KNOWLEDGE MEDIA

KMi

INSTITUTE

Learning Analytics for Epistemic Commitments in Collaborative Information Seeking Environment

Technical Report kmi-13-04

December 2013

Simon Knight

The Open University

Learning Analytics for Epistemic Commitments in a Collaborative Information Seeking Environment

Candidate: Simon Knight

Supervised by: Simon Buckingham Shum and Karen Littleton

Submitted: 10/08/2013

Acknowledgements

In my MPhil thesis I began the acknowledgements by saying:

> *In acknowledging there's always a risk of leaving people out, and I have no doubt had more useful (and interesting) conversations and feedback than those mentioned here, and indeed I have no doubt in some instances I have received useful guidance indirectly.*

I started acknowledging early in the drafting and writing process, and while the point was a valid one, it was (perhaps patently) unfinished insofar as I intended to give *explicit* general acknowledgements. It was an accidentally unfinished thought, but in some ways an apt one (the beauties of post-hoc justification) – acknowledgements are continuous, evolving, snowballing, and certainly insofar as they capture my gratitude they are fairly inadequate too. It is also a useful point to raise here because this work substantially builds on my MPhil work at Cambridge, and to a lesser extent my MA work at the Institute of Education, University of London – thus the gratitude expressed in those theses extends to this work too.

In relation to this year's work, I am particularly grateful to a few people, most notably my supervisors Simon Buckingham Shum and Karen Littleton, for their support from before I arrived at the OU right through; no doubt I would not have managed to get off to such a running start, and been engaged in such interesting activities over the year had they not engaged with me so fully and so fruitfully.

I have also been lucky enough to attend a number of conferences this year, and I am grateful to the reviewers and participants of CSCW13 and LAK13 and their corresponding workshops for productive feedback and conversations. I am also grateful to the participants of the Learning Analytics Summer Institute (LASI) which was held at Stanford this year, and to the Lytics lab there which kindly hosted me for a week following LASI. In particular conversations with Emily Schneider, Brian Perone, and René Kizilcec at the Lytics lab, and Sean Lip (Google), Gene Golovchinsky (FXPAL), Kirill Kireyev (Instagrok), and some folks at the Wikimedia Foundation, made the week an intellectually stimulating, and incredibly productive period.

David Williamson Shaffer and Golnaz Arastoopour at University of Wisconsin Madison also deserve particular thanks for productive conversations at both LAK13 and subsequently for inviting me to visit UW and spend some focussed time thinking about Epistemic Network Analysis and epistemic commitments. The continued conversation has been incredibly productive for my research directly, and for thinking about career trajectories and professional development – for which I owe a big debt of gratitude.

Finally it would be remiss to neglect my colleagues at the OU – across KMi, CREET, IET, library services, and beyond – I have had fantastically useful technical, theoretical, practical, and social conversations with them which have informed this work, and my PhD experience thus far.

Publications and online material contributing to this report

This report builds on prior articles, presentations, and my MPhil work, including:

The EdFutures Wiki article on Learning Analytics (which I am the primary contributor to) http://www.edfutures.net/Learning_Analytics

My blogs - and useful comments on/arising from them – here http://people.kmi.open.ac.uk/knight/

Knight, Simon. "Epistemology, Assessment, Pedagogy - Learning Analytics?" presented at the SoLAR Storm, online presentation, 2012. http://www.slideshare.net/sjgknight/so-lar-epistemology-pedagogy-assessment-v2

Knight, Simon, Buckingham Shum, Simon, and Littleton, Karen. "Collaborative Sensemaking in Learning Analytics." In *CSCW and Education Workshop*. San Antonio, Texas, USA, 2013.

———. "Epistemology, Pedagogy, Assessment and Learning Analytics." Leuven, Belgium: ACM Press, 2013. doi:10.1145/2460296.2460312.

———. "Tracking Epistemic Beliefs and Sensemaking in Collaborative Information Retrieval." In *Collaborative Information Seeking Workshop*. San Antonio, Texas, USA, 2013. http://collab.infoseeking.org/events/cscw2013workshop.

Knight, Simon, and Littleton, Karen. "Discourse, Computation and Context – Sociocultural DCLA Revisited." In *1st International Workshop on Discourse-Centric Learning Analytics*. Leuven, Belgium, 2013. http://oro.open.ac.uk/36640/.

Knight, Simon, Siemens, George, and Williamson Shaffer, David. "Analytics: Epistemology and Pedagogy." presented at the LAK13 online course, March 19, 2013. http://www.youtube.com/watch?v=c16ggOMNSvM&feature=youtu.be

This work also builds to a lesser extent on my MA thesis, and to a greater extent on my MPhil thesis and assignments submitted for fulfilment of that degree. That thesis is:

Knight, Simon. "Finding Knowledge - The Role of Dialogue in Collaborative Information Retrieval in the Classroom." Master's, University of Cambridge, 2012. (in draft for publication - URLs will follow) – publications arising are in draft as:

Knight, Simon, and Mercer, Neil. "The Role of Collaborative, Epistemic Discourse in Classroom Information Retrieval Tasks." *In Draft*

———. "The Role of Exploratory Talk in Classroom Search Engine Tasks." *Accepted subject to major revisions* at Technology, Pedagogy and Education (n.d.).

Contents

Abstract

This report argues that information seeking – the searching, frequently conducted on search engines such as Google, in order to retrieve information for some needs – should be of interest to education. It further suggests that such interest should focus on information commitments, which are implicated in the ways that people find, and process, information. Building on literature researching collaboration in both education and information seeking research, I claim that Collaborative Information Seeking (CIS) is a good lens through which to research information commitments. Indeed, a key component of this report is a preliminary proposal for a new theory of epistemic commitments, which addresses some concerns with prior research on epistemic beliefs, epistemic cognition, and epistemic (or information) commitments. Two novel components of that new theory are a focus on information trace as the core of epistemic activity, and a focus on the 'dialogic space' as particularly epistemically relevant. The report goes on to propose a technological solution for the analysis of epistemic commitments in the form of a Computer Supported Collaborative Learning/Work (CSCL/W) environment. This proposal includes an analysis of trace for epistemic commitments, and a discussion of relevant discourse centric learning analytics for the analysis of chat data around epistemic activity. While discourse data has received some analysis in CIS research, the analysis has generally been somewhat shallow in its focus; the proposal made in this report is for a deeper analysis, both as an extension of CIS research, and as of interest to learning analytics (and indeed learning researchers generally) who are interested in the collaborative context of learning. The report is thus proposes a relevant research theory to investigate the ways in which people make commitments in online information seeking environments, and how they might be supported.

The report comprises three sections:

1. I start with a literature review, which begins with an overview of some relevant theoretical, and philosophical, literature relating this particularly to learning analytics (section 2). I then introduce the relevant literature on information seeking and epistemic cognition (or, as I propose, commitments) (sections 3-8). Section 9 then introduces some relevant literature on tasks to probe epistemic commitments, while section 10 introduces some core software to do so.
2. Section 11 then discusses my preliminary practical work, with respect to pilot empirical work and some key skills gained
3. Section 12 then introduces my formal research proposal – including research questions, proposed tools, and practical experiments to be conducted.

1. Organisation of the Literature Review

The work presented here foregrounds the potential of learning analytics in relation to the related fields of computer support cooperative learning (CSCL); educational technology; psychology in education; and indeed education research in general. It will address learning analytics and these fields in a specific context – that of collaborative information seeking (CIS) – a context which has not typically been studied in education, but which:

1. Provides an analysis paradigm which may be useful for probing the sharing and joint construction of knowledge through CIS contexts
2. Is an important component of education, being a frequent occurrence in educational and work contexts, yet understudied particularly in educational contexts, despite the known value of collaboration in many educational settings
3. Provide a useful pedagogic tool for building epistemic capabilities – skills and dispositions such as those towards evaluation and critique of sources, fact checking behaviours, peer teaching and so on

Each part of the literature review is distilled into a summary section, with the primary line of argument as follows:

- I first discuss some of the relevant history, and purposes of learning analytics, contextualising it as a field (section 2).
- I briefly note the importance of epistemology and ontology to the field of education, and this work (section 2.1) before introducing the relationship of these concepts, and pedagogy, to learning analytics (section 2.2).
- The final parts of that section (sections 2.2.4) introduce an example of a particular policy – the inclusion of internet tools in Danish exams – which seems to implicate a particular epistemology, and relates this policy to some important educational concepts around information search and evaluation.
- These ideas are further developed in section 3 which explicitly relates the Danish example to the concept of epistemic cognition.
- Models of information seeking are then described (section 4), foregrounding some key analytic elements in various models, and some key factors for educational contexts, including the epistemological assumptions built in to models of information seeking.
- Section 5 then raises a key point – that search does not happen in isolation, and that in many cases information seeking is a collaborative activity. This section thus relates information seeking to this collaborative context, highlighting issues of awareness and communication as particularly important.
- Having further explicated the information seeking process, I return to epistemic cognition in section 6, discussing research on epistemic cognition in information seeking environments.
- Section 7 takes a closer look at this concept of 'epistemic cognition', focussing in on its assumptions and the various models of epistemic cognition, beliefs, and commitments. This section relates this broad area of research back to the earlier discussion of epistemology; where the earlier discussion (see 3 and 4 above) focussed on information seeking in exams as implicating a particular epistemology, this section argues that researcher perspectives on what is happening when engaged in information seeking (e.g. perspectives on epistemic cognition) also implicate particular epistemologies.
- Section 8 returns to learning analytics, exploring some notions of 'trace' and relating learning analytics to important factors in researching epistemic commitments – including the important role of dialogue in understanding student activity, and student collaboration.
- Section 9 then foregrounds tasks which might evoke epistemic activity in collaborative information seeking contexts.

- Building on this discussion, section 10 highlights tools to both support tasks, and capture data for use by collaborators and analysts.

This detailed review provides the rationale for the remaining sections:

- a report on this year's preliminary empirical work which required the learning of new software analytics tools, on a range of datasets, plus associated research collaborations (and at least one draft paper in preparation for journal submission)
- the articulation of three top level research questions, refined into a set of sub-issues…
- …which thus focus the proposal going forward, with an associated work plan and risk analysis.

2. Introducing Learning Analytics, Epistemology and Ontology[1]

The first Learning Analytics and Knowledge conference (LAK11, 2011) defined learning analytics as **"**the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs". The second (LAK12, 2012) made clear that this was not only in educational contexts, but informal and workplace learning too, and that the interest was in learning experiences and success – expanding the scope of what might have been perceived as a focus on 'learning environments' in the formal, and VLE/LMS sense of that term. Recent debate at LAK13 and the LASI13 Learning Analytics Summer Institute has continued to probe whether a more specific definition is needed to differentiate broader educational and educational technology research, from the field's distinctive focus on digital, often real-time data, and computational analysis/visualization techniques.

In "The State of Learning Analytics in 2012: A Review and Future Challenges" Ferguson (2012) tracks the progress of analytics for learning, with a broad progression. That paper described an increasing emphasis and interest in 'big data' in business intelligence, for example through targeted advertising, and collaborative filtering. Opportunity for learning from business intelligence in education has come from an increased interest in Virtual Learning Environments (VLEs), Content Management Systems (CMSs), and Management Information Systems (MIS) for education, marking an increase in digital data regarding student background (often held in the MIS) and learning log data (from VLEs). However, this interest in applying business intelligence techniques to VLEs has raised the question of how to optimise these systems to support learning, particularly when visual signals are absent in online education – how do we know a student is engaged/understanding if we can't see them? This concern is particularly pressing given the pressure to evidence 'progress', show good professional practice, and evidence rational pedagogic decision making also placed pressure on analytics for use by management in both internal and external accountability systems. Ferguson thus notes that analytics become particularly interesting at this stage, not only as a tool for the top levels of a stakeholder hierarchy (governments or institutions) but as something in which teachers hold an important stake given the potential role of analytics in accountability structures. There has therefore been an increasing interest in the pedagogic affordances of learning analytics, for example with social constructivist learning models exploring the use of Social Network Analysis (SNA) for

[1] This section is adapted from the Creative Commons (CC) licensed EdFutures page on Learning Analytics http://edfutures.net/Learning_Analytics (to which I was the primary contributor). Some of this material may also appear on the (also CC licensed) main Wikipedia learning analytics page, to which I have contributed http://en.wikipedia.org/wiki/Learning_analytics - all material here was authored by me.

supporting learning, alongside an increasing pressure to engage with analytics from economic concerns. Yet, despite these pressures and a burgeoning use of analytics in institutions, more thought should be given to the purpose of those analytics and the objectives they relate to (Powell & MacNeil, 2012) including whether they are:

1. for individual learners to reflect on their achievements and patterns of behaviour in relation to others;
2. as predictors of students requiring extra support and attention;
3. to help teachers and support staff plan supporting interventions with individuals and groups;
4. for functional groups such as course team seeking to improve current courses or develop new curriculum offerings; and
5. for institutional administrators taking decisions on matters such as marketing and recruitment or efficiency and effectiveness measures. Powell and MacNeil (2012)

One of the contributions of this thesis is to provide theoretical discussion of learning analytics, relating them to epistemology, assessment and pedagogy in particular, as well as emphasising the potential of the multi-vocality in LA research for transformative practices. The third Learning Analytics conference (LAK13, 2013) raised a focus on 'Dialectics in Learning Analytics', bringing the many voices around learning analytics into the 'middle space' of learning and analytics – a space for the exploration of the link between analytic tools, and learning (LAK13, 2013; Suthers & Verbert, 2013). This middle space served as a 'boundary object' for multivocality (Rosé et al., 2011; Suthers, Lund, Rosé, Teplovs, & Law, Accepted) of approaches – old and new. It is within this 'theme' that my work is situated, bringing together a number of disciplines to understand – within a particular, CIS, context – how our tools and techniques might address learning, and how we can characterise learning. Appropriately enough, the next section is adapted from my paper at that conference (S. Knight, Buckingham Shum, & Littleton, 2013a), and discusses in particular the latter question – regarding what learning, and knowledge, look like (the learning side of the middle space), and why this should matter to learning analytics (the analytics side) – a question to which I now turn.

2.1 The Relevance of Epistemology and Ontology

A fundamental question in education is, how do we know, when someone knows something? This question raises further questions regarding the status of our own access to that information, what it means to 'know' something, and what that knowledge is – how it manifests, what its structure is, and so on. These are fundamentally epistemological concerns. Central to the field of epistemology are questions regarding the nature of truth, the nature of justification, the interrelatedness or complexity of knowledge (as propositions or otherwise), and types of knowledge, e.g. knowing *how* (skills), or knowing *that* (facts). Whatever 'knowledge' is, *"it is uncontroversial, pre-philosophically, that education aims at the imparting of knowledge: students are educated in part so that they may come to know things."* (Siegel, 1998, p. 20). Thus, pedagogy may be seen in part to be the study of *how* to impart this knowledge to students – the science and development of approaches to teaching and learning for knowledge. However, epistemology's relationship to the more familiar concepts of pedagogy and assessment is a topic of educational debate (Davis, 1999; Dede, 2008; Kelly, Luke, & Green, 2008; K. Williams, 1998), and consideration of this in relation to LA is important.

As Greene, Azevedo and Torney-Purta (2008) point out: "If education were merely the act of transferring knowledge from the learned to the learner, it would be a logistical problem at best. Instead, education is a constructive process where learners come to know in their own ways with

their prior experiences, theories and frameworks shaping how knowledge is formed (Phillips, 1995)" (Greene et al., 2008, p. 142). How we understand this process, and the 'objects' it involves (actors, learning resources, 'tokens' of knowledge) is fundamentally a question of epistemology.

In this case, the epistemological stance – which will be further elucidated in the PhD thesis – is not only a positioning act with respect to the methods chosen. It is, beyond that, a principled stance with respect to the sort of education we should work towards. The epistemological stance taken has implications for the sort of empirical work we should do, but it also has implications for: how we understand what we are 'getting at' when we do empirical work; how we understand the results of that empirical work; of what it means both for us – as researchers – to know, and what it means for our students to know.

