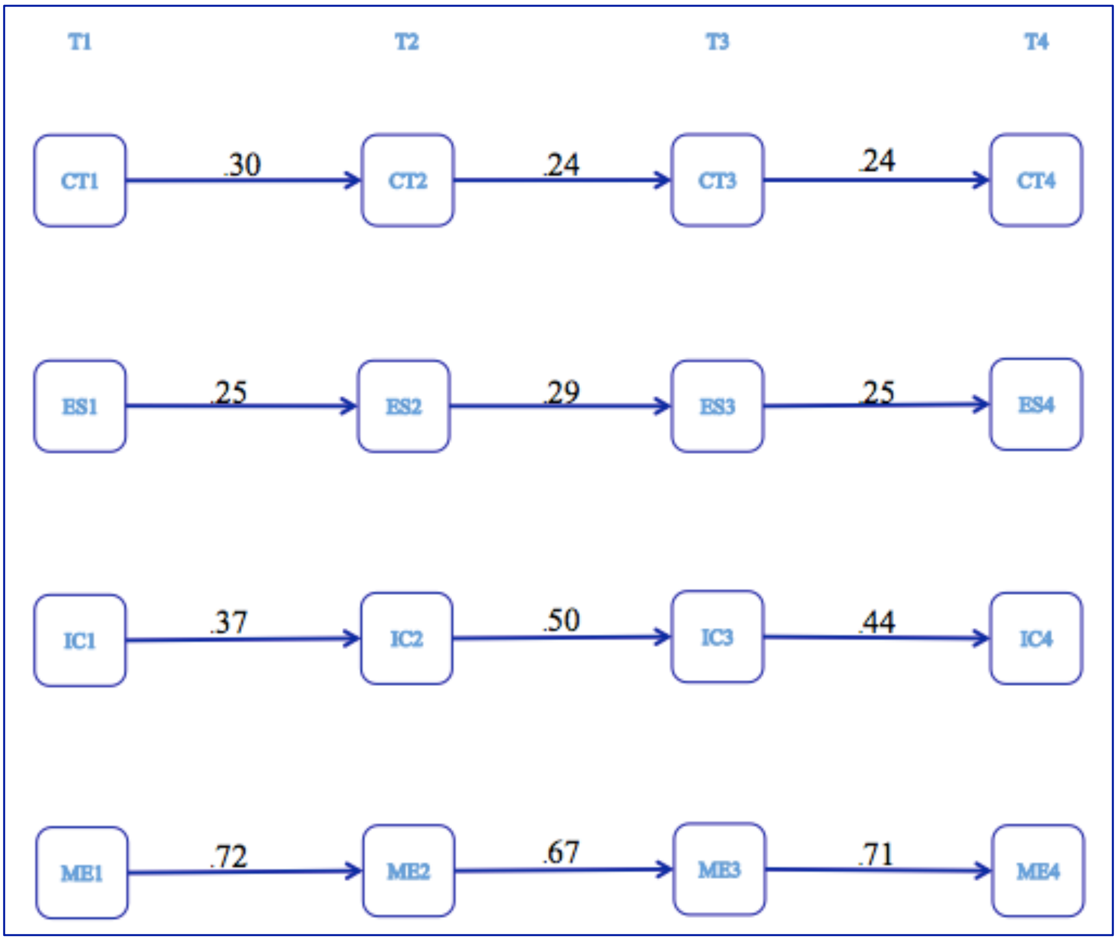


Figure 6-18 First order panel analysis (constrained): standardised estimates
CT: critical thinking; ES: elaboration strategies; IC: inhibitory control; ME: metacognitive experiences

The hypothesised model specified for repeated-measures critical thinking to be related to repeated-measures elaboration strategies, and for repeated measures inhibitory control to be related to repeated measures metacognitive experiences. Constraining temporal pathways, correlating residuals at Time 2, Time 3 and Time 4, and accepting theoretically justified modification indices (ME3 ON ME1, $\beta = .19$; ES3 ON ES1, $\beta = .27$; ME3 ON IC1, $\beta = .20$; ES2 WITH IC1, $\beta = .26$) resulted in an acceptable goodness of fit: χ^2 (83) = 105.944; $p = .0455$; RMSEA= .043; CFI= .948; TLI= .929. This model (Figure 6-19) was therefore chosen as the most appropriate one to model the data for this research question. Parameter estimates are reported in Table 6-25.

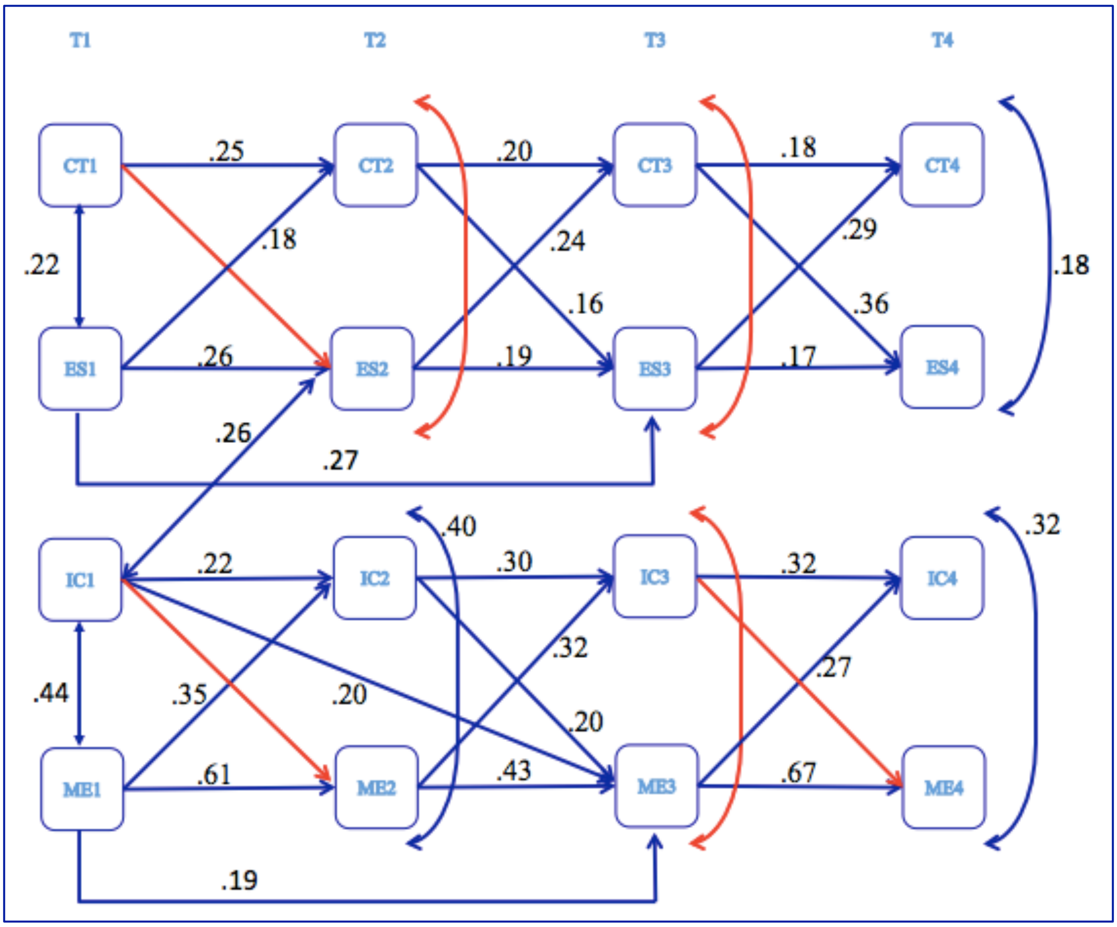


Figure 6-19 Final panel analysis for RQ6: standardised estimates
CT: critical thinking; ES: elaboration strategies; IC: inhibitory control; ME: metacognitive experiences

As hypothesised, two clear sets of relationships can be discerned from this model. Critical thinking (CT) appears to be related to the perceived control of elaboration strategies (ES), whilst inhibitory control (IC) relates to metacognitive experiences (ME).

Table 6-25 Parameter estimates for autoregressive panel analysis

Parameters	B	β	Critical ratio β
CT1 to CT2	.19	.25	4.58
CT2 to CT3	.19	.20	4.26
CT3 to CT4	.19	.18	4.24
ES1 to ES2	.16	.16	3.38
ES2 to ES3	.16	.19	3.56
ES3 to ES4	.16	.17	3.38
IC1 to IC2	.26	.22	4.20
IC2 to IC3	.26	.30	3.85
IC3 to IC4	.26	.32	4.23
ME1 to ME2	.53	.62	9.99
ME2 to ME3	.53	.43	8.84
ME3 to ME4	.53	.67	11.00
CT2 to ES3	.16	.16	1.98
CT3 to ES4	.35	.36	4.89
ES1 to CT2	.15	.18	2.26
ES2 to CT3	.19	.24	3.25
ES3 to CT4	.29	.29	3.73
ME3 to IC4	.07	.27	2.71
ME1 to ME3	.24	.19	2.27
CT1 to ES2	.13 [a]	.15 [a]	1.76 [a]
ES1 to ES3	.23	.27	3.39
ME1 to IC2	.12	.35	3.96
ME2 to IC3	.12	.33	3.88
IC1 to ME2	.06 [a]	.02 [a]	.24 [a]
IC1 to ME3	.72	.20	2.44
IC2 to ME3	.61	.20	2.75
IC3 to ME4	.39 [a]	.14 [a]	1.91
ES2 with IC1	1.028	.26	3.05
ME1 with IC1	2.82	.44	5.33
ME2 with IC2	1.82	.40	4.67
ME3 with IC3	.70 [a]	.19 [a]	1.77 [a]
ME4 with IC4	.88	.32	3.28
CT1 with ES1	2.17	.22	2.74
CT2 with ES2	-.33 [a]	-.05 [a]	-.58 [a]
CT3 with ES3	.11 [a]	.02 [a]	.27 [a]
CT4 with ES4	.87	.18	2.17

NB: A critical ratio of more than |1.96| indicates a two sided significance level of < .05 (Hox and Bechger, 1998). CT: Critical thinking; ES: Elaboration strategies; IC: Inhibitory control; ME: Metacognitive experiences; [a]Non-significant estimates

6.7.2.1 Critical thinking and perceived use of elaboration strategies

The perceived control of elaboration strategies (a metacognitive skill) was a consistent predictor of critical thinking achievement from one week to the next, with an increasing effect size, from a small to medium effect, over the course of the intervention: Week One to Week 2: $\beta = .18, p < .01$; Week Two to Week Three: $\beta = .24, p < .01$; Week Three to Week Four: $\beta = .29, p < .01$.

Critical thinking achievement similarly predicted ES, but not from Week One to Week Two. However, the effect increased over time, with Week Two to Week Three: ($\beta = .16, p < .05$) and Week Three to Week Four: ($\beta = .36, p < .001$). The two variables were synchronously related to each other only during the first week ($\beta = .22, p < .05$) and last week ($\beta = .18, p < .05$) of the intervention.

6.7.2.2 Inhibitory control and metacognitive experiences

A consistent reciprocal relationship was also detected between inhibitory control (a volitional strategy which controls the extent to which a participant avoids distraction) and metacognitive experiences (learning satisfaction, task-related self-confidence, response to challenge, effort and task motivation). The predictive effect of metacognitive experiences, on subsequent inhibitory control had a significant but decreasing effect over time, with Week One ME having a medium effect on Week Two IC ($\beta = .35, p < .01$); Week Two ME having a medium effect on Week Three IC ($\beta = .32, p < .01$) and Week Three ME having a small effect on Week Four IC ($\beta = .27, p < .05$). In contrast, the only week when ME was predicted by IC was in Week

Three, when it was predicted by both Week One IC ($\beta = .20$, $p < .05$) and Week Two IC ($\beta = .20$, $p < .05$).

However, in contrast to the pattern of relationships between critical thinking and elaboration strategies, where cross-lagged relationships (from one week to the next) tended to be marginally stronger than synchronous ones (during the same week), the relationship between metacognitive experience and inhibitory control was characterised by stronger synchronous relationships than cross-lagged ones. The relationship between metacognitive experiences and inhibitory control was a stronger one in general than that between critical thinking and elaboration strategies.

Thus the extent to which a participant perceived that they avoided distraction was consistently related to their metacognitive experiences (learning satisfaction, self-confidence, response to challenge and task-related motivation). It appears that the less they were distracted, the greater their metacognitive experiences (with consistent effect). Similarly, the better their metacognitive experience, the less they were inclined to be distracted (with decreasing effect). For three of the four weeks, inhibitory control and metacognitive experiences were synchronously related to each other (with decreasing effect).

Summary of Research Question Six

In sum, the hypothesis that metacognitive experiences would be related to metacognitive skills and critical thinking was partially supported by these data. Critical thinking (a metacognitive skill), was reciprocally related to the perceived use of elaboration strategies (a metacognitive skill), but not consistently related to inhibitory control (a metacognitive skill), or to metacognitive experiences. However, inhibitory control (a metacognitive skill) was consistently related to metacognitive experiences. The two clear sets of relationships differed in that the relationship between elaboration strategies and critical thinking was generally weaker than the relationship between inhibitory control and metacognitive experiences. However, critical thinking was consistently predicted by use of elaboration strategies, and inhibitory control was consistently predicted by self-reported metacognitive experiences.

6.8 Research Question Seven

What is the relationship between self-regulated learning dispositions (metacognitive knowledge, motivational beliefs and efficacy for learning) and self-regulated learning processes (metacognitive activity) during this intervention?

This research question employed path analysis to explore the relationships between self-regulated learning dispositions (pre- and post-test metacognitive knowledge, motivational beliefs and efficacy for learning) and self-regulated learning processes (metacognitive activity during the intervention).

Analysis employed data from Experimental Condition IC4 only, the group that recorded metacognitive activity in their weekly learning journals, and explored the following hypotheses:

> It was hypothesised that pre-test self-regulated learning dispositions (metacognitive knowledge, motivational beliefs and efficacy for learning) would predict self-reported metacognitive activity during the intervention.
>
> It was further hypothesised that metacognitive activity during the intervention, as recorded in students' learning journals would predict SRL dispositions (metacognitive knowledge, motivational beliefs and efficacy for learning) at post-test.

The measures are described in detail in Chapter Four, and consisted of pre- and post-test measures of a) aggregated self-regulated learning scores, b) metacognitive knowledge, c) motivational beliefs and d) efficacy for learning. Metacognitive activity was represented by: the sum of the four observed assessments in critical thinking (CT); the sum of the four measures of perceived control of elaboration strategies (ES); the sum of the four measures of inhibitory control (IC); and the sum of the four measures of metacognitive experiences (ME).

Preliminary correlational analysis (Table 6-26) established that there were positive and significant relationships between total inhibitory control (the extent to which participants self-reported that they were able to avoid distraction), total metacognitive experiences (learning satisfaction, task-specific self-confidence, effort and response to challenge) and all of the pre- and post-test SRL scores. Therefore the metacognitive activity variables that were carried forward to subsequent path analysis were inhibitory control and metacognitive experiences.

Table 6-26 Correlational analysis for RQ7

	PreMK	PoMK	PreMB	PoMB	PreEfL	PoEfL	CT	ES	IC	ME
PreMK	1									
PoMK	**0.45**	1								
PreMB	**0.53**	**0.36**	1							
PoMB	**0.31**	**0.54**	**0.47**	1						
PreEfL	**0.68**	**0.41**	**0.46**	**0.20**	1					
PoEfL	**0.44**	**0.65**	**0.40**	**0.64**	**0.54**	1				
CT	-0.15	-0.04	-0.01	-0.06	-0.07	-0.05	1			
ES	-0.04	-0.13	0.05	-0.04	0.01	0.02	**0.49**	1		
IC	**0.36**	**0.44**	**0.35**	**0.53**	**0.28**	**0.55**	-0.14	0.10	1	
ME	**0.30**	**0.51**	**0.33**	**0.62**	**0.32**	**0.67**	-0.04	-0.05	**0.59**	1

PreMK: pre-test metacognitive knowledge; PoMK: post-test metacognitive knowledge; PreMB: pre-test motivational beliefs; PoMB: post-test motivational beliefs; PreEfL: pre-test efficacy for learning; PoEfL: post-test efficacy for learning; CT: critical thinking; ES: elaboration strategies; IC: inhibitory control; ME: metacognitive experiences; bold figures indicate statistical significance p < .05

The data conformed to the assumptions required for structural equation modelling (previously addressed in Section 6.7.2). Scrutiny of descriptives (Table 6-27) revealed no violations of the assumption of normality.

Table 6-27 Descriptive statistics for RQ7

	PreMK	PoMK	PreMB	PoMB	PreEfL	PoEfL	IC	ME
N	117	120	118	123	105	124	117	120
M	62.61	69.80	59.23	65.59	58.98	64.06	62.61	69.80
SD	7.29	6.27	5.82	4.25	5.69	5.09	7.29	6.27
Skew	0.09	-0.70	-0.32	-1.24	-0.20	-0.88	0.09	-0.70
Kurt	-0.77	0.55	-0.43	1.63	-0.78	0.15	-0.77	0.55

PreMK: pre-test metacognitive knowledge; PoMK: post-test metacognitive knowledge; PreMB: pre-test motivational beliefs; PoMB: post-test motivational beliefs; PreEfL: pre-test efficacy for learning; PoEfL: post-test efficacy for learning; IC: inhibitory control; ME: metacognitive experiences

A path analysis was constructed to model the set of relationships identified in Table 6-26. Thus pre-test metacognitive knowledge, motivational beliefs and efficacy for learning were set as predictors of both inhibitory control and metacognitive experiences. Similarly, both inhibitory control and metacognitive experiences were hypothesised to predict all post-test SRL components. The correlational analysis featured in Table 6-26 informed the inclusion of other parameters. Thus, parameters added to the model included the correlations between pre-and post-test SRL variables, and between inhibitory control and metacognitive experiences. Confirmatory factor analysis (the measurement model) for this analysis is reported in Section 6.4.1. The estimated model (Figure 6-20) provided a good fit to the data: χ^2 (5) = 5.195; p = .3926; RMSEA = .017; CFI = .984; TLI = .997.

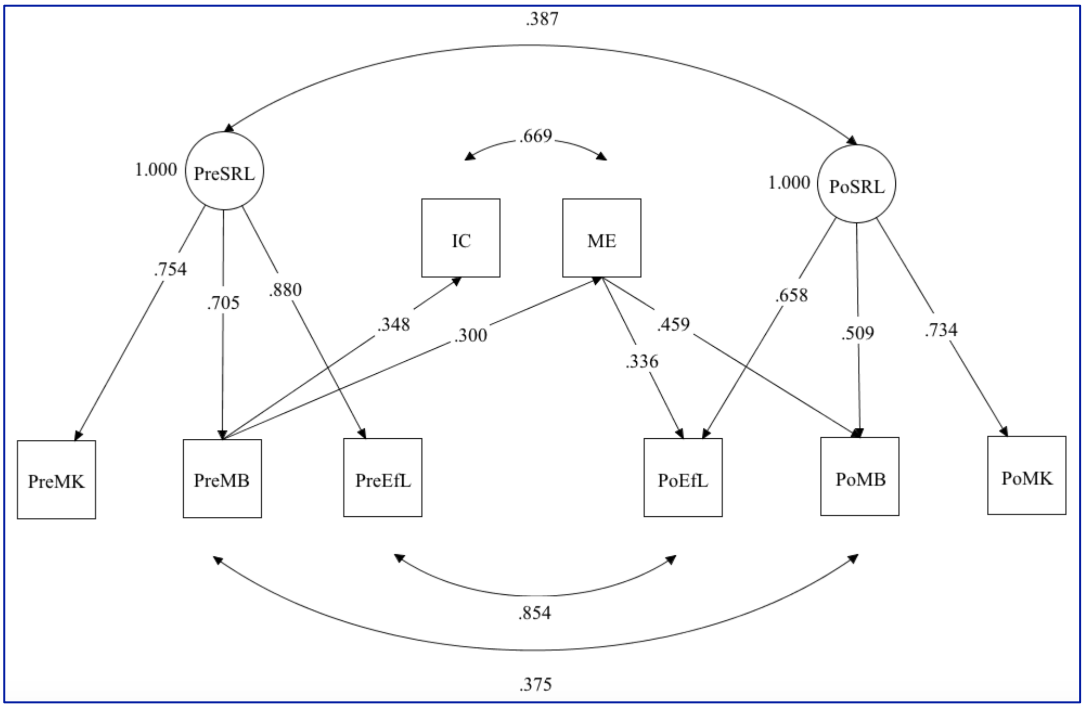


Figure 6-20 Path analysis model for RQ7
Significant paths shown ($p < .05$); standardised estimates; PreMK: Pre-test metacognitive knowledge; PoMK: Post-test metacognitive knowledge; PreMB: Pre-test motivational beliefs; PoMB: Post-test motivational beliefs; PreEfL: Pre-test efficacy for learning; PoEfL: Post-test efficacy for learning

For clarity, the non-significant paths are not included in the estimated model above (Figure 6-20), but they are listed in the table of parameter estimates (see Table 6-28).

Hence, the result of this path analysis provided some support for the conceptual model, proposed in Section 2.3.6, where it was proposed that pre-test SRL dispositions would influence SRL processes (metacognitive activity) during the intervention. In the final model, one component of pre-test self-regulated learning predicted both aspects of metacognitive activity. That is, pre-test motivational beliefs predicted both inhibitory control and the metacognitive experiences self-reported by the participants during the intervention, with medium effect sizes. It was also noted that metacognitive experiences during the intervention predicted both post-test efficacy for learning and post-test motivational beliefs, also with medium effect sizes.

Table 6-28 Parameter estimates for RQ7 path analysis

Parameters	B	β	Critical ratio β
PreMK to ME	.669[a]	.252[a]	1.914
PreMB to ME	**.983**	**.300**	**2.156**
PreEfL to ME	-.159[a]	-.051[a]	0.333
PreMK to IC	.245[a]	.258[a]	1.729
PreMB to IC	**.408**	**.348**	**2.275**
PreEfL to IC	-.330[a]	-.293[a]	-1.578
ME to PoMK	.087[a]	.267[a]	1.532
IC to PoMK	-.058[a]	-.064[a]	-.412
ME to PoMB	**.103**	**.459**	**3.821**
IC to PoMB	.023[a]	.037[a]	.316
ME to PoEfL	**.087**	**.336**	**2.646**
IC to PoEfL	.163[a]	.226[a]	1.788
PreSRL with PoSRL	**10.113**	**.387**	**3.207**
PreMK with PoMK	2.897[a]	.157[a]	1.149
PreMB with PoMB	**4.405**	**.375**	**3.503**
PreEfL with PoEfL	**6.422**	**.854**	**5.312**
ME with IC	**76.560**	**.669**	**10.816**

NB: A critical ratio of more than |1.96| indicates a two sided significance level of < .05 (Hox and Bechger, 1998); PreMK: Pre-test metacognitive knowledge; PoMK: Post-test metacognitive knowledge; PreMB: Pre-test motivational beliefs; PoMB: Post-test motivational beliefs; PreEfL: Pre-test efficacy for learning; PoEfL: Post-test efficacy for learning IC: Inhibitory control; ME: Metacognitive experiences; [a]Non-significant estimates

6.9 Summary of findings

Chapter Six has presented the results of the analyses. These are now summarised in the order of the research questions:

***RQ1**: What are the main and experimental condition effects of the intervention on change over time in participants' critical thinking?*

The first research question explored the effects of the intervention on change in participants' critical thinking (CT) scores over time. As hypothesised, repeated measures analysis of variance detected a statistically significant main effect for time, with a medium effect size ($r = .40$, $p < .001$), across four critical thinking assessment events. This suggests that an intervention such as this can have a positive effect on progress in critical thinking. However, this effect was not differentiated by group, suggesting that the effect was not due to specific experimental conditions. Analysis revealed an increasing effect for time, but a decreasing effect for group over the four learning events. All three experimental groups participated in argumentation activities, which could explain why the main effect was not differentiated by group. However, not all participants engaged in the intervention activities to the same extent. Therefore an analysis of the effect of actual participation in intervention activities, on progress in critical thinking was further explored in RQ2.

Key findings: Repeated measures analysis of variance detected that the intervention as a whole had a medium effect on participants' progress in critical thinking. The intervention effect was not however differentiated according to membership of groups. The rate of progress in critical thinking consistently increased from one measurement to the next, whilst the effect of the group decreased over time.

RQ2*: To what extent does participation in dialogic argumentation predict progress in critical thinking over time?*

Simple regression explored the predictive effect of the extent to which students participated in dialogic argumentation on change in their critical thinking scores over time. The analysis detected a small effect size ($r = .22$; $p < .001$).

Key finding: Participation in dialogic argumentation predicted change in critical thinking, with a small effect size.

RQ3*: What is the main effect of the intervention on change in participants' self-regulated learning (metacognitive knowledge, motivational beliefs and efficacy for learning) from pre-test to post-test, on groups of young adolescents, after one month, compared to a group who had no intervention experience?*

This research question estimated the intervention effect on change in participants' self-reported self-regulated learning over time. Specifically it employed latent change score modelling, a technique of structural equation modelling, and detected a significant main intervention effect, with a medium effect size, on change in overall self-regulated learning, and with small effect sizes on all of the self-regulated learning components (metacognitive knowledge, motivational beliefs and efficacy for learning). The group a student was placed in (experimental or control) accounted for 9.2% of the variance in the aggregated self-regulated learning change scores, 7.6% of the variance in metacognitive knowledge change scores, 7.4% of the variance in motivational beliefs change scores, and 6.6% of the variance in the efficacy for learning change scores. Thus, taking all intervention students together, the intervention had small to medium effects on change in their overall self-reported self-regulated learning dispositions and all SRL components, with the greatest individual effect being on change in participants' metacognitive knowledge.

On average, the lower a participant scored at pre-test, the greater their gain between pre-test and post-test; that is the lower the initial score, the steeper the gradient of increase. This suggests that students whose initial self-regulated learning scores were lower than those of others, benefitted the most from this intervention.

Key Findings: The intervention as a whole had a medium effect on learners' overall self-regulated learning dispositions and small effects on each of their self-regulated learning dispositional components (metacognitive knowledge, motivational beliefs and efficacy for learning). On average, the lower a participant scored at pre-test, the greater their gain between pre-test and post-test; that is the lower the initial score, the steeper the gradient of increase. This latter effect was consistent across all self-regulated learning dispositions, with medium effect sizes.

RQ4: *What are the effects of the separate experimental conditions on change in self-regulated learning and its components over time?*

A comparison of path estimates from each intervention condition to the latent change score for self-regulated learning, provided estimates of the effects of the interventions conditions, relative to the control condition. The intervention condition a participant was placed in accounted for 16.7% of the variance in self-regulated learning (SRL) change scores, 13.8% of the variance in the metacognitive knowledge (MK) change scores, 12.3% of the variance in motivational beliefs (MB) change scores, and 15.8% of the variance in efficacy for learning (EfL) change scores. Hence the group a participant was placed in made the greatest difference to their efficacy for learning scores, followed by metacognitive knowledge, followed by motivational beliefs. All experimental conditions (i.e., members of IC4, IC3 and IC2) experienced the Devil's Advocate activity and the critical thinking assessments. IC3 additionally had the opportunity to engage in daily dialogic argumentation, whilst IC4 additionally experienced the dialogic argumentation activity and the learning journal activity.

The main finding of this research question was that the most effective intervention condition, in terms of its predicted effect on change in overall self-regulated learning, metacognitive knowledge, motivational beliefs and efficacy for learning, with medium effect sizes, was membership of IC4: (SRL: $\beta = .49, p < .001$; MK: $\beta = .45, p < .001$; MB: $\beta = .42, p < .001$; EfL: $\beta = .46, p < .001$). This was the group which experienced all four aspects of the intervention: the learning journal, dialogic argumentation, Devil's Advocate and the critical thinking tests. Inspection of the difference in the group means, comparing each group with the control group, revealed that participants in IC4 made more than twice as much progress as those in any other group. This was the case for overall self-regulated learning, metacognitive knowledge, motivational beliefs and efficacy for learning, all with medium to large effect sizes.

Membership of IC2, the group which engaged in the Devil's Advocate activity and the critical thinking tests, in comparison with the control group (CG), was a significant predictor of change in overall self-regulated learning, metacognitive knowledge and motivational beliefs (.16 < β < .22) but not in change in efficacy for learning.

Membership of the learning journal group (IC4), in comparison with the non-learning journal group (IC3), was a significant predictor of change in SRL and all of its components, with small and medium effect sizes (.24 < β < .32).

Key findings: The most effective intervention condition, in terms of its effect on all self-regulated learning learner dispositions, was membership of IC4 (the group which experienced all four intervention conditions), with medium effect sizes. Students who were placed in the IC4 group reported at least twice as much change in all self-regulated learning dispositions, compared with students in any other groups. Thus the main intervention effects were more than doubled if a student additionally completed a learning journal. Whether learners completed learning journals or not was a significant predictor of change in SRL and all of its components, with small and medium effect sizes (.24 < β < .32). Whether learners engaged with the critical thinking activities (Devil's Advocate and critical thinking tests) or not was a significant predictor of change in overall self-regulated learning, metacognitive knowledge and motivational beliefs, with small effect sizes (.16 < β < .22), but not in efficacy for learning.

***RQ5**: To what extent does participation in dialogic argumentation predict change in self-regulated learning and its components over time?*

RQ5 used latent change score modelling to explore the effect of actual *participation* in dialogic argumentation, rather than *membership* of the intervention groups. The intervention activity explored in this research question was a) the count of participation in dialogic argumentation and b) the count of passive participation (browsing). Participation in dialogic argumentation predicted change in aspects of self-regulated learning, with medium and large effect sizes: self-regulated learning: $\beta = .56$, $p < .001$; metacognitive knowledge: $\beta = .52$, $p < .001$; motivational beliefs: $\beta = .42$, $p < .001$; efficacy for learning: $\beta = .43$, $p < .001$. Participation in dialogic argumentation was predicted in turn by a participant's motivational beliefs at pre-test ($\beta = .25$, $p < .01$).

The extent to which a participant browsed discussions, without actually contributing a post, was a predictor of change in efficacy for learning ($\beta = .18$, $p < .05$) and overall self-regulated learning ($\beta = .21$, $p < .05$). Browsing activity was predicted in turn by learners' metacognitive knowledge at pre-test ($\beta = .21$, $p < .01$).

Key findings: Engagement in dialogic argumentation predicted progress in all self-regulated learning dispositional components, with medium and large effect sizes. The greater a participant's motivational beliefs at pre-test, the more they tended to engage in dialogic argumentation, with a small effect size. Browsing activity predicted progress in efficacy for learning and overall self-regulated learning, with small effect sizes. The greater a participant's metacognitive knowledge, the more they were inclined to browse the discussions, also with a small effect size.

RQ6: *What are the relationships between aspects of metacognitive activity and critical thinking over time?*

A process analysis, using autoregressive cross-lagged panel analysis, explored the hypothesis that metacognitive skills (critical thinking, elaboration strategies and volitional strategies) and metacognitive experiences would be reciprocally related over time.

Two distinct sets of relationships were identified:

- **Critical thinking and the perceived use of elaboration strategies.** Critical thinking (CT) was reciprocally related to the perceived use of elaboration strategies (ES), with small and medium effect sizes ($.18 < \beta < .36$), over four learning events, such that elaboration strategies one week predicted critical thinking the next, and vice versa. Reciprocal relationships between critical thinking and elaboration strategies increased in magnitude of effect over time, with the strongest effect being the predictive effect of elaboration strategies on subsequent critical thinking. Critical thinking was related to elaboration strategies in the same week, during the first and the last weeks of the intervention, with small effect sizes.
- **Volitional strategies (inhibitory control) and metacognitive experiences.** Inhibitory control, the perceived extent to which participants avoided distraction, was synchronously related to metacognitive experiences (learning

satisfaction, self-confidence, experience of challenge and task-related motivation), with medium effect sizes. Self-reported metacognitive experiences consistently predicted subsequent inhibitory control, with small and medium effect sizes.

Key findings: Two distinct sets of relationships were identified. Firstly, consistent reciprocal cross-lagged relationships were detected between critical thinking and the perceived use of elaboration strategies, with small and medium effect sizes. Reciprocal relationships increased in magnitude of effect from one time to another. Secondly, volitional strategies, represented by participants' self-reported inhibitory control, were synchronously related to metacognitive experiences in three of the four learning events, with medium effect sizes. Metacognitive experiences consistently predicted subsequent inhibitory control, with small and medium effect sizes.

***RQ7**: What is the relationship between self-regulated learning dispositions and self-regulated learning processes during this intervention?*

The final research question modelled intervention data in a path analysis, in order to explore the conceptual model first proposed in the literature review (Section 2.3.6). The original conceptual model had proposed a structure which related SRL dispositions to SRL processes. Thus the hypothesised model for this research question explored the relationships between the SRL dispositions measured at pre-and post-test and the SRL processes measured during the intervention. Relationships were detected between pre- and post-test SRL dispositions (metacognitive knowledge, motivational beliefs and efficacy for learning) and two of the self-regulatory processes recorded by students in their weekly journals (metacognitive experiences and inhibitory control).

The estimated path analysis model provided some support for the conceptual model first proposed (Section 2.3.6). It was noted that perceived inhibitory control was related to motivational beliefs at pre-test, with a medium effect size, whilst self-reported metacognitive experiences were related to motivational beliefs at pre-test, and also to both motivational beliefs and efficacy for learning at post-test, also with medium effect sizes.

Key Finding: Motivational beliefs significantly predicted both metacognitive experiences (learning satisfaction, response to challenge and task-related motivation) and inhibitory control (the perceived capacity to avoid distraction), with medium effect sizes. Furthermore metacognitive experiences, reported during the intervention, predicted post-test efficacy for learning, also with a medium effect size.

7 Chapter Seven: Discussion

7.1 Introduction

The principle that motivated and guided the rationale for this study was that the development of both self-regulated learning and critical thinking skills should be essential objectives in the contemporary educational curriculum (Zimmerman, 2011; Marin and Halpern, 2011). The aim of this study was thus to evaluate the effectiveness of an educational intervention, in terms of its effectiveness to influence the specific learning outcomes of self-regulated learning and critical thinking.

The effects of intervention on aspects of self-regulated learning are well-documented, especially in terms of advantageous effects on academic achievement, motivation and the development of cognitive and metacognitive strategies (Schwinger, Steinmayr and Spinath, 2009; Schwinger and Stiensmeier-Pelster, 2012; Zimmerman 1990, 2009). Additionally, critical thinking is seen as an essential metacognitive skill for success at school and beyond, especially in view of its important role in the cognitive development of young adolescents (Inhelder and Piaget, 1958; Piaget, 1964; Flavell, 1979; Kuhn, 1992). Programmes that develop such skills for young adolescents however are rare (Perry, Phillips and Dowler, 2004; Marin and Halpern, 2011), particularly rare in England, with evaluation studies even more so. Thus there are few studies which are available to assist teachers in their choice of effective teaching and learning strategies. This study has attempted to address this issue by evaluating changes in self-regulated learning and critical thinking of young adolescents (N = 617) who have participated in an internet-mediated four-week extra curricular critical thinking course. The results of the study suggest that engagement in dialogic argumentation can have a beneficial effect on the development of critical thinking and self-regulated learning, whilst the completion of a learning journal can significantly enhance the intervention effects, and contribute to improved self-regulation of learning.

Thus the findings have provided answers to research questions that explored main and group effects of the intervention, and the effects of specific intervention activities, on change in aspects of self-regulated learning and critical thinking over time. A call for the exploration of relationships between aspects of metacognitive activity, including in particular metacognitive experiences, was also addressed in this study. Thus a process analysis explored such relationships as they were captured over four learning events during the intervention.

The following hypotheses were supported:

RQ1*:* It was hypothesised that there would be a repeated measures effect of the intervention in terms of students' progress in critical thinking from the first to the last assessment of the intervention. Repeated measures analysis of variance detected a medium effect size (r = .40).

RQ2*:* It was hypothesised that the extent to which participants engaged in dialogic argumentation would predict change in their critical thinking over time. Regression analysis detected a small effect size (r = .22).

RQ3*:* It was hypothesised that there would be a main effect of the intervention on change in a) overall self-regulated learning, b) metacognitive knowledge, c) motivational beliefs and d) efficacy for learning, when comparing the experimental conditions (Groups IC4, IC3 and IC2 together) with the control condition. Latent change score modelling detected small and medium main effects: overall SRL: β = .30; metacognitive knowledge: β = .28; motivational beliefs: β = .27; efficacy for learning: β = .26.