2.1.1 Section Summary

The pragmatic focus is thus on how meaning is made, and understood *in action,* as such, pragmatism rejects a focus on 'ontology' - the study of the nature of reality and its components – as a route to 'truth' (marking correspondences between things in the world and statements), in favour of its action-oriented perspective. This will be outlined further in the next section in the curricula context, and in Sections 7-8 in the context of methodology. A core claim of this brief introduction is thus that epistemology informs our understanding of:

- Our methods
- Our interpretation of results
- Our curricula, and assessment (through learning analytics)

2.2 Epistemology, Assessment and Pedagogy – the middle space of Learning Analytics

"Assessment is one area where notions of truth, accuracy and fairness have a very practical purchase in everyday life" (K. Williams, 1998, p. 221), assessment sits at the heart of learning, but is controversial. Learning analytics – I argue (S. Knight, Buckingham Shum, et al., 2013a) – implicitly or explicitly implicate particular stances towards epistemology and assessment regimes. Presently, many education systems are predicated on assessment regimes seeking to accredit knowledge and skills gained by students through formal assessments – often exam- based. Proponents of such exams suggest they are the fairest way to assess competence and learning under controlled, reliable, conditions. Assessment, pedagogy and curriculum are fundamentally related (Harlen, 2007), but many regimes of what has come to be termed 'high stakes' testing are criticised. For example, standardised assessments, including the Programme for International Student Assessment (PISA), American Standardised Assessment Tests (SATs) and English National Curriculum assessments (Sats), face myriad problems. Not least among these is that the exams are criticised comprehensively (e.g. (Davis, 1999; Gardner, 2011; Hopmann, Brinek, & Retzl, 2007)) for failing to represent adequately the types of problem people are likely to face in their everyday lives (external validity), and that they fail to represent an adequate conceptualisation of what it means to know – of what knowledge is (internal validity). The latter claim is that, while assessments clearly measure something, a good grade does not necessarily reflect mastery (Davis, 1999). These fundamental issues are highlighted in a significant body of research in the philosophy and psychology of education (e.g. (Davis, 1999; Gardner, 2011; Hopmann et al., 2007)).

It should be noted that, while there may well be an empirical concern here – we have not, yet, managed to reach levels of 'reliability' suitable for high stakes assessment – there is also a philosophical, epistemological claim being made. This claim is that:

1. High stakes testing motivates a desire for highly 'reliable' assessments
2. Highly reliable assessments have to tightly constrain the concepts of assessment, their warrants, and their contexts
3. Such a set of constraints is a serious impediment to the validity of the claim that such assessments measure *knowledge*
4. Because *knowledge* as properly understood is not simply a set of tightly defined concepts, but rather a set of interrelating notions, an understanding of links and of the *use* of knowledge, a 'holist' perspective on epistemology rejects the separation of individual tokens which high stakes testing strives for in a drive for reliability.

2.2.1 Why Worry About Epistemology?

This epistemological concern is strongly related to the wider educational curricula as articulated here by Greene, Azevedo and Torney-Purta (2008) with respect to the history curriculum:

> *most history experts and educators hold the perspective that historical facts are subject to scrutiny. This requires an understanding of the complex nature of narratives and the use of analytic tools such as sourcing (Vansledright & James, 2002) in an effortful attempt to determine which claims are sufficiently justified to be considered knowledge. Unfortunately, policies that govern what students will be taught during primary and secondary instruction often reject a complex view of knowledge and the necessity of justification. For example, a recent Florida state law declared 'American history shall be viewed as factual, not constructed, shall be viewed as knowable, teachable, and testable' (Florida State, 2006)*

this has implications for how assessment and pedagogy are conceived.

Harlen (2007) depicted a triadic relationship between pedagogy, assessment, and practice. Influenced by this, and Katz's (2000) description of "competency, epistemology and pedagogy: curriculum's holy trinity" I depict the triad as in Figure 1[2].

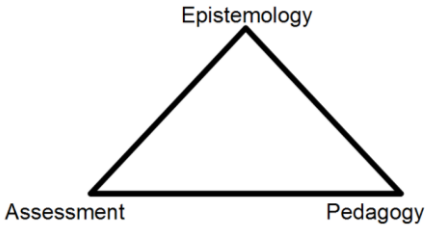

Figure 1: The Epistemology–Assessment–Pedagogy triad

In this picture, epistemology *could* be seen as driving assessments that are aimed at uncovering student knowledge, and driving pedagogy to build high quality knowledge to that end. In this view, assessment is targeted at the learning of high level knowledge – it is assessment *for* learning. However, these relationships are not fixed; neither pedagogies nor epistemologies necessarily *entail*

[2] We could also introduce the notion of 'folk psychology' as a mediating factor between teacher's views on knowledge, and pedagogy – for example, if we hold that some (particular) children will *never* learn x, we are unlikely to attempt to teach it (a pedagogical 'move') regardless of our epistemological stance regarding the nature of 'x' (Olson & Bruner, 1996). Although, in that paper (Olson & Bruner, 1996) Olson and Bruner implicate epistemology in a number of their points regarding 'folk pedagogy'.

the other (Davis & Williams, 2002) (although they may implicate). Furthermore, as I note earlier, and discuss throughout this work, a fundamental issue of assessment is the extent to which, and ways in which, our methods for assessment offer 'access' to students' knowledge states – around which there are philosophical (epistemological) and methodological issues. Some epistemological stances hold that it is not possible to 'map' the knowledge that students hold onto their responses in assessments in reliable and valid ways. This issue is further confounded by the methodological limitations of all assessment methods, and by extension LA. The situation is, therefore, a complex one – which facet of the triad has primacy over the others is not clear in either theory or practice, and may be dynamic according to need and circumstance. However, relationships between the three can certainly be identified, and throughout this work I draw out some of these with respect to LA – which may be conceptualised as a component of assessment. Furthermore, I suggest that, although the relationship may not be a necessary one, assessment regimes do implicate particular epistemological stances.

2.2.2 Our Learning Analytics Are Our Pedagogy

Buckingham Shum (2012) has highlighted the assessment implications of learning analytics for pedagogy when saying "our LA are our pedagogy", arguing that the ways we gather data, interpret them, and act on them implicate, enshrine, and play a role in our pedagogy in action, our understanding of how students learn. The relationship between LA and pedagogy is important because they are both bound up in epistemology – what knowledge is. This section explicitly introduces the relationship between a number of established pedagogic approaches and LA. These are not intended as comprehensive reviews, but rather as brief overviews of how the relationship between pedagogy and LA might be conceptualised.

2.2.2.1 Transactional or instructionalist approach

Transactional approaches hold that learning entails the transfer of knowledge from the knower (teacher) to the learner (student). They are characterized by a perspective on assessment in which success is 'out there', in the degree of correspondence between the claims that learners make, and the facts that they have been taught.

Analytics Implications: LA based on transactional approaches – both in learning, and more broadly – will tend to focus on very simple metrics such as test scores and hit counters, as opposed to any deeper analysis of project outputs or processes.

2.2.2.2 Constructivist approach

Constructivist models hold that learning occurs in the guided experimentation of the learner (student) on the world, typically in classrooms in which such experimentation is age-targeted, and guided by a teacher. Constructivist models are likely to hold a notion of success which highlights construction, with learners experimenting with their environment, and being capable of using tools which are appropriate for their given age.

Analytics Implications: LA with a focus on constructivist approaches of learning will focus on *progress,* particularly through a set of materials, resources or tools selected and arranged by the teacher.

2.2.2.3 Subjectivist or affect based approach

Subjectivist perspectives can be characterised as deemphasizing learning qua academia, in pursuit of personal affect. While individual affect is a concern for educationalists, it is rarely if ever the *overarching* concern in the consideration of learning. However, for example in information seeking contexts, subjectivist approaches are more interested in whether the user is 'satisfied' with the information they have found, than whether the information is 'good'.

Analytics Implications: In tandem with other approaches, LA based on 'subjectivist' approaches are likely to provide motivation assessments for understanding *why* someone is (or is not) undertaking particular actions (see, e.g. (R. Ferguson, Buckingham Shum, & Deakin Crick, 2011)). Such analytics may focus on explicit moves (feedback forms, affect-based semantic markup such as blog tagging) alongside more implicit analysis such as sentiment analysis of communication data.

2.2.2.4 Apprenticeship approach

Apprenticeship approaches are sometimes used in LA with an interest in whether the learner has become part of a community of activity. In this view, success is about 'being part of' a given group; it is bound up in notions of communities of practice – that 'to know x' is to act towards x in some way that is defined by (or reflected in) the behaviours of some community or other.

Analytics Implications: Analytics based on apprenticeship approaches are likely to focus on classifying expert and novice users, and the shift from novice to expert. Such analysis may explore behavioural markers which mirror those made by 'experts', but may not explore the reasons or meanings implicated in such moves.

2.2.2.5 Connectivist approach

Connectivism (2006) claims to highlight a perspective on epistemology which translates into a LA framework. Within this view, learning is about understanding how to connect ideas appropriately, and where to find such information. The suggestion is that in the case of the connectivist knower "the act of knowing is offloaded onto the network itself" (Siemens, 2006, p. 33). Within this perspective then, success is about building connections between ideas.

Analytics Implications: Connectivist approaches use network analysis to explore the 'connectedness' of a learner's knowledge –in terms of both concepts, and social connections. Analytics look at how networks' size, quality and changes over time can serve as proxies for effective learning.

2.2.2.6 Pragmatic, sociocultural approach

Pragmatic approaches (building on for example, Dewey (1998)) hold that learning occurs in the development of – and negotiation of – a mutually shared perspective between learners. Such approaches focus less on truth – where truth reflects facts about the world – than how meaning is co-constructed, and used in context. Pragmatists suggest that, as human knowers, our conception of some given thing is bound up in our understanding of its practical application – and that is all. When we attempt to understand truth beyond such a conceptualisation of practical activity, we are likely to fail. Thus, success is in *use* – the measure of success is how useful the information is for the purposes it is employed; it is socioculturally embedded and mediated, and may be in flux as activities are defined and redefined.

Analytics Implications: Pragmatic approaches have traditionally focused less on assessing the products of learning (except where they are being used *for* something), and more on the process.

Analytics tools in sociocultural approaches encourage learners to reflect on their own activity, in an attempt to understand how they can develop their skills in information processing, in their own particular contexts. Analytics within this approach might attend particularly to quality of discourse for learning, for creating a mutuality of perspectives (D. Edwards & Mercer, 1987) including in collaborative information seeking tasks (Foster, 2009; Hertzum, 2008; Lazonder, 2005). This research foregrounds *how* students interact with information; make sense of it in their context; and co-construct meaning in shared contexts. These are on-going processes which highlight the question of how LA fits into the context of AfL and pedagogy.

2.2.3 Epistemology and LA

The stance we take with regard to the relationship between epistemology, assessment and LA relates to the issue of whether we envisage analytics as a form of diagnosis on the one hand or a kind of biofeedback on the other – is LA (and assessment) the end point of, or a component of, pedagogy. In the former we seek to accredit learning through defining behavioural proxies taken as evidence of knowledge and competencies. LA may also be used to support learners in their own self-regulated learning activities, giving them feedback on changes they make and their impact on learning outcomes, but without – generally – making strong evaluative judgments regarding such changes. The former is thus more closely aligned with assessment *of* learning – often instantiated in high stakes summative assessment, while the latter is closer to Assessment *for* Learning – in which assessment is a continuous process through which formative feedback may be given to further develop the students' learning (see e.g. (Black & Wiliam, 2001; Gardner, 2011)).

The relationships highlighted above serve as general pointers to the sorts of relationships we might see between pedagogy and LA. There I also highlight views on learning, alongside notions of how success may be defined within these approaches; that is, when these systems might accredit knowledge to the student. Fundamentally, this accreditation implicates epistemological stances regarding when knowledge may be claimed (or not). These are general claims, but illustrative of how such notions relate to those of LA, in particular notions of:

- **Mastering curriculum content:** this is the dominant focus of analytics approaches at present, seeking behavioural markers using e-assessment technologies of varying sophistication, in order to generate summaries at varying granularities, for both individuals and cohorts. (Particularly transactional and some constructivist approaches)
- **Evidencing membership and processes:** this approach to LA looks for behavioural proxies which indicate a student is part of a particular subgroup; positive feedback is given towards moving students into 'successful' subgroups, but little attention is paid to the qualities of those groups except instrumentally. (Particularly affect-based, apprenticeship, and possibly connectivist approaches)
- **Success is use:** this approach looks for students developing personal and collective representations of curriculum content, and engagement in sensemaking about not only this material, but also their own analytics. Social Learning Analytics (Buckingham Shum & Ferguson, 2012; R. Ferguson & Buckingham Shum, 2012) in which students are encouraged and supported to do so may work towards this end. (Particularly pragmatist approaches).

These three broad conceptualisations of LA relate to the issue of whether or not we are deemed to consume, discover, or create (internally or/and externally) knowledge – is it 'out there' for us to take, do we need to investigate to find it, or is it formed in our developing understandings of the relationships between entities and the new representations we create in such activities? This is not a

claim about the learning or pedagogy, but a related claim about the status of knowledge, and its assessment.

2.2.3.1 Pragmatism and sociocultural approaches to assessment

The nuance of claims surrounding epistemology and assessment is important. In the introduction I referred to research arguing that conventional exams are designed to maximise the reliability of results, at the cost of straitjacketing what can be defined as learning (poor internal or construct validity) and thus what constitutes evidence of learning (poor external validity). Moreover, if we are to argue that individual tokens of knowledge cannot be identified (and 'owned'), then we should accept that "the content of a specific item of knowledge depends in part on how it is related to other knowledge" (Davis, 2006). Thus, sociocultural setting, interaction, and the purposes for which any artefact or knowledge – in the broadest sense – is being used, are all of fundamental importance in understanding how people make meaning, and learn. Contextual sensitivity is thus a key facet of pragmatist approaches.

Pragmatic approaches, broadly, are likely to focus on the dynamic nature of information needs, and the discourse and other tools and artefacts which mediate our relationship with information in the world. It is not a postmodern approach, in the sense that postmodern approaches take either a relativist approach (there is no fixed truth) or a normative one (the dominant theme is correct at that time) to knowledge, but rather one which focuses on use, and meaning, over mapping of facts to things in the world.

As described briefly above, pragmatic approaches have traditionally focused less on assessing the products of learning, and more on the process. LA in these approaches might encourage learners to reflect on their own contextualised activity, in order to instil an ethos and capacity to become reflective. The key development with the emergence of digital LA is that previously ephemeral processes are now persistent, not just for researchers studying those processes, but for the learners and educators co-constructing those processes. Moreover, the process traces are now amenable to computational analysis which opens new possibilities for assessment and feedback, both formative, and possibly even summative (e.g. where the assessment regime defines those process skills to be an important form of student evidence).

Given the salience of context in this approach, it deserves further explication. As with LA generally, context may be taken as very mechanistic, for example the claim that a person in place/course/role/ability band 'x' should see resource 'y', or other approaches which would include time, topic, or social-group resource discovery. No doubt some of these features will prove useful, and indeed the use of semantic web technology in social learning analytics (R. Ferguson & Buckingham Shum, 2012) may be particularly interesting. However, in addition to temporal, linguistic, aptitude, and geo-spatial markers, I draw attention to the following:

1. The discourse in which, and through which, context is constituted (A. D. Edwards & Furlong, 1978; Potter & Edwards, 2003). That is, I take the discourse to have a multifaceted role in constituting, and helping learners make sense of, the context.
2. Discourse is fundamentally associated with the sensemaking which occurs in respect of any particular task being undertaken; the *use* being targeted is fundamental to context. Stark examples highlight this importance, for example where we ask students to critique versus

summarise a paper we expect rather different outcomes. Assessment regimes which make this explicit may facilitate capture of context around 'doing x for purpose y' LA

3. These assessment systems (2, above), and the broad range of tools, technological and otherwise, which people utilise also act as mediating artefacts impacting on how people perceive their task, and its solution – mediating the context of use.

2.2.4 Epistemology in Action – Policy, and Practice

Consider the following example from Denmark to illustrate the argument that implicitly or explicitly, epistemological assumptions fundamentally shape pedagogy and assessment, and hence, the kinds of LA that one deploys to achieve those ends. In Denmark, a pilot project was conducted permitting the use of the internet (but not communication sites) to support students in five of the school leaver subject exams[3]. This made it possible to set questions requiring the use of multimedia and individual internet search. For example, a student might be asked to write about a poet whom they have not studied (and rote learned about), based on a poem by them and that of a contemporary, a short biography and perhaps an image from the time. They may be given unfamiliar resources, and permitted to source information for themselves from the internet. Thus, while Danish students are expected to evidence 'knowledge-that' – knowledge of facts – they must also exhibit a higher level of 'knowing-how', for example around information processing, synthesis, and metacognitive abilities – which remain unassessed in countries restricting access to external resources which might enhance the student's capability. While this is of course simply one other (controlled) context, the example illustrates how even within a system reliant on exams, those exams might be conducted on a rather different epistemological grounding. Assessment regimes such as the Danish example may be taken to reflect a holistic epistemology in which *how* one comes to know is as important as *what* one comes to know, and in which it makes little sense to pick out individual tokens of knowledge in decontextualized ways (Davis & Williams, 2002; Davis, 1998, 2005; Katz, 2000).