***RQ4**:* It was hypothesised that there would be a differentiated effect of the intervention conditions (on overall self-regulated learning, metacognitive knowledge, motivational beliefs and efficacy for learning) when comparing each of the separate experimental conditions with the control condition, with the most effective experimental condition predicted to be that experienced by IC4 (learning journal + debates + Devil's Advocate + critical thinking tests), followed by IC3 (debates + Devil's Advocate + critical thinking tests), followed by IC2 (Devil's Advocate + critical thinking tests). Latent change score modelling of the effects of intervention condition, in comparison to the control group, yielded small and medium effects. For overall self-regulated learning, the group effects were: IC4: β = .49; IC3: β = .24; IC2: β = .18. For metacognitive knowledge, the group effects were: IC4: β = .45; IC3: β = .22; IC2: β = .22. For motivational beliefs, the group effects were: IC4: β = .42; IC3: β = .22; IC2: β = .16. For efficacy for learning, the group effects were: IC4: β = .46; IC3: β = .19; IC2: non-significant. Whether learners completed learning journals or not was a significant predictor of change in SRL and all of its components, with small and medium effect sizes ($.24 < \beta < .32$).

RQ5: It was hypothesised that the extent to which a participant engaged in dialogic argumentation would predict gains in self-regulated learning and all of its components over time, from pre-test to post-test. Latent change score modelling detected medium and large effect sizes: SRL: $\beta = .56$; metacognitive knowledge: $\beta = .52$; motivational beliefs: $\beta = .42$; efficacy for learning: $\beta = .43$.

RQ6: It was hypothesised that reciprocal relationships would be detected between metacognitive skills and metacognitive experiences over time. Autoregressive cross-lagged panel analysis detected significant reciprocal relationships, with small and medium effect sizes, between critical thinking and use of elaboration strategies ($.16 < \beta < .36$), and also between volitional strategies (inhibitory control) and metacognitive experiences ($.20 < \beta < .44$).

RQ7: It was hypothesised that the metacognitive activity, self-reported in students' learning journals, would predict aspects of post-test self-regulated learning scores. Regression analyses detected large effects ($.54 < r < .70$) of the predictive effect of metacognitive activity on all post-test self-regulated learning variables.

The following hypothesis was rejected:

RQ1: It was hypothesised that the repeated measures effect on critical thinking would be differentiated by intervention condition, such that the most effective experimental condition was predicted to be IC4 (critical thinking tests + Devil's Advocate + dialogic argument + learning journal), followed by IC3 (critical thinking tests + Devil's Advocate + dialogic argument), followed by IC2 (critical thinking tests + Devil's Advocate). The analysis detected no such effect.

The discussion focuses first on the main findings relating to the main intervention effects, the effects of the separate intervention activities, the process analysis of the interaction between aspects of metacognitive activity over time, and the relationship between SRL dispositions and SRL processes. Limitations will be discussed, followed by a consideration of the extent to which this study could influence policy and practice in the relevant areas. There will be a discussion of the implications for future research and the thesis will end with a summary of its potential contribution to knowledge.

7.2 Aspects of the intervention

7.2.1 Main effects

There was a main effect of the intervention on change in aspects of learners' self-regulated learning dispositions (metacognitive knowledge, motivational beliefs and efficacy for learning). The main effect on change in self-regulated learning was differentiated by intervention condition. Whether students completed a learning journal or not predicted small and medium effects on change in self-regulated learning and all of its components ($.24 < \beta < .32$). Furthermore, the intervention was more than twice as effective, in terms of effecting change in all aspects of self-regulated learning, if participants completed a learning journal in addition to the other intervention activities. Critical thinking activities (the Devil's Advocate and the critical thinking tests) had a small but significant effect on change in overall self-regulated learning, metacognitive knowledge and motivational beliefs ($.16 < \beta < .22$), but not on change in efficacy for learning. Links between self-regulated learning intervention and changes in self-regulated learning dispositions are well documented (Boekaerts and Corno, 2005; Dignath, Buettner and Langfeldt, 2008) and the main finding of this study corroborates such research. The latter authors, for instance, in a meta-analysis of 263 studies of primary school SRL interventions, calculated a mean effect size of $d = .73$ for the effect of self-regulated learning interventions on pre- to post-test change in the use of cognitive and metacognitive strategies and $d = .79$ for the mean effect on motivational aspects. Both of these results represented medium effect sizes. These results are comparable with the findings of this study, which also detected a medium effect size for the main intervention effect.

Higgins, Baumfield, Lin, Moseley, Butterworth, Downey, Gregson, Oberski, Rockett, and Thacker (2004) in their Institute of Education meta-analysis of thinking skills interventions, concluded that, ‘there is some evidence that there may be greater impact on low attaining pupils, particularly when using metacognitive strategies’ (p. 4). This study appears to have provided evidence to support that notion, since students who scored themselves lowest at the beginning of the intervention consistently made the most progress in terms of change over time in each of their self-regulated learning dispositions. Whilst the self-report questionnaire on self-regulated learning does not produce an ‘attainment’ score, the finding does however, suggest that this intervention may be of particular relevance for students who self-report themselves as low on self-regulated learning questionnaires. This finding has relevance therefore, for programmes which seek to raise the self-perceptions of the less confident young learner. Such programmes currently exist in the range of offerings from university outreach programmes such as that run by the University of Oxford (Oxford, 2015), whose aims (p.1) include encouraging ‘awareness of the learning journey and the steps on it’ (i.e., metacognitive awareness), inspiring ‘intellectual personal curiosity’ (i.e., developing motivational beliefs), and enriching and enhancing ‘students’ self-esteem’, (which according to Bong and Skaalvik (2003) is related to self-efficacy). Thus, it appears that the intervention has the potential to continue to fulfil its original intention to provide an effective programme for the development of self-regulated learning amongst young adolescents of above average intellectual ability.

The intervention as a whole had a main effect, with a medium effect size, on students’ critical thinking scores. This is an important finding considering the importance placed upon the development of critical thinking in young adolescents (Piaget, 1964; Flavell, 1979; Kuhn, 1992), and even more so considering that critical thinking does not feature as a specific course objective on the English National Curriculum until As Level (i.e., for 17 year olds). The finding that the critical thinking skills of young adolescents can be enhanced through argumentation intervention is therefore one which has the potential to have an impact on curriculum policy, curriculum design and teacher practice.

In both the pilot and the main study, the main effect for change in critical thinking was not differentiated by intervention condition. This lack of group effect could be explained by the similarity between the content of the Devil’s Advocate activity and the questions posed in the critical thinking assessment. Both required participants to explore both sides of ethical dilemmas about which there was no clear solutions. Findings reported in a recent paper by Shehab and Nussbaum (2015) may shed further light on this lack of group effect. The authors explored the comparative effectiveness of two argumentation strategies: the explicit and simultaneous exploration of both sides of an argument; and the strategy of proposing an argument which attempts to minimise the opportunities for an opponent to offer a rebuttal. They reported on the superiority of the former strategy in terms of effects on measures of critical thinking. In the current study, students in all experimental groups were required to complete the core intervention task, which involved that former strategy. It follows that it is reasonable to assume that the effect of this strategy did indeed contribute to the lack of group effect, i.e., to the finding that all intervention participants made significant progress in critical thinking.

It is also possible that the completion of the critical thinking tasks produced a method effect, whereby the measurement procedure itself may have contributed variance to the repeated measures scores (Maul, 2013). In other words, the improvement in critical thinking could have been effected simply through the practice of the critical thinking tests. This is unsurprising; both Kurfiss (1988) and Ennis (1993) suggested that critical thinking could be developed through questions and answer exercises - and it is feasible to assume that the practice of reasoning leads to increased reasoning capacity. However, this does represent a limitation of the findings; whilst a repeated measures effect for improvement in critical thinking was detected (r = .40), it may well have been an outcome of the test itself (Ennis, 1993).

Much has been written about the need to develop critical thinking programmes, with a common rationale referencing the exponential increase in new technologies, new knowledge and need for new skills (e.g., Halpern, 2003; Ku, 2009). Furthermore, cognitive scientists and developmental psychologists have been claiming for several decades that a learner's capacity for critical thinking emerges and develops rapidly in early adolescence (Inhelder and Piaget, 1958; Piaget, 1964; Flavell, 1979; Kuhn, 1992, 2006; Casey, Giedd and Thomas, 2000). It follows that there is a perceived need for critical thinking programmes for this age group. However, critical thinking objectives, whilst being woven through the English National Curriculum, are not specifically taught until As Level, and then only if students opt for the course. Previous research has reported that critical thinking is best taught through a *mixed* approach consisting of specific critical thinking instruction alongside activities that involve critical thinking (Ennis, 1998; Abrami *et al.,* 2008). The evaluation of the current intervention, which used such a mixed approach, appears to support that notion, reporting as it does, large main effects for change in critical thinking. It is argued therefore, that the main critical thinking finding, that forum dialogic argumentation activity can significantly increase the critical thinking of young English adolescents, is one which has the potential to make a contribution to knowledge.

7.2.2 The learning journal

The latent change score analysis of intervention conditions compared with the control condition detected that the most effective condition, in terms of its effect on change in aspects of self-regulated learner dispositions, was membership of IC4 (the group which experienced all four of the intervention conditions). For instance, analysis detected that the membership of IC4 was twice as effective as that of IC3, in terms of precipitating change in self-regulated learning, in comparison with the control group: (IC4: $\beta = .49$, compared with IC3, $\beta = .24$), metacognitive knowledge (IC4: $\beta = .45$, compared with IC3, $\beta = .22$), motivational beliefs (IC4: $\beta = .42$, compared with IC3, $\beta = .22$) and efficacy for learning (IC4: $\beta = .46$, compared with IC3, $\beta = .19$). An estimation of the specific effects of the learning journal, effected by comparing the predictive effect of membership of IC4 (the group which completed the learning journal), with that of IC3 (the group which did not), detected that whether learners completed learning journals or not was a significant predictor of change in SRL and all of its components, with small and medium effect sizes ($.24 < \beta < .32$).

The finding reflects that of Efklides (2008) who reported that metacognitive knowledge can be improved through self-reflection. It is also in accord with the findings of Zimmerman (2011) and Schmidt, Maier and Nückles (2012), who all reported that the use of learning journals could enhance students' learning motivation. In the latter study, they found that using a learning journal had an effect on students' enthusiasm for the topic they were studying at the time, as well as on their longer term motivational beliefs, resulting in an increased expenditure of effort on the subject of the moment. There is also evidence that the process of articulating goals, as the students were required to do in their learning journals, can result in a significant increase in motivation (Locke and Latham, 2002). This is an important finding therefore, albeit a corroborative one, as improved motivation and effort are considered important conditions for effective learning (Efklides, 2011; Ryan and Deci, 2000).

The aspect of self-regulated learning for which the learning journal made the most difference however, was efficacy for learning, where again, those who completed the learning journal reported more than twice as much progress in the self-reported efficacy for learning scores, compared with those who did not. It is possible that the action of self-reflection had an effect on the learner's belief in their capacity to learn. This is a reasonable assumption since self-reflection of mastery experiences is a key self-regulatory activity (Zimmerman, 2008), and mastery experiences are a powerful predictor of self-efficacy beliefs (Bandura, 1997, 2006). However, it is an important result in the current context since few studies have so far explored the relationship between self-reflection in a learning journal and its specific effect on change in self-efficacy. It is argued therefore that this finding makes a contribution to knowledge.

The main effects of the intervention were doubled if learners also made use of a learning journal to record their metacognitive reflections. Why should the completion of a learning journal enhance the effect of an intervention? It is suggested that the completion of the learning journal serves as a cue for the learner to monitor and review their self-regulatory habits. Webber, Scheuermann, McCall and Coleman (1993) argue for instance that improvement in self-regulation is likely because of the learner's increased awareness of their metacognitive activity. In the journals used in the current intervention, students were asked about, and therefore reminded about, a number of their self-regulatory habits. Learners were prompted to articulate their goals and examine the extent to which they had attained them, their metacognitive awareness of the cognitive and metacognitive learning strategies employed during their learning process, their volitional strategies and their metacognitive experiences (learning satisfaction, task-related motivation, response to challenge, self-confidence etc). Thus, prompting students to evaluate their metacognitive knowledge of their elaboration strategies *during the intervention* may have played some part in the increase (with medium effect size) in that corresponding metacognitive knowledge over the course of the intervention (i.e., from pre-test to post-test). Similarly, following the assertions of Efklides (2009), it could be argued that prompts that encouraged learners to consider their metacognitive experiences may have contributed to the increase (with medium effect size) in change their efficacy for learning over time and/or their motivation (also with a medium effect size). As the latter two findings have not been widely replicated elsewhere, further research would be necessary to *isolate* these possible effects. The important hypotheses to test would be the effects over time, on participants' motivational beliefs and efficacy for learning, of the completion of learning journals which focus specifically upon learners' metacognitive experiences.

The experience of completing the learning journal may have contributed to the significant increases over time in the specific variables measured in that journal, with a large effect in change over time in elaboration strategies and small effects in volitional strategies and metacognitive experiences. Being reminded for instance about their strategies for avoiding distraction may have led students to consider ways to enhance those strategies, possibly accounting for the significant positive trend in volitional strategies over the four week period.

Whether participants used a learning journal or not was a significant predictor of change in self-regulated learning and all of its components, with small and medium effect sizes. The students in this study completed their learning journals at the end of a learning episode, i.e., at the end of every week, for four weeks. Schmitz and Perels (2009) on the other hand, framed a learning event as one day, for 49 days, detecting similar effects on general self-regulated learning and self-efficacy. However, the analytical methods for these analyses were different, therefore an exact comparison is inappropriate. Very little research has been conducted into the optimum number of journal events required for its overall effectiveness, either as a learning tool or as a research tool, nor the intervals between each journal event. This presents a potential future line of inquiry in this emerging field of study.

Schmitz and Wiese (2006) suggest that the use of a learning journal is likely to have an effect on self-regulation not only because its main purpose is to stimulate metacognitive activity, but also because the repeated practice of completing the journal is a self-regulatory habit in itself, and practice is likely to lead to improvement by its very nature. This phenomenon, known as the 'reactivity of self-monitoring', has been further documented by Schmitz and Perels (2011, p. 256), with the principal idea being that the act of self-recording precipitates improved self-regulation. This is a promising notion for educational policy and practice - that simply recording one's self-regulatory activity has the potential to change self-regulatory behaviour. However, Schmitz and Perels (*op cit.* p.256*)* go on to say that 'literature does not present data' which explores this phenomenon, either in specific settings, or over successive periods of time. It is argued therefore, that this study contributes to the body of knowledge by offering evidence of the effectiveness of the use of a learning journal over a specific length of time (four weeks), over a specific number of learning episodes (four) and for a specific educational setting (young English adolescents who participate in an intervention provided via the internet). The relevance of self-regulated learning as an essential life skill is growing (De Bruin and Van Gog, 2012; Schwonke, Ertelt, Otieno, Renkl, Aleven and Salden, 2013), since teachers cannot always be at the side of the learner as they venture into new learning territories. Thus a learning method which assists in that process can be seen as a promising tool.

Alongside the educational advantages already mentioned, there are other advantages based upon the method of delivery explored in this study. The internet-based learning journal is particularly easy and cost-effective to administer. The platform chosen to administer the one in this study, for instance, uses a free cloud-based *Google* system for accessing, recording and presenting the journals. Participants only need an internet link to be able to access it, and any number of students can complete it, simultaneously or asynchronously, 24 hours a day. Thus whilst it is being used, it requires no teacher-presence and no paper resources. One teacher can supervise the journals of hundreds of students, since reviewing the content requires the teacher only to download only one file, which is in the form of an Excel spreadsheet. Learning journals are presented already transcribed and can be searched, categorised, sorted and assessed, with much less effort and expense than it takes to transport and evaluate paper journals.

A second, more interactive learning journal, can be delivered via a learning platform such as *Moodle* (Passey and Higgins, 2011). Here, the student's journal is an interactive file that is shared online with the teacher. Thus the teacher receives an email the moment the journal is submitted, and the teacher can respond at a time to suit their own schedule. The journal can be viewed, amended and resubmitted at any time. Again the cost of the platform is free, as *Moodle* is an open-source software. It is assumed that one teacher can supervise many students, thus making a significant difference to costs. However, very little research exists that evaluates how many students can effectively be supervised by one teacher in such a situation. Thus future research should focus upon this line of inquiry.

7.2.2.1 How can schools implement this learning method?

The importance of the learner's involvement in their self-development has been emphasised by several writers (e.g., Boekaerts and Corno, 2005; Zimmerman, 2008). It is generally seen as crucial that individuals have the capacity to take control of their own learning process, not only during formal education, but throughout life. The learning journal provides a means of taking pro-active, self-directed control, through self-reflection of the learning experience. Barclay (1996) suggested that learning journals would become an integral part of many future learning programmes. But more than a decade later it is reported that the implementation of the learning journal in the English school setting is rare and that methodological challenges (such as allocating time for completion by students, or for monitoring of journals by teachers) present themselves in the traditional classroom context (Stephens and Winterbottom, 2010). However, it is argued here that an internet-based learning platform offers an innovative and effective medium for the implementation of the learning journal, not least because it has been demonstrated that recording student reflections has the potential to improve their learning experiences (Mehmet, 2010), but even more since the internet provides unprecedented opportunities for low cost and ubiquitous provision to large numbers of students across a wide geographical field and across a broad spectrum of settings.

It is argued that policy makers and educators should consider implementing learning journal methodology into curriculum practice. Based upon the effectiveness of the learning journal created for this study, a journal would have the potential to be an effective self-regulatory tool if it prompted learners to reflect upon the whole cycle of self-regulation (Zimmerman 2008, 2011). Thus an effective learning journal would include a number of key features, and would be presented in a language suitable to the learners' age and their specific educational setting. It would be completed at regular intervals and would be structured such that the participants were prompted to metacognitively reflect upon their learning experience. Using either text responses or multiple choice, learners would be prompted to explore their learning experiences across three cyclical aspects of self-regulated learning - the stages of forethought, performance and self-reflection.

In the forethought phase of the journal, the student would be encouraged to articulate learning goals, justifying particular choices. This would set the scene for later metacognitive exploration. Depending upon the context, the student would be encouraged to explore the cognitive strategies (such as memorisation, imagery, mnemonics, research strategies, self-testing or mind-mapping) they would be expecting to use, in order to address the learning task, and their specific learning goals. They would be prompted to explore a number of sources of motivation: their belief in their capacity to succeed in the task (i.e., self-efficacy); their outcome expectations; their level of interest in the task; the amount of effort they intended to expend; and the value they attached to the task. Finally, they would be encouraged to explore their goal orientations (Boekaerts, Smit and Busing, 2012) - in terms of whether their purpose in achievement was to satisfy certain goals (such as intellectual growth goals or well-being goals).

At intervals during the performance phase of the learning event, for example once a day, as in the Schmitz and Wiese (2006) study, or once a week as in the current study, the learner would be prompted to reflect upon the action phase of their learning. The purpose of such prompts would be to highlight strategies which were having the intended effect or otherwise, and thus lead the learner towards the deduction of causal attributions, that might in turn precipitate beneficial change in learning behaviour during subsequent learning events. For instance, the learner would be encouraged to report on their volitional strategies - the specific strategies that they had used to avoid distraction, such as switching off the TV, studying in a separate room, or taking steps to avoid interference from social networking. They would reflect upon exactly which cognitive strategies were being put to use, what effect they might be having, and whether as a result they might carry on using them. They would report on the effectiveness of help-seeking, whether it be from peers, teachers or parents.

At the end of the learning event, students would complete the last part of their journal. This would be structured such that students would be guided through a self-evaluation, intended to inform their strategy use in subsequent learning tasks. Thus examining the efficacy of particular cognitive strategies, in terms of their comparative success, would be a useful exercise at this stage - this would encourage the evaluation of strategic outcomes. Metacognitive experiences would be expressed in the form of learning satisfaction, difficulties experienced and levels of confidence. Difficulties, lack of understanding and weaknesses would be identified, remedial strategies would be explored, and unanswered questions articulated. An important component of this last phase of the learning journal therefore would be prompts that would draw the learner's attention to causal attributions which could be deduced from the analysis of the contents of the learning journal. This would be important for re-motivating the student to carry on with the process of self-evaluation and for ensuring that they saw the value in continuing with the completion of the learning journal itself. Finally revised goals would be voiced before the self-regulatory cycle began once more.

Such a learning journal could be integrated into current teaching practice using innovative technology, without creating extra burdens for teachers. After the initial effort of creating the journal, teacher-involvement would be minimal. The pro forma would reside on a cloud-based or locally-based server (such as Google or a school intranet or learning platform). Students would access the journal via an internet-link at set times during the working week (e.g., in tutor-time, following registration periods[12], or as a homework activity). For instance this activity might substitute for other 'tutor-time' activities which the tutor usually has a duty to provide. Once completed the journals would be available for teachers to download in their own time. Thus, it is proposed that implementation of the learning journal into the curriculum, alongside existing commitments to curriculum activities, need not impact

[12] Once or twice a day, students in English secondary schools return to their class tutors for 'registration' during which time there are also opportunities for 'tutor-time' activities. It is generally the duty of the tutor to provide educational activities for the students during this time.

deleteriously on curriculum management and indeed has the potential to make a positive contribution to learning outcomes.

This section has discussed the aspect of the intervention which involved students completing a weekly learning journal. The significant and positive effects, on self-regulated learning and all of its components, of using the learning journal or otherwise, have been discussed in the context of how educators can integrate this learning method into existing curriculum practices.

The following section discusses the second major learning method of the intervention: dialogic argumentation.

7.2.3 Dialogic argumentation

Analysis of intervention effects revealed that the more a participant engaged in discussion (by contributing posts to an online forum), the greater the gain in their self-reported self-regulated learning, from pre-test to post-test. This effect was consistent for all of the self-regulated learning dispositions (metacognitive knowledge, motivational beliefs and efficacy for learning). Effects were of medium and large size with the greatest effect being seen on participants' metacognitive knowledge.

This is an important result that has the potential to contribute knowledge to the field of self-regulated learning. The relationship between dialogic argumentation and metacognition has already been explored (Kuhn, 2005; Cavagnetto, 2010) and the findings of this study corroborate the assumption that participation in dialogic argumentation can have a significant effect over time, on learners' metacognitive knowledge. However very few studies have explored the notion that engaging in dialogic argumentation can have an effect on self-regulated learning aspects other than metacognition - such as an individual's motivational beliefs or their self-efficacy. This is surprising since Kuhn (1992) hinted at such possible relationships when she suggested that, in argument, the reactions of other protagonists in the discussion act as a mirror, reflecting a learner's self-belief, thus enabling that learner to discover aspects of the self. Furthermore, theoretical links have also been posited between 'collaborative activity' (students working together, as in, for instance, group discussion) and motivational beliefs (Wolters, Benzon and Arroyo-Giner, 2011). The inter-relationship between metacognitive processes, self-efficacy and motivation are also well documented by a number of social-cognitive writers (Bandura, 1986;

Zimmerman, 2000; Zimmerman and Cleary, 2006). It follows that an activity that promotes metacognitive awareness may also promote self-efficacy and motivation. This study has explored that notion and has found that, for the participants here involved, participation in dialogic argumentation did indeed have an effect on both their motivational beliefs and their efficacy for learning, with medium effect sizes. Considering the potential effects on academic performance of increasing a learner's capacity for self-regulated learning, such as the mean effect sizes (d = .62) reported in the meta-analysis of Dignath, Buettner and Langfeldt, (2008), it is argued here that dialogic argumentation has the potential to impact upon learning outcomes via increasing a range of self-regulatory mechanisms.

The extent of participation in dialogic argumentation was a predictor, not only of change in self-reported self-regulated learning dispositions, but also of change in observed critical thinking skills (with a small effect size). Analysis revealed that the more a student participated in daily dialogic argument, the more progress they made in their end of week critical thinking scores. This confirms the notion that critical thinking is a function of direct engagement with issues that exercise the higher order thinking skills of evaluation, inference, deduction and reflective judgement (Dwyer, Hogan and Stewart, 2014).

Actual participation in dialogic argumentation accounted for 12% of overall argumentation activity, with the browsing of posts (without posting replies) accounting for the remaining 88%. This corroborates the findings of Preece, Nonnecke and Andrews (2004), who suggest that actual posting accounts only for a small fraction of overall discussion activity in online communities. The browsing of course discussions (without actual participation) had effects, with small effect sizes, on change in aspects of self-regulated learning (notably efficacy for learning), but not on change in critical thinking. This contributes to emerging notion that the activity of browsing discussions can have effects on learning outcomes (Dennen, 2008). It also supports the assumption however, that browsing is not as effective as actual posting participation (Preece *et al., op.cit.*). Whilst these effects have also been hypothesised elsewhere (Beaudoin 2002; McKendree, Stenning, Mayes, Lee and Cox, 1998), they have not been extensively empirically tested; therefore it is argued that this finding is one that merits corroborative study, especially across different settings, such as age groups other than those around the undergraduate range.

A number of writers have suggested that argumentation is an essential skill that has the potential to develop critical thinking (Kurfiss, 1988; Bell, 1991; Sanders, Wiseman and Gass, 1994) and the finding that participation in dialogic argumentation predicted gains in critical thinking appears to confirm that claim. However, few such studies have reported effect sizes, and it is therefore suggested that this study makes a contribution to the field by addressing this important gap in knowledge.

7.2.3.1 How can schools implement this learning method?

Based upon the effect sizes detected in this study, it is argued that the promotion of dialogic argumentation has the potential to be of benefit to both learners' critical thinking and aspects of their self-regulated learning. This intervention was delivered on an internet-based learning platform, but the essential components of dialogic argumentation, as implemented in this study, are arguably transferrable to the classroom setting. Of prime importance is the subject matter, which should be of a controversial nature and which should explore matters of social, moral or ethical dilemmas, about which more than one perspective could be assumed to exist. Teachers should be trained to ensure that their questioning technique involves open-ended questions which encourage participants to make claims, back up their claims with evidence, negotiate their way through logical thinking and make conclusions based upon evaluative reasoning. Importantly, they should be aware of how to move students from one level of debating skill to the next. Suggestions about how to implement open-ended questioning in a dialogic forum are detailed in Section 2.2.4.6. Students should be encouraged, in particular, to explore both sides of an argument, and ways should be found (such as by the tutor offering alternative perspectives) to encourage students to engage in counter-argument.

Of particular note is the finding that the *extent* of a student's participation predicted gains in critical thinking and aspects of self-regulated learning. Therefore, teachers should explore ways to encourage students to participate (rather than passively listen) in discussions and debates. For instance, awards (or electronic 'badges') could be awarded for levels of contribution, as a means of motivating students to actively participate.

7.2.4 Metacognitive activity

This section will discuss the findings of the process analysis. Autoregressive cross lagged panel analysis estimated relationships between aspects of metacognitive activity, recorded as a) critical thinking assessments and b) self-reports in students' learning journals at the end of each week.

The process analysis identified two clear sets of relationships. Critical thinking appeared to be related to the perceived use of elaboration strategies, whilst inhibitory control was related to metacognitive experiences. Zimmerman (2008) highlighted the possibility of 'potentially causal influences of self-regulated learning processes across phases' (p.178), (by which he is referring to the cyclical phases of the self-regulated learning process), and this finding appears to exemplify such a notion.

Magno (2010) suggests that it is important to evaluate the extent to which metacognitive skills such as elaboration strategies are related to the process of critical thinking. This is in the light of his reporting a number of studies that assert a relationship between the two constructs, but which fail to adequately experimentally test that assertion (e.g., Halpern, 1998; Choy and Cheah, 2009). In the current study, critical thinking was reciprocally related to the perceived use of elaboration strategies, with small and medium effect sizes, over four learning events, such that elaboration strategies one week predicted increased critical thinking the next, and vice versa. The finding can be logically explained: critical thinking was measured in this study as the student's capacity to perceive the strength of an argument; whilst the elaboration strategies measure referred to items such as 'the extent to which I considered a dilemma before forming my argument'. The result is however interesting since the critical thinking variable was an observed one, whilst the elaboration strategies one was self-reported, thus confirming an element of methodological triangulation, reportedly so often missing from self-report studies (Denzin, 2006; Hopfenbeck and Maul, 2011).

This suggests that an intervention that focuses upon critical thinking has the potential to positively impact on the acquisition of metacognitive strategies (and vice versa). Sanders, Wiseman and Gass (1994) speculated as such when they suggested that argumentation intervention should enhance both critical and complex thinking; their argument proposed that if effective arguing engaged both critical thinking and the use of cognitive/metacognitive strategies, then argumentation intervention should enhance both outcomes. The finding is interesting however, since other writers (such as Brown, 2004 and Phan, 2006) posited that such strategies *preceded* critical thinking, but not the reverse, (i.e., critical thinking predicting the use of elaboration strategies) as has additionally been found in the current study. Despite the contention that 'the two theoretical orientations operate in a cyclical pattern' (Phan, 2010, p. 289), this reciprocal relationship does not feature widely in empirical studies and therefore merits further investigation. Research is clearly required to disentangle this relationship further.

The cross-lagged reciprocal effect between critical thinking and the use of elaboration strategies also appeared to increase in magnitude, thus suggesting that the longer the intervention, the stronger the potential relationships between these variables. However, this suggestion should be tested in terms of the optimum length for such an intervention, as the strength of the relationship may wane due, for instance, to boredom or over-use. This notion is related to the question of the most effective length of time for journal use, and also the duration of the intervals between each journal event, discussed in Section 7.2.2.

The critical thinking assessment had asked participants to assess the strength of an argument, whilst the elaboration strategies questionnaire items asked participants to determine the extent to which they reflected upon a dilemma before forming an argument, so the relationship is a logical one. However, it is interesting that the relationship was cross-lagged, more often than it was synchronous, which suggests that encouraging a learner to reflect upon their use of elaboration strategies at one time, can have a measurable, and delayed, effect on later critical thinking achievement. The finding has implications for teaching programmes that take place prior to examination periods, such that exercises in the use of elaboration strategies may have the potential to impact, at least in the short term, on later measured learning outcomes. Further research is required however, to determine whether this finding is replicable in other settings, and whether such effects would hold up under the scrutiny of a longitudinal study.

The second discernible pattern suggested that metacognitive experiences (Efklides, 2008) were consistently related to the volitional strategy of inhibitory control, a metacognitive skill (*ibid.*), with small and medium effect sizes ($.20 < \beta < .44$). However, neither variable was consistently related to the other metacognitive skills such as the control of elaboration strategies and critical thinking. This could partly be explained by the notion that attention-related skills are conceptually different to other self-regulatory skills - and relate to each other in unique ways (Barriga, Doran, Newell, Morrison, Barbetti and Robbins, 2002).

In this study *inhibitory control* represented the volitional strategy of controlling the extent to which the student avoided distraction. ‘Metacognitive experiences’ referred to the extent to which learners were a) satisfied that they had made progress, b) were happy to give up their free time in order to study, c) enjoyed the intellectual challenge, d) made a valuable contribution to the course, and e) felt confident in themselves as learners (Efklides and Petkaki, 2005). It seems reasonable to assume that students with positive metacognitive experiences will be motivated to control any predilection for distraction. Similarly, the less they were distracted the more likely it might be that they were having a positive learning experience. The finding has implications for the design of learning programmes that are delivered via media that may have the potential to distract, such as the internet. The relationship between inhibitory control and metacognitive experiences is a particularly interesting one; these two constructs are, according to Efklides (2008), theoretically related because metacognitive experiences provide the ‘trigger’ that not only controls metacognitive decisions such as regulation of effort, but also feed self-appraisal - thus contributing to the cycle of self-regulation (p. 280). Studies which test this notion however are rare, especially in the context of English children, and also in the contemporary context of internet mediated learning. Furthermore, despite calls from researchers such as Selmeczy and Dobbins (2013) for further exploration of the relationship between inhibitory control and other aspects of metacognitive activity, such studies are also hard to find. Hence it is argued that this finding contributes to the knowledge-base which is emerging with regard to the nature of metacognitive experiences (Efklides, 2001, 2008, 2009, 2011, 2012; Efklides, Kourkoulou, Mitsiou and Ziliaskopoulou, 2006). However, whilst there may be merit in exploring the relationships between inhibitory control and other aspects of metacognitive activity, research should first establish the extent to which *perceived* (self-reported) inhibitory control is related to *actual* inhibitory control behaviour, in an educational context.

7.2.5 The relationship between SRL dispositions and SRL processes

The final research question sought to revisit aspects of original conceptual model proposed towards the end of the literature review (Section 2.3.6). Specifically the model summarised the notion that SRL dispositions influence SRL processes. Thus, the pre- and post-test SRL data and the SRL process data (i.e., metacognitive activity) gathered during the intervention provided an opportunity to explore the conceptual model with intervention data. The result of this analysis partly corroborated the conceptual model, providing limited support for the notion of a relationship between SRL dispositions and SRL processes.

Path analysis thus explored a) the extent to which pre-test SRL dispositions predicted the metacognitive activity which was self-reported by participants in their learning journals, during the intervention, and b) the extent to which that metacognitive activity predicted post-test SRL dispositions. The path model detected that motivational beliefs significantly predicted both metacognitive experiences (learning satisfaction, response to challenge and task-related motivation) and inhibitory control (the perceived capacity to avoid distraction), with medium effect sizes. Furthermore metacognitive experiences, reported during the intervention, predicted post-test efficacy for learning, also with a medium effect size.

The result provides empirical evidence for theoretical assumptions made by Vollmeyer and Rheinberg (2006), co-authors of the motivational beliefs scale adapted for this study (Rheinberg, Vollmeyer and Burns, 2001). They, along with other prominent authors such as Boekaerts (1996), theorise that initial motivational beliefs influence performance motivation. The results of this research question detected a relationship between these constructs, such that pre-test motivational beliefs had an effect on the metacognitive experiences (which included task-related motivation) reported during the intervention, with a medium effect size. The finding also extends Vollmeyer and Rheinberg's assumption by providing evidence that motivation during the learning process continues to influence a learner's disposition towards motivational beliefs, once the learning event is over, also with a medium effect size.

7.3 Limitations

Despite all practical implications for educational policy and practice, the interpretation of these results is subject to three general limitations and two context-specific ones, in addition to those already discussed elsewhere (Critical thinking assessment: Section 2.2.3; The learning journal approach: Section 2.3.7.5; Sampling: Section 4.4; Critical thinking tests: Section 4.5.1; Repeated measures designs: Section 4.7.4; and Methodological limitations: Section 4.9).

The general limitations refer to sample, the duration and the time on task. Firstly, the study sampled students from state schools in England, who were also of above average literacy ability - the students who most closely matched the outreach target population. Whilst the findings appear to have relevance for university outreach programmes which target children who have the potential to participate in higher education, and for schools which aim to provide extension activities for the more able, the results cannot currently be generalised to the wider population of young adolescents (i.e., a population which includes those of average and below average literacy ability). However, as has been already noted, the intervention was most effective for those students who scored *lower* on self-regulated learning competencies at pre-test, so further research is necessary to establish whether indeed this intervention may be suitable for a wider audience.

The second general limitation concerns both length and context of the intervention. Significant effects were detected after just four weeks, but longer-term effects were not analysed as part of this study. It would remain to be seen whether or not there were longer-lasting effects, and whether those effects were transferable across domains of study.

Thirdly, time on task was not held constant for each of the groups, since the learning journal group was allocated additional learning time compared to other groups, specifically in order to complete the journal. This additional learning time could have contributed to the difference in effects between the groups. Further studies should isolate the effect of the learning journal whilst compensating for this limitation.

Context-specific limitations refer to the characterisation of critical thinking (CT) and self-regulated learning (SRL) as they are referenced in this study. It is acknowledged that in a quantitative study such as this, the interpretation of the findings must be limited not only to the characterisation of the constructs, but also to the specific measures used to evaluate them. Firstly, the characterisation of the construct of CT, which underpinned both the creation of the CT elements of the intervention and the CT assessments designed to evaluate those elements, was restricted to a narrow definition:

> *'… a metacognitive process that, through purposeful, reflective judgement, increases the chances of producing a logical conclusion to an argument or solution to a problem' (Dwyer, Hogan and Stewart, 2014, p.43).*

Researchers and educators refer to a spectrum of CT definitions, ranging from a broad reference to 'thinking skills' (e.g., Higgins, Baumfield, Lin, Moseley, Butterworth, Downey, Gregson, Oberski, Rockett and Thacker (2004)) to the narrow analysis of logical fallacies (as in the Critical Thinking As Level syllabus in England). However, the most cited CT tests focus upon the evaluation of argument (e.g., inference, analysis and deduction) as an indicator of CT (Ennis and Weir, 1985; Watson-Glaser, 1980; Ennis, Millman and Tomko, 1985). Furthermore, seminal studies such as Kuhn's *Thinking as Argument* (1992) have argued that the capacity for argument *is* the key indicator of the skill of critical thinking. For reasons discussed in Section 2.2.4, the current study has adopted the latter characterisation of critical thinking; therefore the conclusions of this study must be delimited by a characterisation of critical thinking that focuses upon a learner's capacity for argument, and cannot be extrapolated beyond this definition.