We can contrast such assessments with high stakes testing regimes whose construct validity and external validity have been questioned. For instance, Davis (Davis, 1999, 2006) argues that such instruments neither assess those facets of learning they set out to test, nor those facets of learning which would likely be utilized in the everyday deployment of knowledge in any particular domain. Davis has argued that high stakes testing is inadequate for understanding learning, in so far as its construal of that learning is necessarily restricted by a desire for highly reliable metrics of success. As such, it must exclude the nuanced understanding of student meaning-making, and the social context in which learning occurs, and how knowledge is constituted and enacted. He argues that this, as opposed to acquisition, is the appropriate way to talk about knowledge. Davis draws on notions of situated cognition (Salomon, 1996) and sociocultural approaches (Säljö, 1999) – particularly Säljö's "Literacy, Digital Literacy and Epistemic Practices: The Co-Evolution of Hybrid Minds and External Memory Systems" (Säljö, 2012). Säljö highlights that:

> *From the learning and literacy points of view, such tools [memory aides and knowledge management systems of various sorts] imply that users' knowledge and skills, as it were, are parasitic on the collective insights that have emerged over a long time and which have been entered into the instrument in a crystallized form: algorithms, grammatical rules and concepts, etc. The user will manipulate the artificial memory system in a*

[3] Steen Lassen (a Danish Education Minister) on the piloting of internet access in exams: http://vimeo.com/8889340 subsequently adopted by some Danish universities (Cunnane, 2011).

> *number of ways in order to see what comes out of the processing that goes on in the machine (Säljö, 2012, p. 14)*

However,

> *Engaging with external memory systems thus requires familiarity with a varied set of epistemic practices that range from deciphering letters on a page through familiarity with meaning-making in relation to discourses and genres of texts and other media, to meta-knowledge about how such resources may be used. (Säljö, 2012, p. 12).*

Säljö is making an epistemological claim, specifically, a sociocultural, pragmatist claim: that there are important literacies and practices to be mastered in learning; that those should themselves be objects of assessment; and language and discourse are critically implicated in our grasp of the world. Such an epistemology has implications for how we teach, what we assess, and which analytics techniques might be deployed. 'Success' can no longer be defined as a matter of regurgitating, unaided, the correct information in a two hour exam. Such an epistemology also offers a perspective on why it is that, even in those technologically advanced societies which assess knowledge in less abstracted, socially embedded ways – such as Denmark – information seeking and processing via the internet and search engines is a significant area of difficulty for students (Undervisningsministerie (Ministry of Education) & Afdelingen for Gymnasiale Uddannelser (Department of Secondary Education), 2010, p. 15); namely, that although this provides some wider access to information, this does not equate to knowledge. Student engagement with information should consider both the kinds of knowledge which we might call transferable competencies or skills – including those higher order skills often known as metacognitive abilities – and more propositional or fact based knowledge. In this context, we might consider information management and seeking not only as a means to an ends, but as a way to encourage interaction with a complex network of information. As argued by Tsai, as not only:

> *…a cognitive tool or a metacognitive tool; rather, it can be perceived and used as an epistemological tool. When the Internet is used as an epistemological tool for instruction, learners are encouraged to evaluate the merits of information and knowledge acquired from Internet-based environments, and to explore the nature of learning and knowledge construction. (C. Tsai, 2004, p. 525)*

In this conception, learners are encouraged to think about the context, reliability, validity, certainty, and connectedness of knowledge. As Lloyd pointed out "Understanding information literacy as a catalyst for learning necessitates a move away from exploring textual practices towards incorporating an understanding of the sociocultural and corporeal practices that are involved in coming to know an information environment." (Lloyd, 2007). If there is a link between poorer ability to find and access information on the web, and income, race and education (DiMaggio, Hargittai, Celeste, & Shafer, 2001) – which as Hargittai (2010, p. 4) points out, is a consistent finding internationally[4] – then the concern is that access alone is not enough when it comes to reaping the potential payoffs of being online (Hargittai, 2008). It is evident that use of the internet is not without difficulty. It is also apparent that there is a "second level digital divide" (Hargittai, 2002)

[4] See Bonfadelli (2002) for Switzerland; Livingstone and Helsper (Livingstone & Helsper, 2007) for the UK, and Zillien and Hargittai (Zillien & Hargittai, 2009) for Germany. The first of these in particular considers the 'Knowledge Gap' – the finding that more educated users tend to use the internet more for information, while less educated users tend to use it for entertainment purposes; this is an interesting finding in light of the concept of knowledge offered here.

between skilled and less skilled users of the internet, such that it is more useful to some than others – and that this is related to socioeconomic factors. Thus:

> *…rather than perceiving the digital divide as a problem of equal access to technology, an alternative construction defines the digital divide as a literacy issue. From this perspective, information technologies are viewed as cognitive and cultural tools used to manipulated symbols and share meaning* (Ba, Tally, & Tsikalas, 2002, p. 4).

The systems within which these issues exist are also important. Issues regarding access to the internet and its resources relate at least to the purposes for which the internet is employed, and the ethical concerns regarding unequal access to such resources.

With regard to the types of knowledge present under analysis of search engine use, I take it that there are at least two types – fact based, propositional, or 'knowing-that' knowledge; and what is variously called procedural knowledge, 'knowing-how' or skills. These map roughly to the 'knowledge-about' and 'knowledge-of' respectively, discussed by Scardamalia and Bereiter (2006). Broadly, the type of knowledge retrieved from any given search is propositional, and can be analysed in terms of suitability for answering any particular question.[5] The accessing itself can be characterised as the 'knowing-how'. While certainly the former is important for understanding the latter, many educational contexts – and indeed, our assessment system – seem to focus more on this propositional, fact based, knowledge over and above often rather complex metacognitive strategies involved in the seeking and manipulation of such facts. Research should explore what skills are actually utilised in search tasks in classroom contexts. This is particularly important if we are to aim towards:

> *Instead of a system aiming at the reproduction of knowledge, new learning [is] aiming at learning outcomes that are durable, flexible, functional, meaningful, and applicable. Active pedagogical methods, in which students learn by doing instead of listening and in which the teacher has a guiding role, fit this new learning (Walraven, Brand-Gruwel, & Boshuizen, 2008, p. 624)*.

The second philosophical concern regards the 'digital divide', and the intuitive ethical issue it raises. The importance of research into the digital divide is particularly pertinent given that – not just access, but meaningful 'successful' access – confers benefits (Hacker & Mason, 2003). Internet use is not simply about access to a sort of propositional knowledge. It is also about knowing *how* to access that knowledge, and induction into a different kind of community of practice (Lave & Wenger, 1991):

> *The Internet boosts immeasurably our collective capacity to archive information, search through large quantities of it quickly, and retrieve it rapidly…Internet access is an important resource and inequality in Internet access is a significant concern for social scientists who study inequality.*
>
> *…*
>
> *We agree that inequality of access is important, because it is likely to reinforce inequality in opportunities for economic mobility and social participation. At the same time we argue that a more thorough understanding of digital inequality requires placing Internet access in a*

[5] In subsequent (thesis) work I will briefly discuss the concern here that the internet is somehow 'taking away' our native capacity to remember in light of my MA in philosophy of education (S. Knight, 2011b) and an article I wrote for the Psychology Postgraduate Affairs Group magazine – The Quarterly (S. Knight, 2011a).

> *broader theoretical context, and asking a wider range of questions about the impact of information technologies and informational goods on social inequality.* (DiMaggio et al., 2001, p. 2)

Within our context, the aim is education. Of course, there is no easy way to define this field, and in particular the types of knowledge which a system wishes to impart. However, if the use of the internet, and in particular information seeking systems such as search engines are implicated in the appropriation of various types of knowledge – both propositional and skill based – then it seems there is an ethical duty to study them. This claim is independent of, although strengthened by, concerns regarding the digital divide.

2.2.4.1 Critical Skills for Information Seeking

Crucially, pre-moderated reading lists and library materials are no longer the resource of choice for either students or educators, who expect, and want, to be able to find and use information on the web. However, despite the prevalence of internet use, many experience difficulties in their web based information-seeking activities:

> *Searching and processing information is a complex cognitive process that requires students to identify information needs, locate corresponding information sources, extract and organize relevant information from each source, and synthesize information from a variety of sources….However, IPS [information-problem solving] has been given little attention in schools, and instruction in this skill is rarely embedded in curricula. And yet, by giving students assignments in which students have to solve an information-based problem, teachers assume that their pupils have developed this skill naturally. (Walraven et al., 2008, p. 623)*

Information seeking, in particular via search engines – can be seen both as a tool in its own right, and a way to encourage further interaction with a complex network of information; the World Wide Web, as noted above, this has important implications. Although search engines are commonly used by young children and teenagers, many report some issues with finding information (Livingstone, Bober, & Helsper, 2005, p. 10). A recent review (Bartlett & Miller, 2011) paints a bleak picture, indicating that a quarter of 12-15 year olds make no information quality checks at all, that they tend to emphasise aesthetics over quality, and that they take the inclusion of websites on search engine results to be an indicator of their veracity. Worryingly, in a broader age range (9-19 year olds) two thirds also claimed to have never been taught how to judge the reliability of the information they find, while over half of teachers were concerned that their students did not understand how to conduct searches (Bartlett & Miller, 2011).

Again with a broader age range – from school to university students – a 2007 report by the UCL based Centre for Information Behaviour and the Evaluation of Research (CIBER) reviewed 86 peer review articles, including age comparison studies, observational accounts and historic (pre-internet) studies, totalling 86 papers, including 7 reviews (P. Williams & Rowlands, 2007). A few pertinent themes emerged from this literature on young people's information seeking behaviours, including that "contrary to the popular view, there is little evidence that young people are expert searchers, or even that their search prowess has improved with time" (P. Williams & Rowlands, 2007, p. 9). They report studies – from 1970 on – that find students struggle to find appropriate terms to use, tend to use terms which are obtained directly from task instructions, and fail to 'open' or analyse results which do not contain those search terms. The studies also suggest that students tend to have

difficulties in reformulating searches – failing to see search engine use as an iterative process and instead attempting the same search more than once. Further down the 'search stream' issues were still present; students – particularly younger ones – tend not to evaluate sites effectively, yet may still use sites that they do not understand in order to claim they have 'completed' tasks.

Similarly, Walraven, Brand-Gruwel, and Boshuizen (2008) reviewed the literature, since 1995, on problems that children encounter when engaging in information seeking. Importantly, they found that "children, teenagers and adults have trouble with specifying search terms, judging search results and judging source[s] and information. Regulating the search process is also problematic." (Walraven et al., 2008, p. 623). It is striking that despite the slightly different focus of these two reviews (with only 5 of the 15 studies Walraven et al., review focussing on children covered in the Williams and Rowlands review) the key findings are similar across reviews. These findings may be age and income related, with older children, and those from a middle class background tending to have better searching skills, and information checking behaviours – leading them to also have higher levels of trust in internet sources (Livingstone et al., 2005, p. 10), however, fundamentally the literature suggests students of all ages experience issues in their information seeking and evaluative practices.

2.2.5 Section Summary

To summarise, this section has argued that a consideration of epistemology is important to LA in two related senses:

- The ways that we assess, the sorts of tasks we set and the kinds of learning we believe to take place (and aim for) are bound up in our notions of epistemology. LA are not objective or neutral: data does not "speak for itself" but has been designed by a team who, implicitly or explicitly, perpetuate the pedagogical and epistemological assumptions that come with any assessment instrument.
- The Danish example shows concretely how epistemology relates to assessment regimes. When knowledge is seen as something that can only be evidenced in contextualised activity, and when it is embedded in one's physical and digital environment, the role of the internet is redefined as a metacognitive tool which cannot be excluded in assessment.

These epistemological considerations foreground the quality of a student's enquiry processes as important, not just whether they get the right answer. Analytics that offer insight into these higher level processes are likely to be significant levers for change in the educational landscape. Even in Denmark, information seeking is seen as a difficulty for students. This is an important skill generally, for example as outlined above Säljö and Tsai both argue for the importance of epistemic practices (Säljö) and web environments as an "epistemological tool" (Tsai) – that is, there are theoretical reasons for exploring the use of internet tasks as epistemically interesting, and empirical ones relating to the utility of such tools and the relative deficit in students skills in the domain. Despite this the use of internet based tasks as epistemological tools has not been well explored, particularly in collaborative contexts – this work addresses that gap. It is thus that this report now turns to this issue of information seeking as an epistemic act, as the focal point of my empirical research, and a good example of the kinds of process learning analytics could focus on.

The following sections thus:

1. Briefly introduce epistemic beliefs in the context of information seeking

2. Explain the relationship of epistemic beliefs to the broader notion of epistemology
3. Describe the significance of epistemic beliefs in the context of the triad outlined above

3. Epistemic Cognition and Information Seeking – an Introduction

One facet of students' dynamic interaction with the world of information relates to how they conceptualise the information they require to answer any particular question – their *epistemic cognition* regarding the nature of the question, and how it may be answered. Source seeking, selection, evaluation, and decision making regarding task completion implicate the actor's epistemic beliefs – their beliefs about knowledge and knowing – which must be brought to bear both on individual items of information, and their relevance to task completion (Bromme, Pieschl, & Stahl, 2009). They have thus been conceptualised as 'internal conditions of learning' embedded into self-regulation as facets of metacognition (Bromme et al., 2009). Indeed – as shall be discussed further below – recent evidence suggests that students do spontaneously reflect about knowledge, and knowing, in online information searching (Mason, Ariasi, & Boldrin, 2011).

This is particularly relevant given that, students have difficulties evaluating information (Van Strien, Brand-Gruwel, & Boshuizen, 2012) and even teachers with more advanced epistemological beliefs utilise more sophisticated search strategies (P.-S. Tsai, Tsai, & Hwang, 2011). Furthermore, in information seeking activities those students with more advanced beliefs do engage in spontaneous sourcing. In particular, more advanced students are more likely to gather trustworthy sources in controversial contexts, and – when they engage in evaluative behaviours – are more likely to trust unbiased and less likely to trust biased sources (Anmarkrud, Bråten, & Strømsø, In Press). Furthermore, even while controlling for prior knowledge and text comprehensibility, students who believe in personal interpretation are less likely to trust documents, and those who believe claims should be evaluated are more likely to trust scientific documents than those relying on experience (Helge Ivar Strømsø, Bråten, & Britt, 2011). In addition, better learners engage in more sense-making on reliable sites than unreliable and by a larger margin than poorer learners (Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012). Indeed, there is evidence that more sophisticated epistemic beliefs are generally associated with more productive learning strategies (Bromme et al., 2009; Schreiber & Shinn, 2003).

The sorts of assessment, and pedagogy, which students are exposed to will relate to the types of epistemic challenge they encounter in their education – systems with a focus on 'right answerism' and limited access to external epistemic resources offer fewer opportunities for challenging knowledge claims (Davis, 1999; Katz, 2000). This paper thus talks about two related concepts:

1. **Epistemology:** Which I introduced above, and is related to the philosophical analysis and conceptualisation of curriculum content and assessment for knowledge
2. **Epistemic Beliefs:** Which I now introduce, and relates to the intrapersonal, psychological conceptualisations that individuals hold regarding knowledge

Indeed, a key component of Assessment for Learning (AfL) may be the disambiguation of the epistemic requirements of questions – in terms of understanding the question, its context, and the knowledge required to answer the question (Black & Wiliam, 2009).

With respect to the construct I have referred to as epistemic beliefs above, it should be noted that the literature refers various to: epistemological beliefs; epistemic beliefs; epistemic cognition; and

epistemic commitments. In later sections I shall draw out some nuance between the approach I take – around epistemic commitments – and other labels. For the purposes of this report I have generally opted to refer to 'epistemic beliefs' when talking about epistemic cognition and epistemological beliefs research (although quotations retain their original terminology); epistemic commitments, however, are discussed as such throughout the report.

Table 1 indicates four dimensions of epistemic beliefs, for which there is general agreement across the various models of belief[6]. These dimensions are useful to consider in relation to student understanding of knowledge domains. For example, in the context of search engine tasks, "epistemological beliefs are a lens for a learner's views on what is to be learnt" (Bromme et al., 2009, p. 8). In such tasks, student search activity may be analysed using the dimensions in Table 1 (e.g. (Mason et al., 2011)), providing a lens onto students' understanding of their own learning, task demands, and how to meet those demands.