The critical thinking *measure* is similarly subject to limitations. As discussed in Section 4.5.1, the assessment was designed to relate to the content of the intervention, which focused upon one particular critical thinking skill - argumentation. Therefore it had been necessary to identify a CT test that was capable of measuring such a skill. Upon analysis of existing measures, it was concluded that the Defining Issues Test (DIT2) (Rest *et al.,* 1999) most closely matched the construct to be assessed (the evaluation of argument on topics of moral, social and/or ethical issues). The DIT2 assessment was, however, designed for older participants (as were other most-cited CT tests), and therefore had to be adapted for the purposes of this study. Thus, following recommendations from Ennis (1993) regarding the construction of such a test (see Section 2.2.3), the existing DIT2 test was modified for use by younger students. Whilst its reliability was tested using alternate form analysis, its validity was limited, especially since it was adapted from an adult measure. Therefore the analysis of the CT findings in this study must be limited to the interpretation of a construct which is narrow in its measurement. This limitation informs the rationale for one of the main conclusions of this study (see Section 7.5) – that there is a need for the construction of a critical thinking measure suitable for younger students.

The second conceptual limitation of this study refers to the choice of competencies and their measurements (metacognitive knowledge, motivational beliefs and efficacy for learning) chosen to represent the construct of self-regulated learning. Self-regulated learning had been chosen as a major focus for the intervention, since it reflected the Widening Participation aims which the intervention was originally designed to address (see Section 2.3.5.1.3, Section 2.4, Section 3.1 and Section 7.2.1). However, self-regulated learning, as has been discussed throughout the study, and in particular in Section 2.3, is a multi-faceted construct that is characterised differently across and within fields of study. Indeed a number of writers have emphasised the difficulties in attempting to propose a common definition (Boekaerts, 2006; Greene and Azevedo, 2007; Lajoie, 2008). Thus the interpretation of findings relating to the concept of self-regulated learning in this study must be limited to its specific characterisation as represented here. The three elements of SRL competencies (metacognitive knowledge, motivation to learn and efficacy for learning) were chosen for pre-to-post-test analysis, and a range of SRL processes (e.g., goal setting and analysis, task analysis, inhibitory control, elaboration strategies, metacognitive experiences) were featured in the students' learning journals. Metacognition was chosen, among other reasons, because of its *sine qua non* predominance in SRL studies and its recognition as a key driver of self-regulation as a process (see Section 2.3.5.1.1). The construct of motivational beliefs was chosen because of the significant relationships between it and a number of SRL competencies and SRL processes (see Section 2.3.5.1.2), and efficacy for learning was chosen because of its relationship to learning outcomes. However, it is acknowledged that a number of SRL competencies were absent from the study, and therefore the findings of this study must be limited to the three that were featured.

Limitations also apply to the measures employed for the assessment of SRL. The assessments for metacognitive knowledge and motivation to learn had been well-established and were, for the purposes of this study, subject to little alteration. The former required no amendment, as it had been validated across populations which included students from the UK (Marsh *et al.,* 2006a). The latter was amended only insofar as was required to make the language relevant for the context (see Section 4.5.2; Table 4-3). The measure for efficacy for learning, however, was constructed by the researcher, following Bandura's (2006) guidelines. Thus any inferences regarding self-efficacy, which might be taken from this study, should be limited to one particular strand - the learner's belief in their capacity to self-regulate their learning in a collaborative setting – and should be limited to the particular measure used in this study.

The final limitation refers to the learning journal. It was originally designed, for the purposes of the intervention, as a learning tool only, with its primary function being self-evaluation of the particular learning strategies which were employed by the students during each week. Thus students used the learning journal to set goals, evaluate them, and then identify subsequent ones. They also reflected upon a number of self-regulatory processes such as the use of metacognitive strategies, learning satisfaction, motivation, response to challenge, value of contribution, confidence, academic resilience, inhibitory control, effort expenditure and approach to planning. However, the learning journal did not cover all aspects of self-regulatory processes and whilst its effect has been evaluated as a learning tool, it cannot be assumed to be a comprehensive research tool. Thus, whilst the learning journal provided variables which represented elements of self-regulated learning in action, the tool was by no means a comprehensive one, and thus cannot be assumed to represent self-regulated learning as a whole. This limitation has informed the discussion in Section 7.2.2.1, where a more comprehensive learning journal is proposed.

7.4 Implementation into policy and practice

The results of this study have practical implications for national policy and for educators. This section will present a consideration of the implications of the principal findings relating to this study – the effectiveness of dialogic argumentation and the learning journal.

The findings of this study support a dialogic argumentation approach to the development of self-regulated learning and critical thinking. The merits of dialogic argumentation in the classroom are already demonstrated elsewhere (e.g., Kuhn and Crowell, 2011). However, the current study adds to what we already know in specific ways: it appears that engagement in dialogic argumentation can have an effect on learners' efficacy for learning and their motivational beliefs, as well as their metacognitive knowledge. Furthermore, an even greater effect can be seen on gains in critical thinking skills. Since it has long been established that self-regulated learning and critical thinking are strong predictors of academic success (Zimmerman, 2008; Dwyer, Hogan and Stewart, 2014), it follows that the provision of opportunities for dialogic argumentation has the potential to produce beneficial educational outcomes, for both students and their schools. However, as has been argued, organising the modern classroom around the constraints of syllabus prescription renders the provision of dialogic argumentation a timetabling luxury that many teachers cannot afford. It has also been suggested that many teachers are shy of attempting dialogic argumentation approaches because they are not confident in their own argumentation skills (Graff, 2003; Coffin and O'Halloran, 2009). This study has demonstrated that effective and productive dialogic argumentation can be implemented via internet forums, on secure learning platforms, run by specialist teachers, outside the normal school hours. Thus the advantages are manifold, with a major potential benefit being cost-effectiveness. Here, the 'classroom' is situated online and students from more than one school can participate in the same discussion, with one specialist teacher, serving the needs of many schools. Thus it is argued that an intervention such as this can be implemented without major restructuring or disruption to the curriculum, and can be offered as an extra-curricular, out of hours school activity.

It can also be argued that the provision of dialogic argumentation programmes via the internet has the potential to be particularly applicable to university outreach programmes, which aim to motivate students from far flung regions of England, to situate the prospect of higher education on their personal educational radar. Outreach programmes tend to target students from state schools, and participating universities are being advised to provide programmes for younger students, not just those who are about to make their university choices (Harris, 2010). Eyre, in her (2011) report for the Policy Exchange, is critical of the capacity for current English curriculum practice to promote high performance, and suggests that 'There is a place for extra or advanced provision to complement what takes place in normal schools' (p.23), noting the potential effectiveness of online communities. She proposes an interesting equation: high achievement = potential + opportunities + support + motivation. It is suggested that in view of the findings of this study, a programme of online dialogic argumentation has the potential to offer the latter three elements of that equation, as an extra-curricular activity, and to a wide population. Such a programme could be offered in the form of an after-school debating club, run by expert teachers, made available online to students across England.

Evidence from this study has also suggested that encouraging students to use a learning journal to record the evaluations of their learning experience, has the potential to increase the effectiveness of argumentation intervention. Students who used a learning journal reported significantly more 'progress' over time in their self-reported SRL scores than those who did not. Whilst learning journals are often used in the context of further and higher education, currently they are rarely reported as being used in the school context in England. This is despite increasing reports of their effectiveness in promoting the development of self-regulated learning skills, which in turn improve the effectiveness of study (e.g., McCrindle and Christensen, 1995; Hübner, Nückles and Renkl, 2010). One reason for their lack of use in schools may be to do with the difficulty of delivery and administration of the journals. Another may be because of time-constraints attached to the curriculum. The few UK studies requiring students to complete journals which self-report their metacognitive activity, have been presented using paper and pencil methods (e.g., Stephens and Winterbottom, 2010). However, findings from the current study suggest that the learning journal can be effectively administered using an electronic format. There are advantages to this method of delivery. For instance, the electronic medium is more cost-effective than paper and pencil. One teacher can supervise the journals of large numbers of students, since the journals themselves are stored electronically and can be downloaded for scrutiny in spreadsheet format. The journal is accessible outside school hours and is available for download by teachers at any time. Importantly for researchers, the data thus collected are already transcribed and presented in a spreadsheet, ready for transfer to data analysis programmes. Such journals can be set up using either cloud technology (e.g., Google docs) or, if the journal is required to be interactive between student and teacher, using a learning management system (such as Moodle). Both systems are free of charge.

The sample in this study came from the population of above average young adolescents, not only because of the intended outreach audience, but also because it was concluded that students require a benchmarked level of literacy to be able to interpret the intervention material and the questionnaire items (Hopfenbeck and Maul, 2011). However, analysis of pre-to post-test change in self-regulated learning scores, revealed that on average, the lower a participant scored at pre-test, the greater their gain between pre-test and post-test - that is the lower the initial score, the steeper the gradient of increase. This suggests that students, whose initial self-regulated learning scores were lower than those of others, benefitted the most from this intervention.

This finding has implications for those educators who aim to raise the self-regulated learning attributes of students who are less metacognitively aware, less motivated and less self-efficacious, thus implying that an intervention such as this has the potential to be of particular benefit to such students.

7.5 Summary of implications for future research

The analysis of the study findings indicates that three further avenues of research are merited. These relate to the use of the learning journal, the potential breadth of SRL study, and the conceptualisation and assessment of CT.

The study has detected positive and significant effects of the repeated use of a learning journal on gains in self-regulated learning over time. Whilst the design of this study was constructed such that students completed journals once a week for four weeks, similar studies (such as Schmitz and Wiese, 2006) have also detected significant gains in self-regulated learning, employing daily diaries over a longer period. There remains, therefore, the question of the optimum frequency, timing and duration of learning journals, in the context of levels of effectiveness, age groups and settings. Analyses should focus upon whether increasing the frequency of journal use improves learning outcomes, and whether there is a maximum number of journal events, beyond which learner interest may wane. Thus the important research question to ask would be: *Does treatment intensity, frequency, duration, interval and context mediate the effects of learning journal use on specific learning outcomes?* Further characteristics such as the scope of variables, the context and the importance attached to the task by teachers, may well also mediate the effects of the learning journal.

Potential longitudinal effects should also be explored, investigating trajectory effects of self-regulated learning outcomes in relation to the duration, interval and frequency of the original journal intervention. As learning journal studies are still relatively much in their infancy, and considering the positive and significant effects detected in this study, it is suggested that there is merit in exploring such aspects further.

The second suggested line of research relates to the conceptualisation of self-regulated learning (SRL). In this study, and for reasons previously explained, SRL learning outcomes have been limited to three specific SRL *competencies* (metacognitive knowledge, motivational beliefs and efficacy for learning), whilst the study of SRL *processes* has been restricted to those that were presented in student learning journals. As has been explored in Section 7.2.2, the use of learning journals has the potential to encourage students to explore the *whole cycle* of SRL, as conceptualised by Zimmerman (2008) and Schmitz and Wiese (2006). Whilst such journals have been used elsewhere as *research tools*, their use as *learning tools* in the English context is less widespread, and therefore presents a promising line of future study. Such a learning journal would enable the learner to reflect upon aspects of SRL beyond those studied in the current research. The potential content of such a journal is described in Section 7.2.2. In such a study, the use of the learning journal would be the independent variable, whilst specific learning outcomes would be identified as dependent variables.

The third suggested avenue of research focuses upon critical thinking. Critical thinking has been conceptualised in this study, for reasons explained in Section 2.2, as a *skill* relating to that of argumentation. However, CT has also been characterised by others in terms of *dispositions* that may pre-dispose a learner to develop critical thinking skills (dispositions such as inquisitiveness, an analytical approach, self-confidence, open-mindedness, the search for truth, systematic organisation, perseverance, willingness to self-correct and cognitive maturity) (Facione, Sánchez, Facione and Gainen, 1995). Studies which explore the relationship between CT dispositions and CT skills are however rare, particularly in the context of school-age children. It is suggested that a potentially productive line of research should more closely explore the relationships between CT skills and CT dispositions, especially in the context of school-age students. Such studies would have the potential not only to broaden the conceptualisation of CT, but also the opportunities for assessment (e.g., for diagnostic screening).

In view of the importance placed on the development of CT in young people, it follows that is should not only feature in national curricula, but, for reasons discussed in Section 2.2.3, it should also be assessed. The CT test used in this study had been adapted from one devised originally for adults in the US, since CT tests for young school-age students in the UK context are rare. Both the content and the language had to be adapted, for linguistic and cultural reasons. Therefore, future research should be directed at the creation of a critical thinking test suitable for young students. The most prominent existing assessments (e.g., California CT Test: Facione, 1990a; Ennis-Weir CT Test: Ennis and Weir, 1985) focus upon aspects of argumentation (inference, analysis and deduction) and are commonly tested in terms of evaluating the learner's capacity to analyse the strength of an argument (as in Rest *et al.,* 1999) or to demonstrate logical reasoning (as in the Cornell CT Test: Ennis, Millman and

Tomko, 1985). More recent studies however, (e.g., Nussbaum, 2005; Shehab and Nussbaum, 2015) have concluded that the Need for Cognition scale (Cacioppo, Petty and Kao, 1984) is effective as a CT measure, since it correlates highly with the CT *dispositions* mentioned above, and indeed to the capacity to evaluate the strength of an argument. Thus a more comprehensive evaluation of CT would include such a range of CT dispositions. Since CT is considered to hold prime importance in the modern curriculum (Dwyer, Hogan and Stewart, 2014), since this study has detected that CT has the potential to be significantly affected by intervention, and since there is currently a dearth of assessments which can evaluate the CT skills of young people, it is suggested that there is an urgent need to develop a CT test for school-age children – one that covers a broad range of CT skills and dispositions, and one that reflects the importance, relevance and application of CT in the modern curriculum. An important argument to register, in bringing this study to a conclusion, is that the identification and articulation of a significant limitation of this study (i.e., the limitation of the CT measure, discussed in Section 7.3) also translates into one of its key outcomes, insofar as it has highlighted the need to create a critical thinking measure suitable for school-age children in the UK.

7.6 Summary and contribution to knowledge

The primary goal of this study was to report on the development, implementation and evaluation of an intervention on the self-regulated learning and critical thinking skills of young English adolescents. As reflected upon throughout the study, an important aim of education is to develop self-regulated, independent critical thinkers. The results of this study suggest that an intervention such as this (an online argumentation intervention supported by the use of a learning journal) can potentially contribute towards that goal.

The findings reveal that the intervention had a small to medium main effect on participants' self-regulated learning and on their critical thinking. The effects of the intervention, on individuals' self-regulated learning, were doubled if students also used the learning journal. Those who completed learning journals reported significantly more change in their self-regulated learning over time, compared with those who did not. Whilst the most effective intervention *condition* was membership of IC4 (the group that experienced all four conditions), the most effective intervention *activity* was participation in dialogic argumentation. Participation in dialogic argumentation predicted change in both self-regulated learning and critical thinking. The main findings are thus attributed to a) the use of a learning journal to promote the development of self-regulated learning skills and b) dialogic argumentation to promote both critical thinking and self-regulated learning. The process analysis revealed that critical thinking achievement one week significantly predicted the use of elaboration strategies the next, and vice versa, with increasing effect over time. A participant's capacity to avoid distraction was directly and consistently related to their metacognitive experiences, within each learning event.

It is proposed that the study makes four major contributions. The first three relate to the findings in relation to the theory, and the fourth to the methodology. Firstly, the aim of this study was to explore the effectiveness of an intervention that was designed to have an effect on the self-regulated learning and critical thinking of young adolescents in England. Following the analysis, it is suggested that dialogic argumentation can have a positive effect not only on learners' metacognitive knowledge, as has already been established by other authors, but also on their motivational beliefs and efficacy for learning. The latter two potential effects have not so far been extensively studied, especially in the context of UK children. Furthermore, the notion that the browsing of online discussion, rather than the activity of actually posting replies, can affect learning outcomes, has been assumed, but not extensively tested empirically. It is therefore argued that the finding that discussion browsing had a small, significant and positive effect on change in efficacy for learning and overall self-regulated learning, has the potential to make a new contribution to knowledge in the field of self-regulated learning.

Secondly, the analysis also revealed that the effect of the intervention on all three aspects of learners' self-regulated learning was doubled if participants also used a learning journal. Whether learners completed learning journals or not was a significant predictor of change in SRL and all of its components, with small effect sizes for change in efficacy for learning and motivational beliefs, and medium effect sizes for change in metacognitive knowledge. Studies have already established links between the use of a learning journal and metacognitive development, but few have so far detected such an effect on learners' motivational beliefs or their efficacy for learning. In view of the body of knowledge that already exists regarding the predictive effect of both motivational beliefs and self-efficacy on educational achievement, it is suggested that this finding also makes a contribution to knowledge.

Thirdly, it is proposed that this study makes a contribution to the field of study that focuses upon metacognitive experiences (an individual's sense of self-efficacy with regard to a specific task, their feelings of difficulty, challenge, motivation, effort, success and their learning satisfaction). The concept of metacognitive experiences was first mooted by Flavell (1979) and has been explored empirically by Efklides and her team (e.g., Efklides, 2001, 2008, 2009; Efklides and Petkaki, 2005; Efklides, Kourkoulou, Mitsias and Ziliaskopoulou, 2006). However, whilst the relationships between metacognitive skills (e.g., critical thinking and control of cognitive strategies) and metacognitive experiences are assumed in the literature, they have rarely been empirically tested in the context of English school children. This study has thus responded to the call of Efklides (2008) who suggested the need to explore the concept of metacognitive experiences in specific populations, in collaborative situations, and alongside other aspects of self-regulated learning and performance measures. The process analysis in this study suggested a strong relationship between metacognitive experiences and volitional strategies (inhibitory control). However, since the latter two variables were self-reported, further research is necessary to relate metacognitive experiences to actual inhibitory control, rather than self-reported.

Finally, it is proposed that the study makes three methodological contributions. Firstly, it has been demonstrated that the use of latent change score modelling has the potential to be an effective method for the analysis of pre- to post-test change in behavioural variables. Secondly, it is argued that this study has the potential to make a methodological contribution to the study of diary data, by applying a structural equation modelling technique (auto-regressive cross lagged panel analysis) to the analysis of repeated measures learning journal data. In particular, it is proposed that this has the potential to be an effective methodology for the estimation of relationships between multiple variables over multiple times. Thirdly, research studies that empirically explore the relationship between argumentation and critical thinking, whilst reporting statistically significant results, rarely report effect sizes. This study makes a contribution to the field by addressing this important gap in knowledge.

In summary, this study has drawn attention to a number of matters which, it is argued, are highly relevant to the education of the young students of today – the leaders of tomorrow. In particular, it is argued that two specific research outcomes have the potential to make a practical contribution to educational practice. The effectiveness of journal writing has been the subject of studies across disciplines, including psychology and education, but its practical application does not currently feature widely in English educational practice. It has been demonstrated in this study, that the use of a learning journal has the potential to be particularly effective in raising aspects of self-regulated learning which many contemporary authors consider to be of particular importance. Furthermore it has been suggested here and elsewhere that internet technology renders the delivery and management of such a journal much easier than pencil and paper methods, thus enabling educators to grasp this opportunity to supplement their arsenal of teaching and learning techniques.

As discussed throughout the study, it is universally agreed that critical thinking is a vital skill, the development of which should be of prime importance to educators. Yet, for reasons previously discussed, it is rarely mentioned in National Curriculum literature, and, perhaps in view of the time constraints put upon today's teachers, it does not enjoy the prominence which its importance arguably deserves. Yet this study has demonstrated that the capacity to evaluate argument, which it has been argued is an essential aspect of critical thinking, as currently defined, can be affected by providing opportunities for engagement in tutor-led debate of moral, social or ethical issues. The study has demonstrated that not only does this learning strategy have the potential to influence the development of critical thinking, but is deliverable as an extra-curricular activity to large numbers, across a wide geographical area, using free, ubiquitous and ever-improving modern technology, without interfering with the existing curriculum provision. It is suggested, therefore, that educators and policy makers consider the potential benefits of the creation of an after-school network of debating clubs, using teaching techniques and methods of delivery employed in this study. In consideration of these final conclusions, it is argued that the strength of this study is that it adds to the body of effective interventions that seek to improve the educational lives of the young people of today.

8 References

Abrami, P.C., Bernard, R.M., Borokhovski, E., Wade, A., Surkes, M., Tamim, R. and Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A Stage 1 meta-analysis. *Review of Educational Research.* 78 (4), 1102 - 1134.

Ahmed, W., Van der Werf, G., Kuyper, H. and Minnaert, A. (2013). Emotions, self-regulated learning and achievement in mathematics: A growth curve analysis. *Journal of Educational Psychology,* 105, 150 - 161.

Alagoz, E. (2013). Social argumentation in online synchronous communication. *Computer-Supported Collaborative Learning.* 8, 399 - 426.

Alexander, R.J. (2006). *Towards dialogic teaching* (3rd Edn). Cambridge: Cambridge University Press/Dialogos.

Allan, B., Cook, M. and Lewis R. (1996). *The independent learner: developing independence in learning*. Humberside: University of Humberside Press.

Alloway, T.P. and Alloway, R.G. (2012). The impact of engagement with social networking sites (SNSs) on cognitive skills. *Computers in Human Behavior*, 28, 1748 - 1754.

American College Testing Program. (1992). *Clarifying and Assessing General Education Outcomes of College, College Outcome Measures Program: Technical Report 1982 - 1991,* Iowa City, IA: American College Testing Program.

Anderson, R.C., Nguyen-Jahiel, K., McNurlin, B., Archodidou, A., Kim, S.Y. and Reznitskaya, A. (2001). The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition and Instruction,* 19 (1), 1 - 46.

Andriessen, J., Baker, M. and Suthers, D. (2003). *Arguing to learn: Confronting cognitions in computer supported collaborative learning environments.* Dordrecht, NL: Kluwer.

Arbaugh, J.B. and Benbunan-Finch, R. (2005). Contextual factors that influence ALN effectiveness. In S.R. Hiltz and R. Goldman (Eds.), *Learning together online. Research on asynchronous learning networks* (pp. 123 - 144). Mahwah: Lawrence Erlbaum Associates, Inc.

Aron, A.R. (2007). The neural basis of inhibition in cognitive control. *The Neuroscientist.* 13 (3), 1 - 15.

Azevedo, R., Cromley, J.G. and Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology,* 29, 344 - 370.

Azevedo, R. and Hadwin, A.F. (2005). Scaffolding self-regulated learning and metacognition - implications for the design of computer-based scaffolds. *Instructional Science,* 33 (5 - 6), 367 - 379.

Azevedo, R. (2014). Issues in dealing with sequential and temporal characteristics of self- and socially-regulated learning. *Metacognition Learning,* 9, 217 - 228.

Bai, H. (2009). Facilitating Students' Critical Thinking in Online Discussion: An Instructor's Experience. *Journal of Interactive Online Learning,* 8 (2), 156 - 164.

Bailin, S. and Siegel, H. (2003). Critical thinking. In N. Blake, P. Smeyers, R. Smith and P. Standish (Eds.), *The Blackwell Guide to the Philosophy of Education* (pp. 181 - 193). Oxford, UK: Blackwell.

Baker, L. and Brown, A.L. (1984). Metacognitive Skills and Reading. In P.D. Pearson, R. Barr, M.L. Kamil and P. Mosenthal (Eds.), *Handbook of Reading Research* (pp. 353 - 394). New York: Longman.

Bandura, A. (1977a). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review,* 84, 191 - 215.

Bandura, A. (1977b). *Social learning theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist,* 37, 122 - 147.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (Ed.). (1995). *Self-efficacy in changing societies.* New York: Cambridge University Press.

Bandura, A., Barbaranelli, C., Caprara, G.V. and Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development,* 67, 1206 - 1222.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. U.S.A.: W.H. Freeman and Company.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares and T. Urdan (Eds.). *Self-efficacy beliefs of adolescents,* 5, 307 - 337. Greenwich, CT: Information Age Publishing.

Bangert-Drowns, R.L., Hurley, M.M. and Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: a meta-analysis. *Review of Educational Research,* 74, 29 - 58.

Banich, M.T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science,* 18 (2), 89 - 94.

Bannert, M., Reimann, P. and Sonnenberg, C. (2014). Process mining techniques for analyzing patterns and strategies in students' self-regulated learning. *Metacognition and Learning,* 9, 161 - 185.

Baron, J. (1991). Beliefs about thinking. In J.F. Voss, D.N. Perkins and J.W. Segal (Eds.), *Informal reasoning and education,* pp. 169 - 186. Hillsdale, NJ: Lawrence Erlbaum.

Baron, J. (1995). Myside bias in thinking about abortion. *Thinking and Reasoning*, 1, 221 - 235.

Barriga, A.Q., Doran, J.W., Newell, S.B., Morrison, E.M., Barbetti, V. and Robbins, B.D. (2002). Relationships between problem behaviors and academic achievement in adolescents: The unique role of attention problems. *Journal of Emotional and Behavioral Disorders,* 10 (4), 233 - 40.

Barclay, J. (1996). Assessing the benefits of learning logs. *Education and Training,* 38 (2), 30-38.

Barton, K. and McCully, A. (2007). Teaching controversial issues where controversial issues really matter. *Teaching History,* 127, 13 - 19.

BBC (2012). *No evidence for schools falling behind*. Retrieved on 26.3.12 from: http://www.bbc.co.uk/news/education- 160546547.

Beaudoin, M.F. (2002). Learning or lurking? Tracking the ''invisible'' online student. I*nternet and Higher Education,* 5, 147 - 155.

Bell, E.A. (1991). Debate: a strategy for teaching critical thinking. *Nurse Educator,* 16, 6 - 7.

Benbunan-Fich, R., Hiltz, S.R. and Turoff, M. (2002). A comparative content analysis of face-to-face vs. asynchronous group decision making. *Decision Support Systems,* 34, 457 - 469.

Bentler, P.M. (1980). Comparative fit indexes in structural models. *Psychological Bulletin,* 107, 238 - 246.

Berthold, K., Nückles, M. and Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction,* 17, 564 -577.

Black, T.R. (2005). *Doing Quantitative Research in the Social Sciences.* London: Sage.

Blakemore, S-J. and Frith, U. (2005). *The learning brain: lessons for education.* London: Blackwell.

Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H. and Krathwohl, D.R. (1956). *Taxonomy of educational objectives: The classification of educational goals: Handbook 1: Cognitive domain.* New York: David McKay.

Blunch N.J. (2012). *Introduction to Structural Equation Modelling.* London: Sage.

Boekaerts, M. (1996). Self-regulated learning at the junction of cognition and motivation. *European Psychologist,* 1 (2), 100 - 112.

Boekaerts, M. (1997). 'Self-regulated learning: a new concept embraced by researchers, policy makers, educators, teachers, and students'. *Learning and Instruction,* 7, 161 - 86.

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research,* 31 (6), 445 - 457.

Boekaerts, M. and Niemivirta, M. (2000). Self-regulated learning: Finding a balance between learning goals and ego-protective goals. In M. Boekaerts, P.R. Pintrich and M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 417 - 451). San Diego, CA: Academic Press.

Boekaerts, M. (2002). Bringing about change in the classroom: Strengths and weaknesses of the self-regulated learning approach. *Learning and Instruction,* 12, 589 - 604.

Boekaerts, M. and Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology,* 54, 267 - 99.

Boekaerts, M. and Cascallar, E. (2006). How far have we moved towards the integration of theory and practice in self-regulation? *Education Psychology Review,* 18, 199 - 210.

Boekaerts, M. (2010). Motivation and self-regulation: Two close friends. In: Urdan, T.C. and Karabenick, S.A. (Eds.), *The decade ahead: Applications and contexts of motivation and achievement.* (*Vol. 16B Advances in Motivation and Achievement*), pp. 69 - 108. Bingley, UK: Emerald.

Boekaerts, M., Smit, K. and Busing, F. (2013). Salient Goals Direct and Energise Students' Actions in the Classroom. *Applied Psychology: An International Review*, 61 (4), 520 - 539.

Bollen, K.A. (1989). *Structural Equations with Latent Variables*. New York: John Wiley and Sons.

Bollen, K.A. and Curran, P.J. (2004). Autoregressive latent trajectory (ALT) models: A synthesis of two traditions. *Sociological Methods and Research,* 32, 336 - 383.

Bong, M. and Skaalvik, E.M. (2003). Academic self-concept and self efficacy: How different are they really? *Educational Psychology Review*, 15, 1 - 40.

Bouffard-Bouchard, T., Parent, S. and Larivee, S. (1991). Influence of self-efficacy on self-regulation and performance among junior and senior high-school aged students. *International Journal of Behavioral Development,* 14, 153 - 164.

Bowell, T. and Kemp, G. (2010). *Critical Thinking: A Concise Guide.* Abingdon: Routledge.

Brookfield, S.D. (1987). *Developing Critical Thinkers.* Milton Keynes: Open University Press.

Brown, T. (2004). *Critical thinking and learning: An encyclopedia for parents and teachers: Bloom's taxonomy and critical thinking.* Westport: Greenwood Press.

Bryman, A. (2004). *Social Research Methods (2nd Edition).* Oxford: Oxford University Press.

Bullen, M. (1998). Participation and critical thinking in online university distance education. *Journal of Distance Education,* 13 (2), 1 - 32.

Cacioppo, J.T., Petty, R.E. and Kao, C.F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment,* 48, 306 - 307.

Cacioppo, J.T., Petty, R.E., Feinstein, J.A. and Jarvis, W.B. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin,* 119, 197 - 253.

Calderwood, C., Ackerman, P.L. and Conklin, E.M. (2014). What else do college students "do" while studying? An investigation of multitasking. *Computers and Education,* 75, 19 - 29.

Campbell, D.T. and Stanley, J.C. (1966). Experimental and quasi-experimental designs for research. In N.L. Cage (Ed), *Handbook of research on teaching* (p.1080). Chicago: Rand, McNally.

Cantrell, J., Fusaro, J. and Dougherty, E. (2000). Exploring the effectiveness of journal writing on learning social studies: A comparative study. *Reading Psychology,* 21, 1 - 11.

Caprara, G.V., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G.M., Barbaranelli, C. and Bandura, A. (2008). Longitudinal Analysis of the Role of Perceived Self-Efficacy for Self-Regulated Learning in Academic Continuance and Achievement. *Journal of Educational Psychology,* 100 (3), 525 - 534.

Caprara, G.V., Vecchione, M., Alessandri, G., Gerbino, M. and Barbaranelli, C. (2011). The contribution of personality traits and self-efficacy beliefs to academic achievement: a longitudinal study. *British Journal of Educational Psychology,* 81, 78 - 96.

Casey, B.J., Giedd, J. and Thomas, K. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology,* 54, 241 - 257.

Cavagnetto, A.R. (2010). Argument to foster scientific literacy: A review of argument interventions in K- 12 contexts. *Review of Educational Research,* 80, 336 - 371.

CEOP. (2012). *Child Exploitation and Online Protection Centre.* Retrieved on 19.8.12 from: http://ceop.police.uk/.

Chang, K.E., Sung, Y.T. and Chen, S.F. (2001). Learning through computer-based concept mapping with scaffolding aid. *Journal of Computer Assisted Learning,* 17, 21 - 33.

Cho, K.L. and Jonassen, D.H. (2003). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development,* 50 (3), 5 - 22.

Choy, S.C. and Cheah, P.K. (2009). Teacher perceptions of critical thinking among students and its influence on higher education. *International Journal of Teaching and Learning in Higher Education,* 20 (2), 198 - 206.

Clark, D.B., Sampson, V., Weinberger, A. and Erkens, G. (2007). Analytic frameworks for assessing dialogic argumentation in online learning environments. *Educational Psychology Review,* 19 (3), 343 - 374.

Clark, D.B. and Sampson, V. (2008). Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality. *Journal of Research in Science Teaching,* 45, 293 - 321.

Cleary, T.J. and Chen, P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology,* 47, 291 - 314.

Clipa, O., Ignat, A-A. and Stanciu, M. (2012). Learning diary as a tool for metacognitive strategies development. *Procedia - Social and Behavioral Sciences,* 33, 905 - 909.

Coe, R. (2002). *It's the effect size, stupid: What effect size is and why it is important.* Paper presented at the annual conference of the British Educational Research Association. Retrieved on 24.10.13 from: http://www.leeds.ac.uk/educol/documents/00002182.htm.

Coffin, C. (2007). *The language and discourse of argumentation in computer conferencing and essays:* Full research report. Vol. ESRC End of Award Report, RES- 000 - 22 - 1453. Swindon, UK: Economic and Social Research Council.

Coffin, C. and O'Halloran, K.A. (2009). Argument reconceived? *Educational Review,* 61 (3), 301 - 313.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement,* 20, 37 - 46.

Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. NY: Academic Press.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement,* 7, 249 - 253.

Cohen, J. (1992). A power primer. *Psychological Bulletin,* 112 (1), 155 - 159.

Cohen, J. (1994). The Earth is round (p < .05). *American Psychologist.* 49 (12), 997 - 1003.

Cohen, J. (1995). Things I have learned (so far). *American Psychologist.* 45 (12), 1304 - 1312.

Colbert, K.R. (1987). The effects of CEDA and NDT debate training on critical thinking ability. *Journal of the American Forensic Association,* 23 (4), 194 - 201.

Connor-Greene, P.A. (2000). Making connections: evaluating the effectiveness of journal writing in enhancing student learning, *Teaching of Psychology*, 27, 44 - 46.

Cook, T.D. and Campbell, D.T. (1979). *Quasi-experimentation: design and analysis issues for field settings.* Chicago, IL: Rand-McNally.

Corno L. (1993). The best-laid plans: modern conceptions of volition and educational research. *Educational Research,* 22, 14 - 22.

Corno, L. (1994). Student volition and education: Outcomes, influence, and practices. In B.J. Zimmerman and D.H. Schunk (Eds) *Self-regulation of learning and performance.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Corno, L. (2001). Volitional aspects of self-regulated learning. In B. J. Zimmerman and D. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical Perspectives* (pp. 191 - 227). Mahwah, New Jersey: Lawrence Erlbaum.

Cowan, J. (1992). 'Learning journals': What are they for? *British Journal of Educational Technology,* 23 (2), 139 - 140.

Creswell, J.W. (2007). *Educational research (3rd edition).* Thousand Oaks, CA: Sage.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,* 16, 297 - 334.

Cross, D., Taasoobshirazi, G., Hendricks, S. and Hickey, D.T. (2008). Argumentation: A strategy for improving achievement and revealing scientific identities, *International Journal of Science Education,* 30 (6), 837 - 861.

Darabi, A., Liang, X., Suryavanshi, R. and Yurekli, H. (2013) Effectiveness of Online Discussion Strategies: A Meta-Analysis, *American Journal of Distance Education*, 27 (4), 228 - 241.

DCSF. (2007). *Social and emotional aspects of learning for secondary schools. Department for Children, Schools and Families.* Retrieved on 27.3.12 from: http://dera.ioe.ac.uk/6663/1/f988e14130f80a7ad23f337aa5160669.pdf .

Debowski, S., Wood, R. and Bandura, A. (2001). Impact of guided exploration and enactive exploration on self-regulatory mechanisms and information acquisition through electronic search. *Journal of Applied Psychology,* 86, 1129 - 1141.

De Bruin, A. and van Gog, T. (2012). Improving self-monitoring and self-regulation: From cognitive psychology to the classroom. *Learning and Instruction,* 22, 245 - 252.

Dembo, M.H. and Eaton, M.J. (2000). Self-regulation of academic learning in middle-level schools. *Elementary School Journal,* 100, 473 - 490.

Dennen, P.V. (2008). Pedagogical lurking: Student engagement in non-posting discussion behavior. *Computers in Human Behavior,* 24, 1624 - 1633.

Denzin, N.K. (2006). *Sociological Methods: A Sourcebook.* New Jersey: Aldine Transaction.

Dewey, J. (1910). *How We Think.* Boston: D.C. Heath.

DfE. (2007). *Making good progress: How can we help every pupil to make good progress at school?* Retrieved on 24.12.14 from: http://bit.ly/13XW7Cm.

Dignath, C., Buettner, G. and Langfeldt, H.P. (2008). How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes. *Educational Research Review,* 3, 101 – 129.

Dinsmore, D.L., Alexander, P.A. and Loughlin, P.M. (2008). Focusing the conceptual lens on metacognition, self-regulation and self-regulated learning. *Educational Psychology Review,* 20, 391 - 409.

DirectGov. (2013). *Schools, pupils, and their characteristics, January 2013.* Retrieved on 10.9.13 from: https://www.gov.uk/government/publications/schools-pupils-and-their-characteristics-january- 2013.

Doise, W., Mugny, G. St. James, A. and Emler, N. (1984). *The social development of the intellect.* Oxford: Pergamon Press.

Driver, R., Newton, P. and Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84, 287 - 312.

Duckworth. K., Akerman, R., MacGregor. A., Salter. E and Vorhaus, J. (2009). Self-regulated learning: a literature review. *Centre for Research on the Wider Benefits of Learning Research Report No. 33,* Institute of Education.

Duncan, T.E., Duncan, S.C. and Stryker, L. (2006). *An Introduction to Latent Variable Growth Curve Modeling (2nd edition).* Mahwah NJ: Lawrence Erlbaum.

Duncan, T.G. and McKeachie, W.J. (2005). The making of the motivated strategies for learning questionnaire. *Educational Psychologist,* 40, 117 - 128.

Durston, S., Mulder, M., Casey, B.J., Ziermans, T. and van Engeland, H. (2006). Activation in ventral prefrontal cortex is sensitive to genetic vulnerability for attention-deficit hyperactivity disorder. *Biological Psychiatry,* 60, 1062 - 1070.

Dweck, C.S. (2006). *Mindset: the New Psychology of Success.* New York: Random House.

Dwyer, C.P., Hogan, M.J. and Stewart, I. (2014). An integrated critical thinking framework for the 21st Century. *Thinking Skills and Creativity,* 12, 43 - 52.

Dwyer, C.P., Hogan, M.J., Harney, O.M. and O'Reilly, J. (2014). Using interactive management to facilitate a student-centred conceptualisation of critical thinking: a case study. *Education Technology Research Development,* 62, 687 - 709.

Eccles, J.S. and Wigfield, A. (2002). Motivational beliefs, values, and goals, *Annual Review of Psychology,* 53, 109 - 132.

Educational Testing Service. (1990). *The Academic Profile: A User's Guide, Part 1.* Princeton, NJ: Educational Testing Service.

Educational Testing Service. (1994). *Tasks in Critical Thinking: Core Scoring Manual.* Princeton, NJ: Educational Testing Service.

Efklides, A. (2001). Metacognitive experiences in problem solving: Metacognition, motivation and self-regulation. In A. Efklides (Ed.), *Trends and Prospects in Motivation Research,* (pp. 297 - 323), Netherlands: Kluwer Academic Publishers. Retrieved on 3.11.14 from: http://site.ebrary.com/lib/bodleian/reader.action?docID=10053382&force=1&ppg=322.

Efklides, A. and Petkaki, C. (2005). Effects of mood on students' metacognitive experiences. *Learning and Instruction.* 15, 415 - 431.

Efklides, A. (2006). Metacognitive Experiences: The Missing Link in the Self-Regulated Learning Process. *Education Psychology Review,* 18, 287 - 291.

Efklides, A., Kourkoulou, A., Mitsias, F. and Ziliaskopoulou, D. (2006). Metacognitive knowledge of effort, personality factors, and mood state: their relationships with effort-related metacognitive experiences. *Metacognition Learning,* 1, 33 - 49.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist,* 13 (4), 277 - 287.

Efklides, A. (2009). The role of metacognitive experiences in the learning process. P*sicothema,* 21 (1), 76 - 82.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL Model. *Educational Psychologist,* 46 (1), 6 - 25.

Efklides, A. (2012). How readily can findings from basic cognitive psychology research be applied to the classroom? *Learning and Instruction,* 22, 290 - 295.

Ellis, G. and Sinclair, B. (1989). *Learning to Learn English,* Cambridge University Press.

Ennis, R.H. (1962). A concept of critical thinking. *Harvard Educational Review*, 32, 81 - 111.

Ennis, R.H. and Weir, E. (1985). *The Ennis-Weir critical thinking essay test.* Pacific Grove, CA: Midwest Publications.

Ennis, R.H., Millman, J. and Tomko, T.N. (1985). *Cornell Critical Thinking Tests Level X and Level Z— Manual.* Pacific Grove, CA: Midwest Publications.

Ennis, R.H. (1989). Critical Thinking and Subject Specificity: Clarification and Needed Research. *Educational Researcher,* 18 (3), 4 - 10.

Ennis, R.H. (1993). Critical thinking assessment. *Theory into Practice,* 32 (3), 179 - 186.

Ennis, R.H. (1998). Is critical thinking culturally biased? *Teaching Philosophy,* 21 (1), 15 -33.

Erwin, T.D. (2000). U.S. Department of Education, National Center for Education Statistics. The NPEC Sourcebook on Assessment, Volume 1: Definitions and Assessment Methods for Critical Thinking, Problem Solving, and Writing, NCES 2000 - 172, *Council of the National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development.* Washington, DC: U.S. Government Printing Office.

Eynon, R., Fry, J. and Schroeder, R. (2008). The Ethics of Internet Research. In Fielding, N., Lee, R.M. and Blank, G. (Eds.) *The Sage Handbook of Online Research Methods.* London: Sage.

Eyre, D. (2011). Room at the Top: Inclusive education for high performance. *The Policy Exchange.* Retrieved on 17.6.11 from http://www.policyexchange.org.uk/ images/publications/pdfs/ Room_at_the_Top_-_Apr__11.pdf.

Fabriz, S., Ewijk, C.D-V., Poarch, G. and Buttner, G. (2014). Fostering self-monitoring of university students by means of a standardized learning journal - a longitudinal study with process analysis. *European Journal of Psychology Education,* 29, 239 - 255.

Facione, P.A. (1990). *Executive summary - critical thinking. A statement of expert consensus for purposes of educational assessment and instruction.* Millbrae, CA: California Academic Press.

Facione, P.A. (1990a). *The California Critical Thinking Skills Test.* Millbrae, CA: The California Academic Press.

Facione, N.C. and Facione, P.A. (1992). *The California Critical Thinking Skills Test: Forms A and B— Test Manual.* Millbrae, CA: The California Academic Press.

Facione, P.A, Sánchez, C.A, Facione, N.C. and Gainen, J. (1995). The disposition toward critical thinking. *Journal of General Education.* 44 (1), 1 - 25.

Facione, P.A. and Facione, N.C. (2001). Analyzing explanations for seemingly irrational choices: Linking argument analysis and cognitive science. *International Journal of Applied Philosophy,* 15 (2), 267 - 286.

Ferla, J., Valcke, M. and Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences,* 19, 499 - 505.

Fernandez-Duque, D., Baird, J.A. and Posner, M.I. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition,* 9, 288 - 307.

Field, A. (2009, 2013). *Discovering Statistics Using SPSS,* London: Sage.

Figueredo, A.J., Olderbak, S.G., Schlomer, G.L., Garcia, R.A., Wolf, P.S.A. (2011). Program Evaluation: Principles, Procedures, and Practices. In Little T. D. (Ed). *The Oxford Handbook of Quantitative Methods.* Oxford: Oxford Library of Psychology.

Finkel, S. (1995) *Causal Analysis with Panel Data.* Sage publications.

Fisher, R.A. (1935). *The Design of Experiments.* New York: Hafner.

Fisher, R. (2007). Dialogic teaching: developing thinking and metacognition through philosophical discussion. *Early Child Development and Care,* 177, 6 - 7.

Flavell, J. (1979). Metacognition and cognitive monitoring: A new area of cognitive development inquiry. *American Psychologist,* 34, 906 - 11.

Forrest-Pressley, D.L. and Waller, T.G. (1984). *Cognition, Metacognition and Reading.* New York: Springer-Verlag.

Garcia, T. and Pintrich, P.R. (1992). *Critical thinking and its relationship to motivation, learning strategies, and classroom experience.* Paper presented at the annual meeting of the American Psychological Association, Washington, DC.

Garner, R. and Alexander, P.A. (1989). Metacognition: answered and unanswered questions. *Educational Psychologist,* 24, 143 - 158.

Georghiades, P. (2004). From the general to the situated: three decades of metacognition. *International Journal of Science Education,* 26 (3), 365 - 383.

Gibb, N. (2014). *The Fruits of Autonomy*. Retrieved on 14.11.14 from: https://www.gov.uk/government/speeches/nick-gibb-the-fruits-of-autonomy.

Gillies, R. and Khan, A. (2007). Promoting reasoned argumentation, problem‐solving and learning during small‐group work. Cambridge Journal of Education, 39 (1), 7 - 27.

Glass, G.V., McGaw, B. and Smith, M.L. (1981). *Meta-analysis in social research.* London: Sage.

Glaveneau, V-P. (2011). How are we creative together? Comparing sociocognitive and sociocultural answers. *Theory Psychology,* 21, 473 - 492.

Glogger, I., Schwonke, R., Holzäpfel, L., Nückles, M. and Renkl, A. (2012). Learning strategies assessed by journal writing: prediction of learning outcomes by quantity, quality and combinations of learning strategies. *Journal of Educational Psychology,* 104 (2), 452 - 468.

Göb, R., McCollin, C. and Ramalhoto, M. F. (2007). Ordinal methodology in the analysis of Likert scales. *Quality and Quantity*, 41, 601 - 616.

Gorard, S. (2001). *Quantitative methods in educational research: the role of numbers made easy.* London: Continuum.

Gough, D. (1991). Thinking about thinking. *Research Roundup,* 7 (2). National Association of Elementary School Principals.

Graff, G. (2003). *Clueless in academe: How schooling obscures the life of the mind.* New Haven: Yale University Press.

Graham, S., Harris, K.R., MacArthur, C.A. and Schwartz, S. (1991). Writing and writing instruction students with learning disabilities: Review of a research program. *Learning Disability Quarterly,* 14, 89 - 114.

Green, C.S. and Klug, H.G. (1990). Teaching critical thinking and writing through debates: An experimental evaluation. *Teaching Sociology, 18*, 462 - 471.

Greene, J.A. and Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: new perspectives and directions. *Review of Educational Research.* 77 (3), 334 – 372.

Greene, J.A., Moos, D.C. and Azevedo, R. (2011). Self-regulated learning with computer-based learning environments. *New Directions for Teaching and Learning,* 126, 107 - 115.

Greene, J.A., Costa, L-J. and Dellinger, K. (2011). Analysis of self-regulated learning processing using statistical models for count data. *Metacognition Learning,* 6, 275 - 301.

Guardian. (2014) *Guardian Jobs: Graduate Jobs.* Retrieved on 29.9.14 from: http://jobs.theguardian.com/jobs/graduate/

Habermas, J. (1991). *The Theory of Communicative Action.* Cambridge: Polity Press.

Hadwin, A. and Oshige. M. (2011). Self-Regulation, Coregulation, and Socially Shared Regulation: Exploring Perspectives of Social in Self-Regulated Learning Theory. *Teachers College Record,* 113 (2), 240 - 264.

Halpern, D.F. (1998). Teaching critical thinking across domains: dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist,* 53 (4), 449 - 455.

Halpern, D.F. (2003). The "how" and "why" of critical thinking assessment. In D. Fasko (Ed.), *Critical thinking and reasoning: Current research, theory and practice.* Cresskill, NJ: Hampton Press.

Harnishfeger, K. (1995). The development of cognitive inhibition: Theories, definition, and research evidence. In F. Dempster and C. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175 - 204). San Diego, CA: Academic Press.

Harris, M. (2010). *What more can be done to widen access to highly selective universities?* Report for the Office For Fair Access April 2010 http://www.offa.org.uk/wp-content/uploads/2010/05/Sir-Martin-Harris-Fair-Access-report-web-version.pdf (accessed 4.11.14).

Hattie, J.A.C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* London: Routledge.

Heijltjes, A., van Gog, T., Leppink, J. and Paas, F. (2014). Improving critical thinking: Effects of dispositions and instructions on economics students' reasoning skills. *Learning and Instruction.* 29, 31 - 42.

Hertzog, C. and Nesselroade, J.R. (1987). Beyond autoregressive models: Some implications of the trait-state distinction for the structural modeling of developmental change. *Child Development,* 58, 93 - 109.

Hettinger, G. (1995). Raising the level of the debate: The effects of computer mediated communication on group dynamics and critical thinking skills. In: *Proceedings of the 1995 Annual National Convention of the Association for Educational Communications and Technology (AECT)* ERIC, ED 383300.

Hidi, S., Berndorf, D. and Ainley, M. (2002). Children's argument writing, interest and self-efficacy: an intervention study. *Learning and Instruction,* 12, 429 - 446.

Hidi, S. and Ainley, M. (2008). Interest and self-regulation: Relationships between two variables that influence learning. In D.H. Schunk and B.J. Zimmerman (Eds), *Motivation and self-regulated learning: Theory, research and application*. New York: Erlbaum.

Higgins, S., Baumfield, V., Lin, M., Moseley, D., Butterworth, M., Downey, G., Gregson, M., Oberski, I., Rockett, M. and Thacker, D. (2004). Thinking skills approaches to effective teaching and learning: what is the evidence for impact on learners? In: *Research Evidence in Education Library.* London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Higgins, S., Hall, E., Baumfield, V. and Moseley, D. (2005). A meta-analysis of the impact of the implementation of thinking skills approaches on pupils. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Higgins, S., Katsipataki, M., Kokotsaki, D., Coleman, R., Major, L.E. and Coe, R. (2013). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit.* London: Education Endowment Foundation.

Hillman, D. (1999). A new method for analysing patterns of interaction. *The American Journal of Distance Education,* 13 (2), 37 - 47.

Hofmann, W., Schmeichel, B.J. and Baddeley, A.D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, 16 (3), 174 - 180.

Hollingshead, A.B. and McGrath, J.E. (1995). Computer assisted groups: a critical review of the empirical research. In Guzzo, R. A. and Salas, E. (Eds.), *Team effectiveness and decision making* (pp. 46 - 78). San Francisco, CA: Josey Bass.

Holsti, O. (1969). *Content analysis for the social sciences and humanities.* Don Mills: Addison-Welsley.

Honey, P. and Mumford, A. (1989). *The Manual of Learning Opportunities.* Ardingley House: Maidenhead.

Hopfenbeck, T.N. and Maul, A. (2011). Examining Evidence for the Validity of PISA Learning Strategy Scales Based on Student Response Processes. *International Journal of Testing,* 11 (2), 95 - 121.

Hox, J.J. and Bechger, T.M. (1998). An introduction to structural equation modelling. *Family Science Review,* 11, 354 - 373.

Hu, L. and Bentler, P.M. (1999). Cutoff criteria in fix indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling,* 6 (1), 1 - 55.

Hübner, S., Nückles, M. and Renkl, A. (2010). Writing learning journals: instructional support to overcome learning-strategy deficits. *Learning and Instruction,* 20 (1), 18 - 29.

Humphrey, N., Lendrum, A. and Wigelsworth, M. (2010). Social and emotional aspects of learning (SEAL) programme in secondary schools: national evaluation. *Department for Education Research Report,* DFE RR- 049.

Inch, E. S. and Warnick, B. (2002). *Critical thinking and communication: The use of reason in argument* (4th ed.) Needham Heights: Allyn and Bacon.

Inhelder, B. and Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Jacobson, M. and Archodidou, A. (2000). The design of hypermedia tools for learning: Fostering conceptual change and transfer of complex scientific knowledge. *Journal of the Learning Sciences,* 9 (2), 149 - 199.

James, W. (1890). *The Principles of Psychology.* New York: Dover.

James, N., Hughes, C. and Cappa, C. (2010). Conceptualising, developing and assessing critical thinking in law. *Teaching in Higher Education,* 15 (3), 285 - 297.

Jerrim, J. (2013). *Family background and access to 'high status' universities.* Sutton Trust: London.

Joiner, R. and Jones, S. (2003). The effects of communication medium on argumentation and the development of critical thinking. *International Journal of Educational Research,* 39 (8), 861 - 971.

Jonassen, D.H. and Kim, B. (2010). Arguing to learn and learning to argue: design justifications and guidelines. *Educational Technology Research and Development,* 58 (4), 439 - 457.

Joo, Y-J, Bong, M. and Choi, H-J. (2000). Self-efficacy for self-regulated learning, academic self-efficacy, and internet self-efficacy in web-based instruction. *Education Technology Research and Development,* 48 (2), 5 - 17.

Jöreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika,* 34, 183 - 202.

Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika,* 39, 31 - 36.

Kamin, C., O'Sullivan, P. and Deterding, R. (2002). Does project L.I.V.E. case modality impact critical thinking in PBL groups? *Paper presented at the annual meeting of the American Educational Research Association,* New Orleans, LA.

Kaplan, A. (2008). Clarifying Metacognition, Self-Regulation, and Self-Regulated Learning: What's the Purpose? *Educational Psychology Review.* 20, 477 - 484.

Kass, R.A. and Tinsley, H.E.A. (1979). Factor analysis. *Journal of Leisure Research,* 11, 120 - 138.

Keefer, M. W., Zeitz, C.M. and Resnick, L.B. (2000). Judging the quality of peer-led student dialogues. *Cognition and Instruction,* 18 (1), 53 - 81.

Kenny, D.A. (1976). An empirical application of confirmatory factor analysis to multitrait-multimethod matrix. *Journal of Experimental Social Psychology*, *12*, 247 - 252.

Kenny, D.A. (1979). *Correlation and Causality.* New York: Wiley.

Kline, R.B. (2010). *Principles and Practice of Structural Equation Modeling.* New York: Guilford.

Komarraju, M. and Nadler, D. (2013). Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? *Learning and Individual Differences,* 25, 67 - 72.

Ku, K.Y.L. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking Skills and Creativity*, 4 (1), 70 - 76.

Kuhl J. (1987). Action control: the maintenance of motivational states. In F. Halisch and J. Kuhl, (Eds), *Motivation, Intention, and Volition,* pp. 279 - 307. Berlin: Springer-Verlag.

Kuhn, D. (1991). *The skills of argument,* Cambridge, UK: Cambridge University Press.

Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review,* 62 (2), 155 - 178.

Kuhn, D., Shaw, V. and Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and instruction,* 15, 287 - 315.

Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher,* 28 (2), 16 - 46.

Kuhn, D. and Dean, D. (2004). Metacognition: A Bridge Between Cognitive Psychology and Educational Practice. *Theory Into Practice,* 43 (4), 268 - 273.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science,* 9 (5), 178 - 181.

Kuhn, D. (2000a). Theory of mind, metacognition, and reasoning: A life-span perspective. In P. Mitchell and K. J. Riggs (Eds.), *Children's Reasoning and the Mind* (pp. 301 - 326). Hove, UK: Psychology Press.

Kuhn, D. (2001). Why development does (and doesn't) occur: Evidence from the domain of inductive reasoning. In R. Siegler and J. McClelland (Eds.), *Mechanisms of cognitive development: Neural and behavioral perspectives* (pp. 221 - 249). Mahwah, NJ: Erlbaum.

Kuhn, D. (2005). *Education for thinking.* Cambridge, MA: Harvard University Press.

Kuhn, D. (2006). Do Cognitive Changes Accompany Developments in the Adolescent Brain? *Perspectives on Psychological Science,* 1, 59 – 67.

Kuhn, D. and Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking and Reasoning,* 13, 90 - 104.

Kuhn, D. (2010). Teaching and learning science as argument. *Science Education,* 94, 810 - 824.

Kuhn, D. and Crowell, A. (2011). Dialogic Argumentation as a Vehicle for Developing Young Adolescents' Thinking. *Psychological Science,* 22, 545 - 552.

Kuhn, D, Zillmer, N., Crowell A. and Zavala, J. (2013). Developing Norms of Argumentation: *Metacognitive, Epistemological, and Social Dimensions of Developing Argumentive Competence, Cognition and Instruction,* 31 (4), 456 - 496.

Kurfiss, J.G. (1988). *Critical thinking theory, research, practice and possibilities. (ASHE-ERIC Higher Education Report no. 2).* Washington DC: Association for Study for Higher Education.

Lajoie, S.P. (2008). Metacognition, Self Regulation, and Self-regulated Learning: A Rose by any other Name? *Educational Psychology Review.* 20, 469 – 475.

Lane Fox, R. (2008). Tutorials in Greats and history: the Socratic Method. In D. Palfreyman (Ed) *The Oxford Tutorial: Thanks, you taught me how to think.* Oxford: Blackwell.

Langer, A.M. (2002). Reflecting on practice: Using learning journal in higher and continuing education. T*eaching in Higher Education, 7*, 338 - 351.

Larson-Hall, J. (2010). *A Guide to Doing Statistics in Second Language Research Using SPSS.* New York: Routledge.

Lee, W., Lee, M-J. and Bong, M. (2014). Testing interest and self-efficacy as predictors of academic self-regulation and achievement. *Contemporary Educational Psychology,* 39, 86 - 99.

Lee, H.W., Lim, K.Y. and Grabowski, B.L. (2010). Improving Self-Regulation, Learning Strategy Use, and Achievement with Metacognitive Feedback. *Educational Technology Research and Development,* 58 (6), 629 - 648.

Leitão, S. (2000). The potential of argument in knowledge building. *Human Development*, 43 (6), 332 - 360.

Leitão, S. (2003). Evaluating and selecting counterarguments. *Written Communication,* 20 (3), 269 - 306.

Leitão, S. (2007). Argumentation and development of reflective thinking. *Psicologia: Reflexão e Crítica,* 20 (3), 454 - 462.

Lieberman, M.D. (2011). Why Symbolic Processing of Affect Can Disrupt Negative Affect. In A. Todorov, S. Fiske and D. Prentice (Eds) *Social Neuroscience: Toward Understanding the Underpinnings of the Social Mind.* Oxford University Press. Retrieved on 9.9.14 from: http://ezproxy- prd.bodleian.ox.ac.uk:2153/view/10.1093/ acprof:oso/ 9780195316872.001.0001/acprof-9780195316872 -chapter- 13.

Lin, T-J. (2011). Review of Moodle 2.0. *Language Learning & Technology,* 15 (2), 27 – 33.

Little, R.J.A. and Rubin, D.B. (1987). *Statistical analysis with missing data.* New York: Wiley.

Locke, J. (1689). *An Essay Concerning Human Understanding.* New York: Prometheus.

Locke, E.A. and Latham, G.P. (2002). Building a practically useful theory of goal setting and task motivation: a 35 year odyssey. *American Psychology,* 57, 705 - 717.

Lodewyk, K.R. and Winne, P.H. (2005). Relations Among the Structure of Learning Tasks, Achievement, and Changes in Self-Efficacy in Secondary Students. *Journal of Educational Psychology,* 97 (1), 3 - 12.

Luna, B., Garver, K., Urban, T., Lazar, N. and Sweeney, J. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development,* 75, 1357 - 1372.

Magno, C. (2010). The role of metacognitive skills in developing critical thinking. *Metacognition and Learning,* 5 (2), 137 - 156.

Mair, C. (2012). Using technology for enhancing reflective writing, metacognition and learning. *Journal of Further and Higher Education,* 36 (2), 147 - 167.

Marin, L. and Halpern, D.F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity,* 6, 1 - 13.

Marsh, H.W. and Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist,* 20, 107 - 123.

Marsh, H.W. (1989). Confirmatory factor analysis of multitrait-multimethod data: Many problems and few solutions. *Applied Psychological Measurement*, 13, 335 - 361.

Marsh, H.W. and Yeung, A.S. (1997). The causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology,* 89, 41 - 54.

Marsh, H.W., Trautwein, U., Ludtke, O., Koller, O. and Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct validation and relations to well-being and achievement. *Journal of Personality,* 74, 403 - 456.

Marsh, H.W., Hau, K-T, Artelt, C., Baumert, J. and Peschar, J.L. (2006a). OECD's Brief Self-Report Measure of Educational Psychology's Most Useful Affective Constructs: Cross-Cultural, Psychometric Comparisons Across 25 Countries. *International Journal of Testing,* 6 (4), 311 - 360.

Martin, A. Papworth, B., Ginns, P. and Liem, G.A.D. (2014). Boarding School, Academic Motivation and Engagement, and Psychological Well-Being: A Large-Scale Investigation. *American Educational Research Journal,* 51 (5), 1007 - 1049.

Matell, M.S. and Jacoby, J. (1971). Is there an optimal number of alternatives for Likert scale items? Study 1: reliability and validity. *Educational and Psychological Measurement,* 31, 657 - 674.

Maul, A. (2013). Method effects and the meaning of measurement. *Frontiers in Psychology,* 4 (169), 1 - 13.

Maurer, T.J. and Pierce, H.R. (1998). A Comparison of Likert Scale and Traditional Measures of Self-Efficacy. *Journal of Applied Psychology*, 83, 324 - 329.

Mayer, J. (1986). Teaching Critical Awareness in an Introductory Course. *Teaching Sociology,* 14, 249 - 256.

McArdle, J.J. and Prindle, J.J. (2008). A latent change score analysis of a randomized clinical trial in reasoning training. *Psychology and Aging,* 23, 702 - 719.

McArdle, J.J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology,* 60, 577 - 605.

McCarthy-Tucker, T.S.N. (1998). Teaching logic to adolescents to improve thinking skills. *Korean Journal of Thinking and Problem Solving*, 8 (1), 45 - 66.

McCrindle, A. and Christensen, C. (1995). The impact of learning journals on metacognitive and cognitive processes and learning performances. *Learning and Instruction,* 5, 167 - 185.

McDonald, J. (1998). Interpersonal group dynamics and development in computer conferencing: The rest of the story. In *Wisconsin Distance Education Proceedings.* Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. Retrieved on 27.7.11 from: http://eric.ed.gov/PDFS/ED422864.pdf.

McGeown, S.P., Putwain, D., Simpson, E.G., Boffey, E., Markham, J. and Vince, A. (2014). Predictors of adolescents' academic motivation: Personality, self-efficacy and adolescents' characteristics. *Learning and Individual Differences,* 32, 278 - 286.

McInerney, D.M. (2008). The motivational roles of cultural differences and cultural identity in self-regulated learning. In D.H. Schunk and B.J. Zimmerman (Eds), *Motivation and self-regulated learning: Theory, research and application*. New York: Erlbaum.

McKendree, J., Stenning, K., Mayes, T., Lee, J. and Cox, R. (1998). Why observing a dialogue may benefit learning. *Journal of Computer Assisted Learning,* 14, 110 - 119.

McMillan, J.H. and Schumacher, S. (2010). *Research in Education (7th Edition),* Boston: Pearson.

McPeck, J. (1981). *Critical Thinking in Education.* New York: St Martin's Press.

Means, M.L. and Voss, J.F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. *Cognition and Instruction,* 14 (2), 139 - 178.

Meece, J.L. and Painter, J. (2008). Gender, self-regulation, and motivation. In D.H. Schunk and B.J. Zimmerman (Eds), *Motivation and self-regulated learning: Theory, research and application*. New York: Erlbaum.

Mega, C., Ronconi, L. and De Beni, R. (2014). What Makes a Good Student? How Emotions, Self-Regulated Learning, and Motivation Contribute to Academic Achievement. *Journal of Educational Psychology,* 106 (1), 121–131.

Mehmet, B. (2010). Making learning visible in kindergarten classrooms: pedagogical documentation as a formative assessment technique. *Teaching and Teacher Education,* 26 (7), 1439 - 1449.

Meyer, B., Haywood. N., Sachdev, D. and Faraday, S. (2008). *Independent Learning: Literature Review* (DCSF RR051). DCSF: London.

Milburn, A. (2009). Panel on Fair Access to the Professions, *Unleashing aspiration: the final report of the Panel on Fair Access to the Professions,* Cabinet Office, London, Retrieved 24. 12.2014, http://www.cabinetoffice.gov.uk/media/227102/fair-access.pdf.

Milburn-Curtis, C.J. (2011). *Digital-Thinkers: The impact of teaching presence on the higher-order thinking outcomes of an online learning community of high-ability young teenagers in the UK.* Unpublished MSc dissertation.

Millar, K. (1983). Clinical trial design: the neglected problem of asymmetrical transfer in cross-over trials. *Psychological Medicine,* 13 (4), 867 - 73.

Miller, E.K. and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience,* 24 (1), 167 - 202.

Miller, P.H., Kessel, F.S. and Flavell, J.H. (1970). Thinking about people thinking about people thinking about...: A study of social-cognitive development*. Child Development,* 41, 613 - 623.

Molenaar, I, van Boxtel, C.A.M. and Sleegers, P.J.C. (2011). Metacognitive scaffolding in an innovative learning arrangement. *Instructional Science,* 39, 785 - 803.

Molenaar, P.C.M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika,* 50, 181 - 202.

Morse, B. (2012). Why critical thinking is overlooked by schools and shunned by students. *The Guardian Newspaper,* 12th September 2012. Retrieved on 14.11.14 from: http://www.theguardian.com/teacher-network/2012/sep/12/critical-thinking-overlooked-in-secondary-education.

Moshman, D. and Geil, M. (1998). Collaborative reasoning: Evidence for collective rationality. *Thinking and Reasoning,* 4, 231 - 248.

Moshman, D. (2004). From inference to reasoning: the construction of rationality. *Thinking and Reasoning,* 10 (2), 221 - 239.

Muijs, D. (2010). *Doing quantitative research in education with SPSS. 2nd edition.* London, GB: Sage Publications.

Mun, E.Y., von Eye, A. and White, H.R. (2009). An SEM approach for the evaluation of intervention effects using pre-post-post designs. *Structural Equation Modeling,* 16 (2), 315 - 337.

Muthén, B.O. and Muthén, L.K. (2013). *MPlus User's Guide.* Retrieved on 12.12.13 from: http://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver_7_r6_web .pdf.

Nassar, D. (2012). Spinoza in Schelling's early conception of intellectual intuition. In E. Forster and Y.Y. Melamed (Eds) *Spinoza and German Idealism*. Cambridge: Cambridge University Press.

National Curriculum. (2014). *The National Curriculum in England: English.* Downloaded from https://www.gov.uk/government/publications/national-curriculum-in-england-english-programmes-of-study/national-curriculum-in-england-english-programmes-of-study.

National Curriculum. (2014a). *English Programmes of Study Key Stages 1 and 2.* Retrieved on 21.8.14 from: https://www.gov.uk/government/uploads/system/ uploads/attachment_data/file/284162/Primary_English_curriculum_to_July_2015_RS.pdf.

National Curriculum. (2014b). *Department for Education. National Curriculum Framework Document for Teaching.* Retrieved on 14.11.14 from: https://www.gov.uk/government/ uploads/ system/uploads/attachment_data/file/339805/MASTER_final_national_curriculum _until_sept_2015_11_9_13.pdf.

Nett, U.E., Götz, T., Hall, N.C. and Frenzel, A.C. (2012). Metacognitive strategies and test performance: an experience sampling analysis of students learning behavior. *Education Research International,* 1 - 16.

Newman, D., Webb, B. and Cochrane, C. (1995). A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century,* 3 (2), 56 - 77.

Niemivirta, M. (1999). Motivational and cognitive predictors of goal setting and task performance. *International Journal of Educational Research (Oxford, UK),* 31, 499 - 513.

Nückles, M., Hübner, S. and Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning and Instruction,* 19, 259 - 271.

Nückles, M., Schwonke, R., Berthold. K. and Renkl, A. (2004). The use of public learning diaries in blended learning. *Journal of Educational Media,* 29 (1), 49 - 66.

Nunnally J.C. and Bernstein, I.H. (1994). *Psychometric Theory (3rd Edition).* New York: McGraw Hill.

Nussbaum, E.M. (2005). The effect of goal instructions and need for cognition on interactive argumentation. *Contemporary Educational Psychology,* 30, 286 - 313.

Nussbaum, E.M. and Kardash, C.M. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology,* 97 (2), 157 - 169.

Nussbaum, E.M. and Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *The Journal of Experimental Education,* 76, 59 - 92.

OECD. (2010), *PISA 2009 Results: Learning to Learn - Student Engagement, Strategies and Practices (Volume III)* Retrieved on 20.4.12 from http://www.oecd.org/document/2/0,3746,en_32252351_46584327_46609794_1_1_1_1,00.html

OECD. (2012). *Programme for International Student Assessment (PISA). Results from 2012.* Retrieved on 13.11.14 from: http://www.oecd.org/unitedkingdom/PISA-2012 -results-UK.pdf

OFFA. (2015). Promoting Fair Access to Higher Education. *Office for Fair Access.* Retrieved on 1.1.15 from: http://www.offa.org.uk.

OFSTED. (2012). *OFSTED News*. Retrieved on 23.3.12 from: http://ofstednews.ofsted.gov.uk/issue/33/schools.

O'Keefe, D.J. (1999). How to handle opposing arguments in persuasive messages: a meta-analytic review of the effects of one-sided and two-sided messages. *Communication Yearbook,* 22, 209 - 249.

Olson, D.R. and Astington, J.W. (1993). The role of concepts in perception and inference. *Behavioral and Brain Sciences,* 16 (1), 65 - 66.

O'Malley, J.M., Chamot, A.U., Stewner-Mazanares, G., Kupper, L. and Russo, R. (1985). Learning strategies used by beginning and intermediate ESL students. *Language Learning,* 35, (1), 21 - 46.

ONS. (2013). Office for National Statistics. *Household Internet Access.* Retrieved on 10.12.13 from: http://www.ons.gov.uk/ons/rel/rdit2/internet-access-households-and-individuals/2013/stb-ia-2013.html#tab-Household-Internet-access.

ONS. (2014). Office for National Statistics. *Self-employed workers in the UK - 2014.* Retrieved on 29.9.14 from: http://www.ons.gov.uk/ons/dcp171776_374941.pdf

Oxford. (2015). *Widening Participation, University of Oxford.* Retrieved on 1.1.15 from http://www.ox.ac.uk/local-community/outreach-and-education/widening-participation.

Pajares, F. and Miller, M.D. (1994). The role of self-efficacy and self-concept beliefs in mathematical problem-solving: A path analysis. *Journal of Educational Psychology,* 86, 193 - 203.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research,* 66, 543 - 78.

Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into Practice,* 41 (2), 116 - 125.

Pajares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. In D.H. Schunk and B.J. Zimmerman (Eds), *Motivation and self-regulated learning: Theory, research and application*. New York: Erlbaum.

Paris, S.G. and Byrnes, J.P. (1989). The constructivist approach to self-regulation and learning in the classroom. In B.J. Zimmerman and D.H. Schunk (Eds.), *Self-regulated Learning and Academic Achievement: Theory, Research and Practice.* New York: Springer-Verlag.

Passey, D. and Higgins, S. (2011). Learning platforms and learning outcomes - insights from research. *Learning, Media and Technology,* 36 (4), 329 - 333.

Pastorelli, C., Caprara, G.V., Barbaranelli, C., Rola, J., Rozsa, S. and Bandura, A. (2001). The structure of children's perceived self-efficacy: A cross-national study. *European Journal of Psychological Assessment,* 17, 87 - 97.

Paton, G. (2014). Teachers 'cheat' to hit exam targets. *Daily Telegraph,* April 13th, 2014. Retrieved on 14.11.14 from: http://www.telegraph.co.uk/education/educationnews/ 10763017/Teachers-cheat-to-hit-exam-targets.html.

Paul, R. (1990). *Critical thinking*. Rohnert Park, CA: Center for Critical Thinking and Moral Critique.

Paul, R. (1993). *Critical thinking: What every person needs to survive in a rapidly changing world.* Rohnert Park, CA: Foundation for Critical Thinking.

Paul, R. and Elder. L. (2009). Critical thinking concepts and tools (6th Edition). *Foundation for critical thinking,* Retrieved on 9.8.12 from: www.criticalthinking.org.

Pekrun, R., Goetz, T., Titz, W. and Perry, R.P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist,* 37, 91 - 105.

Perels, F., Dignath, C. and Schmitz, B. (2009). Is it possible to improve mathematical achievement by means of self-regulation strategies? Evaluation of an intervention in regular math classes. *European Journal of Psychology of Education,* 14 (1), 17 - 31.