Table 1: Dimensions of epistemic belief*

Dimension	Description
Certainty of knowledge	The degree to which knowledge is conceived as stable or changing, ranging from absolute to tentative and evolving knowledge
Simplicity of knowledge	The degree to which knowledge is conceived as compartmentalised or interrelated, ranging from knowledge as made up of discrete and simple facts to knowledge as complex and comprising interrelated concepts
Source of knowledge	The relationship between knower and known, ranging from the belief that knowledge resides outside the self and is transmitted, to the belief that it is constructed by the self
Justification for knowing	What makes a sufficient knowledge claim, ranging from the belief in observation or authority as sources, to the belief in the use of rules of inquiry and evaluation of expertise

*adapted from (Mason, Boldrin, & Ariasi, 2009, p. 69)

3.1 Section Summary

Epistemic beliefs are thus one example of the type of construct which sociocultural LA may probe. However, they are also a particularly *good* example given epistemic beliefs' relationship to our everyday dealings with the world of information, and their relationship to pedagogy, assessment, and classroom practices (B. K. Hofer, 2001).

In section 2.2.4 I noted that, despite an exam system which emphasises the finding and use of information – which I suggested was grounded in a different epistemology to our own assessment system – students in Denmark still struggle with these skills in the context of search engine tasks. Furthermore, I noted that – as Black and Wiliam, amongst others – have stated, while we should seek to ground our assessment in epistemological theorising, we should also understand that, on the intrapersonal level, students will also have particular perspectives – epistemic beliefs – regarding the tasks they are asked to engage with. Furthermore, I noted some strong theoretical reasons for thinking that this sort of construct – epistemic cognition, or beliefs – is a good candidate for study under the sort of epistemology I put forward in the introduction.

In this section I have introduced the construct of epistemic beliefs, as one of interest to learning analytics, to a particular epistemology, and to 21st century learning in the context of the important skill of information seeking. I will now outline some theories of information seeking, before

[6] See (Schraw, 2013) for an extensive review of the multiple theoretical frameworks

indicating an issue with some of these theories, an issue which also raises a complication for epistemic beliefs research – that of collaboration information seeking.

4. Information Seeking; Defining Needs[7]

When we seek information, and make judgements about whether information we find meets our needs, we are making epistemic judgements. Within information and computer science disciplines various theories of information retrieval and seeking exist, with one recent theory noting that, "while computer science sees information retrieval as an information – or answer-finding system, focused on the user finding an answer, an information science or user oriented theory of information need envisages a knowledge formulation/acquisition system" (Cole, 2011, p. 1216). The 'middle space' notion is helpful here since our interest is in learning analytics (not simply search analytics), it is important to consider models of both systems and users insofar as they are related to learning, and analytic techniques which might lend themselves to the exploration and support of that learning. This section thus introduces a number of models of various types, ending with a summary of useful facets explored by each.

4.1 Information Seeking Models

4.1.1 The Classical Account

As various authors have described (see for example, Hearst, 2009b, Chapter 3), the classic model of information retrieval involves:

1. Identify an information need
2. Query system
3. Examine retrieved results
4. If needed, modify and repeat

At its most basic, such a model is not user-centric, but rather based on how systems model documents and queries – results are associated with keywords or metadata, these are indexed, and that indexed can be searched (or traversed). The model is technocentric because the steps are deemed unproblematic, and its description ends at the point when results are returned: a simple search system has no access to what sense the user makes of the results (other than logging that they clicked on a search result).

4.1.2 Cognitive Models of Information Seeking

However, more cognitively oriented frameworks develop this notion further, with Marchionini and White's (2007) description involving users:

1. Recognising a need for information
2. Deciding to try to fulfil that need
3. Formulating the problem
4. Expressing the information need in some search system

[7] There are a number of comprehensive overviews of, for example, the literature around information search models and: psychological relations (Dinet, Chevalier, & Tricot, 2012); search interface and design (Hearst, 2009a) – whose structure I partially adopt in this section – and (Russell-Rose, 2013); and modern developments and new evaluative frameworks (Lewandowski, 2012).

5. Examining the results returned by the system
6. Reformulating the problem and its expression, and finally
7. Using the results

Other models further emphasise users as agents in the process, for example Sutcliffe and Ennis (1998) describe four activities in information seeking:

1. Problem identification – in terms of complexity, intended target, specificity of expression
2. Articulation of needs – which takes the problem identified and restricts it to high level concepts and semantic propositions
3. Query formulation – which takes the articulation and transforms this into keywords and query syntax
4. Evaluation of results – which takes the returned results and evaluates as triggered by the volume, relevance and precision of the returned results

While these models might seem similar to the classic model described above, it differs in at least two key respects: firstly, in all actions the users' domain knowledge and system knowledge impact on their behaviour; secondly, Sutcliffe and Ennis proposed strategies within each activity for efficient search – although these are idealised, and thus do not model actual non-expert searcher's strategies. However, while attempting to elucidate some of the user-activities in interaction with the system, it is still a very linear, and system-oriented, model.

4.1.3 The Berry Picking Model

In contrast, the 'berry-picking model' (Bates, 1989) has the same basic model as those described above, but in contrast to a linear process, it describes a traversal path through queries and reformation – often depicted as a meandering line, with queries and document collections offset at various points along the way. As Hearst (2009a) points out, there are two key benefits to this model over more traditional linear models:

1. It provides a description of searcher's process throughout searches in which information needs are altered and partially fulfilled by information encountered in the process of search – and thus the nature of shifting information needs
2. It does not seek to model the end point of search processes in the form of a final retrieved set of documents, but rather as the bits of selected information found along the way.

4.1.4 Information Needs

As described in the introduction to this section, models such as the 'berry picking' model mark a shift from command-answer oriented systems analysis, to theory in which we assume user's needs are manifested in, and formulated through, the asking of questions based on a starting need (Cole, 2011). Cole (2011, 2012) has thus proposed recasting the issue of information seeking in the light of a theory of 'information needs'. In this theory, the process describes users who must: consider their need; relate this need to concepts; and then consider the mapping of these concepts to search engine key-terms which might return results to satisfy their need (Cole, 2011).

> *Unlike the need for food, water, or shelter, or any of the other primary human needs, what is required to satisfy an information need is often not known to the individual concerned. This makes important the context or information-situation of the user from which the information need arises (where the primary needs for social, economic, and physical survival are being played out). And there is a question as to whether it is a*

primary human need at all, but rather only a secondary need, and must therefore be contextualised in the user's situation in order to be meaningful" (Cole, 2011, p. 1216).

It is thus a perspective which understands the situated context within which information needs arise, and are in mediation – it has a focus on "a user's innate mechanism to generate knowledge formation while seeking, finding, and using information during information search" (Cole, 2011, p. 1220).

4.1.5 Information Seeking Process (ISP) Kuhlthau

Related to this focus on users as contextualised, task-oriented and acting with 'needs' in mind is the work of Kuhlthau (1991) who identified six stages of information seeking – crucially beginning with recognising a lack of knowledge, and thus an information need (as opposed to taking the information need as a given). Kuhulthau developed her model from field and case studies with student library users – including a field study of 385 students. Kuhulthau's model thus divides the process into:

1. Initiation – recognising a lack of information, and thus seeking to address this in the context of prior knowledge and task understanding. Feelings of uncertainty.
2. Selection – Selecting the topic or approach to take, dependent upon requirements and constraints. Feelings of optimism.
3. Exploration – Explore the topic, which may lead to further confusion with the aim of extending understanding and – at the next stage – refining. Feelings of confusion, frustration or doubt.
4. Formulation – The turning point in exploration comes when some resolution is met with respect to specifying needs and addressing conflicting information. Feelings of clarity.
5. Collection – Searches are used to collect relevant information. Feelings of confidence.
6. Presentation – Final searches conducted and further searches become less productive. Feelings of relief, satisfaction or disappointment.

Kuhulthau's analysis of this process found that in particular many students struggle at the 'exploration' stage, and find 'formulation' challenging, confusing, and frustrating. While this study is somewhat older and tools have advanced since then, the relevance of *search* here as an exploratory, needs-defining stage is key. It is also important to note that, although the affective element of the model is part of its 'user-centric' focus (and derived from Kuhlthau's research) these specific feelings may not be reflected by all users, and indeed may not be related to (e.g., motivational for) progression through the stages of information search.

Indeed, in a study on Kuhlthau's information search process (ISP) and epistemological beliefs Whitmire (2003) showed that epistemological beliefs impact on the various stages of ISP. In particular, Whitemire suggests that epistemological beliefs did not impact stage 1 (task initiation – operationalized as essay topic selection), but did impact stage 2 (topic selection – writing a sentence for topic approval from supervisor), stage 3 (basic exploratory search), stage 4 (asked to identify a point where the project 'turned'), stage 5 (consolidation of conflicting sources), and stage 6 (presentation of paper) – although many were deemed not to reach this stage. However, it should be clear that the operationalization of stages as related to epistemic constructs was closely matched here – and it is thus unsurprising they were somewhat related. Conversely, elements of the operationalization may have restricted the perceived impact of epistemic constructs – in particular,

the first stage is so closely aligned with the second that it may have been challenging for researchers to distinguish separate epistemic impacts in these two stages. The retrospective self-report nature of the this study also raises a note of caution given that students may have found it challenging to remember particular activities or cognitions at the identified stages, and indeed may not be able to accurately describe all such relevant information in any case.

4.2 Information Problem Solving Model

4.2.1 An introduction to the model

Building on information seeking work, Brand-Gruwel, Wopereis and Walraven (2009) developed a model of information-problem solving (IPS) on the internet (IPS-I), validated via the analysis of 48 participant talk-aloud protocols from four groups of participants (psychology freshmen and PhD students, trainee teachers, and secondary school students – mean age 14.22). They suggest (Figure 2) that "the IPS-process consists of five constituent skills: (a) defining information problem, (b) searching information, (c) scanning information, (d) processing information, and (e) organizing and presenting information" (Brand-Gruwel et al., 2009, p. 1207). Furthermore they suggest that effective IPS involves self-regulation, "During the process they have to monitor, steer, and check whether the proposed plan is still the right one, or decide if changes in the approach are needed" (Brand-Gruwel et al., 2009, p. 1209)

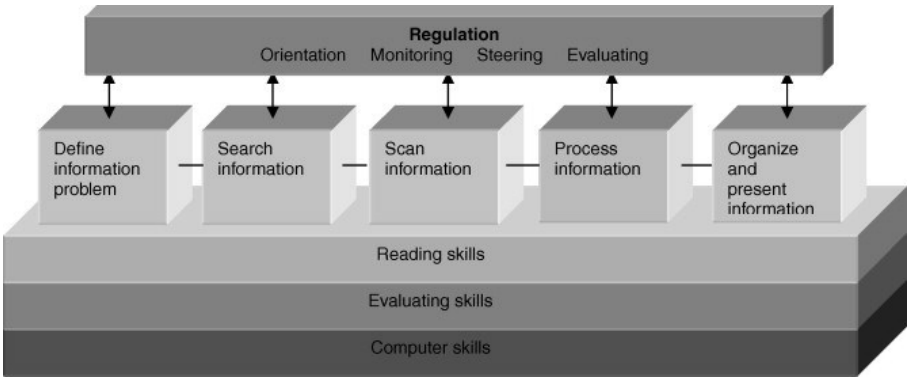

Figure 2 - The Information Problem Solving using the Internet Process (IPS-I) (Brand-Gruwel et al., 2009)

In 2011 a special section of *'Learning and Instruction'* was published on "Solving information-based problems: Evaluating sources and information", in which Brand-Gruwel and Stadtler point out the importance of being able to define problems, and search, select and synthesise information towards those problems in both educational and non-educational contexts such as looking for health information (Brand-Gruwel & Stadtler, 2011). They cite earlier work (Brand-Gruwel, Wopereis, & Vermetten, 2005) in which a model of information problem solving skills, to be executed in iterative cycles, was defined in which students:

a) Define information problem
b) Search information
c) Scan information
d) Elaborate information
e) Organise and present information (Brand-Gruwel & Stadtler, 2011, p. 176)

In the same issue it was noted that, the heart of the problem is "evaluation of information" (Wopereis & van Merriënboer, 2011) – however, of particular interest is that in their original IPS research, it was noted that "[t]he main difference between the experts and the novices is that

experts pay frequent attention to the (re)formulation of the problem while this is completely ignored by novices" (Brand-Gruwel et al., 2005, p. 503) in addition to engaging in more regulatory strategies, and making more inter-textual links. Again, the implication is that more advanced epistemic strategies – rather than simply better search techniques – are associated with improved outcomes.

4.3 Information Seeking – Additional Factors

4.3.1 Exploratory Search

A key element of understanding information seeking in learning contexts is understanding not just the processes of using the system, but the sorts of tasks the system is being used to accomplish. One factor here is the type of search users engage in:

> *A hierarchy of information needs may also be defined that ranges from basic facts that guide short-term actions (for example, the predicted chance for rain today to decide whether to bring an umbrella) to networks of related concepts that help us understand phenomena or execute complex activities (for example, the relationships between bond prices and stock prices to manage a retirement portfolio) to complex networks of tacit and explicit knowledge that accretes as expertise over a lifetime (for example, the most promising paths of investigation for the seasoned scholar or designer)* (Marchionini, 2006, p. 42).

Some of these notions are illustrated in Figure 3 below.

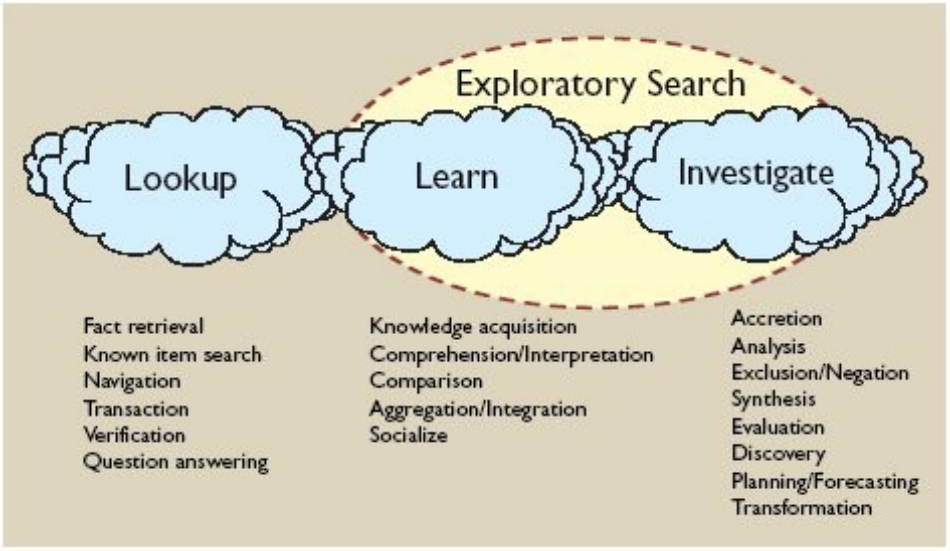

Figure 3 - A Taxonomy of Search Task Types (Marchionini, 2006, p. 42)

Marchionini associates these search-types with levels of Bloom's taxonomy of educational objectives (Bloom, 1956) – a common tool for conceptualising classroom questions in a hierarchy of complexity in educational settings. In particular, he notes the relationship between relatively low level questions – e.g. when was Freud born? – and 'lookup' search tasks where a single query can deliver a single correct result, as opposed to more complex learning and investigation based questions, which relate to 'exploratory search'[8]. Marchionini notes that 'learning' searches, where learning is

[8] See http://people.kmi.open.ac.uk/knight/edusearch-tips/ for some example questions and search tasks organised around Bloom's taxonomy, alongside some further tips for using search engines in school classrooms.

taken in the broad sense to include lifelong and self-directed learning, involve iteration, managing multiple resources and making judgements on those, and comparison:

> *Much of the search time in learning search tasks is devoted to examining and comparing results and reformulating queries to discover the boundaries of meaning for key concepts. Learning search tasks are best suited to combinations of browsing and analytical strategies, with lookup searches embedded to get one into the correct neighborhood for exploratory browsing* (Marchionini, 2006, p. 43).

Thus, these are "searches that support learning aim to achieve: knowledge acquisition, comprehension of concepts or skills, interpretation of ideas, and comparisons or aggregations of data and concepts" (Marchionini, 2006, p. 42).