Perels, F., Otto, B., Landmann, M., Hertel, S. and Schmitz, B. (2007). Self-regulation from the process perspective. *The Journal of Psychology,* 215, 149 - 204.

Perepletchikova, F., Treat, T. and Kazdin, A.E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology*, 75, 829 - 841.

Perkins, D.N. (1985). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology,* 77, 562 - 571.

Perkins, D.N., Farady, M. and Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. Voss, D.N. Perkins and J. Segal (Eds.), *Informal reasoning* (pp. 83 - 105). Hillsdale, NJ: Lawrence Erlbaum.

Perry, N.E., Phillips, L. and Dowler, J. (2004). Examining features of tasks and their potential to promote self-regulated learning. *Teachers College Record,* 106 (9), 1854 - 1878.

Phan, H.P. (2006). Examination of student learning approaches, reflective thinking, and epistemological beliefs: A latent variables approach. *Electronic Journal of Research in Educational Psychology,* 10 (4), 577 - 610.

Phan, H.P. (2010). Students' academic performance and various cognitive processes of learning: An integrative framework and empirical analysis. *Educational Psychology,* 30, 297 - 322.

Phan, H.P. (2014). Self-efficacy, reflection, and achievement: a short-term longitudinal examination. *The Journal of Educational Research,* 107 (2), 90 - 102.

Piaget, J. (1932). *The moral judgment of the child.* New York: The Free Press.

Piaget, J. (1964). Development and Learning. In R.E. Ripple and V.N. Rockcastle: *Piaget Rediscovered.* pp. 7 - 12. New York: Freeman.

Piantadosi, S. (2005). Crossover Designs. In: Piantadosi, S. *Clinical Trials: A Methodological Perspective (2nd Edition).* Hobaken, NJ: John Wiley and Sons.

Pinto, M.A., Iliceto, P. and Melogno, S. (2012). Argumentative abilities in metacognition and in metalinguistics: a study on university students. *European Journal of Psychology of Education*, 27 (1), 35 - 58.

Pintrich, R.R. and De Groot, E.V. (1990). Motivational and self-regulated learning components of classroom academic performance, *Journal of Educational Psychology,* 82, 33 - 40.

Pintrich, P.R., Smith, D.A., Garcia, T. and McKeachie, W.J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)* (Technical Rep. No. 91 -B-004). Ann Arbor: University of Michigan, School of Education.

Pintrich, P.R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*. 31, 459 - 470.

Pintrich, P.R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P.R. Pintrich and M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451 - 502). San Diego, CA: Academic Press.

Pintrich P.R. (2004). A Conceptual Framework for Assessing Motivation and Self-Regulated Learning in College Students. *Educational Psychology Review,* 16 (4), 385 - 407.

Preece, J., Nonnecke, B. and Andrews, D. (2004). The top 5 reasons for lurking: Improving community experiences for everyone. *Computers in Human Behavior*, 20, 201 - 223.

Preiss, D.D., Castillo, J.C., Flotts, P. and San Martín, E.S. (2013). Assessment of argumentative writing and critical thinking in higher education: Educational correlates and gender differences. *Learning and Individual Differences,* 28, 193 - 203.

Preiss, D.D., Castillo, J.C., Grigorenko, E.L. and Manzi, J. (2013) Argumentative writing and academic achievement: A longitudinal study. *Learning and Individual Differences,* 28, 204 - 211.

Pressley, M. (1995). More about the development of self-regulation: Complex, longterm and thoroughly social. *Educational Psychologist*, 30 (4), 207 - 12.

Pressley, M. and Ghatala, E.S. (1990). Self-regulated learning: monitoring learning from the text. *Educational Psychologist,* 25, 19 - 33.

Preston, C.C. and Colman, A.M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica,* 104, 1 - 15.

Puustinen, M. and Pulkkinen, L. (2001). Models of Self-regulated Learning: A review, *Scandinavian Journal of Educational Research,* 45 (3), 269 - 286.

Rest, J.R., Narvaez, D., Thoma, S.J. and Bebeau, M.J. (1999). DIT2: Devising and testing a revised instrument of moral judgment. *Journal of Educational Psychology,* 91 (4), 644 - 659.

Reznitskya, A., Anderson, R.C., McNurlin, B., Nguyen-Jahiel, K., Archodidou, A. and Kim, S.Y. (2001). Influence of oral discussion on written argumentation. *Discourse Processes,* 32 (2 and 3), 155 - 175.

Rheinberg, F., Vollmeyer, R. and Burns, B.D. (2001). QCM: A questionnaire to assess current motivation in learning situations. *Diagnostica*, 47, 57 - 66.

Riesenmy, M.R., Mitchell, S., Hudgins, B.B. and Ebel, D. (1991). Retention and transfer of children's self-directed critical thinking skills. *Journal of Educational Research,* 85 (1), 14 - 25.

Robson, C.R. (2002). *Real World Research,* Oxford: Blackwell.

Rourke, L., Anderson, T., Garrison, R. and Archer, W. (2001). Methodological Issues in the Content Analysis of Computer Conference Transcripts. *International Journal of Artificial Intelligence in Education* (IJAIED), 12, 8 - 22.

Ryan, R. and Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology,* 25, 54 - 67.

Ryan, A. (2008). A Liberal Education: and that includes the Sciences! In D. Palfreyman (Ed). *Thanks, you taught me how to think.* Oxfordshire: Oxford Centre for Higher Education Policy Studies.

Salmon, G. (2000). *E-moderating: The key to teaching and learning online.* London: Kogan Page.

Sanders, J.A., Wiseman, R.L. and Gass, R.H. (1994). Does teaching argumentation facilitate critical thinking? *Communication Reports,* 7 (1), 27 - 35.

Sanetti, L.M.H., Dobey, L.M. and Gallucci, J. (2014). Treatment integrity of interventions with children in School Psychology from 1995 – 2010. *School Psychology International,* 35, 370 - 383.

Schafer, J.L. and Graham, J.W. (2002). Missing Data: Our View of the State of the Art. *Psychological Methods,* 7 (2), 147 - 177.

Schellens, T. and Valcke, M. (2006). Fostering knowledge construction in university students through asynchronous discussion groups. *Computers and Education,* 46, 349 - 370.

Scherer, R. (2013). Further evidence on the structural relationship between academic self-concept and self-efficacy: On the effects of domain specificity. *Learning and Individual Differences,* 28, 9 - 19.

Schmidt, K., Maier, J. and Nückles, M. (2012). *Writing about the Personal Utility of Learning Contents in a Learning Journal Improves Learning Motivation and Comprehension.* Retrieved on 8.7.14 from: http://www.hindawi.com/journals/edri/2012/319463/ref/.

Schmitz, B. (2006). Advantages of studying processes in educational research. *Learning and Instruction,* 16, 433 - 449.

Schmitz, B. and Wiese, B.S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time series analyses of diary data. *Contemporary Educational Psychology,* 31, 64 - 96.

Schmitz, B., Schmidt, M., Landmann, M. and Spiel, C. (2007). New developments in the field of self-regulated learning. *Journal of Psychology,* 215 (3), 153 - 156.

Schmitz, B. and Perels, F. (2011). Self-monitoring of self-regulation during math homework behaviour using standardized diaries. *Metacognition Learning,* 6, 255 - 273.

Schmitz, B., Klug, J. and Schmidt, M. (2011). Assessing self-regulated learning using diary measures with university students. In B.J. Zimmerman and D.H. Schunk (Eds). *Handbook of Self-Regulation of Learning and Performance*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schraw, G. and Dennison, R.S. (1994). Assessing metacognitive knowledge. *Contemporary Educational Psychology,* 19, 460 - 475.

Schraw, G., Crippen, K.J. and Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education,* 36, 111 - 139.

Schunk, D.H. (1983). Progress Self-Monitoring: Effects on Children's Self-Efficacy and Achievement. *Journal of Experimental Education,* 51, 89 - 93.

Schunk, D.H. (1984). The self-efficacy perspective on achievement behavior. *Educational Psychologist,* 19, 199 - 218.

Schunk, D.H. (1989). Social cognitive theory and self-regulated learning. In B.J. Zimmerman and D.H. Schunk (Eds) *Self-regulated learning and academic achievement: Theory, research, and practice.* New York: Springer-Verlag.

Schunk, D.H. (1990). Goal Setting and Self-Efficacy During Self-Regulated Learning. *Educational Psychologist,* 25 (1), 71 - 86.

Schunk, D.H. (2005). Self-regulated learning: The educational legacy of Paul R. Pintrich. *Educational Psychologist,* 40 (2), 85 - 94.

Schunk, D.H. (2008). Metacognition, self-regulation and self-regulated learning: Research recommendations. *Educational Psychology Review,* 20, 463 - 467.

Schunk, D.H. and Ertmer, P. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. In J. Boekarts, P. Pintrich and M. Zeidner (Eds.) *Handbook of Self-Regulation.* Burlington, MA: Elsevier Academic Press.

Schunk, D.H. and Zimmerman, B. J. (2008). *Motivation and self-regulated learning: Theory, research, and applications.* New York: Taylor and Francis.

Schunk, D.H., Pintrich, P.R. and Meece, J.L. (2008). *Motivation in education: Theory, research, applications.* (3rd Ed). Upper Saddle River, NJ: Pearson Education.

Schwartz, B.B., Neuman, Y., Gil, J. and Ilya, M. (2003). Construction of collective and individual knowledge in argumentative activity. *The Journal of the Learning Sciences,* 12 (2), 219 - 256.

Schwinger, M., Steinmayr, R. and Spinath, B. (2009). How do motivational regulation strategies affect achievement: Mediated by effort management and moderated by intelligence. *Learning and Individual Differences,* 19, 621 - 627.

Schwinger, M. and Stiensmeier-Pelster, J. (2012). Effects of motivational regulation on effort and achievement: A mediation model, *International Journal of Educational Research,* 56, 35-47.

Schwonke, R., Hauser, S., Nückles, M. and Renkl, A. (2006). Enhancing computer-supported writing of learning protocols by adaptive prompts. *Computers in Human Behavior,* 22, 77-92.

Schwonke, R., Ertelt, A., Otieno, C., Renkl, A., Aleven, V. and Salden, R.J.C.M. (2013). Metacognitive support promotes an effective use of instructional resources in intelligent tutoring. *Learning and Instruction,* 23, 136 - 150.

Sechrest, L., Davis, M.F., Stickle, T.R. and McKnight, P.E. (2000). Understanding 'method' variance, in L. Bickman (Ed) *Research Design: Donald Campbell's Legacy*, p.p. 63 - 87, Thousand Oaks, CA: Sage Publications.

Selmeczy, D. and Dobbins, I.G. (2013). Metacognitive Awareness and Adaptive Recognition Biases. *Journal of Experimental Psychology: Learning Memory and Cognition,* 39 (3), 678 - 690.

Shadish, W.R., Cook, T.D. and Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin.

Shapiro, A. (2008). Hypermedia design as learner scaffolding. *Educational Technology Research and Development,* 56, 29 - 44.

Shavelson, R.J., Hübner, J.J. and Stanton, G.C. (1976). Self concept: Validation of construct interpretations. *Review of Educational Research*, 46, 407 - 441.

Shea, P. and Bidjerano, T. (2010). Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments. *Computers and Education,* 50, 1721 - 1731.

Shehab, H.M. and Nussbaum, E.M. (2015). Cognitive load of critical thinking strategies. *Learning and Instruction,* 35, 51 - 61.

Simon, K.G. (2001). *Moral questions in the classroom: How to get kids to think deeply about real life and their schoolwork.* New Haven: Yale University Press.

Simpson, E. and Courtney, M. (2002). Critical thinking in nursing education: Literature review. *International Journal of Nursing Practice*, 8, 89 - 98.

Sitzmann, T. (2010). Self-regulating online course engagement. *American Society for Training and Development,* 64 (3), p.26. Retrieved on 16.12.13. from: http://www.ilr.cornell.edu/cahrs/news/upload/BellResearchT-DMarch10.pdf.

Slagter van Tyron, P.J. and Bishop, M.J. (2009). Theoretical foundations for enhancing social connectedness in online learning environments. *Distance Education,* 30, 291 - 315.

Sperling, R.A., Howard, B.C., Miller, L.A. and Murphy, C.M. (2002). Measures of Children's Knowledge and Regulation of Cognition. *Contemporary Educational Psychology,* 27, 51 - 79.

Spitzberg, B.H. (2011). The Interactive Media Package for Assessment of Communication and Critical Thinking (IMPACCT): Testing a Programmatic Online Communication Competence Assessment System. *Communication Education,* 60 (2), 145 - 173,

Spörer, N. and Brunstein, J.C. (2006). Assessing self-regulated learning with self-report measures: a state of the art review. *German Journal of Educational Psychology,* 20, 147 - 160.

SPSS. (2011). *IBM SPSS Statistics for Windows, Version 20.0.* Armonk, NY: IBM Corp.

Stahl, G., Cress, U., Ludvigsen, S. and Law, N. (2014). Dialogic foundations of CSCL. *International Journal of Computer-Supported Collaborative Learning*, 9, 117 - 125.

Stanovich, K.E. and West, R.F. (2008). On the failure of cognitive ability to predict myside and one-sided thinking biases. *Thinking and Reasoning,* 14, 129 - 167.

Stedman, N., Irani, T., Friedel, C., Rhoades, E. and Ricketts, J.C. (2009). Relationships between critical thinking disposition and Need for Cognition among undergraduate students enrolled in leadership courses. *NACTA Journal,* 53 (3), 62 - 70.

Steiger, J.H. and Lind, J.C. (1980). Statistically based tests for the number of common factors. *Paper presented at the Annual Meeting of the Psychometric Society,* Iowa City, IA.

Stein, N.L. and Bernas, R. (1999). The early emergence of argumentative knowledge and skill. In J. Andriessen and P. Corrier (Eds.), *Foundations of argumentative text processing* (pp. 97 - 116). Amsterdam: Amsterdam University Press.

Stephens, K. and Winterbottom, M. (2010). Using a learning log to support students' learning in biology lessons. *Journal of Biological Education,* 44 (2), 72 - 80.

Sungur, S. (2007). Modeling the relationships among students' motivational beliefs, metacognitive strategy use, and effort regulation. *Scandinavian Journal of Educational Research*, 51 (3), 315 - 326.

Swanson, H.L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology,* 82 (2), 306 - 314.

Tabachnick, B.G. and Fidell, L.S. (2007, 2013). *Using Multivariate Statistics,* London: Allyn and Bacon.

Taylor, J. (2012). Philosophical teaching will get students thinking for themselves again. *Guardian Newspaper, 14th August, 2012.* Retrieved on 14.11.14 from: http://www.theguardian.com/teacher-network/2012/aug/14/philisophical-teaching-students?guni=Article:in%20body%20link.

ThinkUKnow. (2011). Retrieved on 21.6.11 from www.thinkuknow.co.uk.

Toplak, M.E. and Stanovich, K.E. (2003). Associations between myside bias on an informal reasoning task and amount of post-secondary education. *Applied Cognitive Psychology,* 17, 851 - 860.

Toulmin, S.E. (1958). *The uses of argument.* Cambridge, England: Cambridge University Press.

Tucker, L.R. and Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika,* 38, 1 - 10.

Valle, A., Cabanach, R.G., Nunez, J.C., Gonzalez-Pienda, J., Rodriguez, S. and Pieniro, I. (2003). Cognitive, motivational, and volitional dimension of learning. *Research in Higher Education,* 44, 557 - 580.

Van den Brink-Budgen, R. (2010). *Critical thinking for students.* Oxford: How to Books.

Van Eemeren, F.H. and Grootendorst, R. (2003). A pragma-dialectical procedure for a critical discussion. *Argumentation,* 17, 365 - 386.

Van Laerhoven, H., Van der Zaag-Loonen, H.J. and Derkx, B.H.F. (2004). A comparison of Likert scale and visual analogue scales as response options in children's questionnaires. *Acta Paediatra,* 93, 830 - 835.

Vollmeyer, R. and Rheinberg, F. (2006). Motivational Effects on Self-Regulated Learning with Different Tasks. *Educational Psychology Review,* 18, 239 - 253.

Von Aufschnaiter, C., Erduran, S., Osborne, J. and Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation related to their scientific knowledge. *International Journal of Science Education,* 45, 101 - 131.

Voss, J.F. and Means, M.L. (1991). Learning to reason via instruction in argumentation. *Learning and Instruction,* 1, 337 - 350.

Voss, J.F., Perkins, D.N. and Segal, J.W. (1991). Preface. In F. Voss, D. N. Perkins and J. W. Segal (Eds.), *Informal reasoning in education.* Hillsdale, NJ: Erlbaum.

Walton, D.N. (1989). Dialog Theory for Critical Thinking. *Argumentation,* 3, 169 - 184.

Walton, D.N. (1995). Argument. In (Ed) T. Honderich, (1995). *The Oxford Companion to Philosophy,* Oxford: Oxford University Press.

Wang, J-S., Pascarella, E.T., Nelson Laird, T.F. and Ribera, A.K. (2014). How clear and organized classroom instruction and deep approaches to learning affect growth in critical thinking and need for cognition. *Studies in Higher Education,* 1 – 22.

Waring, M. (2010). *14 - 19 Alliance Discussion Paper. Post- 14 Research and Innovation, Institute of Education, University of London.* Retrieved on 18.4.12 from http://www.ioe.ac.uk/Diploma_discussion_May2010.pdf.

Warwick University (2015). *Learning with IGGY.* Retrieved on 16.3.15 from: https://www.iggy.net/about/content/learning-with-iggy.

Watson, G. and Glaser, E.M. (1980). *Watson-Glaser Critical Thinking Appraisal Manual: Forms A and B.* San Antonio, TX: The Psychological Corporation.

Webber, J., Scheuermann, B., McCall, C. and Coleman, M. (1993). Research on self-monitoring as a behavior management technique in special education classrooms: A descriptive review. *Remedial and Special Education*, 14, 38 - 56.

Wecker, C. and Fischer, F. (2014). Where is the evidence? A meta-analysis on the role of argumentation for the acquisition of domain-specific knowledge in computer-supported collaborative learning. *Computers and Education,* 75, 218 - 228.

Weinstein, C.E. and Mayer, R.E. (1983). *The Teaching of Learning Strategies*. Retrieved on 28.7.14 from: http://www.scribd.com/doc/79036011/Weinstein-and-Mayer- 1983.

Welser, H.T., Smith, M., Fisher, D. and Gleave, E. (2008). Distilling digital traces: computational social science approaches to studying the internet. In (Eds) N. Fielding, R.M. Lee and G. Blank, *The SAGE Handbook of Online Research Methods,* pp.116 - 141. London: Sage.

West, S.G., Taylor, A.B. and Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling.* New York: Guilford.

Wigfield, A., Eccles, J.S., Schiefele, U., Roeser, R.W. and Davis-Kean, P. (2006). Development of achievement motivation. In N. Eisenberg (ed.) *Handbook of child psychology: Vol. 3, Social, emotional, and personality development.* New York: Wiley.

Wigfield, A., Hoa, L.W. and Klauda, S.L. (2008). The role of achievement values in the regulation of achievement behaviors. In D.H. Schunk and B.J. Zimmerman (Eds), *Motivation and self-regulated learning: Theory, research and application*. New York: Erlbaum.

Williams, R.L. (2003). *Critical thinking as a predictor and outcome measure in a large undergraduate educational psychology course.* University of Tennessee, Knoxville. Retrieved on 20.11.14 from: http://files.eric.ed.gov/fulltext/ED478075.pdf.

Wimmer, R.D. and Dominick, J.R. (2006). *Mass media research: an introduction.* USA: Thomson.

Winne, P.H. (1995). Inherent details in self-regulated learning. *Educational Psychologist,* 30, 173 - 187.

Winne, P.H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences,* 8, 327 - 353.

Winne, P.H. and Perry, N.E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich and M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531 - 566). Burlington: Elsevier.

Winne, P.H. (2005). Key Issues in Modeling and Applying Research on Self-Regulated Learning. *Applied Psychology.* 54 (2), 232 - 238.

Winne, P.H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition Learning,* 9, 229 - 237.

Winters, F.I., Greene, J.A. and Costich, C.M. (2008). Self-regulation of learning within computer-based learning environments: a critical analysis. *Education Psychology Review,* 20, 429 - 444.

Wise, A.F., Hausknecht, S.N. and Zhao, Y. (2014). Attending to others' posts in asynchronous discussions: Learners' online "listening" and its relationship to speaking. *International Journal Computer-Supported Collaborative Learning,* 9, 185 - 209.

Wolfe, C.R. and Britt, M.A. (2008). The locus of the myside bias in written argumentation. *Thinking and Reasoning*, 14, 1 - 27.

Wolfe, C.R., Britt, M.A. and Butler, J.A. (2009). Argumentation schema and the myside bias in written argumentation. *Written Communication,* 26, 183 - 209.

Wolters, C.A. and Rosenthal, H. (2000). The relation between students' motivational beliefs and their use of motivational regulation strategies. *International Journal of Educational Research,* 33, 801 - 820.

Wolters, C.A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist,* 38(4), 189 - 205.

Wolters, C.A., Benzon, M.B. and Arroyo-Giner, C. (2011). Assessing strategies for the self-regulation of motivation. In B.J. Zimmerman and D.H. Schunk (Eds), *Handbook of Self-Regulation of Learning and Performance.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Wong, B.Y.L., Kuperis, S., Jamieson, D., Keller, L. and Cull-hewitt, R. (2002). Effects of guided journal writing on students' story understanding. *Journal of Educational Research*, 95, 179 - 191.

Xu, Y., Iran-Nejad, A. and Thoma, S.J. (2007). Administering Defining Issues Test Online: Do response modes matter? *Journal of Interactive Online Learning,* 6 (1), 10 - 27.

Yang, Y-T.C., Newby, T.J. and Bill, R.L. (2005). Using Socratic questioning to promote critical thinking skills through asynchronous discussion forums in distance learning environments. *The American Journal of Distance Education,* 19(3), 163 - 181.

Yip, M.C.W. (2013). Learning strategies and their relationships to academic performance of high school students in Hong Kong. *Educational Psychology: An International Journal of Experimental Educational Psychology,* 33 (7), 817 - 827.

Zimmerman, B.J. (1986). Becoming a self-regulated learner: which are the key sub-processes? *Contemporary Educational Psychology,* 11, 307 - 313.

Zimmerman, B.J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology,* 81, 329 - 338.

Zimmerman, B.J. (1990). Self-regulating academic learning and achievement: The emergence of a social cognitive perspective. *Educational Psychology Review,* 2, 173 - 201.

Zimmerman, B.J. and Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment*. American Educational Research Journal,* 31, 845 - 862.

Zimmerman, B.J. (1995). Self-regulation involves more than metacognition: A social cognitive perspective. *Educational Psychologist,* 30(4), 217 - 21.

Zimmerman B.J. (2000). Attaining self-regulation. A social cognitive perspective. In M. Boekaerts, P.R Pintrich and M Zeidner (Eds.), *Handbook of Self-regulation.* (pp.13 - 39). London: Academic.

Zimmerman, B.J. (2000a). Self-efficacy: an essential motive to learn. *Contemporary Educational Psychology,* 25, 82 - 91.

Zimmerman, B.J. (2001). Theories of self-regulated learning and academic achievement: an overview and analysis. In B.J. Zimmerman and D.H. Schunk (Eds.), *Self-regulated learning and academic achievement. Theory, research and practice,* pp1 - 37. New York: Springer-Verlag.

Zimmerman, B.J. (2002). Becoming a self-regulated learner: an overview. *Theory into Practice,* 41 (2), 64 - 72.

Zimmerman, B.J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal,* 45 (1), 166 - 83.

Zimmerman, B.J. (2009). Theories of self-regulated learning and academic achievement: an overview and analysis. In Zimmerman and Schunk (Eds.) *Self-regulated Learning and Academic Achievement (2nd Edition),* London: Routledge.

Zimmerman, B.J. (2011). Barry Zimmerman Discusses Self-Regulated Learning Processes, *Emerging Research Fronts Commentary,* December 2011. Retrieved on 5.6.13 from: http://archive.sciencewatch.com/dr/erf/2011/11decerf/11decerfZimm/.

Zimmerman, B.J. (2011a). Motivational sources and outcomes of self-regulated learning and performance. In B.J. Zimmerman and D.H. Schunk (Eds), *Handbook of Self-Regulation of Learning and Performance.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Zimmerman, B.J., Bonner, S. and Kovach, R. (2002). *Developing self-regulated learners: Beyond achievement to self-efficacy.* Washington, DC: American Psychological Association.

Zimmerman, B.J. and Cleary, T.J. (2006). Adolescents' development of personal agency: The role of self-efficacy beliefs and self-regulatory skill. In F. Pajares and T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 45 - 69). Greenwich, CT: Information Age.

Zimmerman, B.J., Greenberg, D. and Weinstein, C.E. (1994). Self-Regulating Academic Study Time: A Strategy Approach. In D.H. Schunk and B.J. Zimmerman (eds.), *Self-Regulation of Learning and Performance: Issues and Educational Applications.* Hillsdale, N.J.: Erlbaum, 1994.

Zimmerman, B.J. and Kitsantas, A. (2005). Homework practices and academic achievement: the mediating role of self-efficacy and perceived responsibility beliefs. *Contemporary Educational Psychology,* 30, 397 - 417.

Zimmerman B.J. and Kitsantas, A. (2007). Reliability and validity of Efficacy for learning Form (SELF) scores of college students. *Journal of Psychology,* 215, 157 - 163.

Zimmerman, B.J. and Kitsantas, A. (2014). Comparing students' self-discipline and self-regulation measures and their prediction of academic achievement. *Contemporary Educational Psychology*, 39,145–155.

Zimmerman, B.J. and Martinez-Pons, M. (1990). Student differences in self-regulated learning. *Journal of Educational Psychology,* 82, 51 - 59.

Zimmerman, B.J. and Paulsen, A.S. (1995). Self-monitoring during collegiate studying: an invaluable tool for academic self-regulation. *New Directions for Teaching and Learning,* 63, 13 - 27.

Zimmerman, B.J. and Schunk, D.H. (Eds.) (2001). *Self-regulated learning and academic achievement. Theory, research and practice.* New York: Springer-Verlag.

Zion, M., Michalsky, T. and Mevarech, Z. (2005). The effects of metacognitive instruction embedded within an asynchronous learning network on scientific inquiry skills. *International Journal of Science Education,* 27 (8), 957 - 983.

Zohar, A., Weinberger, Y. and Tamir, P. (1994). The effect of the biology critical thinking project on the development of critical thinking. *Journal of Research in Science Teaching,* 31 (2), 183 - 196.

Zuffianò, A., Alessandri, G., Gerbino, G., Kanacri, B.P.L., Di Giunta, L., Milioni, M. and Caprara, G.V. (2013). Academic achievement: The unique contribution of self-efficacy beliefs in self-regulated learning beyond intelligence, personality traits, and self-esteem. *Learning and Individual Differences,* 23, 158 - 162.

9 Appendices

9.1 Appendix 1: Scales

Table 9-1 Questionnaire items

Questionnaire item	Aspect of metacognition	When
When I study I start by figuring out exactly what I need to learn.	Metacognitive knowledge (control strategies)	Pre-test and post-test
When I study I force myself to check to see if I remember what I have learned.	Metacognitive knowledge (control strategies)	Pre-test and post-test
When I study I make sure that I remember the most important things.	Metacognitive knowledge (control strategies)	Pre-test and post-test
When I study I try to figure out which concepts I still haven't understood.	Metacognitive knowledge (control strategies)	Pre-test and post-test
When I study I try to relate new material to things I have learned in other subjects.	Metacognitive knowledge (Elaboration strategies)	Pre-test and post-test
When I study I try to understand the material better by relating it to things I already know.	Metacognitive knowledge (Elaboration strategies)	Pre-test and post-test
When I study I figure out how the material fits in with what I have already learned.	Metacognitive knowledge (Elaboration strategies)	Pre-test and post-test
When I study I figure out how the information might be useful in the real world.	Metacognitive knowledge (Elaboration strategies)	Pre-test and post-test
I am confident about asking other students to explain their ideas.	Efficacy for learning	Pre-test and post-test
I am confident in my capacity to work independently.	Efficacy for learning	Pre-test and post-test
I am confident that my hard work will lead to academic success.	Efficacy for learning	Pre-test and post-test
I am confident that I can learn from my mistakes.	Efficacy for learning	Pre-test and post-test
I am confident about communicating my ideas.	Efficacy for learning	Pre-test and post-test
If at first I don't succeed, I try again.	Efficacy for learning	Pre-test and post-test
When I start something, I try to finish it.	Efficacy for learning	Pre-test and post-test
I am looking forward to the challenge of this course.	Motivational beliefs	Pre-test and post-test
While doing this course, I will enjoy learning how to improve my thinking skills.	Motivational beliefs	Pre-test and post-test

After having read the course guidelines, the course seems to be very interesting to me.	Motivational beliefs	Pre-test and post-test
I am eager to see how I will perform on this course.	Motivational beliefs	Pre-test and post-test
I'm really going to try as hard as I can on this course.	Motivational beliefs	Pre-test and post-test
For activities like this course, I don't need a reward; they are lots of fun anyway.	Motivational beliefs	Pre-test and post-test
If I complete this course, I will feel proud of myself.	Motivational beliefs	Pre-test and post-test
When I think about this course, I feel excited.	Motivational beliefs	Pre-test and post-test
I will be happy to give up my free time to work on this course.	Motivational beliefs	Pre-test and post-test
I was able to relate this week's topics to other areas of my interest.	Metacognitive skills (elaboration strategies)	End of each week
I tried to consider several ways of looking at this week's dilemma before forming my argument.	Metacognitive skills (elaboration strategies)	End of each week
I was able to put aside all of my negative thoughts and/or worries this week while I was studying.	Metacognitive skills (inhibitory control)	End of each week
I coped well with distraction this week.	Metacognitive skills (inhibitory control)	End of each week
I am satisfied with my learning progress today.	Metacognitive Experiences	End of each week
I was happy to give up my free time for this course this week.	Metacognitive Experiences	End of each week
I enjoyed the challenge of intellectual activity this week.	Metacognitive Experiences	End of each week
I felt that I made a valuable contribution to this week's activities.	Metacognitive Experiences	End of each week
I feel confident this week.	Metacognitive Experiences	End of each week

9.2 Appendix 2: Learning journal

The live version of this journal can be found at: http://bit.ly/1hivddF

The Learning Journal
This journal is part of your coursework. It is a self-reflection exercise which will help you to develop effective learning strategies throughout this course and beyond. It will be presented to you at the end of each week of the course. It is not a test and you are not graded on it. Answering honestly will help you to think about the way you learn.
*Required

What is your name? * ☐

Please consider the following statements and decide how well they describe your learning experience TODAY.

3. I was able to relate this week's topics to other areas of my interest.

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.	○	○	○	○	○	○	○	○	○	○	True. This describes me very well.

5: I tried to consider several ways of looking at this week's dilemma before forming my argument. *

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.	○	○	○	○	○	○	○	○	○	○	True. This describes me very well.

8. I achieved my learning goals this week.

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.	○	○	○	○	○	○	○	○	○	○	True. This describes me very well.

4. I am satisfied with my learning progress this week.

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.	○	○	○	○	○	○	○	○	○	○	True. This describes me very well.

6. I was happy to give up my free time for this course this week.

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.	○	○	○	○	○	○	○	○	○	○	True. This describes me very well.

7. I enjoyed the challenge of intellectual activity this week.

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.	○	○	○	○	○	○	○	○	○	○	True. This describes me very well.

10. I felt that I made a valuable contribution to this week's activities.

1 2 3 4 5 6 7 8 9 10

False. Not at all like me.											True. This describes me very well.

11. I feel confident today.

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.											True. This describes me very well.

12: I was able to put aside all of my negative thoughts and/or worries today while I was studying.

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.											True. This describes me very well.

17. I was able to cope with distraction this week.

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.											True. This describes me very well.

23. I tried as hard as I could this week.

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.											True. This describes me very well.

28. I planned my week well, so that I could be free for the activities.

	1	2	3	4	5	6	7	8	9	10	
False. Not at all like me.											True. This describes me very well.

Looking at your responses above, write down ONE learning strategy that you will focus on, as your goal for next week. *

Thank you for taking the time to reflect upon your learning in this way!

Please click SUBMIT below.

9.3 Appendix 3: Funding

Benefactor	Funding	Used for:
University of Oxford, an Outreach grant	£7200	Tutor fees
Francis Terry Foundation	£1500	Hire of learning platform
Green Templeton College	£514	Schools database
Coral Milburn-Curtis (private funds)	£480	Insurances

9.4 Appendix 4: Ethical application material

University of Oxford
CENTRAL UNIVERSITY RESEARCH ETHICS COMMITTEE (CUREC)
CUREC/1A Checklist for the Social Sciences and Humanities

The University of Oxford places a high value on the knowledge, expertise, and integrity of its members and their ability to conduct research to high standards of scholarship and ethics. The research ethics clearance procedures have been established to ensure that the University is meeting its obligations as a responsible institution. They start from the presumption that all members of the University will take their responsibilities and obligations seriously and will ensure that their research on human subjects is conducted according to the established principles and good practice in their fields and in accordance, where appropriate, with legal requirements. Since the requirements of research ethics review will vary from field to field and from project to project, the University accepts that different guidelines and procedures will be appropriate. Please check the CUREC website to ensure that you have the correct form for your project.

This form does not cover research governance, satisfactory methodology, compliance with the requirements of publishers when administering their tests or questionnaires, or the health and safety of employees and students. As principal investigator, it is your responsibility to ensure that requirements in these areas are met. Please carry out a risk assessment of the project, in consultation with all researchers involved, using the checklist and CUREC's other documentation.

The use of an asterisk in this form indicates a phrase defined in the glossary. The glossary and further information on the University's research ethics procedures are available from the CUREC website:

www.admin.ox.ac.uk/curec

This form is designed largely for research that falls within the Divisions of Social Sciences and Humanities and which does not involve a high level of risk to the subjects. Elite interviews, field work and oral history are included in the CUREC process. Please take a moment to read through it and if you have any questions or doubts as to whether it is the appropriate form, please review Section A or consult the CUREC website.

Note on anonymised data and audit: you do not need to obtain ethical approval for your study if:

you are using previously collected **anonymised data** about people which neither you nor anyone else involved in your study can trace back to the individuals who provided them (e.g. census data, administrative data, secondary analysis). Please refer to the definition of *personal data in the glossary and FAQ A4 for further guidance; or
you are conducting research on behalf of or at the request of a service provider that matches the definition of *audit in the glossary.

If your research is audit or uses prior-anonymised data, please check this box: ☐

You do not need to seek ethical approval from CUREC, and you do not need to complete any more of this form. However, please check with your department's requirements, as some departments require you to lodge this form with them.

Office use only: IDREC Ref. No. ___________________
Date of confirmation that checklist accepted on behalf of IDREC: // //

SECTION A	**Yes**	**No**
1) Are you using research methodologies commonly used in biomedical or behavioural laboratory sciences?		✔
2) Is there a significant risk that the research will induce anxiety, stress or other harmful psychological states in participants that might persist beyond the duration of any test or interview in which they are participating?		✔
3) Will the research involve human participants recruited by means of their status as present or past NHS patients or their relatives or carers?		✔
4) Does the research involve *human participants aged 16 and over who do not have *capacity to consent for themselves? See the Mental Capacity Act 2005		✔
5) Is the study to be funded by the US National Institutes of Health or another US federal funding agency?		✔

If you answered 'yes', please **stop** work on this checklist and
for questions 1 and 2, complete CUREC/1 instead (available from www.admin.ox.ac.uk/curec/);

for questions 3 and 4, submit your proposal to the appropriate NHS ethics committee (see www.nres.npsa.nhs.uk and www.admin.ox.ac.uk/researchsupport/ctrg for further information);
for question 5, or **if you answered 'yes' to questions 1, 2 or 4 and your research will take place outside the EU and is a biomedical study**, submit your proposal to OXTREC, which uses separate documentation. **Applications to OXTREC using this form will not be accepted.** If your research is not a biomedical study and does not have US funding, but will take place outside the EU, you may use this form to submit your application for approval to the Social Sciences and Humanities IDREC.

If you have answered 'no' to all questions in Section A, please complete Sections B-E. This form and any supporting materials should be typewritten.