4.3.2 Information Seeking – 'Success' Involves Epistemic Assumptions…

Marchionini's work on information seeking raises interesting issues regarding 'success' in search activities. If one seeks some token of factual information, ascertaining success – whether determining one's own success, or systems determining the success of their users – this task may be relatively easy. In the context of exploratory search this is not so.

Sundin and Johannisson (2005b) offer a review and description of the broad approaches and epistemological models used in the study of information retrieval, (see Table 2), to which studies focussing on affective aspects of user experience – a 'subjectivist' epistemological approach – have been added. This table illustrates some theoretical advances that have been made in the field, and some issues with these approaches.

The table should draw readers to Sundin and Johannisson's pragmatic position influenced by pragmatists such as Dewey (1916), and neo-pragmatists such as Rorty (1981) and McDowell (1996) – which, as should be clear from the earlier sections of this report is strongly related to my own approach. This approach recognises that consideration of the usefulness of knowledge and language 'in action' at work in the world, is preferable to trying to get at the 'real world'. As such, the focus shifts from verification of correspondences between linguistic labels and 'things in the world', to the ways in which knowledge and language acts on and in the world.

Table 2 – An overview of approaches to the study of Information Seeking (adapted from Sundin & Johannisson (2005a))

Approach	Epistemology	Theoretical Perspective	Methodological interest	Methodology	Methods	Implication for Information Seeking	Limitations
Structure Approach	Realist	Positivist	"…emphasizes the shaping of human behavio[u]r as the result of social structures in society, such as class, education, gender or ethnicity."	Survey data, forms of experiment (particularly focussed on the computer/server side systems as opposed to the user)	Statistical analysis of queries matched by results. Analysis for 'noise' in search results. User analytics including service use statistics. Log data. Experiments. Eye-tracking.	Information needs are expressions of the structures from which those needs arise; they are objective to the domain, and can be met. Users are aware of their needs. Context (structures) provide the frame/limits of action	Too much focus on systems at expensive of individuals. Assumptions regarding objectivity of information needs and access to these.
Individual Approach	User-centred, closer to constructivism	Broadly interpretivist	How do users 'make sense'/construct meaning.	Ethnography, action research, discourse analysis, experimental research	Log data, experiments, surveys, structured interviews, eye-tracking, survey & (cognitive) psychometric instruments	Information needs seen as a 'deficit' with an interest in a) ability to meet this and b) affective aspects of motivation and satisfaction. Needs develop over time.	Too much focus on individuals rather than their sociocultural context. Could be considered 'idealist'.
Affective/ motivational approaches	Subjectivist	Phenomenological, introspection, post-modernist,	Interest in when users are *satisfied* & frustrated by their knowledge.	Some ethnography, phenomenology, discourse analysis, survey methods	Measures of positive affect, self-report measures of satisfaction, self-efficacy measures, studies of workplace motivation, observations,	Information needs should be 'satisfied' by the systems designed to meet them. It is in a sense pragmatist – needs are met when positive affect is high; however, the focus on individual affect is subjectivist.	Too much focus on affect over other aspects of cognition, system design, and 'truth' or 'good' knowledge.
Communicative Approach	Pragmatist/neo-pragmatist	Interactionism, activity theory. sociocultural	How is knowledge useful, to whom and for what purpose – for what actions is it deployed	Discourse analysis, experimental approaches, ethnography,	Discourse analysis, eye tracking, quasi and field-experiments, observations	Information needs should be considered as they relate to communities of justification, and the purposes for which knowledge is deployed (e.g. practical v. medical nursing knowledge). Information not seen as transferred, it is part of a sociocultural, dialogic, toolkit.	Too broad - sometimes we may wish to look at only individuals, or structures. Some concerns with how it defines 'truth'.

Importantly, "…information seeking is not carried out for its own sake but to achieve an objective that lies beyond the practice of information seeking itself." (Sundin & Johannisson, 2005b, p. 107). Within this pragmatist socio-cultural epistemology:

> *…judging the truth of an idea becomes a question of whether the idea makes any difference to practice or not, whether the idea provides us with a useful tool or not. (Sundin & Johannisson, 2005a, p. 27)*.

That is, analysis cannot focus solely on whether some clearly defined need which reflects a deficit in the 'real' world, is plugged because in exploratory information seeking which goes far beyond the verification of uncontentious facts, this is not the nature of knowledge construction or use. Information needs arise from, and are addressed through activities in which knowledge is distributed, bi-directional, and in constant negotiation – it is through this process that our information needs are defined, and addressed.

4.4 Section Summary – Key aspects of information seeking

In a key work on the state of research in 'Web search' Knight and Spink (2004) outline a "Web Search Information Behavio[u]r Model" (in a chapter of the same name), highlighting that:

> *Information retrieval entails the integration of a number of complex processes within the context of three major factors or entities:*
>
> - *An information Need (Broder, 2002)*
> - *An information Searcher (Kuhlthau, 1991)*
> - *An information Environment* (J. D. Johnson & Meischke, 2006)
>
> (S. A. Knight & Spink, 2004, p. 209)

They later also suggest the addition of a fourth element, representing the interaction between the user, and the search system. This focus on understanding the construction of information needs, by searchers, within a particular environment, and utilising the available tools available is important – particularly in the educational context. This section has thus highlighted:

1. The need to understand information needs and search as processes, mediated by systems and environmental factors – but correspondingly, some concerns around both more computing-based, and process models of information seeking
2. The importance of exploratory search particularly around ill-structured problems
3. The need to focus on 'in-action' and 'use' in understanding success around information seeking, in contrast to assessing the efficiency or 'accuracy' of search

In particular, this section has foregrounded that the sorts of search described as 'exploratory', and the ways information needs are created and mediated in those searches are a) likely to be of significant interest to educators, and b) may allow us to probe particular facets of student's understanding – their epistemic cognition. Needs are dynamic and defined, and exploratory – and their exploration is both a part of, and created by, their information environment. This is true both of dynamic interaction with the tools – search engines (the fourth element) – and in understanding tasks. Searchers are also collaborative, and this collaboration also is also both part of, and creating, context – and it is part of our understanding of 'success' in exploratory, collaborative information seeking tasks. I will now turn to introducing collaborative search, its incidence, and models to represent it. I will then return to discussing epistemic beliefs, before raising a key issue with both

epistemic beliefs and collaborative information seeking research – that neither recognises the constitutive role of dialogue as a creator of context (rather than simply an effect of collaborative contexts).

5. Collaborative Information Seeking

While in many cases information seeking is an individual task, this is not always so. Indeed, a whole section of the edited book "Web Search Engine Research" from the Library and Information Science (LIS) perspective, is devoted to the various forms of "searching together" (Lewandowski, 2012), while Shah (2012) has devoted a whole book to the study of collaborative information seeking.

This attention is warranted by the empirical evidence. Morris' early Microsoft study indicated that just over half of participants said they had collaborated on search activities, with a substantial percentage (87.7%) involved in tasks involving watching over someone's shoulder and making suggestions for alternative query terms – a kind of co-located synchronous search (Morris, 2008). Within a subgroup of 109 self-identified co-operators, "22% indicated they were always co-located when cooperatively searching, 11.9% indicated they always collaborated remotely, and 66.1% reported engaging in both remote and co-located collaborative searches." (Morris, 2008). More recent research (Morris, 2013) has indicated that such practices are still very common, although technological changes mean more collaboration happens in mobile contexts now.

Similarly, Amershi and Morris' (2009) conducted a small scale diary based study on 20 Microsoft employees which tracked co-located collaborative search at home and work for a week, 90% reported at least one occurrence of collaborative search. They also noted that there was a fairly even split between home and work collaboration, but with more 'informational' or exploratory searches at home than work, where users were more likely to be seeking specific items.

Using a broad definition of collaboration, that included implicit and explicit interaction, via physical interaction, texts or documents, Hansen and Järvelin (2005) found collaboration to be remarkably common at all stages of the IR process – overall, just over half involved collaboration. This included a rate of 44% in direct, and 43% via document collaboration at the 'information seeking' stage – which roughly corresponds to our notion of search. While certainly this lends support to our analysis of both collaboration, and a focus on the search process, it should be noted that this study was in a highly specific workplace – the Swedish Patent Office – and that data were mostly collected via interview and diary methods with some focussed observations taking place. Certainly it is interesting that the participants clearly engage in, and recognise their engagement in CIS, however the applicability of this work based collaboration to learning environments is not clear, nor are the socio-cultural practices surrounding such activities. Given that users often want to collaborate on searching, and sharing information, it is particularly concerning that most modern browsers do not facilitate this activity (Twidale, Nichols, & Paice, 1997) despite the longstanding nature of the concern. I now turn to further elaborating this notion of CIS, and its components – in doing so I will offer more empirical evidence for the prevalence, nature, and role of CIS.

5.1 What Is Collaborative Information Seeking?

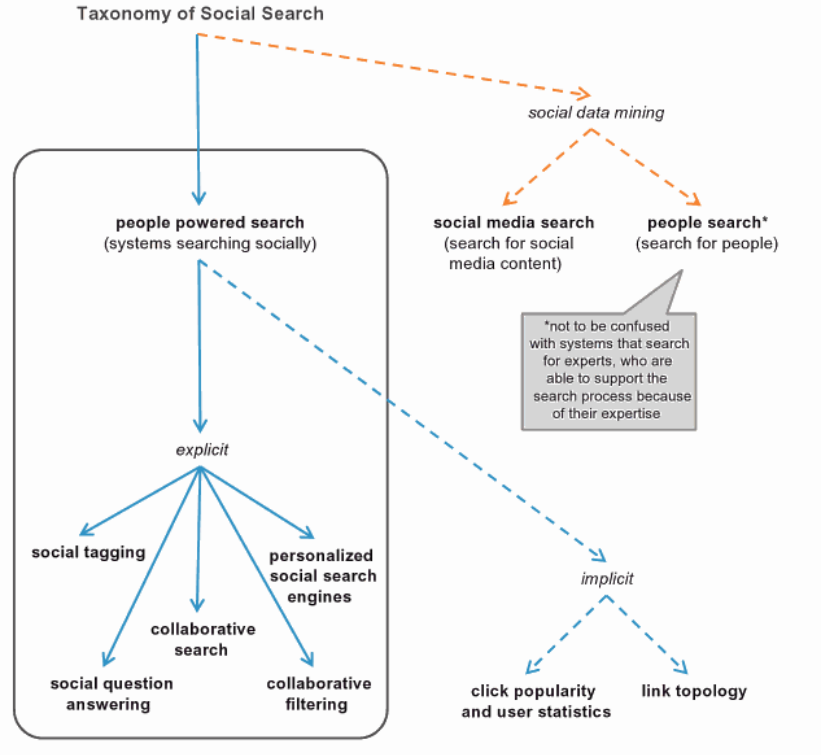

Figure 4 - Taxonomy of Social Search Approaches (Burghardt, Heckner, & Wolff, 2012, p. 26)

Figure 4 (Burghardt, Heckner, & Wolff, 2012, p. 26) offers an overview of social search approaches – highlighting collaborative search as explicit, and separate from some other related activities. Although there are other ways of conceptualising collaborative search (and, the figure may exclude some activities, or example collaborative browsing)[9] it provides a useful starting point for this section. It highlights the explicit nature of collaborative search (as contrasted with algorithmic mediation of results based on other's selections). Of course some systems will support more than one of these activities, and in particular the exploration of social question answering (Q&A forums, etc.) may also be of interest to us alongside collaborative information seeking. While there are a number of factors at play in collaborative information seeking (see below), two key conceptualisations are required: The nature of collaboration; and the nature of information seeking.

5.1.1 What is the Collaboration in CIS?

In his review of CIS theory, Shah (Shah, 2012, pp. 62–63) highlights the importance of collaboration – as a clearly defined concept – in CIS activities, noting the following hierarchy in which each element is a pre-requisite for the next (culminating in collaboration):

- Communication. This is the process of sending or exchanging information, which is one of the core requirements for carrying out collaboration, or maintaining any kind of productive relationship.
- Contribution. This is an informal relationship by which individuals help each other in achieving their personal goals.

[9] See also Chapter 3, Shah (2012) for a comprehensive overview of social information seeking and the activities different systems support.

- Coordination. This is a process of connecting different agents together for a harmonious action. This often involves bringing people or systems under an umbrella at the same time and place. During this process, the involved agents may share resources, responsibilities, and goals.
- Cooperation. This is a relationship in which different agents with similar interests take part in planning activities, negotiating roles, and sharing resources to achieve joint goals. In addition to coordination, cooperation involves all the agents following some rules of interaction.
- Collaboration. This is a process involving various agents that may see different aspects of a problem. They engage in a process through which they can go beyond their own individual expertise and vision by constructively exploring their differences and searching for common solutions. In contrast to cooperation, collaboration involves creating a solution that is more than merely the sum of each party's contribution. The authority in such a process is vested in the collaboration rather than in an individual entity.

Noting elsewhere (Shah, 2012, p. 23) that, for effective CIS:

1. A CIS system should provide effective ways for the participants to communicate with each other.
2. A CIS system should allow (and encourage) each participant to make individual contributions to the collaborative.
3. A CIS system should coordinate participant actions, information requests, and responses to have an active and interactive collaboration. This collaboration could be synchronous or asynchronous, and co-located or remote.
4. Participants need to agree to and follow a set of rules to carry out a productive collaboration. For instance, if they have a disagreement on the relevancy of an information object, they should discuss and negotiate; they should arrive at a mutually agreeable solution rather than continuing to dispute it. The system needs to support such a discussion and negotiation process among the participants.
5. A CIS system should provide a mechanism to let the participants not only explore their individual differences, but also negotiate roles and responsibilities. There may be a situation in which one participant leads the group and others follow (cooperate), but the real strength of collaboration lies in having the authority vested in the collective.

In highlighting that the user – and explicit user collaboration – is key in collaborative information seeking (as opposed to other information seeking activities) Shah limits the focus of information seeking, in a way conducive to our educational aims, to explore "*intentional,* and *interactive* [collaboration] among users with the *same information goal*" (Shah, 2012, p. 67). Users thus must address not only the affective nature of the information seeking (as Kuhlthau's model (1991) suggests), but also of their interactions with collaborators, "in this sense the dynamics of emotions, feelings, or moods are much more complex to explain than in individual settings" (Shah, 2012, p. 76).

5.1.2 What is the Information Seeking in CIS?

Work has thus investigated Kuhlthau's (1991) Information Search Process model in the collaborative context, indicating that groups cannot simply be modelled as individuals in another sense (Hyldegård, 2006). In this work, (Hyldegård, 2006) a small sample of five library and information

science graduate students (in two groups) self-reported during and after a group-based project assignment. Their reports suggested that although the broad process of moving from weak formulation to more focussed at the end was followed, aligning activities to particular stages was more challenging. Furthermore, although both groups submitted their assignments, their feelings of satisfaction, and movement through affective stages indicated that they may still have been in the 'exploratory' stage, experiencing frustration and uncertainty. They also note that, although the general stages were followed, intra and inter-group comparisons suggest different ways of working within and between groups which relate to the task setting and social factors.

A subsequent study (Hyldegård, 2009), again with a small sample size (n=10, split into three groups) of mostly female (n=9) library and information science students again found support for Kuhlthau's model in CIS, finding that students: moved through a broad progression of ISP stages; aligned with cognitive stages from vague thoughts to more focussed; and that their writing increased and searching decreased as they progressed.

Subsequent work by Shah and González-Ibáñez (2010) on a much larger group (n=84, in 42 pairs), finds similar support for a broad accord between CIS and individualised information search processes. In this study, the authors used trace data to delineate the search process – from: chat messages to greet each other and check-up between stages (initiation); chat messages discussing strategy (selection); number of queries used (exploration); number of webpages looked at (formulation); number of webpages or snippets collected (collection); to number of moving actions teams perform on collected snippets (presentation). As discussed in section 4.1.5, this operationalisation of stages, collection of data, and subsequent claim for the existence of stages is perhaps circular, however in terms of exploring intra and inter-group differences such an approach is effective. In particular, Shah and González-Ibáñez (2010) agreed with earlier work (Hyldegård, 2006, 2009), that affect was less clearly delineated, and that search stages often blended, in this case with exploration, formulation and collection involving multiple iterations and participants going "back and forth between trying search queries, exploring various sources, and collecting relevant information as they worked through the task while interacting with their collaborators" (Shah & González-Ibáñez, 2010, p. 8).