SECTION B

*Principal investigator/ supervisor/student researcher (title and name):	Mrs Coral Milburn-Curtis
Name of supervisor (STUDENT RESEARCH PROJECTS ONLY):	Dr Rebecca Eynon
Degree programme, e.g. DPhil, MPhil, MSc (STUDENT RESEARCH PROJECTS ONLY):	DPhil
Department or institute:	Department of Education, Norham Gardens, Oxford OX2 6PY
Address for correspondence (if different):	
Email and phone contact:	coral.milburn-curtis@education.ox.ac.uk; 07875471128
Title of research project:	Learning how to learn: A study of the impact of a thinking skills intervention on the self-regulated learning dispositions and learning outcomes of young students in the UK.

Brief description of research methods and goals plus description of the nature of participants (including the criteria for inclusion/exclusion, method of recruitment), explanation of how professional guidelines and/or CUREC protocol(s) will be applied (if relevant) and expected use to which the results/data will be put. Please describe how you will obtain informed consent. Approx 400 words.

This application refers to an educational intervention which forms one element of a DPhil research project exploring the learning strategies employed by young adolescents in the UK. The principal activity for which CUREC approval is sought is based around 11 - 12 year old students having access to an online critical thinking course (i.e., virtually, not in person). This will take place over three months, and is based upon a set of learning activities which are informed by research on self-regulated learning (e.g., Zimmerman, 2008; Boekaerts and Corno, 2005). Activities include tutor-led discussion, collaborative research, the maintenance of an online learning log and interactive learning quizzes.

Students will be surveyed before and after the intervention, using a self-report survey instrument developed for PISA, the *Programme for International Student Assessment* (Marsh *et al.*, 2006a). This measures elements of self-regulated learning. The data, which is captured from a GoogleDocs survey and then securely stored on a single stand-alone computer, will be backed up on an external hard-drive for the duration of the project, and will be used solely for the purposes of this research. Two types of data will be collected: a) Pre-test and post-test survey data and b) records of students' contributions to the course. Data will be anonymised before analysis. Survey data analysis will include a comparison of means (ANOVA), correlational analysis, latent change score modelling, regression and structural equation modelling, in an attempt to explore the nature of the construct and measure the impact of the intervention on participants' self-regulated skills and critical thinking skills.

The researcher will recognise the contribution of all stakeholders in the research study, maintaining honesty, academic integrity and respect for all participants. The researcher's responsibility will include compliance with Oxford University ethics requirements, with particular reference to Protocol P.2.1.2. The project will also adhere closely to BERA and ESRC guidelines. Identities will be protected and confidentiality observed. Researcher—Participant relationships will be handled with care and participants will be informed of the objectives of the research. Procedures for collecting and evaluating data will be explained, and assurances given about anonymity, security, storage of data and publication. It is important to note that Child Protection measures will be observed at all times, with particular reference to the 'ThinkUKnow' (Child Exploitation and Online Protection (CEOP) Centre: www.thinkuknow.co.uk) guidelines for safe internet use. These guidelines will be displayed prominently at all times, to all participants. Data will be coded and anonymised on collection, stored securely on the researcher's computer and backed up on an external hard-drive.

References:

Boekaerts, M. and Corno, L. (2005) Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, 54, 267 - 99.

Marsh, H.W., Hau, K-T, Artelt, C., Baumert, J. and Peschar, J.L. (2006) OECD's Brief Self-Report Measure of Educational Psychology's Most Useful Affective Constructs: Cross-Cultural, Psychometric Comparisons Across 25 Countries, *International Journal of Testing*, 6:4, 311 -360.

Zimmerman, B. (2008) Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166 -83.

List actual or probable location(s) where project will be conducted, if known:	Online, in a virtual classroom
Anticipated duration of project:	__3__ months
Anticipated start date:	1 / 1 / 13
Anticipated end date:	1 / 9 / 14 (to include a possible re-visit to the field, if necessary)
Name and status (e.g. 3rd year undergraduate; post-doctoral research assistant) of others taking part in the project:	
Please indicate what training on research ethics the researchers involved with this study have received, e.g. the title of the online or in-person course, and date completed (online training available at ***www.admin.ox.ac.uk/researchsupport/integrity***/):	Ethics seminars at the Department of Education

SECTION C
Methods to be used in the study (**tick** as many as apply: this information will help the committee understand the nature of your research and may be used for audit).

	Please tick
Interview	
Questionnaire (survey)	✔
Analysis of existing records	
Participant performs verbal/paper and pencil/computer based task (course assignments)	✔
Measurement/recording of motor behaviour	
Audio recording of participant	
Video recording or photography of participant	
Physiological recording from participant	
Participant observation	
Covert observation	
Systematic observation	
Observation of specific organisational practices (students' use of their learning log)	✔
Other (please specify)	

SECTION D
Have you read one or more of the following professional guidelines and do you undertake to use the principles listed there as a guide for your own work? Please note that this is not intended to be an exhaustive list. Links to the guidelines listed below are included on the CUREC website.

	Please tick
British Society of Criminology: Code of Ethics for Researchers in the Field of Criminology [***www.britsoccrim.org/codeofethics.htm***]	
British Educational Research Association Ethical Guidelines for Educational Research [***www.bera.ac.uk/guidelines***]	✔
Academy of Management's Code of Ethics [***www.aomonline.org/aom.asp?ID=andpage_ID=242***]	
Association of American Geographers Statement on Professional Ethics [***www.aag.org/cs/resolutions/ethics***]	
Oral History Society of the UK Ethical Guidelines [***www.oralhistory.org.uk/ethics/index.php***]	
American Political Science Association (APSA) Guide to Professional Ethics in Political Science (Section H) [***www.apsanet.org/content_9350.cfm***]	
Political Studies Association Guide to Good Professional Conduct (see section on "Research") [***www.psa.ac.uk/Pubs***]	
British Psychological Society Code of Ethics and Conduct [***www.bps.org.uk/what-we-do/ethics-standards/ethics-standards***]	
Ethics Guidelines of the Association of Social Anthropologists of the UK and Commonwealth [***www.theasa.org/ethics/guidelines.shtml***]	
Social Research Association: Ethical Guidelines [***www.the-sra.org.uk/guidelines.htm***]	
Statement of Principles of Ethical Research Practice from the Socio-Legal Studies Association [***www.slsa.ac.uk/content/view/247/270/***]	
Statement of Ethical Practice for the British Sociological Association [***www.britsoc.co.uk/about/equality/statement-of-ethical-practice.aspx***]	
Other professional guidelines (please specify): Association of Internet Researchers (AoIR http://aoir.org/documents/ethics-guide/	✔

SECTION E

Please put a tick in the yes/no column as appropriate to indicate your response.

	Yes	No
1) Will you obtain informed consent according to good practice in your discipline before participation?	✔	
2) Will you ensure that *personal data collected directly from participants or via a *third party is held and processed in accordance with the provisions of the Data Protection Act?	✔	
3) Does the research involve as participants *people whose ability to give free and informed consent is in question? (This includes those under 18 and vulnerable adults.)	✔	
4) As a consequence of taking part in the research, will participants be at serious risk of rendering themselves liable to criminal prosecution (e.g. by providing information on drug abuse or child abuse)?		✔
5) Does the research involve the *deception of participants, as part of the investigation/experiment?		✔
If any of your answers above are in a shaded box, please indicate whether those aspects of your project are fully covered by the following.		
6) Research protocol(s) which has/ve received IDREC/CUREC approval? If yes, please give protocol number(s): Protocol P 2.1.2	✔	
7) Professional guidelines that you will be following, as noted under Section D?	✔	

If any of your answers in Section E are in a shaded box and are not covered by a protocol or by professional guidelines, please complete CUREC/2, available to download from the CUREC website. Then submit both this form (you need not complete section F) and the CUREC/2 to the Social Sciences and Humanities IDREC.

If all your answers in Section E are in the unshaded boxes or your answers in shaded boxes are covered by a protocol or professional guidelines, complete Section F and submit this form and any accompanying documents to the Social Sciences and Humanities IDREC or to the relevant officer/committee at departmental level (see notes and address below).

FINAL CHECK

Please check each of the following before submitting the checklist. Failure to provide this information could delay the start of your research.

Have you completed Sections A-E? ✔

Have you defined all technical terms and abbreviations used? ✔

If you have produced any documentation in support of your application (which might include questionnaires, participant information, consent forms/form or note of procedure for recording oral consent, advertisements and surveys), have you attached a copy of these? ✔

Are all pages (including appendices and attachments) numbered? ✔

SECTION F

You can submit this checklist by email and/or as a signed hard copy; if it is being sent by email only, the checklist, and any email from the head of department or nominee separately endorsing its submission, must be sent from a University of Oxford email address (i.e. as a minimum, the checklist and supporting documents must be submitted by the head of department or nominee indicating his/her approval from a University of Oxford email account). **Complete this section only if you do not need to submit form CUREC/2.**

I understand my responsibilities as principal researcher/supervisor/student researcher as outlined in the CUREC glossary and guidance on the CUREC website.

I declare that the answers above accurately describe my research as presently designed and that I will submit a new checklist should the design of my research change in a way which would alter any of the above responses so as to require completion of CUREC/2 (involving full scrutiny by an IDREC). I will inform the relevant IDREC if I cease to be the principal researcher on this project and supply the name and contact details of my successor if appropriate.

Signed by principal researcher/supervisor/student researcher:……………………
………………

Date:7.6.12
Print name (block capitals)……CORAL MILBURN-CURTIS……………………………………

Signed by supervisor:……………………………………………………………(for student projects)

Date:7.6.12
Print name (block capitals)…………DR REBECCA EYNON……………………………………

I understand the questions and answers that have been entered above describing the research, and I will ensure that my practice in this research complies with these answers, subject to any modifications made by the principal researcher properly authorised by the CUREC system.

Signed by associate/other researcher:
………………………………………………………………….
Print name (block capitals)…………………………………………………………………………
Date ………………

I have read the research project application named above. On the basis of the information available to me, I:

(i) consider the principal researcher/supervisor/student researcher to be aware of her/his ethical responsibilities in regard to this research;

(ii) consider that any ethical issues raised have been satisfactorily resolved or are covered by relevant professional guidelines and/or CUREC approved protocols, and that it is appropriate for the research to proceed without further formal ethical scrutiny at this stage (noting the principal researcher's obligation to report should the design of the research change in a way which would alter any of the above responses so as to require completion of a CUREC/2 full application);

(iii) am satisfied that the proposed project has been/will be subject to appropriate *peer review and is likely to contribute something useful to existing knowledge and/or to the education and training of the researcher(s) and that it is in the *public interest.

(iv) [FOR DEPARTMENTS/FACULTIES WITH A DEPARTMENTAL RESEARCH ETHICS COMMITTEE (DREC) OR EQUIVALENT BODY - PLEASE DELETE IF NOT APPLICABLE] confirm that this checklist (and associated research outline) has been reviewed by the Department's Research Ethics Committee (DREC)/equivalent body, and attach the associated report from that body.

Signed:………………………………………..
(Head of department or nominee e.g Chair of DREC, Director of Graduate Studies for postgraduate student projects)

Print name (block capitals)……………………………………………………………………………

Date:……………………

If your research involves participants recruited by means of their status as current or former NHS staff, or the research will, in whole or in part, be carried out on NHS premises, use NHS facilities or assess NHS facilities or services, please see FAQ B3 (www.admin.ox.ac.uk/curec/faqs/).

Please check with your department about its procedures for the approval of CUREC forms. If your department indicates that the checklist should be submitted directly to the IDREC, please send it, together with any supporting documentation, to the following address(es), keeping a copy for yourself:

Secretary of the Social Sciences and Humanities IDREC Email: ethics@socsci.ox.ac.uk
University of Oxford Social Sciences Division
Hayes House, George Street
Oxford, OX1 2BQ
IDRECs and/or CUREC will review a sample of completed checklists and may ask for further details of any project.

Letter to headteacher, following initial contact with school via telephone

Letterhead
Date

Dear Headteacher,

This is to confirm the invitation for some of your students to participate in a doctoral research project during 2013. The study explores the nature of self-regulated learning and the potential for an online learning programme to develop the learning strategies of young adolescents. The focus is on students in Year 7, who have gained Level 5 Literacy already, but who are not currently working beyond Level 6. The project can accommodate one **boy and one girl** from any one school and you are asked to select students of similar ability who would benefit from an intellectual challenge.

Your students are invited to participate in:

- an online course which is designed to develop the critical thinking skills of argument, reasoning, communication and debate
- a before-and-after survey which evaluates their self-regulated learning skills. Students will be allowed to complete the course whether or not they complete the surveys.

The benefit to students: development of metacognitive and critical thinking skills; an opportunity to engage in an activity which presents significant intellectual challenge.

The benefits to the school: provision of a free, additional extension project for students who have the potential to be high-performing; development of the critical thinking skills of a cohort of Year 7 students.

The benefits to the researcher: access to statistical data (survey results) which will assist in an exploration of the development of self-regulated learning in young adolescents.

What students will do
Students will enrol on the internet-based course and will complete a short survey, which will provide a benchmark for the analysis of their progress. The critical thinking programme, '*Is this the right room for an argument?* is accessed through secure username and password and students will make use of course materials at different times over a period of one month. At the end of the course, they will complete the same survey as they did at the beginning.

The theoretical components of the course include the development of:
- self-belief and motivation
- learning strategies (time management, goal setting etc.)
- metacognition (understanding and regulation of the learning process)

The learning activities include:
- critical thinking activities (the development of argument, reasoning and logic)
- problem-based collaborative research projects
- debate (ethical dilemmas)
- learning logs

Students will access the course during the Spring Term 2013, either mid January to mid-February or mid February to mid-March, depending upon the group to which they are allocated. Survey data will be anonymised before analysis and will be stored securely on a stand-alone computer. The research will form part of my doctorate study. Data will be analysed in 2014 and the study is likely to be published as a thesis in 2015.

How to get involved

There is an information sheet for parents enclosed. If you would like your students to participate, please could you distribute the information sheet to two students in Year 7, one boy and one girl, who have achieved Level 5 Literacy, (but who are not yet working beyond Level 6).

Thank you for considering this. I do hope that I can make my contribution to your extension programme and that you will be able to assist me in my research.

Very best wishes,

Coral

Coral Milburn-Curtis
Green Templeton College, University of Oxford, Woodstock Road, Oxford OX2 6HG
Department of Education, University of Oxford, Norham Gardens, Oxford OX2 6PY
Tel: 07875 471128; email: coral.milburn-curtis@education.ox.ac.uk

Letter to Parents

Letter head
Date

Learning how to learn: an internet-based project
Invitation for your child to participate

Dear Parent,

Your child's school has agreed to take part in a research project exploring the learning strategies of young students who are studying online. Before you decide to let your child participate, it is important to understand why the research is being conducted and what your child's participation would involve. Please take time to read this information carefully. Please ask if there is anything about the project that is unclear, or if you would like more information. Take time to decide whether or not you would like your child to take part in this research.

What is the research project trying to find out?

This research is looking into the learning strategies that young people use when they study. This may help us to improve learning strategies of all children.

Why is your child being invited to take part?

I am inviting your child to participate in this research because they have been identified as a student who might enjoy an intellectual challenge.

What will happen if your child takes part?

If your child takes part, they will:

- participate in an internet-based course which is designed to develop thinking skills and learning strategies. If your child participates, they will access an online learning platform via a secure username and password. There, they will participate in a variety of learning activities designed to improve their study skills, at different times over a period of one month. The activities include tutor-led debate, collaborative research, learning quizzes, short assignments and a learning journal. Your child would need to commit up to two hours per week to the coursework, and they may spend longer than this if they wish. The course may help them in their school work.
- complete a before-and-after survey about their course experience and their use of learning strategies.

Your child will be allowed to complete the internet course whether they complete the surveys or not. If they do complete the whole programme, they will be able to download a Record of Achievement.

Because students will be engaging in discussion online, strict Child Protection procedures will be in place. For example the programme will be monitored by moderators, and behaviour and online safety guidelines will be strictly enforced.

Does your child have to take part?

No, your child's participation is entirely voluntary. In addition, you may take your child out of the project at any time, without needing to give a reason.

What will happen to the information from the project?

I will study the contributions your child makes to the course, as well as the answers they give to the survey. Results for each child will be kept strictly confidential. During the project, only I will have access to the information, and it will be kept in a password-protected file on a password-protected computer, backed up on a password-protected hard drive. I will anonymise the information before I analyse it. At the end of the project, the information will be archived.

This study forms part of my doctoral research. It has been approved by the University of Oxford's ethics committee. The study will be written up as a thesis over the next two or three years. On successful submission of the thesis, it will deposited both in print and online in the University archives, to make it available for future Oxford University researchers only. Parts of the research may also be published in academic journals and presented at conferences. The thesis and other publications may contain quotes from some of the participating children, but the children will not be identified. Your child's name will never be mentioned in the thesis or any other publications or presentations, and neither will the name of their school.

What next?

I do hope that you will want to get involved. If you would like your child to participate, please sign the consent form, and return it to: <name of school coordinator>

If you want ask any questions or make a complaint, please contact me or my supervisor at the address at the bottom of this page.

Very best wishes,

Coral Milburn-Curtis
Email: coral.milburn-curtis@education.ox.ac.uk
Telephone: 07875 471128
Address: Green Templeton College, University of Oxford, Woodstock Road, Oxford OX2 6HG

Supervisor: Dr Rebecca Eynon
Email: rebecca.eynon@oii.ox.ac.uk
Address: Department of Education, Norham Gardens, Oxford OX2 6PY

Letter head

Consent Form

Title of Project: 'Learning how to learn: A study of the impact of a thinking skills intervention on the self-regulated learning dispositions and learning outcomes of young students in the UK'.

Name, position and contact address of Researcher: Coral Milburn-Curtis, DPhil candidate, Green Templeton College, University of Oxford, Woodstock Road, Oxford, OX2 6HG.

Supervisor: Dr Rebecca Eynon, Department of Education, Norham Gardens, Oxford OX2 6PY.

			Please initial box
I confirm that I have read and understood the information sheet for the above study and have had the opportunity to ask questions.			
I understand that my child's participation is voluntary and that he/she is free to withdraw at any time, without giving reason.			
I consent to my child taking part in the above study.			
I agree to the use of anonymised quotes from my child's online participation being published in this study.			
I understand that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee.			
I understand that no personal data will be used in this study; that participation data in the form of my child's anonymised course transcript will be stored online for the purposes of this study only.			
I understand that the research will be written up as a student (DPhil) thesis and that personal data will not be included in that thesis.			
Name of Participant	Date:	Signature:	
Name of Parent	Date:	Signature:	
Parent's contact email address:			
Name of Researcher: Coral Milburn-Curtis	Date:	Signature:	

Please return this form to: <name of school coordinator>

9.5 Appendix 5: Unique random numbers up to 1000

0996	0288	0145	0840	0491	0537	0932	0481	0569	0836	0769	0438	0697	0186	0188	0712	0876
0117	0658	0036	0637	0100	0445	0943	0449	0665	0547	0900	0577	0015	0701	0908	0021	0947
0171	0498	0782	0929	0957	0139	0594	0494	0060	0096	0722	0844	0214	0370	0167	0776	0684
0694	0662	0758	0470	0601	0740	0057	0573	0558	0868	0673	0726	0566	0047	0605	0196	0156
0808	0587	0983	0797	0619	0887	0085	0754	0748	0502	0505	0028	0192	0434	0975	0352	0953
0417	0761	0993	0765	0716	0598	0951	0893	0331	0017	0224	0338	0263	0487	0814	0833	0246
0274	0455	0910	0545	0111	0412	0039	0160	0530	0686	0483	0092	1000	0011	0979	0075	0786
0652	0790	0374	0889	0609	0919	0989	0043	0882	0363	0921	0513	0207	0124	0904	0299	0113
0936	0203	0402	0071	0064	0818	0555	0079	0509	0750	0025	0669	0004	0733	0812	0310	0081
0032	0915	0267	0210	0647	0068	0541	0654	0579	0804	0865	0149	0562	0590	0772	0961	0861
0427	0729	0089	0477	0846	0737	0534	0143	0316	0327	0295	0391	0103	0968	0107	0690	0940
0925	0235	0306	0359	0199	0680	0972	0829	0523	0175	0220	0615	0164	0252	0519	0718	0121
0380	0135	0872	0395	0825	0801	0342	0985	0320	0049	0128	0626	0132	0348	0231	0583	0526
0964	0384	0857	0705	0630	0854	0182	0466	0878	0641	0822	0278	0177	0744	0780	0406	0793
0897	0053	0850	0459	0633	0643	0611	0708	0154	0284	0423	0007	0256	0242	0551	0622	0675
0515	0435	0642	0756	0681	0905	0232	0517	0664	0692	0873	0328	0228	0794	0831	0457	0578
0948	0104	0901	0510	0418	0429	0397	0493	0204	0335	0474	0792	0307	0292	0602	0407	0461
0300	0781	0339	0931	0891	0542	0322	0717	0531	0354	0621	0820	0489	0482	0236	0239	0762
0927	0168	0709	0087	0688	0151	0495	0728	0499	0450	0333	0685	0628	0065	0752	0959	0072
0997	0222	0549	0568	0980	0008	0190	0645	0279	0845	0147	0773	0895	0264	0421	0218	0826
0734	0745	0713	0809	0521	0386	0525	0108	0624	0343	0653	0724	0777	0617	0098	0656	0247
0941	0859	0638	0033	0848	0670	0937	0136	0805	0798	0553	0290	0813	0243	0485	0760	0403
0738	0467	0546	0044	0816	0766	0649	0001	0944	0382	0803	0275	0389	0314	0538	0600	0884
0296	0324	0506	0696	0596	0162	0463	0824	0211	0581	0471	0268	0877	0051	0061	0029	0126
0837	0702	0841	0425	0674	0660	0969	0040	0093	0933	0414	0706	0563	0258	0909	0955	0350
0899	0987	0254	0453	0856	0115	0869	0606	0130	0559	0535	0076	0720	0055	0784	0863	0361
0867	0083	0965	0318	0260	0698	0119	0591	0439	0365	0589	0916	0200	0613	0375	0557	0012
0912	0478	0514	0140	0527	0632	0788	0585	0194	0367	0378	0346	0442	0888	0019	0158	0741
0991	0976	0286	0356	0410	0250	0730	0023	0880	0574	0226	0271	0666	0215	0303	0570	0503
0172	0431	0920	0923	0446	0610	0852	0393	0770	0371	0835	0179	0677	0183	0399	0282	0634
0311	0749	0332	0456	0069	0926	0620	0272	0317	0261	0349	0616	0966	0492	0898	0881	0225
0723	0229	0357	0795	0802	0727	0710	0994	0150	0556	0464	0475	0443	0539	0381	0520	0838
0353	0648	0507	0827	0385	0977	0588	0368	0763	0400	0667	0866	0528	0285	0973	0755	0133
0197	0774	0496	0379	0731	0112	0005	0118	0595	0614	0026	0054	0691	0325	0193	0819	0791
0759	0855	0567	0432	0571	0699	0823	0663	0144	0293	0894	0851	0599	0336	0289	0806	0592
0090	0862	0695	0990	0428	0849	0321	0360	0584	0646	0930	0552	0742	0208	0870	0257	0627
0097	0883	0086	0460	0304	0396	0945	0902	0161	0176	0122	0101	0830	0913	0129	0364	0165
0411	0635	0962	0659	0603	0058	0958	0524	0560	0678	0834	0631	0240	0413	0424	0392	0488
0934	0787	0037	0022	0810	0906	0618	0205	0440	0821	0874	0714	0344	0038	0735	0679	0767
0034	0233	0636	0650	0387	0340	0500	0564	0141	0746	0041	0479	0372	0369	0981	0337	0308
0974	0148	0668	0799	0938	0771	0066	0137	0030	0511	0355	0006	0447	0995	0084	0351	0212
0700	0703	0173	0152	0963	0180	0062	0415	0216	0536	0013	0297	0444	0472	0109	0009	0575
0237	0885	0682	0198	0209	0273	0116	0572	0088	0073	0383	0241	0120	0711	0671	0323	0102
0497	0312	0134	0401	0269	0262	0543	0707	0949	0468	0276	0508	0280	0230	0465	0408	0532
0739	0853	0778	0002	0329	0970	0045	0201	0998	0607	0301	0166	0305	0404	0433	0504	0436
0639	0419	0451	0917	0586	0070	0265	0540	0184	0248	0056	0169	0094	0077	0105	0476	0376
0942	0244	0604	0842	0294	0059	0027	0422	0390	0123	0454	0251	0486	0952	0091	0187	0326
0219	0155	0984	0358	0042	0334	0191	0886	0978	0815	0484	0743	0234	0922	0163	0704	0082
0683	0146	0593	0946	0623	0747	0954	0067	0217	0544	0828	0003	0185	0640	0106	0142	0768
0890	0416	0213	0516	0914	0719	0612	0651	0202	0843	0131	0800	0548	0074	0238	0480	0398
0999	0807	0811	0644	0939	0377	0063	0270	0309	0533	0860	0879	0501	0956	0157	0458	0206
0576	0732	0529	0138	0046	0832	0420	0935	0655	0035	0928	0409	0967	0253	0170	0950	0345
0159	0982	0249	0448	0110	0864	0125	0796	0715	0050	0779	0858	0127	0078	0313	0693	0114
0625	0911	0608	0907	0473	0775	0892	0783	0580	0189	0362	0373	0341	0437	0014	0153	0736
0986	0971	0281	0405	0245	0018	0875	0221	0266	0661	0298	0565	0764	0426	0181	0918	0441
0871	0847	0388	0031	0366	0095	0174	0672	0178	0394	0277	0629	0010	0430	0903	0751	0676
0512	0924	0687	0452	0839	0099	0896	0689	0657	0330	0469	0302	0597	0721	0561	0462	0347
0817	0195	0259	0753	0255	0725	0757	0052	0490	0992	0319	0522	0223	0227	0024	0020	0550
0518	0988	0554	0960	0287	0785	0291	0789	0582	0048	0080	0283	0016	0315			

9.6 Appendix 6: Critical thinking assessment

At the end of each week, students were presented with a scenario and then asked to consider ten statements, presented at random:

Example Scenario: Imagine that you are about to vote for your School Captain. Which of these issues are relevant to your decision about whom to vote for?

1.Your grades have improved a lot over the past year.
Choose one answer.
This is VERY relevant to the issue.
This has SOME relevance to the issue.
This is NOT relevant to the issue.
(Answer: NOT relevant: 2 marks)

2. One candidate has the best ideas for developing the school council. (Answers: VERY relevant: 2 marks; of SOME relevance: 1 mark)

3. One of the candidates is taller than the other. (Answer: NOT relevant: 2 marks)

4. One candidate has won the Bronze Duke of Edinburgh's Award three years ago. (Answer: of SOME relevance: 2 marks; VERY relevant: 1 mark)

5. One of the candidates has sat on the school council for two years. (Answer: VERY relevant: 2 marks; of SOME relevance: 1 mark)

6. One of the candidates has just come top in the maths exam. (Answer: NOT relevant: 2 marks)

7. One candidate raised a lot of money for Oxfam last year. (Answer: of SOME relevance: 2 marks; VERY relevant: 1 mark)

8. One candidate is known for their strong moral character. (Answer: VERY relevant: 2 marks; of SOME relevance: 1 mark)

9. One of the candidates was found cheating in exams. (Answer: VERY relevant: 2 marks; of SOME relevance: 1 mark)

10. One of the candidates was voted 'most popular student' by their friends, four years ago. (Answer: of SOME relevance: 2 marks)

9.7 Appendix 7: Examples of tutor responses to discussions

Why do you think it's part of our animal nature to prioritize our own species above others? .
Hi Matthew, You say "I think that humans are probably more important than apes", what makes you think that? . "but that doesn't give us the right to wipe out their habitats" can you find an example of this?
Why do you think it's animal instinct to prioritize our own species over others? Is that something that we should strive to change or is it a basic law of nature and therefore justified? .
Hi Emma, Do other species of animal protect other species? Why should we do something when other species do not?
Hi Kayleigh, Do you not think we should just let nature take it's course? Like it did with wiping out the dinosaurs? Should we just leave nature to itself or do we have a moral obligation to intervene?
Hi Mohammad, But other species of animal seem to have no problem destroying each others' environment... Why should we, as humans, act differently? . Why do we have a moral obligation to help?
Hi Emily, Do you think all species of animal believe themselves to be more important than all the others, or do we - as humans - stand alone in this matter? .
Hi Adam S, How would you go about doing that? .
Hi Katerina, That's an interesting point. The question is left that way so you can interpret it yourself, but that would be an interesting thing to discuss? How would your answer change depending on the different answers to your questions? .
Hi Sophie and Anna, (And Everyone) You say "all animals are equal" I wonder if either of you have read Animal Farm by George Orwell? There is a quote from the book which goes something like this: "all animals are equal but some are more equal than others." Do you think it's ever possible to achieve true equality... And is true equality even desirable? .
Is it possible to improve our world instantly or is it a process that will take time? .
Hi Laurence, Do you think languages are constantly evolving things, and, unlike Ciphers, are you sure that you could really 'create' one overnight. What do you think?
Hi Everyone, Are we relying too much on technology? Is it slowing the evolutionary process? .
Is money really the pound coins or notes we see or is it more of a concept. Before money as we know it, things like hens and eggs were used to bargain with, but in a way, this was still money. It's the same concept. Would society be able to function without money, what do you think? . P.S: Did you know, that a £5 pound note isn't actually worth £5... It's a piece of cotton. Really it is an IOU saying that if you take the piece of cotton to the vaults then they will swap it for £5 pounds worth of a precious metal. So is money actually worth anything? What is it in money that makes the whole concept work? What do you think? .
George Orwell quite famously said "Sport is War minus the Shooting", do you agree or disagree? Why? What similarities/differences can you draw between sport and war? .
What advantages/disadvantages does reading give us, in terms of improvement in school and the way we are viewed by other people? . What are the advantages/disadvantages of 'critical reading' compared to just reading for fun?
Hi Priyanka, You say they are 'completely unalike' but what about tactics and plans, what about emotional baggage? There must be at least some similarities? Hi Everyone, If you disagree, why do you think George Orwell might have said this? .
Hi Emily, George Orwell was actually using football as an example when he made this statement. :-) Do you think something needs to change in the football community?
Those are some good similarities... What about the emotional baggage of both sport and war?

Are there any similarities there? Imagine if you scaled sport up in intensity... Would that make it more warlike? Do you think if someone had asked you this before the 2012 Olympic Games you would have answered differently? . :-)
Twilight... That raises a question. After the success of Stephanie Meyer's "Twilight" the bookshelves of bookshops started to be filled with stories about vampires and werewolves - writers churning out books to fuel popular demand. The same can often be said about series... Is writing becoming too much about money rather than about the quality? .
Is true communism ever possible to achieve? Is it possible for everyone to be equal? . Is true and complete equality even desirable?
What makes a good film? What about films of books - are these a good or bad thing, and should we read the book or watch the film first? .
Hi Scott, Did you also enjoy the Harry Potter books? Did you prefer the films?:-)
Is any obsession a good thing? .
Has anyone seen the advert, the one which says "We are the creators of normal..." What is your reaction to the advert? . https://www.youtube.com/watch?v=LbcwsSZCKuI Do you think it's possible to 'create' normal?
Hi Annabelle, But if there cannot be a 'normal human' does that mean there can't be a 'normal weather' or 'normal clothing'? Where do we draw the line? .
...Where do we draw the line when it comes to self defence? .
Are the exams we take at the moment a good reflection of what we have learnt? Do they benefit people who work best under pressure and who find it easy to sit still? What other ways could we have of measuring progress? How seriously do you take your 'levels' at school? Are you someone who compares themselves to other people or do you just strive to improve your own levels? .
Why would listening to Mozart help any more than listening to, say, the Sex Pistols? .
Shouldn't the government let people be responsible and make their own decisions... rather than taking the choice completely away from them? .
Aren't most words open to interpretation, Thomas? .
...Maybe it depends on what you're watching? .
Do games like Call of Duty set a good example for the future generation? . Or do they provoke anti-social behaviour?
Hi Matthew, Those are good points. :-) Do you think sport can sometimes take over people's lives: how far should athletes be prepared to go: how much should they sacrifice in favour of training? .
Do you think parents put too much pressure on their children, then? Has anyone heard of 'Tiger Mothers'? Maybe you could research them? :-)
What do people think about emoticons? Do you find that they help convey what you mean? Can anyone find the reason why emoticons were created?,,=^_^=,,
Great debating Everyone! :-) See you all tomorrow. Bye for now.
Can anyone find any example in the real world (Perhaps looking towards the Arab Spring) where the internet has been a force for democracy? .
Lots of things in the world are 'obvious' at a first glance, but once you look into them and start to unpack them a little more, maybe they are not as obvious as you first thought, at all. . What do you think?
I guess it depends on how you define 'World'. By world do you mean life as we know it, the Earth, the Universe, or are you talking about a more spiritual world? . In a way, the conceptual worlds inside our heads will live forever, or the ones that fantasy writers write about in books...? . Depending on you definition of world, are some worlds eternal?
But, on the other hand, how does the premier league benefit the UK? It gives us world round recognition, doesn't it? Maybe you could research about how much money the premier league

actually brings into the UK in terms of fans, merchandise, tourism...?
What does everyone think about the legacy of the London Olympic Games? Do you think the UK will be able to preserve it? "The higher you climb, the further you have to fall." Does this age old quote bear any relevance to 2012? .
Social networking and, in fact, the internet in general are actually relatively recent phenomenon. If Social Networking is such a massive part of many people's lives, today, where do you see us being in 10, 20, 30 years time? . Will we even step outside our houses?
There is an argument, that since the average living age is going up almost every year, the ages that we become legal for things should also be rising as we should all be staying in school for longer. What do you think about this idea? .
I'm from Bristol, the home town of famous graffiti artist Banksy. Here in Bristol it's pretty much everywhere, and I've grown up around it. Do you think it's an art form which is more or less respectable than a lot of other art forms out there? :-) What are the sides to the argument? .
Does it depend on what they use their skills of hacking for... If it's to hack a website as a sort of protest, is it not so bad after all - or just as bad?? .
Hi Sumayyah, So would you feel responsible for the death of the one person? If you had done nothing... . would you be responsible for the death of the five?
Hi Louise, What arguments do you think people defending the captain would use against you? .
Hi Matthew, Why do you think young children and families had the right to be on the lifeboat even if they were the weakest? .
Louise disagrees with you... She thinks the captain did the wrong thing. Maybe you could go read her reasons and then see if your view has been changed, or try to change her view? .
Hello Elliot, Do you think the morally/ethically right thing to do is always to save the most lives or are there ever any exceptions to this? . Louise disagrees with you... What do you think after reading her post? :-)
Hi Louise, Yours is a really interesting view. Are you, in a way, suggesting that the morally/ethically right thing to do is not always the thing that will save the most lives? . :-)
Hi Thomas, Does the fact that some of the people in the lifeboat disagree with the captain change things at all? .
Hi Scarlet, Do you think the captain made the mistake of just not thinking it through entirely? Do you think that people should still go to jail (even for only a short time) if their crime was a 'mistake'? .
Hi Sarah, Do you think the captain thought his actions through entirely? . Should he feel guilty for the deaths of the few whom he made go over the side despite the fact that some people of the lifeboat disagreed with him?
Hi Jade, So if someone is accidentally killed but the blame can be pinned on one person, even though it was an accident, should that one person still go to prison? . Does there always need to be someone to blame for a death?
Hi Jessy, So you think that the captain was right to try and save at least a few lives? . Does the fact that some people on the boat disagreed with his decision change anything?
Hi Jade, It's an interesting point about the way that the captain threw over the weakest and those least likely to survive on their own. I hadn't thought about that. .
Hi Louise, That's a very interesting way of looking at things. :-) But surely lives are more important than the feelings of those who are alive? .
Hi Joey, Do you think that the fact that some people on the lifeboat disagreed with the captain's decision changes this at all? .
Hi Sarah, That's a good point. . But do you think the captain should have sacrificed himself first? ...In fact, what was he even doing on the life boat - surely the captain should go down

with his ship? .
Hi John, You say that the Captain is not guilty of murder, do you think he is guilty of anything at all? Should he feel guilty? . The people's lives were his responsibility, weren't they? .
Hi Kathryn, Why do you think it was morally wrong? Is it not the moral thing to do to save the maximum number of lives? . Or is there more to it than that?
Hi Isobel, But surely we should still take some sort of responsibility for the deaths? .
Hi Kayleigh, Do you feel you are partly responsible just by being there? .
Do you think the fact that some of the others on the lifeboat disagreed with the Captains decision have any impact at all?