Thus analysis using Kuhlthau's model, removing the tightly constraining individual stages, may be appropriate. This is supported by the evidence here, and although early work suffered from sampling issues, a focus on self-report measures and a lack of analysis of interaction between collaborators (Hyldegård, 2006, 2009), subsequent work (Shah & González-Ibáñez, 2010) has used larger samples and found similar results with trace data.

5.1.3 The Role of CIS

Having outlined the role of collaboration, and a possible model for information seeking, the question now turns to what the features of such collaboration might be, and what benefits collaboration might hold – particularly in educational contexts. It is important to note that "collaboration between [these] two users can occur at various levels: (1) while formulating an information request, (2) while obtaining the results, and (3) while organizing and using the results" (Shah, 2012, p. 67).

As

Table 3 below indicates, a number of potential benefits to collaboration information seeking have been highlighted.

Table 3 – Potential benefits of collaborative search

Source	Benefit	Explanation
Adapted from Chi, Pirolli and Lam (2007, pp. 60–61) – based on large scale, implicit (algorithmic) collaboration.	Information flow	Allows individuals to see other things as highlighted by collaborators, which might otherwise have been missed
	Collectivity	Allows a greater confidence that the information seeking process is more exhaustive than if only one user is relied on
	Coordination	Allows the use of experts whose skills and knowledge can be utilised to recognise and interpret information
	Diversity	A range of viewpoints can be utilised and synthesised in order to avoid bias, and cognitive blindspots
Adapted from Evans, Kairam and Pirolli (2009, p. 680) – based on explicit collaboration.	Human-to-human communication assists with the "vocabulary problem"	"…peer support provides the greatest benefit when users are performing informational searches …. This type of exploratory searching can be hard to support from a system's perspective due to the occasional gulf between users' concepts and keywords and the jargon of the problem domain. Human-to-human communication has the potential to address this ''vocabulary problem"" (Evans et al., 2009, p. 680).
Adapted from Amershi and Morris (2009, p. 3638).	Pedagogical value of face-to-face communication	"Students often want to collaboratively search the Web to complete homework assignments for the pedagogical value that a shared context and face-to-face communication provides"
	Social Experience	"Friends and families want to collaboratively search for the social experience of planning activities together"
	Shared information needs	"…colleagues want to collaboratively search the Web to conduct joint research"

Building on such work, Evans and Chi (2010, p. 661), have proposed a model of social information seeking, indicating various ways, and stages at which, collaboration might occur including:

(1) The defining of information needs and exchange of relevant information surrounding those, such as important URLs and keywords,
(2) The search processes itself, such as shared understanding of information found – both the short previews given by search engines, and deeper information from websites
(3) The evaluation, and 'use' stage, such as organising information into various shared tools, and perhaps dissemination

And proposing a model (Figure 5) of such search.

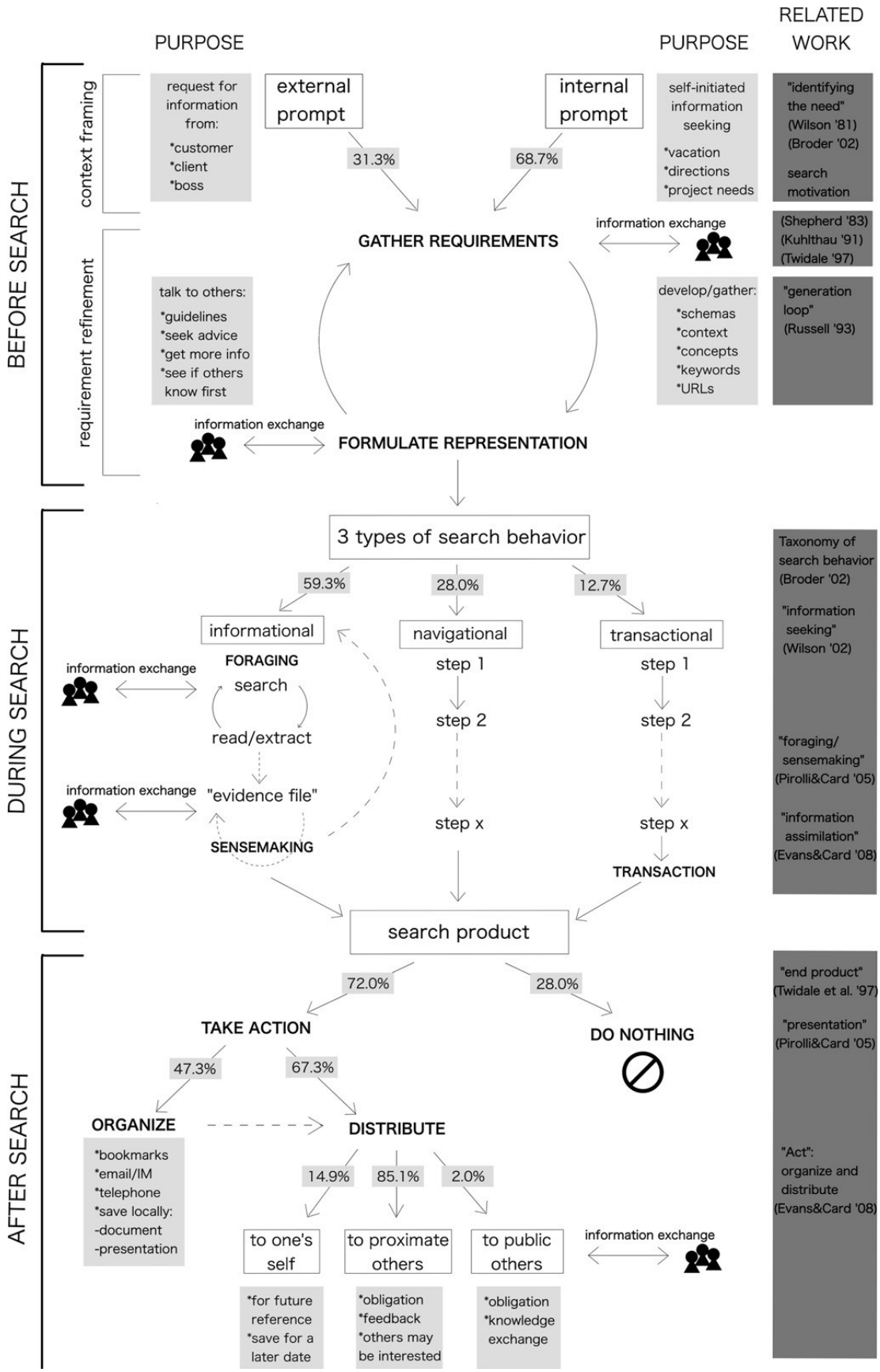

Figure 5 – Canonical social model of user activities before, during, and after a search act, with occurrence (%) indicated, including citations from related work in information seeking and sensemaking behavio[u]r (Evans & Chi, 2010, p. 661). (Examples of situations are given by asterisks)

5.2 Collaborative Information Seeking Models

In order to highlight the salient features of such activity, Shah has suggested that instead of trying to provide an overarching one-size-fits all framework, it is better to highlight, "various elements or dimensions of groupwork/collaborative systems. One could, hopefully, pick and choose the elements needed to study or explain a given context for collaborative systems from the list presented" (Shah, 2012, p. 51) – as summarised in Table 4.

Table 4 - Shah's summary of elements of CIS (Shah, 2012, pp. 51–57)

Element	Key Issue	Instantiation
Intent	How explicit is the collaboration?	From algorithmic mediation (e.g. collaborative filtering) through to working on the same task to the same goal
Activeness	How willing and aware is the user of the collaboration?	From passive collaboration through implicit trace, to explicitly leaving markers (ratings, comments, etc.) for sharing purposes
Concurrency	Is the collaboration occurring at the same time (concurrently) or not?	From synchronous to asynchronous
Location	Are the collaborators co-located?	From co-located to remotely located
System mediation	What role does the system play in mediating collaboration?	From collaboration via algorithmic mediation through to little mediation via tools
Awareness	How aware of the collaborators is the user?	From little awareness (e.g. in divided labor and algorithmic mediation) to more (e.g. from casual office interactions which lead to collaboration, through to formal supported collaboration such as in SearchTogether)
Interaction level	How much interaction with the system does the user have?	From highly interactive systems which support 'back-and-forth' within the system, to less interactive, possibly transactional, systems.
Communication level	How much communication occurs between collaborators?	From no communication, to high levels of communication
User roles	Do the users have defined roles in the system?	From systems which support (encourage?) users to divide labor or take on particular roles within the task to more open ended systems
Strength of connection	How 'connected' are the collaborators socially?	From weak, temporary, general connection to strong, long lasting, or/and task-specific connections
Balance of benefits	Is the collaboration mutually beneficial?	From mutually beneficially and balanced (e.g. co-authoring) to less balanced situations (e.g. sharing curated bookmarks)
Usage of information	How does information flow in the system?	From information flowing between users, to sensemaking and synthesising on retrieved information

Other work in CIS has focussed on a smaller number of key facets, to create a taxonomy of information seeking by system mediation and user location. There are a number of different levels of interaction here, some of which involve no direct interaction, for example, the algorithmic and implicit systems which tools such as 'collaborative filtering'[10] are based on. Of more interest to our context are the 'explicit' collaborative systems, which can be described along the dimensions of concurrency – synchronous or asynchronous – and location – co-located or distributed – as in Table 5.

[10] A good example of collaborative filtering is Amazon's "customers who bought this, also bought…" suggestions. Collaborative filtering is an implicit collaborative tool based on algorithms that try to tailor results – whether for purchases or within search engines – to the user based on the behaviour of similar users. Some concerns have been raised about the implications of this for political freedom and user autonomy, most recently by The Filter Bubble (Pariser, 2011).

Table 5 - Taxonomy of Collaborative Search (adapted from FXPAL (2011) and Golovchinsky, Pickens and Back (2009))

	Co-located	**Distributed/remote**
Synchronous (i.e. searching together)	Collaborative exploratory search; this could be many students on one PC, or spread, as long as they are in the same place, on the same task.	Remote collaborative exploratory search using tools such as chat systems/email or tools designed explicitly for searching, e.g. Microsoft's 'Search Together'[11]
Asynchronous (i.e. searching at different times)	This is less discussed, but feasible examples include emailing results from home, to be used in class, and similar practices.	Asynchronous browsing and search, e.g. www.searchteam.com allows you to send your results to another user, comment on them, rerank them, etc.

5.3 Further Factors in CIS – Awareness and Communication

5.3.1 Awareness[12]

A crucial component of collaboration, particularly in the seeking, sharing, and evaluation of resources, is an awareness of a collaborators activities and resources. As Shah (2013) notes:

> *In an information-seeking situation, it refers to the information seeker being aware of various aspects of the searching and sense-making processes, including the task and its context, past and present actions, and various attributes of the information objects and the system. This may not be very helpful when a single information seeker is doing quick searching that lasts a short session, but it becomes a salient aspect to consider when an information-seeking process lasts several sessions and/or is conducted in collaboration. For instance, when a lawyer is researching a case, collecting as much information from the available literature as possible, the process may span multiple sessions. It is crucial that the lawyer is aware of his past searches and found information (relevant or nonrelevant) and the overall context of the case. If such a project is done in collaboration with other people, then the issue of awareness becomes even more critical as the involved parties may have to keep track of not only their own processes and objects but also that of others* (Shah, 2013, p. 3)

Building on established work, (Liechti & Sumi, 2002) Shah (2012) thus proposes a taxonomy addressing four kinds of awareness:

1. *Group awareness. This kind of awareness includes providing information to each group member about the status and activities of the other collaborators at a given time.*
2. *Workspace awareness. This refers to a common workspace that the group members share and where they can bring and discuss their findings, and create a common product.*
3. *Contextual awareness. This type of awareness relates to the application domain, rather than the users. Here, we want to identify what content is useful for the group, and what the goals are for the current project.*
4. *Peripheral awareness. This relates to the kind of information that has resulted from personal and the group's collective history, and should be kept separate from what a participant is currently viewing or doing.*

[11] http://research.microsoft.com/en-us/um/redmond/projects/searchtogether/

[12] See also section 4.4.3 of Shah (2012) for a review of awareness in CIS systems, and section 10.2 for further description of awareness tools in a CSCL context.

(Shah, 2012, p. 48)

While there is established work in the incidence of CIS, and the building of tools to support CIS (see section 4.4.3 of Shah (2012) and Section 10.1 of this work), little work has explored the ways in which collaborators share their resources when using computers. "To get a better understanding of how computers and displays are shared in real-world collaborative search situations, we re-examined our data by looking at all combinations of the search computers used and shared information display variables." (Crescenzi & Capra, 2013, p. 3) That study reports:

- The highest incidence (74%, n=227) of search using individual computers and sharing information without a shared display (via phone, email, etc.)
- This was followed (15%, n=46) by those who searched used a shared computer and shared information on the shared display
- A mixture of the two primary strategies was also reported (7%, n=22) with searching and information sharing using a mixture of individual and shared computers
- Some participants also reported using individual computers for search but sharing information on a shared display (3%, n=10)
- Finally a couple (1%, n=2) reported taking turns using a shared computer and sharing information on that display

Crucially these results indicate the importance of information sharing, and the ways in which collaborators will make use of 'tools at hand' to do so, and to communicate about their searches and results – I now turn to discuss this combined aspect, of communication and awareness.

5.3.2 Communication

5.3.2.1 Communication and Awareness

This has implications for the nature of CIS, and the systems that support it. For example, in an extension of the Amershi and Morris' (2009) study described above, 12 three person groups (who had a prior relationship) were tested with both pre-assigned and self-defined tasks. In their observations, they found that although there was a high level of communication, it was not always effective. Where it was, it was related to collaborators being actively solicited for suggestions, although even in these instances and even when personalities were relatively well matched, problems occurred. In comparison, when users were asked to engage in parallel search in which they were each provided with a PC to explore the same task – a situation which is common in the educational contexts – they found low levels of communication, poor planning, and frequent "redundant work" (Amershi & Morris, 2009, p. 3640).

There is thus a balance between parallel search – which is problematic – and shared computer use which, although common and often involving useful verbal suggestions, may lead to the PC being controlled by one individual, with little input from other users (Amershi & Morris, 2008). Amershi and Morris attempted to overcome this by creating a tool to allow multiple users input on the same PC via separate mice, with different coloured on-screen pointers which allowed users to 'queue' search results simultaneously on the same PC[13]. Their trial of the tool, using groups of 3 participants

[13] See CoSearch: http://research.microsoft.com/en-us/um/people/merrie/cosearch.html the multi-point aspect of which is built on the Microsoft MultiPoint Mouse SDK http://www.microsoft.com/multipoint/mouse-sdk/

researching an area of mutual interest, suggests that CoSearch facilitated high levels of group communication which were above those of the parallel search, and allowed for a great sharing of labour thus reducing some of the frustration of the shared PC use.

However, users did not feel CoSearch facilitated communication or reduced ignoring over the 'co-located' condition. In addition, users felt more aware of other user's actions in the shared condition, than the CoSearch condition; these findings suggest that there may be a decrease in awareness because users feel less need to engage directly with other users, rather than just interacting with the shared PC directly. Indeed, where people wish to collaborate, they may simply use "tools-at-hand" (such as printers, telephones, text processors, email, and, simple copy and paste) to facilitate their search and sharing processes (Capra et al., 2012; Capra, Marchionini, Velasco-Martin, & Muller, 2010). In that study, 30 academic researchers, corporate workers and people looking for medical information were interviewed about their practices of collecting, managing, organising and sharing results from exploratory search tasks. Similarly to Morris' (Morris, 2008, 2013) results, they found that the incidence of collaborative search was high, with many participants sharing results particularly with their own annotations (to increase awareness and add value). Work to develop a system based on these and other findings resulted in laboratory studies supporting this evidence of results sharing, suggesting: "participants used the collaborative features [of the system] not just to avoid duplication of effort, but also to check and refine collaborators' work, to gain a general understanding of collaborator's actions, and to get ideas for new queries. " (Capra et al., 2012, p. 1).

A further lab-experiment (Shah & Marchionini, 2010) using 42 pairs (n=84) of collaborators explored two awareness conditions and a baseline in exploratory search tasks to find snippets related to a particular problem:

1. The baseline group were given only a chat tool, shared task instructions and a 'saved snippets' area
2. One awareness group had a browser extension which in addition to that in (1) gave a personal history of queries made and links clicked
3. While the second awareness group could see in addition to that in (1) and (2) the links and queries made by their collaborator

In this study it was found that participants in the third condition used significantly more unique queries than those in condition one, and were more engaged, concluding that – although those in the first condition managed well – basic group awareness features do not add a cognitive burden, while offering potentially substantial advantages in multi-session exploratory searches (Shah & Marchionini, 2010).