9.8 Appendix 8 Examples of SPSS syntax

!Little's Test for missingness

```
MVA VARIABLES=CT1 CT2 CT3 CT4 PreMK PoMK PreMot
PoMot PreEfL PoEfL
  /EM(TOLERANCE=0.001 CONVERGENCE=0.0001
ITERATIONS=25).

!Repeated measures ANOVA

SPLIT FILE OFF.
GLM CT1 CT2 CT3 CT4 BY Group
  /WSFACTOR=time 4 Repeated
  /METHOD=SSTYPE(3)
  /SAVE=ZRESID COOK
  /POSTHOC=Group(SCHEFFE)
  /PLOT=PROFILE(time*Group)
  /EMMEANS=TABLES(OVERALL)
  /EMMEANS=TABLES(time) COMPARE ADJ(BONFERRONI)
  /EMMEANS=TABLES(Group) COMPARE ADJ(BONFERRONI)
  /PRINT=DESCRIPTIVE ETASQ HOMOGENEITY
  /CRITERIA=ALPHA(.05)
  /WSDESIGN= time
  /DESIGN= Group.

!Regression

REGRESSION
  /DESCRIPTIVES MEAN STDDEV CORR SIG N
  /MISSING PAIRWISE
  /STATISTICS COEFF OUTS R ANOVA COLLIN TOL
CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT CTChange
  /METHOD=ENTER DebTotNoOutliers
  /SCATTERPLOT=(*ZPRED ,*ZRESID)
  /RESIDUALS DURBIN
  /SAVE MAHAL COOK ZRESID.
```

9.9 Appendix 9: MPlus code for latent change score model

```
TITLE: Change Model for SRL on LJ v non LJ

     DATA: File is 'MPlus/SPSSdatafile4.12.14.csv';

      VARIABLE:
NAMES = ID
!Groups
Group Exp_Cntrl DegGrps

!Critical thinking raw
CT1 CT2 CT3 CT4

!SRL raw
PreMK PoMK PreMB PoMB PreEfL PoEfL PreSRL PoSRL

!Progress in SRL
MKProg MotProg EfLProg;

!SRL Items
Mo11 Mo12 Mo13 Mo14 Mo15 Mo16 Mo17 Mo18 Mo19
E112 E113 E114 E117 E120 E121 E122
Me123 Me124 Me125 Me126 Me127 Me128 Me129 Me130
Mo21 Mo22 Mo23 Mo24 Mo25 Mo26 Mo27 Mo28 Mo29
E212 E213 E214 E217 E220 E221 E222
Me223 Me224 Me225 Me226 Me227 Me228 Me229 Me230

!Learning journal items
LJ1_3 LJ1_4 LJ1_5 LJ1_6 LJ1_7 LJ1_8 LJ1_10
LJ1_11 LJ1_12 LJ1_15 LJ1_17 LJ1_23 LJ1_28
LJ2_3 LJ2_4 LJ2_5 LJ2_6 LJ2_7 LJ2_8 LJ2_10
LJ2_11 LJ2_12 LJ2_15 LJ2_17 LJ2_23 LJ2_28
LJ3_3 LJ3_4 LJ3_5 LJ3_6 LJ3_7 LJ3_8 LJ3_10
LJ3_11 LJ3_12 LJ3_15 LJ3_17 LJ3_23 LJ3_28
LJ4_3 LJ4_4 LJ4_5 LJ4_6 LJ4_7 LJ4_8 LJ4_10
LJ4_11 LJ4_12 LJ4_15 LJ4_17 LJ4_23 LJ4_28


!Metacognitive activity
ES1 ES2 ES3 ES4
IC1 IC2 IC3 IC4
ME1 ME2 ME3 ME4
TotalCT TotalES TotalIC TotalME

!Argumentation activity
DArg1 DArg2 DArg3 DArg4
BArg1 BArg2 BArg3 BArg4
DArg
BArg
```

```
Missing are All (-9);

 Useobs Debgrps eq 1;

  USEVAR = PreMK PoMK PreMB PoMB PreEfL PoEfL PreSRL PoSRL Group;

  ANALYSIS:  ESTIMATOR = ML;

  MODEL:

!The Latent Change Score Model
 !The asterisk (*) frees the first factor loading
PreSRL BY PreMK* (a)
 PreMB (b)
 PreEfL (c);
!(a), (b) and (c) are equating the variables across time
 PoSRL BY PoMK*(a)
 PoMB (b)
 PoEfL(c);
!  The residuals of parallel items at Time 1 and 2 are allowed to be correlated across time to
control for the dependencies in the repeated measures.
PreMK WITH PoMK;
PreMB WITH PoMB;
PreEfL WITH PoEfL;
!@ fixed values. The latent regression path is fixed to 1.
Change BY PoSRL@1;
!  The variance of the change score is estimated.
Change;
![Latent mean of the change score is estimated.]
[change];
[PreSRL@0 PoSRL@0];
PreSRL@1 PoSRL@0;
PoSRL ON PreSRL@1;
PreSRL WITH Change;
[Premk Pomk](d);
[Premb Pomb](e);
[PreEfL PoEfL](f);
Group ON PreSRL;
Change ON Group;

!FROM MODINDICES:


OUTPUT:    STAND MODINDICES (ALL);
```

9.10 Appendix 10: Quality assurance of discussion text

Discussions took place each evening from 6pm to 7pm. The four most prolific discussions were sampled for quality assurance.

Great ape habitat in Africa has dramatically declined.

Great apes, such as gorillas, chimps and bonobos, are running out of places to live, say scientists. They have recorded a dramatic decline in the amount of habitat suitable for great apes, according to the first such survey across the African continent. Eastern gorillas, the largest living primate, have lost more than half their habitat since the early 1990s. Cross River gorillas, chimps and bonobos have also suffered significant losses, according to the study. Does it matter? Aren't humans more important than apes? (This discussion was initiated by a tutor).

How to improve our world?

We will discuss some important issues today and debate the morality of some difficult dilemmas. This is now our opportunity to make suggestions about how to improve the world!
What LAWS need to change to make our world a better place?
What GADGETS need to be invented to improve the lives of those who need help? (This discussion was initiated by a tutor).

Kindle or books?

What do you prefer? If you have a kindle (or something like that), like me, do you prefer it to a book? why? why not? (This discussion was initiated by a student).

Should the school leaving age be reduced?

The UK government has announced plans to raise the school leaving age. Others think that it should be reduced. What's your opinion and why? (This discussion was initiated by a tutor).
Discussions were recorded according to number of participants; number of words per discussion; number of posts; average number of words per post; average number of posts per student; % of posts which expressed opinions; % of posts which expressed empathy, agreement or acknowledgement of others' opinions; % of posts which attempted to engage others in further discussion by asking questions. Also recorded was the number of spelling mistakes per discussion; the number of posts which simply stated 'I agree' or similar statement; and the number of very poor posts (such as those which simply consisted of a 'smiley face').

Analysis

The average length of the sampled discussions was 5700 words, consisting of 131 posts, with an average of 44 words per post and 3.81 posts per student. The average number of participants engaged in any one of these discussions was 37.25.
77% of posts expressed opinions: *Humans and apes are both alive, conscious and have feelings. They are equally important; just because humans are able to build our own homes, that doesn't mean that animals who can't are less important.*
16.6% of posts asked questions: *What does "important" mean in this case? Important to whom and why?*
29.4% of posts acknowledged the views of others, by specifically mentioning names: *That is an interesting point Christopher. We would not say that we are more important than our grandparents so why say that about apes?*

Some posts included all three elements of opinion, asking questions, and acknowledging the views of others: *Hi Tom, I think that although humans are more intelligent, they are also more destructive and are more likely to destroy the environment than make it a better place. Evidence shows that the rainforests are getting smaller every year and the ice caps are melting because humans have been expanding. Why do you think we can make the world more sustainable?*
.34% of words were misspelt. Most spelling mistakes were made by the same students, who misspelled several words wrongly in one post: *decended; speices; exstinct; thier; intelligance; mabey; goverment.*
3.6% of posts simply expressed agreement: *I agree*
3.2% of posts were classified as 'poor': *very right*
.76% of posts were classified as very poor: e.g., ☺
0% of posts were written in text-speak.

Example of a discussion from the intervention

Tutor	Great ape habitat in Africa has dramatically declined Great apes, such as gorillas, chimps and bonobos, are running out of places to live, say scientists. They have recorded a dramatic decline in the amount of habitat suitable for great apes, according to the first such survey across the African continent. Eastern gorillas, the largest living primate, have lost more than half their habitat since the early 1990s. Cross River gorillas, chimps and bonobos have also suffered significant losses, according to the study. Does it matter? Aren't humans more important than apes?
Child	
1	Humans are more intelligent and are more likely to make the world a more sustainable and better place.
2	It does matter. Apes are living creatures just like humans. If we were apes and apes we knew died we would not like it. Neither would they.
3	I think every living thing is equal, no-one is more important than anyone else. Some people think we need to be more cared for than apes - do you think that is reasonable?
4	What alternative ways of looking at this are there? I don't believe humans are more important than apes because we are just another species among many others, but it just happens to be we are the more powerful species.
Tutor	Hi Tom, Isaac and Katerina, Do you think it's part of our animal nature to prioritize our own species above others?
21	I think that humans are probably more important than apes but that doesn't give us the right to wipe out their habitats.
14	Can you expand on that? Surely humans are more important than many things in the world that there are problems with? Should we abandon all of them and just think about ourselves?
15	Humans and apes are both alive, conscious and have feelings. They are equally important; just because humans are able to build our own homes, that doesn't mean that animals who can't are less important.
16	I agree. We evolved from apes so without them we might not exist. Neither species is more important in my opinion.
17	I believe that humans are equally as important as apes as without apes we would not be here today and without humans there would be nobody to protect the apes.
16	I dont think that humans are more important than apes because the evidence shows

	that humans are decended from great apes and so without apes humans might not even exist.
18	I think apes habitats should be protected because we shouldn't allow them to die so that humans can cut down forests for paper and buildings. What do others think? Do you think we should just let the apes die?
16	I think that, yes we will change the world, but we have to remember that people believe once we were apes like them. Instead of us, they might have evolved into something like humans, while we were were still apes. We have to respect these creatures as we are very similar to us, but we are more intelligent. Just because we've evolved, it doesn't mean we have to be disrespectful towards less advanced species.
19	We can't be snotty and say we're the most important race in the world until it is proven that there are no super-advanced alien races out there somewhere.
20	I think that it is unfair that they lose their habitat just because we want more space for our houses. We evolved from apes, I think that we are practically the same as them. They have as much right to the land as us after all we are all animals. We may have evolved further than them but they are so clever and we shouldn't just kick them out when we want wherever they live. People have got to imagine what it would be like if it was the other way round, if we were in their position, would you just let someone take away your home because they felt they were more important than you. Soon the apes will have nowhere to live and it will be our fault when they start dying. We should respect where they live and leave them to their own lives but if we are going to interfere we should help them breed and find places for them to live instead of taking away their homes and just making their lives worse. In other words I feel that we may have evolved further but we are all animals and they are just as important as us.
3	Why do you think people destroy the habitats of apes (and other animals)?
21	But are we protecting the apes by destroying their habitats? You say "I think that humans are probably more important than apes", what makes you think that? "but that doesn't give us the right to wipe out their habitats" can you find an example of this?
22	I agree and do not think apes should die. They are just as important as us and their habitats should not get cut down. Even though we may be more powerful we are all living creatures. What do you think? Do you think it's animal instinct to prioritize our own species over others? Is that something that we should strive to change or is it a basic law of nature and therefore justified?
23	Humans don't make the world a better place as we are polluting it. We are killing the apes habitat and that is unfair. We need to protect them and their environment as well as fixing the problems we have caused.
21	Sorry but I have to disagree with the opinion that humans were descended from apes as i believe that they were created in their own right.
24	Humans are slightly more intelligent than apes however speices are becoming exstinct all the time. I think that any exstinct species should have been helped to survive therefore all endangered species should be helped and conservation projects should be started. Also, we have homes and we are destroying their homes to make room for ours which is unfair.
25	They shouldn't be because we have evolved from them and that is like saying that our grandparents are not important.
18	Hi Tom, I think that although humans are more intelligent, they are also more destructive and are more likely to destroy the environment than make it a better place. Evidence shows that the rainforests are getting smaller every year and the ice

	caps are melting because humans have been expanding. Why do you think we can make the world more sustainable?
5	Human have a higher level of intelligence and are more eco-friendly and we are trying to be more sustainable. So there fore humans are more important. apes are just older versions of ourselves, they are outdated.
4	Interesting point, but do humans really make the world more sustainable? Some of us do make an effort but how many of us contribute to pollution and destruction of the world's environment?
Tutor	Do other species of animal protect other species? Why should we do something when other species do not?
6	If you think about it humans are in the primate family which means that apes are very close relatives to us. As well as this it is animal nature to favour our families and ourselves over any other species.
7	Yes, you are completely right! Also, if they die out this will be another species that has become extinct at the hands of humans.
8	So they have more space to do things in- such as raise livestock, grow crops and build settlements- among others.
14	That is an interesting point Christopher. We would not say that we are more important than our grandparents so why say that about apes?
9	They are both as important as each other however, how important they are isn't relevant because humans will always think highly of themselves and seize land as they wish.
10	Yes Sophie but we also need our habitat because of the fact our population is growing but i agree it is not fair to destroy the apes. Do you not think we should just let nature take it's course? Like it did with wiping out the dinosaurs? Should we just leave nature to itself or do we have a moral obligation to intervene?
12	Humans are definitely more important than apes, but that doesn't give us the authority to destroy their lives and habitats.We should still care for them and respect their existence.
3	This may be true, although if we compared our welfare and an apes which would you choose?
8	I think humans are more important than apes, because they are basically a less evolved version of us- but apes don't pollute the planet, start nuclear wars or destroy other species.
23	It is unfair that humans are destroying apes habitats and expecting them to look after themselves as they cannot rebuild their natural habitat. Humans are supposed to be more intelligent so surely can ensure they survive alongside us as well.
26	I think humans and animals should be valued and treated equally due to the fact that without apes the human revelotion would not have been possible
27	I agree. Scientists believe that we evolved from apes so why do we take their habitats? We as humans, may be smarter than apes in some ways, and apes are smarter than us in others.They are a part of nature as well as us so we shouldn't take their homes.
22	I agree that we are not the most important race. We may think that we are more important than apes but that is no excuse to shorten their lives. We are more evolved than most species but that doesn't mean they are less important and deserve their homes to be ruined.
10	If you think about it we are some times very eco friendly but they are comletly so they deserve this planet more than us allot of the time.
3	How are we more important?

10	But other species of animal seem to have no problem destroying each others' environment... Why should we, as humans, act differently? Why do we have a moral obligation to help?
3	I think it's so that they can use the land for something they need and is useful to them, which I think is selfish. I think we should use land which isn't a home for another species of animal, but if we do, we should help them find another home by donating and raising money.
10	very right
4	I think we automatically put ourselves above other species without thinking about them and base things on our own needs. But then every species would put themselves first wouldn't they? In that way we are no better or worse. As humans we prioritize the situation, yes this situation is bad but the situation of the Polar bears, leopards are higher in this list; however this is no right to kill another race of prey off especially our ansestor's race.
29	Its in our instincts to protect the less fortunate in their battles and if we was a vulgar race how can you explain WWF? Save the children? Because surely we wouldn't be doing that.
30	Although Man has made significant discoveries over the past millenia, every species has its value. It is unfair to judge a species on its intellegence alone, but if we do then Man is slightly more important. But it is unfair so we shall not. Instead we shall judge the species on the skills and ability to survive and thrive. Although we are good at (e.g.) driving cars, Apes are good at climbing trees. Also we share 98.5% of our DNA so in those respects we are equal.
3	How wouldn't have it been possible?
8	Yes, but apes don't make weapons of mass destructuion and start nuclear wars, do they?
31	I believe that apes are suffering from forest fires not lack of land for there WOULD be land if there were no forest fires. I believe that apes are as impotant as humans and they should be treasured and are each precious lives. My reasons are, how the human race has not only created the problems for the apes by hunting them and creating forest fires but also for themselves as there are now far less trees in the world. However, humans have created phenominal sanctuaries for endangered animals and campains/charities to save them. So therefore, I think that we are all equally as important as one another. Humans both create problems yet eliminate them too.
12	We are intelligent and we know what is right and what is wrong.In this case helping apes is the right choice.
32	I think that humans may be less important if they kill innocent creatures because it is selfish and inhuman.
33	Although that is an interesting thought, the scientific evidence in favour of evolution is almost certain, I think?
26	apes have lives too just like us. They have families and children and shouldnt have to give up thier lives to extinction. Do you think all species of animal believe themselves to be more important than all the others, or do we - as humans - stand alone in this matter?
34	Hi, I think that although humans are more important than apes, everything on the earth has a right to a home and other living things shouldn't take that away from them.
21	I believe that god made us to rule over all animals. 'Let us make man in our likeness and let them rule over all the birds of the air, the livestock, the fish in the sea and all

	things that move along the ground.' Having said that i still don't think that we have the right to destroy as we have to protect and care for the animals and not destroy habitats. Surely you wouldn't like your home to be destroyed unexpectedly? Is this about respecting the earth?
10	Yes they destroy other environments but not on such a drastic scale as it threatens their existence .
35	That is quite true but if we treat them equally then where would we live?
36	I think that humans and apes are very similar, and that we should try not to destroy their habitats. Humans and apes are very similar in ideas (eg. If we have a problem, we find the best way to solve it, and so do apes) and in DNA. Humans may be important in the way that they have power over what happens to the land, but in most other ways we are almost the same.
3	You seem to be assuming that mainly we are equal, how could this world, besides not ruining there habitats, make our welfare equal to theres?
20	I agree with you it isn't fair on the apes.
4	What does "important" mean in this case? Important to whom and why?
2	I agree
18	Hi Matthew, could you explain what you mean by created in their own right? Evidence shows that humans are descended from apes so thats not an opinion.
7	It's not as if we are fighting for space and it is survival of the fittest. If it was then perhaps intelligance would matter. There is no reason why humans have to cut down their habitats. We obviously think that we are more important but I think that that should change. It may be part of nature in some ways but we shouldn't destroy their homes.
22	We wouldn't like it if they decided to come and destroy most of our houses to expand their habitats. We are possibly more evolved than apes and that means that we should be clever enough to know we should help them and not allow them to become extinct in the wild.
37	We are the more intelligent species and have a bigger impact with what we do so it is our responsibility to look after the other species which are declining in numbers.
38	I like your idea but it isn't like saying your grandparents are not important because it takes millions of years for an ape to evolve into a human. We should give them more space though because what would we do without other animals?
26	apes are just as relevant as us
39	Hmmm.... I'd have to say both. With humans, we are advancing technology and always crossing new boundaries and improving, however at the same time we are polluting and destroying the Earth but are not capable of acting on it or even noticing. With the theories that humans evolved from apes, perhaps the apes would turn into a better us, so we must preserve their lives as well as ours. Maybe people should just try to stop the population growing by not having too many children. If everyone has just one child by choice (not saying they should but it's an example), the population will either decrease or stay the same. If we don't need new houses for new population, we won't have to destroy the habitats of other animals.
15	The problem is, people are so prone to diseases and therefore can't live in forests like apes. Most animals have a habitat (rivers, seas, deserts, forests etc.) and live happily with other animals in that habitat. However, no other animals live in houses, so humans need more space than them, meaning we need to control our population better.
11	I think that we can learn a lot from the apes, they make their homes without killing too many trees or spoiling the atmosphere. **Why should we cut down there homes**

	for our homes when they have not done anything?
40	Humans are more intelligent but we can learn from animals like apes, because we are closely related and they work as families like humans. Also you see a lot of human behaviours in apes.
26	a
21	Surely if we are similar to apes we would be competing with them for power? That's an interesting point. The question is left that way so you can interpret it yourself, but that would be an interesting thing to discuss? How would your answer change depending on the different answers to your questions?
41	Humans are not more important than apes. No one is any more important than any other. Apes have just as many rights as we do so it is not fair for us to destroy their habitats and their freedom.
42	I believe that all living things are equal. We evolved from apes and without them we wouldn't be where we are today. There homes are being destroyed and they are becoming extinct because of our poachers. I think it is a false accusation to say that we are more important than apes.
30	By establishing trusts to maintain the welfare of other species; in this case apes.
10	me humans are good to animal such as people who work to preserve them but we also are basically getting out of hand exploiting the land and killing the animals. why should we deserve to live, maybe we should die out!
12	But we are a whole new species that can think for itself and rule over the other species. So I think that we shouldn't destroy apes but just rule over them a little.
17	I appreciate what you are saying but because of our ever growing population, we need more space to develop housing, schools etc. If we do not use more 'open space' then where would the humans live? You suggest finding places for the apes to live but where? We could resort to putting them in zoos and safari parks but eventually the zoos will become over crowded. Maybe we could create a 'managed' breeding programme to keep numbers to a specified level? This would mean that we could have the land for housing but also have a specified number of happy apes. Is this a happy medium?
43	Humans are destroying this planet as it is with the factories polluting the atmosphere with their thick black smoke and we destroy the forests for firewood making animals suffer.
44	Without being religiously disrespectful, can I just ask why you think God (if you think there was a God) made it so that we were more superior than animals?
20	So Kathryn do you think that we are more important than them or not?
38	Yes but if we let loads of types of animals go extinct we would have nothing to eat and other animals help get rid of rodents etc.
26	we would still live here but mabey the goverment could make an aniamal shelter
45	I can see where your coming from adam but the human race has done great things like inventing cars and vechiles. And also not all humans cause nuclear wars and some make an effort to stop global warming and other issues
46	In some ways humans are more inportant than apes, but in some other ways you can say apes are more important than humans... Humans could be more important than apes as they have managed to adapt better to living conditions and made full use of what resorces they have. The humans are also at the top of the food chain, which is in a way, important as they will keep certain animals and plants from overpopulation. But on the other hand, apes could be the more important ones as humans have envolved from apes (according to Charls Darwin) and apes can climb, and move around better than humans. However I believe that no animal, human or

	plant is more important than another because sometimes, no life can surive without another type of life, in a way.
26	we all should be valued
47	It is common knowledge that we are direct descendants of the ape, surely it is important to protect our past as well as the future.
3	Are these good enough reasons for the result of climate change?
48	The theory of evolution, proposed by Charles Darwin, says that we descended from apes, and if that is true then apes must be related to humans in some form or another. Some people may think differently but I believe that apes can be just as important as humans. They are living creatures who think, eat and live just as we do.
3	What evidence is there to support what you are saying?
26	i unterstand your point of view but would you like it if the human race ceased to exist its the same think with the apes
10	Yes we may help the world sometimes but most of the time we destroy it and animals help it buy pollinating flowers, make the soil nutritious etc. but they don't destroy it either.
37	No we are not protecting apes because if we are destroying their habitats their numbers will decrease.
44	If you take Charles Darwin's Evolution into account, theoretically, you would notice that apes would be more important than humans. Online, there is somewhere, a picture of Evolution in 24 Hours. If you look at it, humans haven't been around for long. So, no, apes are more important than humans.
26	life is a important part of the world
17	That is a very valid point! We do protect some of the apes as we put them into captivity (zoos, reserves, safari parks) but unfortunately with our ever growing population we do not have enough resources to keep them all safe.
40	I don't agree with your statement as you cannot tell people how many children they can have. I am a third child and would never have been born. I think we should protect apes they are as important as us and we should help protect their habitats for the generations to come.
19	No real proof of evolution. Was Charles Darwin a myth then? Did your school strenuously deny the theory that is proven?
26	if apes becaome extint everything that has apes in thier food chain will cease to exist resulting to less food ans starvation
10	Actually your right we are not descended from apes we have a common ancestor meaning we took different evolutionary paths so it is kind of like killing your brothers
37	Well arn't all areas habitats to some kind of animal?
4	"I believe that god made us to rule over all animals." As an aetheist, I disagree and believe that we have been given no "rights" to rule over other species. we may be superior with intelligence and power, but other creatures rule themselves.
17	If we have the ability to make the right choice, then surely we should be looking after our fellow humans who need somewhere safe, dry and clean to live?
40	No I think we are equal because why should people and animals be ranked differently. Especially higher intelligence animals who have thoughts and feelings like us. I think that some other highly evolved species will think that they are most important but some will know that other animals they need (like part of the food chain) are important as part of their lives.
22	I do not think we fully stand alone but all less evolved species probably just see life as life and just get on with things.

20	Anna i think that is a really good point it just shows us that we should realise that all animals are equal
45	hello cara, its not a matter of giving them a place to live but to keep their environment how they like it and not killing them.
49	Mathew, I happen to believe we are descended from apes. But whether your agree or not, they came into the world with us and even if we are the dominant species they have a right to stay in this world with us.
44	I can see where everyone is coming from, as much as I and other people would like something to happen. Anna, we both would like the government to do something about this to either prevent it or stop the slow death of apes, but I don't think the British Government wouldn't really be able to do anything as Africa's not in the EU and it's abroad. Having saying that, I do hope this will change really soon.
34	We are more important than apes but with that we get the responsibilities of care for them as they have their rights.
35	good idea, maybe it could be somewhere on the middle of a forest I've heard of them
33	Although these are all interesting points, if there was a choice between an ape, or human's life (for example you had the choice to throw one or the other out of a hot air ballon) wouldn't it be ok to throw the ape out? Why? Because we have greater intelligence for example we can do art, science, verbally communicate and many more things an ape can not be taught to do...
44	Sorry, but I also don't disagree. If people still just had one child, the majority of residents would still have more than one child.
tutor	You say "all animals are equal" I wonder if either of you have read Animal Farm by George Orwell? There is a quote from the book which goes something like this: "all animals are equal but some are more equal than others." Do you think it's ever possible to achieve true equality... And is true equality even desirable?
10	Anna's right a lot of famines in history have come from things like over fishing and deforestation killing animals.
37	There is proof of evolution but it is not certain, and is evolution even relevant to this discussion?
40	I agree with you Molly, I think by killing apes and destroying their homes is murder because they are the most closely related animals to us.
12	No but I would say that we are supposed to care for our species not to make it out that we are the most important species.
45	I agree katerina, they are completely different lifestyles
50	Yes humans are more important than apes, they decide what happens in the world and are more intelligent than apes. However, humans also shape the way for the future and we must use our power and intelligence to create a sustainable environment in which all animals can live fairly and unfortunately we can sometimes use our power in the wrong ways. In this case our decisions have lead to a loss of life. What is more the more important use for the wood, a habitat for the apes or an energy source for humans to further pollute the world? Which point do you think is more important?
30	I think that it is natural for a species to think it is the best. Humans seem to have (partially) overcome this when they developed a sense of "self". I think that we should use this evolutionary skill to protect other species whom don't possess this trait.
10	Ah yes, but what if every on e on the planet did the same we would kill the entire race, there are millions less of them than us.
37	I think that some animals may think they are the most important species because

	they only have an understanding of the area they live in and not the whole world. Not all of them might think this because most of them know of a predator that hunts them.
12	That is true but we as humans, with sense, would surely understand that destroying other species is not the right way to go.
19	Most species aren't clever enough to consider that. That's why we are sitting here having this discussion and they're worrying about what's going to eat them.
17	Hi Zoe! I agree, apes should have rights. Are you proposing then, that humans should be living in over crowded housing in over crowded cities whilst apes can have large amounts of open space in which to live?
33	That's very true, but are their rights to things they already have such as their habitats, or things we could improve like better living conditions or medical help?
15	I meant people choosing to not have so many children that deforestation is needed. I agree that people should always be able to decide how many children they have.
20	You make a good point. I think that they all have equal relevance to the argument.
51	Your pretty right we need to find a balanced way of living with apes, this goes with all animals we need to find a way to live with all other life on earth after all the world isn't getting smaller but populations are . How can we keep populations down without committing genocide and killing lots of people?
49	Hi Jesse, I see what you're saying. I believe that as the dominant species we have a huge impact on other species and our population is far too big. It is unfair that because of that we are wiping out their habitat.
3	In what ways will animals shape the future?
33	True, but if it was a choice of all the apes or all the humans wouldn't we still be right to choose to kill all the apes, or would you choose humans?
10	Hi Thomas. Your right but we need animals to survive as we need food and other products. so we would not survive.
38	I'm not sure which but how could you justify killing a whole species.
10	I think One human would make less of a diffirence than one ape.
49	Hi Cari. It depends on what true equality is. If each animal had an equal right to life it would result in a huge adjustment to the food chain!
51	We evolved from apes that makes apes our VERY, VERY great grandparents therefore we should treat them as other humans because they may as well be.
50	Laura I meant humans shape the future, it depends on how you read the sentence, I see why you thought that though. The also wasn't meant to mean we shape the future as well as apes but another one of humans jobs as the most dominating species on the earth. I hope this helps.
33	This is a good point, but if you believe humans to be superior then doesn't that mean we carry the moral obligation to act as the superior species, therefore behave in the right way?
19	I didn't bring it up, I was just voicing my opinion.
52	As we have evolved from them we should also look after them. They need places to live and if so many humans were in their position we would give them a home. If we don't look after them more, they could start becoming endangered.
33	That was just how I was imagining the situation - either them or us? I am not saying that we should kill them off, but if we were faced with the choice, wouldn't we be right to choose humans over apes?
50	What do you mean "to survive" , because we don't need them for food, as vegetarians survive and we don't need their products because vegans get away

	without having anything to do with animals.
19	'Think' is a loose term, but I see what you're getting at.

Summary
148 contributions (posts) from 52 participants, over one hour on 1st January 2013.
5371 words
36 words per post on average
103 words per participant on average
26 spelling mistakes out of 5371 words = .48% (less than 1% spelling mistakes)
11 posts ignored punctuation conventions (7%)
1 post had poor content - probably a mistake (.68%)
64 out of 148 posts (43%) acknowledged another's response
3 out of 148 posts (2%) just expressed agreement with others or repeated someone else's point.
128 out of 148 posts (86%) expressed an opinion
49 out of 148 posts (33%)of posts posed a new question to others
The tutor posed the original question and then made three contributions to the discussion.

9.11 Appendix 11: Inter-rater reliability exercise

Pre-course tutor training Key: 1 = higher order question; 0 = lower order question				
Post	Tutor coding	Researcher coding	Difference	Tutor post
1	Reinforcement	Reinforcement	0	Hi Everyone, You talk about our curiosity to learn which is almost becoming something more.. Do you think our thirst for knowledge outweighs our morals, sometimes? Will there ever be a time when curiosity and morals cross? ..Have we already passed that line?
2	1	1	0	It's great that you're raising awareness of this. :) Another question is whether or not road-safety is a priority matter at the moment.. And how do you know that? Is there anything we should be concerned by first? Perhaps the statistics will tell us more..
3	1	1	0	"We could still live simply and happily." Is this linking into the idea that Ignorance is bliss ? What do you think about this idea? Would the world be a happier, (more moral?) place without the interference of this thirst for knowledge? And how do you know?
4	1	1	0	Are there any stereotypes that could be positive or that people purposefully bring on themselves? What's the different between a stereotype and an image. We could ask this by looking at newspapers - there are certain stereotypes about people who read The Sun, and others about people who maybe read The Guardian. Often in surveys you will be asked what newspaper your family reads - that's because the survey-runner can assume things about you from your answer. Is that fair? On one hand, no, but on the other hand, to get fair results on your survey you have to ask a group of people who represent people at a whole - including both Guardian readers and Sun readers.. so it's necessary? There are even judgements that can be made on the different types of car you drive. People invest in cars because they want the sort of image that car will give them.. These things are all related to stereotypes and you could flip them around "All people who drive such-and-such-s are such-and-such-s" so that it becomes a

				stereotyping.. Where are the lines drawn and how can you justify your argument?
5	0	0	0	Most of you are saying that Mina is the most important friend to Michael. What qualities does Mina possess that make her such a good friend? Can you find any examples in the book of things Mina does that show what a good friend she is?
6	1	1	0	Some of you think the book should have no suggested reading age at all. Is this something you would extend to all books? In what ways can a suggested reading age be helpful? And HOW would you know that it was helpful?
7	1	1	0	Why do you think Shadow was awarded a prize? What criteria do you think the judges of the competition might have been using? How did they come up with those criteria?
8	1	1	0	Why do you think Morpurgo chose Shadow as the title of the book? What's important to consider when choosing a title? Some people have mentioned that "Shadow" may not just be referring to the name of the dog.. Could we explore this further?!
9	0	0	0	There are lots of people who would say that school has a very negative impact on teenagers. There are a lot of statistics out there saying that every child will be bullied at some point during they're school life. A lot of people also believe that school increases anxiety and stress and there are statistics to back this up. What do you think?
10	1	1	0	This depends on how you define bullying. Is it a teasing now and then? Being excluded from a school group? Only a luck few can look back on their lives and say that they've never had a nasty word said to them. Perhaps a lot of people get life lessons out of things like this though? Is the slightly teasing nature of school ever good for people? And how do you know whether it is goodfor people?
11	1	1	0	Do you think the story could have stood up on its own without Skellig? The whole book is called "Skellig", would this imply that Skellig is the most important character in the book? Some of you have started to indicate that Skellig brings certain things - sadness? - to the

				story. Could Skellig be some sort of symbol for something else? What does he represent?
12	1	1	0	Why do you think that? Could you try to justify this opinion?
13	1	1	0	What's the benefit of introducing a supernatural element to the story? Do you think it makes more people interested in the book? How do you know?
14	Reinforcement	Reinforcement	0	That's an interesting idea. What does everyone else think about this idea? Is there anything specific in the story that makes you think that this could be the case?
15	Reinforcement	Reinforcement	0	That's interesting. Can you find examples of this?
16	1	1	0	Does their use outweigh possible moral arguments as to why they shouldn't be used? Is using dogs to check for bombs (something a human could potentially do) in a way saying humans are "more important" than animals. Is that fair - and how do you justify 'fairness' in this situation?
17	1	1	0	How important is Michael's sister's name in the book? Does Almond draw attention to her name? How?
18	1	1	0	What is the significance of Aman's mother's name? If it didn't say in the book, why do you think that was? Is it just unnecessary for us to know?
19	1	1	0	Why do you think Morpurgo chose not to tell us? Does using the word "Grandpa" make us feel closer or make Grandpa's character seem more "home-y"? A lot of people call their grandparents "Grandpa" or other similar names, maybe it is so that some people can relate more with the book? It makes him seem familiar..? What do you think?
20	1	1	0	I'm interested as to why you all seem to think parents wouldn't want to go to "parenting classes", what are you parents' views about this topic? Again and again in the news we see stories of children who have been neglected and about how the social services have let us down. Surely parenting classes, or assessing parents, might prevent these awful cases from happening? What about children who are born into poverty? Some families living in poverty have three or four children and maybe who have more on the way even though they can't really afford it.. Should this be allowed? Is it

				really our place to say who can or can't have children?
21	1	1	0	Here's an article from the Telegraph reporting from back in 2008 when Asda was found to be selling Larger cheaper than bottled water. What do we think about this? Is it okay? Is the readily availability of alcohol fueling alcoholics? Is it moral? What's going on!?
22	0	1	1	But surely all good "heroes" have to be easily relatable to the reader because we have to really get behind them and will them on? (Or do they?) Wolf is an animal, so it's harder for us to understand him. How does Paver make Wolf's character relatable?
23	0	0	0	Is it lazy of Sachar to leave the ending so open? When you read it what was your initial reaction - were you frustrated that in the end you never get to find out the whole ending, or did you like the way Sachar leaves parts to your imagination? It's an interesting technique because instead of simply reading a story it makes you more part of the story, more involved. Have you read any other books that also use this technique? If all books ended like this, is it a novelty that would wear off?
24	1	1	0	However since he is being portrayed this way by Scout looking back at life (and considering we have established that he's based on Harper Lee's real father) is it possible that we're seeing an idealised version, presented free of flaws? How do we know?
25	0	1	1	That's an interesting thought. Do you think all books need to have a clear "Hero", you're suggesting that Wolf Brother doesn't, that instead the characters all work in a team and couldn't have succeeded without each other, does this make it any less of a good book? Does the way it breaks this trend actually add something to the book, is it a purposeful decision on Paver's part and if so what can we learn from it?
26	1	1	0	Does the way all the characters seem like a hero make the book more realistic? Is this decision on Paver's part to teach us something