In subsequent work (Shah, 2013) participants were again asked to collaborate on an exploratory search problem in which they were asked to address particular issues, and to find as many relevant snippets as they could (but not to write these up into a report). In addition, these chat messages were then coded for coordination purpose messages involving asking for a collaborators status, responding to that question, a confirmation or reaction. These were further coded as being past oriented, current status oriented, or future actions or strategies oriented. Shah (2013) reports that most chat was not coordinating in nature, but that the baseline group engaged in more coordinating

talk, significantly more of which was past oriented, in contrast to the third group which was more present and future oriented as Table 6 indicates.

Table 6 - Summary of coded messages for teams in different conditions. Each condition had 14 teams.*

	None	Past	Present	Future	Total
Baseline	817 (29%)	1020 (36%)	461 (16%)	512 (18%)	2810
Personal	751 (39%)	638 (30%)	429 (20%)	277 (13%)	2095
Group	453 (37%)	165 (13%)	355 (29%)	252 (21%)	1225
Overall	2021	1823	1245	1041	6130

*Adapted from Shah (Shah, 2013, p. 1133), percentages indicate the proportion of messages in each row accounted for in the given cell.

On the basis of coordination cost – number of messages exchanged, inaccuracy in reporting status, and time taken to coordinate with teammates – Shah concludes that communication alone (as in the control group) is not enough to facilitate coordination, and thus support effective CIS. Correspondingly, those in the shared awareness group were able to best co-ordinate their efforts, and communicate their activities. While this study provides further support for the consideration of awareness in CIS system designs, the constraints of the task design introduce at least two concerns. Firstly, asking participants to find "as many" snippets as possible emphasises factual recall over sensemaking and, arguably, truly exploratory search. This may impact on the nature of coordination required, and CIS. For example (as noted by Shah) the provision of 'issues' participants were required to cover facilitated task splitting, and reduced the need to 'make sense' of their information needs. Secondly, asking participants not to write up their snippets – may affect results, particularly in light of the discussion above regarding the fluidity of ISP stages in CIS. Finally, the coding of talk as only 'coordinating' or not, limits our understanding of the ways in which coordination occurred, and users made sense of information together. This point is of central importance to this work, and is a key contribution – CIS work has not well explored the role of dialogue in its learning role (focussing on its group or system management role), doing so holds benefits for understanding how to better support CIS. The role of dialogue is an issue to which I return below.

5.3.2.2 Further Evidence for Communication in CIS

Thus CIS studies have noted the use of chat for example: after results are found, but without considering how the chat helps build knowledge (Yue, Han, & He, 2012); as a task oriented tool tending towards the division of labour (Shah, 2013); as a proxy for communication (via a simple message count) without looking at content (Shah & Marchionini, 2010); as an indication of particular stages of the ISP (Shah & González-Ibáñez, 2010, p. 7).

This highlights the need to further explore the particular functions of the dialogue, in this last study, a sentiment analysis (positive or negative) of messages was conducted, and the collaborators own judgements of relevance of chat messages was noted. Interestingly Shah and González-Ibáñez report that users:

> *expressed their feelings regarding the information they found and shared during the ISP, especially within the segments associated to high levels of Exploration, Formulation, and Collection. In the transitions from one task to another, as well as in the Presentation phase the affective relevance practically vanishes. It was also observed that the selection of relevant information was first done by an individual participant and then subjected to the group's judgment and reflection.* (Shah & González-Ibáñez, 2010, p. 7)

This point regarding collaborators sharing and reflecting together has been noted in other work too, and related not only to the consideration of information as it is being traversed, or after it has been saved but also to defining information needs, finding that:

> *At the pre-focus-stage, in particular, they were actively engaged in generating a shared focus and understanding of the problem at hand, e.g. shown in various forms of collaborative information activities and strategies. Information was communicated, discussed, exchanged and shared, primarily to help formulate a collective goal and obtain a shared understanding of the problem in focus. At the focus and post-focus stage information was primarily communicated and discussed according to specific elements of the assignment, e.g. based on the reading of other group members' writings.* (Hyldegård, 2009, p. 155).

5.3.2.3 What Role Communication in CIS?

This is an important point, not only for CIS, but also for information seeking in general – the dialogue oriented context of information needs plays a significant, yet understudied, role in their real world enacting (Savolainen, 2012)[14] [15].

Hertzum (2008) discusses the role of this shared motivation towards knowledge accumulation alongside the types of dialogue which facilitate such IR activities.

> *In a collaborative context, information is typically distributed unevenly across actors, and they may interpret the information known to them in different ways or be unable to make coherent sense of it. On the one hand, this is what makes communication and information seeking worthwhile activities. On the other hand, it also emphasizes the considerable work and constraints involved in making coherent sense of information within a group of actors.* (Hertzum, 2008, p. 958).

Hertzum's suggestion is that, as the collaboration becomes closer, the 'common ground' underpinning both the dialogue of collaboration, and the shared understanding of the information need should also increase, while in looser collaboration, such common ground can be more temporary and may require more continual effort[16].

For example, there was some evidence of this kind of grounding in Hyldegård's work, which suggested that particularly for more effective groups:

> *…group communication formed part of the constructive and cognitive process of the project assignment, each group member also acted as an information source during this process. Through group meetings and email-communication, for example, information was exchanged either as concrete references or as documented comments and suggestions to a group member's written manuscript. This was also a way to ensure or provide for a shared understanding of the project focus* (Hyldegård, 2006, p. 287)

[14] C.f. (Sundin & Johannisson, 2005a, 2005b) also cited below

[15] One theoretical paper has explored models of communication in context of information behaviour, but only of mass communication rather than of collaborative information seeking behaviour (Robson & Robinson, 2013). Similarly, one of the only analyses from an activity theoretic perspective (Hjørland, 1997) – close to the sociocultural one I further outline below, and introduced in the section on epistemology – focused the perspective of discourse qua document histories, not qua active discussion and use.

[16] One model using this suggestion is offered by Karunakaran, Spence and Reddy's model of Collaborative Information Behaviour (2010).

In particular, Ellis et al., (2002) note that the interaction is driven by dialogue which addresses the terminology of search, and the ways in which one might search. Some of the types of responses made by users can be readily matched to the types of dialogue that Mercer and colleagues identify, as in Table 7 (Mercer and colleagues' work will be elaborated in Section 8.4). However, much of the analysis focussed on the *content* of the utterances, rather than the intentions behind them, and the style of talk engaged in. As such, by focussing solely on the dialogue's relevance to tool-mediated action they may miss important information regarding the nature of the 'speech acts' (Grice, 1975) and the sociocultural practices in which they are embedded to create shared meaning (Wells, 2002). In educational contexts, this idea of the shared nature of language to create meaning has been termed by some as interthinking (Littleton & Mercer, 2013; Mercer & Littleton, 2007).

Table 7 - A comparison of Ellis et al., (2002) types of dialogue, and Mercer's (2000) types of talk

Ellis et al., (2002)	**Mercer** (2000)
Monochromatic passive – approval of individual items	Cumulative – summing up, little exploration or contextualising.
Monochromatic active – includes comments on bigger picture	Exploratory – focus on understanding others, lots of 'why' questions
Polychromatic – emphasises non-relevance	Disputational – focus on disagreements with little 'resolution'

*(Adapted from Ellis et al., (2002, p. 890) and Mercer (2000, pp. 97–99))

Although Ellis et al.'s (2002) study does not fully explore the dialogue, and is focussed on the use of expert intermediaries who are not co-located, the comparison presented in Table 7 is indicative of a useful means to analyse such talk in educational contexts. Foster (2009) analysed discourse in an educational context, although his work was on undergraduates studying information management – who one might reasonably expect to display somewhat particular information seeking behaviours. This work is closer again to Mercer's in that it focuses on understanding the motivating problem – part of the shared history of those in the discourse – and considers the nature of the task, with "…users as active constructors rather than passive receivers of information…" (Foster, 2009, p. 85). Although the analysis focuses on only a later stage in the information process – planning a presentation – it is interesting to note that in their analysis, 50.9% of talk was 'exploratory', 33.53% what they describe as coordinating (planning), with the rest disputational or cumulative in nature (Foster, 2009, p. 88).

5.4 CIS in Education

Having established some general evidence of the prevalence of CIS, and some of its conditions I now turn to consider evidence in education. It is important to consider that the information seeking context is of particular interest to education, but challenging to study in formal contexts. Furthermore, given that collaborative incidents may be ad-hoc (such as 'over the shoulder' collaboration) identifying CIS in formal educational contexts is further complicated. Finally, it should be noted that many of the studies reported above were in educational contexts (although, often with library and information science students), below I report on some studies in explicitly formal education contexts.

That noted, evidence does suggest that in educational contexts, CIS is a frequent phenomenon (Amershi & Morris, 2008; Ba et al., 2002; Livingstone et al., 2005; SQW, 2011); however, these studies have focussed on professionals' perceptions – not students' – (Amershi & Morris, 2008) and student self-report measures (Ba et al., 2002; Livingstone et al., 2005; SQW, 2011) as opposed to direct observation. Although certainly self-report measures of collaborative use are important, they

may neglect the specific ways in which collaborators mediate contact with the world of information through discourse. Fundamentally, self-report measures may contain bias – through sampling, interviewer effects, and the subjective nature of understanding one's situation both as an interviewer, and interviewee. By failing to explore collaboration in action we may ignore means to support higher quality collaboration.

One observational study, conducted by Lazonder (2005), explored this epistemic component of information seeking in the context of collaborative educational tasks, suggesting that students are, "largely unable to select appropriate search strategies (planning), check their progress (monitoring) and assess the relevance of search outcomes (evaluating)." (Lazonder, 2005, p. 466). He thus suggests that collaboration may aid in overcoming the "inert knowledge problem" (Lazonder, 2005, p. 466) in that verbalisation to collaborators may contribute to the self-regulatory processes, prompting users into better negotiating the search process. The implication here is that, by encouraging the creation of common ground or knowledge, we facilitate better information seeking processes. However, this was a small scale study, based on older students in which, although talk or 'verbalisation' was deemed important for self regulation, it was not analysed as a data form or co-constructive activity.

Another educational based study (Large, Beheshti, & Rahman, 2002), with a focus on gender, analysed 53 12 year olds in 44 groups of boys and girls engaging in search tasks around an assigned topic. Unfortunately, they did not investigate the one student who worked alone, or the 8 who opted to work in mixed-sex groups. The results of this study are also hard to interpret because, although a large number of findings are reported – some of which are interesting such as the relatively higher rate of 'natural language' (as opposed to keyword) searches in girls, and the relatively faster speed at which boys navigated pages – most results were not statistically significant. This is one reason why qualitative analysis may be of interest in this context; qualitative analysis would offer an insight into the different kinds of discourse surrounding such differences, which might, for example, allow an exploration of the types of language used which produce fewer keywords in searches, versus natural language search, alongside measures of success.

Further support for the finding that students tend toward impulsive searching and have difficulty in navigating search results comes from a case study across three schools, on 92 students with a mean age of 10.6 (Kuiper, Volman, & Terwel, 2009). They also concluded that, "…the conditions for students working collaboratively deserve attention. Our results confirm the importance of collaborative inquiry activities being more than just 'working together'". They suggest that such successful situations, "showed students who helped each other, who knew what everyone else was doing and who all shared the same goals. This resulted in a high motivation and an accumulation of knowledge." (Kuiper et al., 2009, p. 679).

Thus evidence suggests that CIS is not uncommon in formal educational contexts. However, many of these studies fail to explore collaboration *in action* but only indirectly, and they thus may ignore means to support higher quality collaboration. Furthermore, given the importance of collaborative discourse for educational outcomes (see below, and the collection edited by Littleton and Howe (2010)) it is important to understand the ways discourse helps to shape the epistemic properties of particular tasks, including information seeking.

5.5 Salient Elements of CIS for Education

It is interesting then to relate the evidence of CIS in educational contexts to the wider literature around CIS. To return to the list provided by Shah,

Table 8 gives an indication of some of the focal points for CIS, and their relationship to educational interests and contexts – for example, understanding what sorts of collaboration are likely to be of interest to, or/and of high incidence in, educational settings.

Table 8 - Focal Points for CIS Research

Element	Key Issue	Educational Context
Intent	How explicit is the collaboration?	Same task, to same goal
Activeness	How willing and aware is the user of the collaboration?	Explicit collaboration and sharing
Concurrency	Is the collaboration occurring at the same time (concurrently) or not?	Either synchronous (e.g. in class) or asynchronous (e.g. homework) within a constrained timeframe
Location	Are the collaborators co-located?	Either co-located or remote
System mediation	What role does the system play in mediating collaboration?	CSCL tools could mediate search, or tasks could be designed to encourage use of mediating tools such as email
Awareness	How aware of the collaborators is the user?	Collaborators should be aware
Interaction level	How much interaction with the system does the user have?	CSCL and CMC tools could support interactional systems, while email lends itself to more transactional approaches
Communication level	How much communication occurs between collaborators?	Communication should be facilitated – this is a key interest to educationalists.
User roles	Do the users have defined roles in the system?	Roles may be useful (e.g. for differentiation in classrooms) but generally open ended systems may be best suited
Strength of connection	How 'connected' are the collaborators socially?	Connection may depend on learning context – e.g. a mooc v. a classroom. May also be task specific (groups constructed for particular purpose)
Balance of benefits	Is the collaboration mutually beneficial?	In most educational contexts it is expected that there will be mutual benefit, even in peer teaching contexts it is assumed there is benefit to both teacher and student.
Usage of information	How does information flow in the system?	Given the evidence around Kuhlthau's (1991) ISP and CIS, indicating a difficulty in delineating stages it is likely CIS will occur at all stages – and various stages may be of direct interest to education researchers.

5.5.1 'Success' in CIS

As noted in section 4.3.2, the sort of 'success' measure of interest to our pragmatic approach is likely to focus on use of information, on its communicative properties – properties which may be particularly foregrounded in CIS contexts.

Similarly to individualised approaches, some success metrics could include amount of information found (with an expectation that groups find at least twice as much information between them), or speed with which information is found to answer questions. Yet neither of these captures the nature of success in exploratory contexts, nor the sort of joint meaning making we might a) expect and b) desire to happen in CIS. Again, this is a topic of interest to Shah (2012, Chapter 7), who explores 'evaluation' in his CIS book, noting that research in the area has explored and utilised: usability studies; user relevance ratings of collaborator content; collaborative capability surveys;

qualitative descriptions of information traversal; and traditional methods of success particularly precision, recall and F-measures.

This latter method has been popular because it is possible to judge a team's performance relative to the performance of the pooled team – for example, if we are interested in retrieval of pages (recall), we can compare the number of pages retrieved by one group to the set retrieved by the whole group (and indeed, the average of each group) (Shah, 2012). Similar methods can also probe deeper, for example a page's 'likelihood of discovery' might indicate the ease with which information on it could be found; pages with a high likelihood being pages which many groups find, and which are more likely to be near the top of search results, while those with low likelihood have the converse properties. While such metrics might offer insight into the search practices and efforts of groups, they could also indicate factors in task difficulty, or the quality of search results returned in high (and thus, accessible) positions, or the 'lostness' of those who seek less easy to find results. As Shah (ibid) notes, such metrics also give no information regarding pages which no group visits. A similar approach would take queries as the object of inquiry, looking at the 'diversity' of queries issued by computing their 'Lavenshtein distance' – where a distance of 0 indicates a match between two queries, and distances closer to 1 indicate more diverse queries (as measured by character difference).

Each of these methods is problematic in its analysis of 'success' or 'performance' – although each may provide interesting insight. The lack of focus on communicative practices in CIS research as noted above, is also interesting in the context of 'success' and the importance of effective dialogue for collaboration, and for learning more generally. This highlights a need to explore the sorts of dialogue which might be salient in learning contexts for understanding the 'success' of a CIS project – and this will be discussed further in sections 8.3 and 8.4. Given the focus on educational contexts for CIS, educational aims form part of the context for task activity, and judging task success.

5.6 Section Summary

> *…if two people working together can find twice as much information as either of them working independently, was that a good thing? How about the amount of time they spent cumulatively? The participants may not be able to find twice as many results, but what if they achieved better understanding of the problem or the information due to working in collaboration? Then there are other factors, such as engagement, social interactions, and social capital, which may be important depending upon the application, but are usually not looked at in non-interactive or a single-user IR evaluations.* (Shah, 2012, pp. 115–116)

CIS is complex, and multi-faceted – and this complexity is carried over into assessment of its success. However, as the quote indicates, the potential for greater understanding, diversity of results, engagement, and – central to this work – epistemic cognition, is of great interest. CIS provides both a context in which constructs such as epistemic cognition might be probed, and indicates some means through which epistemological assumptions regarding high-level assessments might be represented in tasks that go beyond a simple "factual recall" assessment.