				or get something across? How do we know?
27	0	0	0	Does that mean that the whole book is a bit of a "novelty" or "gimmick"? Without the ending would the book have become so popular?
28	1	1	0	Then again, the point of view of a child is often different to the point of view of an adult. Young people have less life experience (I'm saying this, also as a young person!) so we involuntarily see things as bigger than they our. The problems we face seem gigantic and impossible to solve whereas an adult faced with the same problem would most likely be able to rationalize it and see that the problem is not actually that great, compared with their previous life experience. In this sense, it's impossible for Scout not to make things seem bigger than they might have seemed to, say, Atticus. However, in a way this is one of the things that makes the book great - we are not reading it for a totally accurate representation of events, we are reading it to get an account of a series of events through the eyes of a young person. So in a way, this doesn't matter. I guess you could argue that everyone sees things differently so as long as a story is written in first person there is always going to be differences to how another person would account the story. In terms of Scout's reliability I think you're right, she has strong morals and we have no reason to doubt her retelling of events. I suppose things we could take into account if we were really analyzing her reliability would be things like memory, and how long ago all this happened! Tutor P.S: Speaking of Scout being a young person and having a young person's perspective, how well does Harper Lee get into the mind of a child?
29	1	1	0	Speaking of a child's perspective, how well does Harper Lee embody the view point of a child? Is it believable?
30	0	0	0	This is an interesting idea because it almost suggests that the popularity of a book is nothing to do with how good the book is or isn't. This is sometimes the case, take Twilight for example, (personal opinion) however, To Kill A Mockingbird is a book students study for English Literature up and down the country, to suggest it might have just got lucky feels odd. However, I suppose most books use a

				combination of both to get popular. It does raise some questions about the system we have where books are published and advertised, though.
31	1	1	0	What makes you say Harry Potter doesn't have much philosophical depth? Might it have much more depth than you might see at a first glance?! Time to justify your opinion!
32	Reinforcement	Reinforcement	0	You have to take every opportunity. ;-)
33	Reinforcement	Reinforcement	0	That's a good question. Another one: Is there such thing as an instant classic? Or does it take years and years to discover whether something is a classic or not?
34	1	1	0	Do you think Torak's only reason to say "go in peace" was so that he didn't offend the Gods? Did it work? How do you know? Did saying "go in peace" somehow justify the killing?
35	Reinforcement	Reinforcement	0	These are some good ideas so far, there are lots of reasons why Holes is a good title. (keep them coming!!) Before you picked up the book, what assumptions did you make about it from its title? Were they right?
36	0	0	0	It's interesting to have some vegetarians in this discussion. How is the way Torak kills the buck, to eat it, different to how we kill animals in the modern day? Being vegetarian is an interesting decision because we can tell from the positioning of our eyes and our canine teeth that we have evolved to be predators, however I think in the modern day many people eat more meat than we were originally defined for. For me, one big difference stands out between how we kill animals for meat and how Torak does. What do you think?
37	1	1	0	Perhaps we could start by thinking about the difference between promoting something and just writing about it? In this morning's discussion a lot of you thought Torak did the right thing in honouring the buck. Did honouring it outweigh the cruelty of killing it? How do you know?
38	1	1	0	How do we define a "good" or "bad" person? Is there really such thing or are most people a mix of both? Everyone makes mistakes, digging holes isn't going to change the past and it seems a bit extreme as a punishment. If it could make anyone a better person, surely it would be the self development caused by having a lot of time to think rather than the holes

				themselves, anyway? What do you think?
39	0	0	0	In the end, justice doesn't prevail, but are things any better at the end of the story than they are at the beginning? Was it all futile of did Atticus' actions have some sort of effect even if they didn't save Tom?
40	Reinforcement	Reinforcement	0	Can anyone find any quotes from the book that describe the setting that make it seem realistic or unrealistic?
41	1	1	0	Is it possible to be truly connected with nature without living as Torak lives? Is it possible to be truly connected with nature in the modern world?
42	1	1	0	Does it matter if the setting is not entirely realistic? How does it affect your opinion on the book?
43	Reinforcement	Reinforcement	0	Can you find any quotes or examples of how Stanley's character changed from the start to the end of the story?
44	Reinforcement	Reinforcement	0	Can we find any examples or quotes that back up these themes?
45	1	1	0	What exactly is a hero? Can we find any examples of heroes in literature?
46	1	1	0	Why do you think we're asking this question? What can thinking about the theme of a book tell us about the book? ..Or are we just destroying the book by over analyzing it?
47	1	1	0	Do you think the way a lot of books have a clear "hero" and/or "villain" (especially books for younger people) is helpful? Are we giving young people an unrealistic view of the world? ..and all these happy endings..?
48	org	org	0	'ere 'ere!! It's coming up to the last ten minutes o' discuss'on. Y'all be'er use this 'ere last few minutes to convince the oth'r jury members o' you're way o' t'inkin. ;-)
49	org	org	0	The verdict will be made at 7 o'clock when this 'ere discuss'on is closed the vote will op'n then we'll see the verdict fair an' square.
50	0	0	0	Does the element of magic in this book make it less believable?
51	0	0	0	It is also not something that we have managed to eliminate from society. What examples of discrimination are still a major threat?
52	Reinforcement	Reinforcement	0	That's an interesting point - a lot of trials are open to the public, anyone can walk into the court and sit in the gallery and watch what's going on.. How is live-streaming any different

				to that?
53	Reinforcement	Reinforcement	0	Can I just say how impressed I am with the way you've referenced your post. It's a brilliant habit to get into. :-)
54	1	1	0	Why isn't public transport the answer? Things can't continue as they are, we have to do something, don't we? What are the options?
55	1	1	0	Does having one million pounds make you rich? This article says no.
56	1	1	0	Hello Everyone, Edmund Burke was an Irish political philosopher who said, "All that is necessary for evil to triumph is for good men to do nothing." Do you agree with this quote? How do you interpret it? If you agree for it, what does it mean for you in your day to day life?
57	Reinforcement	Reinforcement	0	Hello, This is a copy of Maslow's Pyramid of Needs. (You can do a quick Google Search to come up with a slightly less blurry copy. :-))
58	org	org	0	This House would ban cosmetic surgery.
59	org	org	0	Trending Topics
60	org	org	0	Trending Topics
61	0	1	1	In 1714 the British government asked the population the biggest problem of the time - how to pinpoint a ship's location at sea by knowing its longitude - offering a reward for the solution. Now in 2014 the Longitude Prize is open again. Submissions aren't open yet but the challenge is on Antibiotics. However, to decide the category there was a vote on what was the biggest problem. The options were: Dementia Flight Food Paralysis Water Antibiotics Find out more about the Longitude Prize here: http://www.longitudeprize.org/ Find out more about the categories here: http://www.bbc.co.uk/news/science-environment-27425224 Which category do you think is the most important?
62	1	1	0	This House would use "positive discrimination" to increase diversity in schools and universities.

63	1	1	0	How does Penelope Lively create tension at the start of The Ghost of Thomas Kempe? How does C.S. Lewis do it?
64	0	0	0	Hi Everyone, I think the three main points everyone has raised so far have been a) escapism: the change to get totally absorbed in another world and be able to forget your own b) relation: many fantasy books are based around familiar characters from our childhoods like wizards, witches, elves and fairies or are based on humans, which makes them easy to relate to. c) we read fantasy books and now, because of technology, can wonder whether they will soon become a reality. We also enjoy observing other worlds and drawing parallels between our own world and the fantasy world. What do you think?
65	0	1	1	We've explored some of the reasons why we enjoy fantasy books so much, but what is it about the specific books given in the original post (The Chronicles of Narnia, Lord of the Rings, Harry Potter and His Dark Materials) that has made them such a massive success? Luck? Or do these books have something special that other books are lacking?
66	0	0	0	I especially like your point about the films. Harry Potter, The Chronicles of Narnia, Lord of The Rings and His Dark Materials all have had films made after them (Coincidence?) and are all massively popular. It's a bit of a chicken and the egg scenario - where the books so popular because they were made into films or were they made into films because they were so popular? What do you think?
67	0	0	0	So a good book opening is one that asks questions and makes us want to read on? Those are two criteria right there! What other criteria can be used to judge a good book opening?
68	0	0	0	What specifically is it about the opening of Harry Potter that makes it stick in your memory?
69	0	0	0	That's an interesting view. Many people would say that it's good that Penelope Lively didn't explain all the things that perhaps needed to be explain because that draws the reader in and creates a sense of tension. What do you think?
70	1	1	0	Interesting that you brought up happy endings. I'm going to throw the gauntlet out and suggest that the vast majority of fantasy books have

				overall happy endings. Do you agree? Is this another reason why we love to read them so much? Unlike in real life, we are almost certainly assured a happy ending? Don't forget to provide evidence for your reasoning!

9.12 Appendix 12: Tutor posts on the Argument course

Post	Coding	Tutor posts on the Argument course 1 = Lower-order; 2 = Reinforcement; 3 = Organisational; 4 = Higher-order
1	2	What is your reasoning Kathryn?
2	4	But is it a 'mistake' Rida-Fatima? Is there a 'right' answer here?
3	4	What's the big theme we are discussing here? The morality of murder? Is it murder to let someone die? Is that just as bad as deliberately killing them? Or are there different grades of murder?
4	3	This discussion is closing in 5 minutes, in order to keep the thread of a reasonable length. Please now move to the next discussion.
5	4	Eleanor, it is interesting that you mention the 'greater good'. Have you heard of utilitarianism? How relevant is that philosophy to this thought experiment? ...some research for you..... :)
6	4	So - this is a thought experiment - which means that we are not allowed to add any extra conditions (e.g., putting boulders in the way of the train) into the dilemma. How does that affect your response?
7	4	As this is a thought experiment, we are not given information about the 'worth' of the protagonists. What is the main theme which is emerging here? And HOW do you know?
8	4	How does the philosophy of utilitarianism apply to the 'solutions' you have been proposing?
9	3	Just to remind everyone that the discussions close at 7pm. See you tomorrow for the final day! :(
10	4	Does the utilitarian perspective take 'responsibility' into account? How do you know?
11	3	Do you mean debts, as in money? The Debating Chamber is actually open for 2 hours per day. Everyone received an email earlier today, to notify you. There have been a lot of threads today, which is actually a good thing. :) Loads of variety! I imagine that the number will dwindle tomorrow, but it is very impressive to see such a variety of discussions.
12	2	We all make mistakes. Life would be pretty boring otherwise! :)
13	3	Give it time , Emily - there are a lot of discussions on the go at the mo. :)
14	4	Good question! What is 'a good education'? What are the arguments which convince you?
15	1	There has been a lot of research into brain size and brain effectiveness. Has anyone come across this? Can you enlighten us? Evidence, evidence, evidence please.
16	2	Good question, Thomas!
17	2	Good question Sarah!
18	4	So what is the answer if we don't know something? Research it perhaps? :) Let us know when you find out. BTW: are there some topics which you are 'too young' for? What are the parameters for such a judgement?
19	2	Might be a good idea to define 'gaming'. Are we talking about Scrabble for instance? :)
20	2	<u>Excellent question.</u> This one is always debated on philosophy courses - and always gets participants animated! Please remember to RESPECT others' opinions. We have students from every faith and none on this course - we all have the right to our

		opinions. Evidence evidence evidence please.
21	4	Should the school leaving age be reduced? The UK government has announced plans to raise the school leaving age. Others think that it should be reduced. What's your opinion and why?
22	2	Perhaps someone can research the scientific evidence on this one?
23	4	<p class="introduction">People or profit? Should we care? Brazil biofuel: Shell axes 'illegal' sugar cane plan. fuels company set up in Brazil by oil giant Shell has signed a landmark agreement giving up plans to buy sugar cane grown on indigenous lands. </p> <p>The company, Raizen, was obtaining some of the raw material for ethanol from farmers who encroached on the lands of the Guarani people in Mato Grosso do Sul state.</p> <p>The deal comes after months of pressure by the Brazilian authorities.</p> <p>Indigenous leaders have welcomed the agreement. </p> <p>Raizen said in a statement that the deal "shows the company's commitment to combine sustainable development with the well-being of the local communities".</p> <p>Campaign group Survival International described Raizen's decision to back away from ancestral Guarani land as "historic".</p> <p>But it warned that the tribe's future continued to be threatened by illegal logging and farming on their lands.</p> <p>Valdelice Veron, an indigenous Guarani in Mato Grosso do Sul state, says their rivers have been polluted by pesticides.</p> <p>"We will be able to drink water from our land again. We will be able to start afresh," she told Survival International.</p> <p>Read the full article here:</p> <p>http://www.bbc.co.uk/news/world-latin-america-18433008</p> <p>Have your say:</p> <p>People or profit? Should we care? </p>
24	4	Another excellent question! This one is HUGE! Do you think that SATs are for the school or the pupils? What is the political debate around needing to know about standards in schools?
25	2	Excellent question!
26	2	Excellent question. Might result in a bias, as most students on this course are from the state sector, but let's see if we can have a balanced debate, with pros and cons.
27	3	Can anyone research the scientific data about the benefits of laughing - and enlighten us? I am sure that it would be relevant to this discussion.
28	4	All animals are equal - but some are more equal than others'. Anyone recognise that quotation? To what extent is it relevant to this debate?
29	2	Evidence, evidence, evidence please.:) Perhaps someone can research what Stephen Hawking has to say on the matter?
30	4	What might be the consequences if we were able to go back in time and change history?
31	2	Well done, Joseph, for demonstrating respectful debate. :) Excellent.
32	4	This is a very interesting question and highlights some of the issues which 'able' students experience. What other issues do you face and how do you deal with them?
33	3	This discussion has branched in more than one direction. The discussion about favourite books might merit its own discussion thread perhaps?
34	4	So what are the pros and cons of this debate?
35	3	The Debating Chamber will be open from 6pm until 8pm again tomorrow. Do call in for another fantastic set of discussions. The Debating Chamber will self-destruct at 8pm precisely, so it might be a good idea to sign off now....
36	3	10 minutes... Please sign off now...
37	3	6 minutes.... Thanks to our amazing Discussion Leaders - and to

		everyone who has set up some very thought provoking debates.
38	2	Good question, Jake!
39	4	Are there any benefits to being home-schooled? What are the arguments on both sides?
40	4	This s a very interesting question. :) Who can list the benefits of alcohol? If there aren't many, what are the disadvantages? If the disadvantages outweigh the advantages, why hasn't it been banned? Is this a political debate? Who does benefit from the sale of alcohol?
41	4	...and if we chose to get rid of them, how would we do this? What is the political debate surrounding the supply of these things?
42	2	Well done Eleanor, for making the parameters of this argument clear.
43	4	This House believes that the internet is the greatest force for democracy in the world. Do you agree? What are the arguments?
44	4	<font size="4">Should zoos be banned? >London Zoo has welcomed its latest inhabitant - Jae Jae the Sumatran tiger, who has travelled nearly 4,000 miles to meet his new mate. It is hoped he will breed with Melati, a female from Australia, and help save the critically endangered species. There are only 300 Sumatran tigers left in the wild. Many people think that zoos should be banned. What are the issues here? Why does this subject evoke such strong opinions on both sides?
45	4	Is there an ethical dilemma surrounding our choice of band? Hmmm....
46	1	What do you think are the three biggest mysteries of the universe?
47	4	Great ape habitat in Africa has dramatically declined. . Great apes, such as gorillas, chimps and bonobos, are running out of places to live, say scientists.</p> <p>They have recorded a dramatic decline in the amount of habitat suitable for great apes, according to the first such survey across the African continent.</p> <p>Eastern gorillas, the largest living primate, have lost more than half their habitat since the early 1990s. </p> <p>Cross River gorillas, chimps and bonobos have also suffered significant losses, according to the study.</p> <p>Does it matter? Aren't humans more important than apes? </p>
48	4	Which particular questions/options got you thinking? And what were those thoughts?
49	3	I am closing this discussion in 5 minutes and have re-opened the discussion (same subject). This is so that we don't get threads which are zillions of posts long.
50	2	Well done Jasmine - I can see that you have read the guidelines! :)
51	4	Does it matter if the subject matter does not relate directly to your experience? Is there any benefit to being pushed out of your comfort zone?
52	3	Hi Oliver, This is the place to discuss the topic and to give each other feedback on your ideas.
53	4	So let's unpick this issue of 'not relating to the topic well'. Do you have to remain within your intellectual comfort zone or is there merit in pushing your boundaries beyond your immediate experience?
54	4	So what do you do, Alex, when the language is challenging? And is it effective? And how do you know if it is effective?
55	4	Asmara - what do you think might have been the reasoning behind asking you to speak for both sides of the argument?
56	3	Just to remind everyone that the discussions close at 8pm. See you tomorrow :)

57	4	So what was your strategy to overcome this Kasey?
58	4	Some of you mention that there were phrases that you didn't understand. What was your strategy for overcoming this obstacle?
59	2	You make a good point, Leela. Can you all agree on a definition for 'too dependent'?
60	1	Have national trends changed regarding the average age of first time parents?</p> <p>Can you find out?</p>
61	4	Hi Aine You weren't the only one who found some of the language difficult. What approach did you take when faced with difficult vocabulary?
62	4	James has raised an interesting concept - that of voting in a world leader.</p> <p>What advantages or disadvantages would there be to such a role?</p>
63	4	A very interesting debate! I think that it is safe to say we are reasonably dependent on technology, but being dependent on it is necessarily a bad thing? What are the arguments? If we have it, why not use it? What do you think???
64	4	Many scientists have agreed that global warming is happening, (whilst others violently disagree on the evidence) and most (but not all) believe that it has been caused as a result of human interference. What do you think - and how do you reason YOUR argument? Where's YOUR evidence? Time for some research!
65	2	Hi Miles I think supernovas only happen to giant stars, much larger than the Sun, so fortunately we are not going to die from a supernova.
66	4	Yes that's a really good point. Countries like China are good examples of this. What does everyone think about social media sites like facebook, should these be controlled by the government? And if so, why?
67	2	Hi Orla How do you think you would impose these restrictions?
68	4	Ah but does it... It would definitely be interesting to see how much the queen spends in comparison to the money recieved from tourism. Who can find ojut for us? Do you think we as a nation are still proud of the Royal family? After all thousands flocked to the streets to witness the diamond jubilee and the wedding. What do you think? What are the arguments surrounding this issue?
69	4	Hi John, You say that the Captain is not guilty of murder, do you think he is guilty of anything at all? Should he feel guilty? The people's lives were his responsibility, weren't they?
70	4	Hi Kathryn, Why do you think it was morally wrong? Is it not the moral thing to do to save the maximum number of lives? Or is there more to it than that?
71	4	Hi Isobel, Are you saying that we should not take any sort of responsibility for the deaths? How do you reason that?
72	4	Hi Kayleigh, Why do you feel you are partly responsible just by being there?
73	4	Is any obsession a good thing? How do we know?
74	1	Do you think the fact that some of the others on the lifeboat disagreed with the Captains decision have any impact at all?
75	4	Hi Sumayyah, So why would you feel responsible for the death of the one person? If you had done nothing... would you be responsible for the death of the five?
76	4	...Where do we draw the line when it comes to self defence?
77	4	Why would listening to Mozart help any more than listening to, say, the Sex Pistols? What are tje reasons behind this argument?
78	4	Do games like Call of Duty set a good example for the future generation? Or do they provoke anti-social behaviour? And how do we

		know?
79	2	...Maybe it depends on what you're watching?
80	4	George Orwell quite famously said "Sport is War minus the Shooting", do you agree or disagree? Why? What similarities/differences can you draw between sport and war?
81	4	Hi Priyanka, You say they are 'completely unalike' but what about tactics and plans, what about emotional baggage? Are there any similarities? Hi Everyone, If you disagree, why do you think George Orwell might have said this?
82	4	Those are some good similarities... What about the emotional baggage of both sport and war? Are there any similarities there? Imagine if you scaled sport up in intensity... Would that make it more warlike? Do you think if someone had asked you this before the 2012 Olympic Games you would have answered differently? :-)
83	4	Hi Matthew, Those are good points. :-) Do you think sport can sometimes take over people's lives: how far should athletes be prepared to go: how much should they sacrifice in favour of training?
84	4	Hi Emily, George Orwell was actually using football as an example when he made this statement. :-) Do you think something needs to change in the football community? And if so - what - and why?
85	4	Twilight... That raises a question. After the success of Stephanie Meyer's "Twilight" the bookshelves of bookshops started to be filled with stories about vampires and werewolves - writers churning out books to fuel popular demand. The same can often be said about series... Is writing becoming too much about money rather than about the quality - and how do we know?
86	4	Is true communism ever possible to achieve? Is it possible for everyone to be equal? Is true and complete equality even desirable?
87	4	Do you think parents put too much pressure on their children, then? Has anyone heard of 'Tiger Mothers'? Maybe you could research them? :-)
88	4	Are the exams we take at the moment a good reflection of what we have learnt? Do they benefit people who work best under pressure and who find it easy to sit still? What other ways could we have of measuring progress? How seriously do you take your 'levels' at school? Are you someone who compares themselves to other people or do you just strive to improve your own levels?
89	4	Hi Annabelle, But if there cannot be a 'normal human' does that mean there can't be a 'normal weather' or 'normal clothing'? Where do we draw the line?
90	4	Has anyone seen the advert, the one which says "We are the creators of normal..." What is your reaction to the advert? https://www.youtube.com/watch?v=LbcwsSZCKuI Do you think it's possible to 'create' normal? And if so, how?
91	2	Aren't most words open to interpretation, Thomas?
92	3	Great debating Everyone! :-) See you all tomorrow. Bye for now.
93	4	There is an argument, that since the average living age is going up almost every year, the ages that we become legal for things should also be rising as we should all be staying in school for longer. What do you think about this idea?
94	4	What does everyone think about the legacy of the London Olympic Games? Do you think the UK will be able to preserve it? "The higher you climb, the further you have to fall." Does this age old quote bear any relevance to 2012?
95	4	But, on the other hand, how does the premier league benefit the UK? It gives us world round recognition, doesn't it? Maybe you could research about how much money the premier league actually brings into the UK in

		terms of fans, merchandise, tourism...? And more importantly -is it worth it?
96	4	I guess it depends on how you define 'World'. By world do you mean life as we know it, the Earth, the Universe, or are you talking about a more spiritual world? Depending on your definition of world, are some worlds eternal?
97	4	Can anyone find any example in the real world (perhaps looking towards the Arab Spring) where the internet has been a force for democracy?
98	2	Maybe it depends what they use their skills of hacking for... If it's to hack a website as a sort of protest, then maybe it is not so bad after all?
99	4	Hi Scarlet, Do you think the captain made the mistake of just not thinking it through entirely? Do you think that people should still go to jail (even for only a short time) if their crime was a 'mistake'? And how do you justify your argument with evidence? Tricky...
100	2	Hi Jade, It's an interesting point about the way that the captain threw over the weakest and those least likely to survive on their own. I hadn't thought about that.
101	1	Hi Joey, Do you think that the fact that some people on the lifeboat disagreed with the captain's decision changes this at all?
102	4	Hi Jessy, So you think that the captain was right to try and save at least a few lives? Does the fact that some people on the boat disagreed with his decision change anything?
103	4	Hi Sarah, Do you think the captain thought his actions through entirely? Should he feel guilty for the deaths of the few whom he made go overside despite the fact that some people of the lifeboat disagreed with him? Should we be considering his point of view - or this this a plain and simple moral issue?
104	2	Hi Sarah, That's a good point. But do you think the captain should have sacrificed himself first? ...In fact, what was he even doing on the life boat - surely the captain should go down with his ship?
105	4	Hi Jade, So if someone is accidentally killed but the blame can be pinned on one person, even though it was an accident, should that one person still go to prison? Does there always need to be someone to blame for a death?
106	4	Hi Matthew, Why do you think young children and families had the right to be on the lifeboat even if they were the weakest? Why not the strongest instead?
107	4	Louise disagrees with you... She thinks the captain did the wrong thing. Maybe you could go read her reasons and then see if your view has been changed, or try to change her view?
108	4	Hi Louise, What arguments do you think people defending the captain would use against you?
109	4	Hi Louise, Yours is a really interesting view. Are you, in a way, suggesting that the morally/ethically right thing to do is not always the thing that will save the most lives? :-) Let's unpick this argument.
110	2	Hi Louise, That's a very interesting way of looking at things. :-) But surely lives are more important than the feelings of those who are alive?
111	4	Hello Elliot, Do you think the morally/ethically right thing to do is always to save the most lives or are there ever any exceptions to this? Louise disagrees with you... What do you think after reading her post? :-)
112	4	Hi Thomas, Does the fact that some of the people in the lifeboat disagree with the captain change things at all?
113	2	Hi Adam S, How would you go about doing that?
114	4	Is it possible to improve our world instantly or is it a process that will take time?

115	1	Hi Laurence, Perhaps languages are constantly evolving things, and, unlike Ciphers, I'm not sure if you could really 'create' one overnight. What do you think?
116	4	Hi Sophie and Anna, (And Everyone) You say "all animals are equal" I wonder if either of you have read Animal Farm by George Orwell? There is a quote from the book which goes something like this: "all animals are equal but some are more equal than others." Do you think it's ever possible to achieve true equality... And is true equality even desirable?
117	4	Hi Emily, Do you think all species of animal believe themselves to be more important than all the others, or do we - as humans - stand alone in this matter?
118	4	Hi Katerina, That's an interesting point. The question is left that way so you can interpret it yourself, but that would be an interesting thing to discuss? How would your answer change depending on the different answers to your questions?
119	4	Hi Mohammad, But other species of animal seem to have no problem destroying each others' environment... Why should we, as humans, act differently? Why do we have a moral obligation to help?
120	4	Hi Matthew, You say "I think that humans are probably more important than apes", what makes you think that? "but that doesn't give us the right to wipe out their habitats" can you find an example of this?
121	4	Do you think it's animal instinct to prioritize our own species over others? Is that something that we should strive to change or is it a basic law of nature and therefore justified?
122	4	Hi Emma, Do other species of animal protect other species? Why should we do something when other species do not?
123	4	Hi Kayleigh, Do you not think we should just let nature take it's course? Like it did with wiping out the dinosaurs? Should we just leave nature to itself or do we have a moral obligation to intervene?
124	4	Hi Tom, Isaac and Katerina, Do you think it's part of our animal nature to prioritize our own species above others?
		1 = Lower-order; 2 = Reinforcement; 3 = Organisational; 4 = Higher-order

9.13 Appendix 13: Tutor feedback

Feedback after: 'This House would ban cosmetic surgery'

This was an excellent debate. For a change the two sides appeared to be reasonably evenly matched. Everyone argued their case really well. It was very pleasing to see so many of you beginning to use evidence to back up your opinions. This is a great step forward in your debating skills.

How to improve your debating skills further

The secret to being able to get your point across is to be sure of your ground in the first place. Preparation is the key. It is up to you to do some investigating and get some sound references. They make all the difference, as we have seen today.

Fantastic debate - well done to everyone! 🙂

Feedback after: 'This House would condone torture to obtain information from suspected terrorists'.

As this was likely to be a very controversial and emotive topic, we all held our breath as the debate started! However, we were VERY impressed today with the standard from start to finish. It appeared that there was an equal split of opinion, and those who contributed had made the effort to do their research in order to offer a reasoned argument. The whole debate was characterised by maturity and a sense of balance. Well done indeed.

It was also clear that many of you have read the resource entitled 'What level of debater are YOU now?' We think that you are all consciously trying hard to take advice on board and improve your debating skills. If you haven't read that resource yet, now is the time to have a look...

Feedback after: 'This House would go vegetarian'.

This debate was even more balanced than usual, as you can see from the voting booth. (NB: have you voted yet?)

It was great to see that everyone's skills of debating are improving. We have moved on from the absolutist stance ('Wikipedia says so, so it must be true!') to a more evaluativist position, where you are offering well-balanced and evidenced reasoning for your opinions. It is clear that you are considering both sides of the debate before forming your opinion.

One way that we can improve even further: It is clear that you are doing your research before this debate - which makes the debate all the richer. Please can you make a note of your sources, so that we can explore them in more detail if necessary? Thanks!

See you tomorrow, for, 'This House would condone torture to obtain information from suspected terrorists'. Hmmmm - we need LOTS of evidence here please. Get researching!

Feedback after: 'This House would restrict advertising to children'.

It was generally agreed that there were dangers attached to advertising to young children, and that parents needed to be vigilant. We spent some useful time debating whether or not we should have a Minister for Children.

Your debating skills are improving! It is clear that you are taking note of the feedback each week.

This week's tip: try to anticipate the counter-argument.

It will make your own argument even stronger. So you could say something like:

It could be argued that...... Evidence could be brought to show that....

However, on balance, I think that.... because....

Feedback after: 'This house believes that homework is a waste of time'.

Well this topic exercised all of our opinions! It was lively, well-balanced (well almost) and full of energy.

Each day, your debating skills are improving. You are now forming your arguments carefully and are providing evidence to support your statements. Would you like to improve your debating skills? Of course you would! This little voting task will help you to evaluate what you need to do to move up a level:

Voting booth

See you tomorrow, when the debate is:

This House would restrict advertising to children. This is particularly pertinent, since there is currently a national debate about whether or not to start allowing the advertising of e-cigarettes.

Bye for now! Enjoy your evening!

....

Feedback after: 'This House would ban animal testing'

This was another lively debate with a more or less equal set of opinions on both sides. It is very good to see so many of you citing your evidence - that is a noticeable improvement since Week One.

<u>There are three levels of debaters in a debate:</u>

1. There are the '**absolutists**'. These folks state their opinion as a 'fact' and don't want to consider the opinions of others. Their contribution tends to try to over-ride the opinions of others - and their aim is to close down the debate by being 'right'.

2. Then there are the **'multiplists'.** These debaters know that there are both sides to the argument and feel that everyone has a right to an opinion - therefore it is likely impossible to reach a conclusion.

3. Finally there is the '**evaluativist**'. This person considers both sides of the argument and comes to a reasoned conclusion, having explored all of the evidence and having considered the strengths and weaknesses of both sides. This person understands that, although everyone has a right to an opinion, there is the possibility that some opinions are better reasoned than others.

Which type of debater are you?

9.14 Appendix 14: Statistical techniques (supplementary information)

9.14.1 Repeated measures ANOVA

Repeated measures is calculated by partitioning variance according to Figure 9-1:

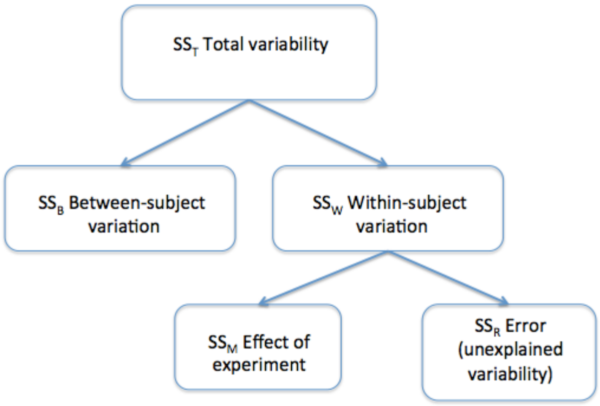


Figure 9-1 Partitioning the variance for repeated measures analysis of variance

Degrees of freedom are calculated thus:

df_W = n(t-1), where n = number of subjects and t = number of repeated measures

df_M = t - 1, where t = number of repeated measures

$df_R = df_W - df_M$

Mean squares are then calculated:

$MS_M = SS_M / df_M$

$$MS_R = SS_R / df_R$$

The *F* ratio then represents the ratio of the variance explained by the model and that which is unexplained, i.e., the ratio of the experimental effect to the performance of the participants due to unexplained factors:

$$F = MS_M / MS_R$$

Therefore values of *F* greater than 1 suggest that the experiment has had some effect.

Effect sizes can be calculated with the equation:

$$r = \sqrt{\ } (F\,(1, df_R\,) / F(1, df_R)) + df_R)$$

Sphericity refers to the assumption that measures from one time point T1 to another T2 are similar from T2 to a further time point T3 and so on, i.e., the equality of the variances of the *differences* between experimental levels. Therefore sphericity refers only to repeated measures where there are more than two time points. It is calculated on the differences between pairs of scores on each manipulation level, and is not violated if:

$$\text{variance}_{T2\,-\,T1} \approx \text{variance}_{T3\,-\,T2} \approx \text{variance}_{T4\,-\,T3} \text{ etc}$$

Statistical adjustments (e.g. Greenhouse Geisser) can be made if sphericity is violated (see Field, 2013; p. 548 for a full explanation of this adjustment).

9.14.2 ANCOVA

The ANCOVA analysis first employs regression methods to account for the variation in the dependent variable which is due to the covariate, then performs analysis of variance (ANOVA) on the adjusted scores. By removing the influence of a relevant variable, the sensitivity of the *F* test is increased, making it more likely that differences can be detected between groups. One-way ANCOVA was used in the pilot study. It involved one independent categorical variable (the group) at multiple levels of the group, one dependent variable (the post-test), and one covariate (the pre-test). The ANCOVA test indicates whether or not the difference between the group mean scores at post-test is significant, after the pre-test scores have been taken into account. The main advantage of using ANCOVA is its capacity to partial out the effect of the covariate; the within-group variance is reduced by allowing the covariate to explain some of the otherwise 'unexplained' error variance (see Figure 9-2). The result is that one is able to more accurately assess the effect of the independent variable.

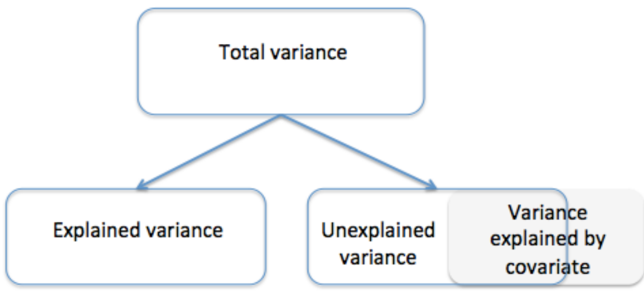


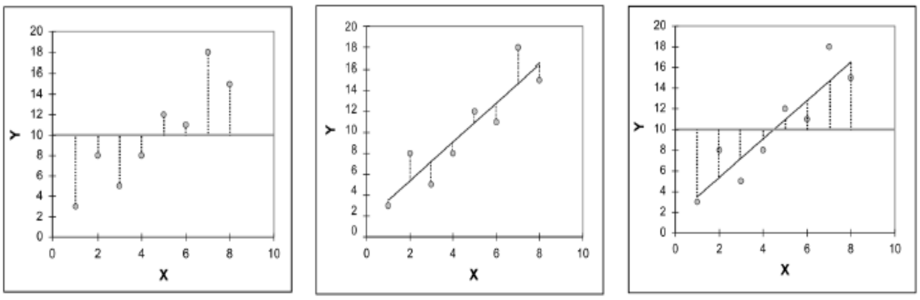
Figure 9-2 The role of the covariate in ANCOVA

Tabachnick and Fidell (*op cit.)* discuss the limitations of the ANCOVA test, making suggestions, all of which are addressed in this study. For instance they point out that the covariate should ideally be collected before the experimental conditions are put into place, so that the covariate is independent of the treatment. Furthermore, they suggest that there should be a minimum number of covariates, that they should not correlate with each other, but that they should correlate with the dependent variable. Discussing the problem of unequal sample sizes, an issue that is relevant to the pilot study, they suggest that a 'Type III' adjustment should be included in the ANCOVA test (p.278). Fortunately, SPSS v. 20 is capable of making this adjustment.

9.14.3 Regression

Estimating the model involves observing how closely the data fits the regression line, with the differences between the data and the model being known as residuals (either positive or negative depending upon whether data points are above or below the line). To estimate this error in the model the residuals are squared and summed, resulting in the 'residual sum of squares'. This calculation (SS_R) contributes to the assessment of how well the line fits the data.

The next required calculation is the total sum of squares, denoted by SS_T. This is calculated as the difference between the observed data and the mean of the outcome variable y. Using the SS_R and SS_T, it is possible to evaluate the extent to which the model is better that a simple model, such as the mean. The difference between SS_R and SS_T represents the improvement in prediction of the model and is represented by SS_M, the model sum of squares, as represented in Figure 9-3.



SS_T SS_R SS_M

Figure 9-3 Sums of squares informing the regression calculation

Two key statistics are produced in the regression process:

a) The variance in the outcome variable explained by the model is represented by R^2, and is calculated as:

$$R^2 = SS_M / SS_T$$

b) The measure of how much the model improves the prediction compared to the error in the model is represented by the ratio *F*. This is calculated as:

$$F = ((N - k - 1)R^2)/(k(1 - R^2))$$

where N is the number of participants and *k* is the number of predictors in the model.