This section has highlighted the prevalence of CIS and some of its features, in particular drawing attention to some understudied areas related to the use of dialogue to mediate CIS processes. I now explore in more detail its relationship to epistemic cognition, before describing: some methods for

probing that construct, particularly in the online context (section 6); some possible advances in tracking epistemic cognition (or, as I will introduce – epistemic commitments) (sections 7 and 8); and concluding with some candidate CIS tasks (section 9) and tools (section 10) which might be used to conduct such an investigation.

6. Information Seeking as an Epistemic Process

Having explored some models of information seeking, it is interesting to note the striking resemblance they have with those of self-regulated learning – which has been associated with epistemic beliefs and metacognition (Muis & Franco, 2009; Pieschl, Stahl, & Bromme, 2013). There is at least a prima facie commonality between models of information seeking which highlight the shift through stages of identifying an information need, seeking information, and evaluating that information and models of epistemic beliefs such as that proposed by Muis, who describes: "four phases of self-regulated learning…(1) task definition, (2) planning and goal setting, (3) enactment, and, (4) evaluation" (Muis & Franco, 2009, p. 307).

Even at this descriptive level, the relationship between this and models of search, as presented below, are no, indicating that we can conceptualise the first phase as that in which the 'need' is defined, the second as that in which those needs are translated into search queries, and enacted (with more or less use of the sophisticated search engine tools), and the third as the ways these results are treated. The fourth could denote the iterative nature of search, both in terms of need redefinition, and refinement. The nature of these similarities may further indicate that information seeking is a good activity through which to explore epistemic beliefs.

This claim is particularly strong in the case of exploratory search tasks. Furthermore, the use of dialogue to mediate the CIS process may provide further insight into the epistemic context of information seeking (as I will discuss in section 8.3).

Indeed, there is some work beginning to operationalize such models in action, suggesting that:

> *learners with absolutistic beliefs will plan and execute different learning processes than those with sophisticated beliefs; these differences are especially pronounced under conditions of high complexity. Given the general superiority of the learning and adaptation processes of more sophisticated learners such beliefs should be a learning goal of their own and should be explicitly addressed in learning scenarios* (Pieschl et al., 2013, p. 53).

Two important points are raised here: that epistemic beliefs are important in learning, and that they are likely to relate to observable differences in action, an important claim given the poor information skills of many young people (see section 2.2.4). A body of research has explored this claim, and I now turn to discuss some of this relevant research.

6.1 The role of epistemic cognition in information seeking[17]

6.1.1 Multiple Document Processing

[17] In literature search here I have particularly limited my focus to epistemic (or epistemological) beliefs (or cognition) and information seeking, or multiple document processing. Many other concepts may be related to information seeking (and I discuss some of these throughout), and epistemic cognition may relate to many other behaviours (and some of these are mentioned) but the particular focus of this thesis is on those two concepts.

One class of research on epistemic cognition has focussed on its role in multiple document processing, including attempting to relate epistemic cognition to multiple document processing models (Bråten, Britt, Strømsø, & Rouet, 2011)[18]. This sort of research is particularly interesting in the context of information seeking, given the need in such activities to deal with multiple websites (documents) and their potentially conflicting, and related, information. A typical pattern in this research involves gathering psychometric data on epistemic beliefs, and then asking students to engage in some task – constructing an argument, or summarising information – using a number of pre-selected documents, selected for their variability in terms of credibility and information. Some of this research has further utilised think-aloud protocols to gather epistemic data.

This research indicates that those with less belief in personal justification, and stronger beliefs in justification by multiple sources perform best on text-comprehension (L. E. Ferguson & Bråten, 2013), that (controlling for topic knowledge) those whose talk-aloud data justifies knowledge claims by corroboration were more likely to include explicit referencing and had better argument structures (Bråten, Ferguson, Strømsø, & Anmarkrud, 2012), that belief in opinion as justification is negatively related to multiple document comprehension while the converse is true of belief in justification by multiple sources (Bråten, Ferguson, Strømsø, & Anmarkrud, in press), that belief in inquiry rules evaluation and integration are important to understanding science claims (Helge Ivar Strømsø & Bråten, 2009), and that students with more sophisticated epistemic beliefs are better at melding information (Braten & Stromso, 2006).

As Aula (2005) notes "considered in cognitive terms, searching is a more analytical and demanding method for locating information than browsing, as it involves several phases, such as planning and executing queries, evaluating the results, and refining the queries, whereas browsing only requires the user to recognize promising-looking links" (Aula, 2005, p. 14) – it is to this that I now turn.

6.1.2 Epistemic Beliefs Online

Recent work has been conducted on the impact of epistemic cognition on comprehension of multiple *online* sources – which may vary radically in the nature of their sources and justifications – on the ba that students who perceive knowledge as simple and finite may conduct brief and perfunctory searches with little recourse to integration or multiple sourcing (Barzilai & Zohar, 2009; Bråten & Strømsø, 2006). As such, "exploring students' thought processes during online searching allows examination of personal epistemology not as a decontextualized set of beliefs, but as an activated, situated aspect of cognition that influences the knowledge construction process" (B. K. Hofer, 2004a, p. 43).

This work suggests that students with more "evaluative stances" on psychometric measures are more likely to evaluate websites meaningfully, while those with more sophisticated perspectives on the multiplicity of knowledge are more likely to integrate and critically evaluate multiple online sources (Barzilai & Zohar, 2009; Bråten & Strømsø, 2006). Further preliminary work suggests that while epistemic cognition was not a significant factor in understanding converging perspectives in online sources, in conflicting sources those with evaluativist beliefs did significantly better in their comprehension (Barzilai & Eshet-Alkalai, 2013).

[18] Bråten (2008) reviews the relevant literature (to 2008) in epistemic beliefs and multiple document processing in the context of learning within internet technologies. A selection of this literature, and the literature since, is discussed throughout this thesis.

A growing body of work associates search and sourcing patterns with particular patterns of epistemic metacognition (Mason et al., 2009), with subsequent think-aloud research indicating that students engaged in web-based learning spontaneously engage in some degree of epistemic reflection, particularly around source selection and credibility (Mason et al., 2011; Mason, Boldrin, & Ariasi, 2010), particularly finding that students who verbalised about source credibility and information veracity significantly outperformed those who evaluated only sources (Mason et al., 2011). Although it should be noted that the use of think-aloud protocols (Schraw & Impara, 2000; Schraw, 2000) may – as an artefact of the method – increase practices such as credibility judgements. Thus, the collaborative context may be more interesting both from an external validity perspective, and in that it provides insight into a group's epistemic practices. These practices might be called 'epistemic commitments' – the implicit or explicit stance taken with respect to regarding information as good, or poor quality, as shall be further outlined in Section 7.1.1.

6.2 Section Summary

The above discussion thus indicates that in tasks requiring online information seeking and the interpretation and synthesis of multiple documents, there is evidence that one's epistemic cognitions shapes how one deals with evaluating sources, finding information, assessing credibility, and so on. This section has thus established the broad notion of epistemic cognition as an area of interest. The next section delves further into what sort of construct epistemic cognition might be (section 7.1), before discussing some methods for the assessment of these constructs (section 7.2). I then turn (section 8) to discussing some alternative methods for exploring epistemic cognition (or commitments), particularly in the context of collaborative, dialogue mediated information seeking.

7. Epistemic Beliefs, Cognitions, Commitments, and Dialogue

Having briefly introduced the relationship between epistemic beliefs and information seeking, and its status as a good candidate for enquiry within our particular approach to learning analytics, I now turn to addressing what it is I mean by epistemic beliefs, and why I come to reject that construct preferring instead to talk about epistemic commitments.

7.1 An overview of theories related to epistemic constructs

As Table 9 illustrates, there are three broad models of epistemic beliefs. The complexity of epistemic cognition suggests a particular perspective on how we are to understand these beliefs. No approach 'mirrors' reality with a true, immutable, incontrovertible perspective on a learner's epistemic cognition. This concern is a dual one. Firstly, it is a methodological concern regarding our access to the world, our ability to 'get at' what is out there. Secondly, it is a conceptual and psychological concern, regarding the nature of epistemic cognition and whether it itself is stable – developmentally, and across domains – or shaped in some way by resources or beliefs. These two concerns are reflected in the epistemic beliefs literature. Firstly, cognitive developmental models (P. M. King & Kitchener, 2004; Kuhn & Weinstock, 2002) suggest that individuals progress through a sequence of increasingly sophisticated epistemic beliefs, while multidimensional perspectives (B. K. Hofer, 2001; Schommer, 1990) suggest that epistemic beliefs can be separated into dimensions, within which levels of sophistication can be identified (Greene, Muis, & Pieschl, 2010, p. 248). However, both of these assume a fixed uni-directional developmental trajectory, where beliefs are seen as global across (and within) domains. The resources view, in contrast, emphasizes the

interaction of believer, with resources, highlighting that at various points in any task a cognizer may invoke differing resources (Hammer & Elby, 2003).

Insofar as each model attempts to understand different facets of epistemic cognition, this research takes an agnostic stance on the particular cognitive model of epistemic beliefs. This is for two reasons – firstly, the research proposed here is not in a position to hypothesis test on models (for example, by conducting longitudinal studies for development of cognition); secondly, because the thrust of this work is to focus on particular aspects of behaviour – trace data of various sorts – as a shift in focus from cognitive models, to discursive properties of activity (see section 8). Thus the models will instead inform the ways in which the data is understood – the ways that epistemic beliefs are brought to bear through activity, and the particular methods that have probed this. Importantly, as introduced in section 3, across the three broad models, there is agreement on two main areas – what knowledge is, and how one comes to know:

> *There are two dimensions within the first area (knowledge):*
>
> - *Certainty of knowledge: the degree to which knowledge is conceived as stable or changing, ranging from absolute to tentative and evolving knowledge;*
> - *Simplicity of knowledge: the degree to which knowledge is conceived as compartmentalized or interrelated, ranging from knowledge as made up of discrete and simple facts to knowledge as complex and comprising interrelated concepts.*
>
> *There are also two dimensions which can be identified within the second area (knowing):*
>
> - *Source of knowledge: the relationship between knower and known, ranging from the belief that knowledge resides outside the self and is transmitted, to the belief that it is constructed by the self:*
> - *The justification for knowing: what makes a sufficient knowledge claim, ranging from the belief in observation or authority as sources, to the belief in the use of rules of inquiry and evaluation of expertise*
>
> (Mason et al., 2009, p. 69)

Table 9 – An overview of models of epistemic beliefs

Model	Summary	Implications	Implications for IS	Issues
Cognitive developmental perspective (P. M. King & Kitchener, 2004; Kuhn & Weinstock, 2002)	Individuals progress through sequence of beliefs surrounding (with broad agreement across models): (1) absolutism/objectivism (e.g., dualism, knowledge is right or wrong), (2) multiplism/ subjectivism (e.g., knowledge is merely opinion), and (3) evaluativism/objectivism-subjectivism (e.g., knowledge is continually evolving and must be critically judged)." (B. Hofer K. .. & Pintrich, 1997)	Age/developmental stage will be key & beliefs will be relatively global – experimental/quasi experimental research may probe.	Abilities will advance in relatively predictable and stable ways as children develop. Epistemic beliefs will be relatively global – tests in one discipline should be generalisable.	Ignores nuance of epistemic beliefs – sometimes it is appropriate to hold an absolutist view of a knowledge token. Assumes development is fixed, uni-directional, and global.
Multidimensional perspective (e.g., Hofer & Pintrich, 1997; Schommer, 1990)	Individuals have various, independent epistemic dimensions, e.g. "…Hofer and Pintrich identified four common belief dimensions about knowledge and knowing: (1) the certainty of knowledge (ranging from knowledge is unchanging to evolving); (2) the simplicity of knowledge (ranging from knowledge is bits of facts to highly integrated and complex); (3) the source of knowing (ranging from an authority to derived through logic and reason); and, (4) justification for knowing (ranging from authority providing warrants to warrants through rational or empirical methods).	Various aspects of epistemic beliefs should be discernible in cognition, and action – experimental/quasi experimental research may probe. Other self-report methodologies may explore particular facets of beliefs.	Task constructs could be designed to separate out particular aspects of epistemic beliefs – such as justification, versus knowledge certainty.	Ignores nuance of epistemic beliefs – sometimes it is appropriate to hold an absolutist view of a knowledge token. Assumes each dimension is fixed and global.
Perspective that considers epistemic beliefs to be more like task-specific resources (Hammer & Elby, 2003).	Epistemic beliefs are activated within the context of task-specific resources. In Hammer and Elby's framework, **the sociocultural setting is key** [emphasis added] to considering what resources may be evoked during learning. They stressed that learners may invoke different resources at varying times throughout a learning task.	More likely to take a naturalistic approach to study, and explore the ways in which meaning is created in particular settings – in particular through the use of dialogue.	The ways in which various resources are brought to bear on a particular task – including epistemic beliefs, which may be co-constructed in any given setting – should be studied.	Difficulties in application to some problems – including (ICT) systems level analysis.

Adapted from Greene et al., (2010, p. 248). Multiple theoretical frameworks have been developed and reviewed (B. Hofer K. .. & Pintrich, 1997; B. K. Hofer & Pintrich, 2002; B. K. Hofer, 2004b; Muis, 2007; Schraw, 2013). This table originally appeared in my MPhil thesis (see acknowledgements).

7.1.1 Recent Developments – Philosophical Commitments

However, more recent work (C. Chinn A., Buckland, & Samarapungavan, 2011; Greene et al., 2008; Muis, Bendixen, & Haerle, 2006) has sought to bring lessons from the philosophical literature on epistemology, into the psychological literature on epistemic cognition. However, while Greene, Azevedo and Torney-Purta (2008) discuss the philosophical literature in the context of epistemic cognition research, they explicitly focus on 'classical' notions of epistemology, thus remaining within an individualistic, cognitivist model, and neglecting the rich literature in social, virtue, and pragmatic epistemology which is of strong relevance to notions of justification, credibility, and 'knowledge' broadly.

Similarly, a comprehensive timeline of philosophical approaches to epistemology is given in Muis, Bendixen and Haerle (2006), but again does not address these more recent epistemological advances. Muis et al., (ibid) attempt to relate particular epistemologies with domains – empiricism with science, rationalism with mathematics. This analysis is conducted as a means to explore the domain specific – and thus, normatively defined in social contexts – epistemic stances held, not by individuals but by communities of practice (or disciplines). Thus the neglect of more recent advances in social, virtue, and pragmatic epistemology is particularly unfortunate given their explicit and specific analysis of such normatively defined epistemic contexts. In any case, the comprehensive review of literature related to domain specificity is interesting, and notes an important point – that it is challenging to generate domain-general, and cross-domain-comparative measures of epistemic (or, in their terms, epistemological) beliefs. Interestingly, while Muis, Bendixen and Haerle (2006) propose a sociocultural approach to understanding development of epistemic cognition, it is not their main focus in the paper itself. Indeed, from the description given it may be that their proposal is sociocultural in the sense of looking at domain context, but not in the sense of using specific sociocultural methodologies and the social epistemology which it might draw well on.

In contrast to many other models, Chinn, Buckland and Samarapungavan (2011) propose a model which, "differs from the current prevalent conceptualization of the structure of knowledge in two ways: (a) We view the structure of knowledge as multidimensional rather than undimensional, and (b) in addition to broad structural dimensions such as simplicity-complexity, we emphasize the importance of more specific structural forms such as mechanisms and causal frameworks" (C. Chinn A. et al., 2011, p. 150).

Chinn, Buckland and Samarapungavan (2011) built on philosophical scholarship to extend beyond the current focus on facets 2 and 3 below, to include:

1. Epistemic aims and epistemic value – what is the aim of knowledge work, and what is its value?
2. The structure of knowledge and other epistemic achievements – is knowledge and its aims complex or simple?
3. The sources and justification of knowledge and other epistemic achievements, together with related epistemic stances – where does knowledge originate, and what reasons are good warrants for knowledge claims? What stances can one hold towards knowledge claims (true/false, tentative belief, entertained possibility, etc.)?
4. Epistemic virtues and vices – the sorts of praiseworthy dispositions (virtues), and dispositions likely to hinder achievement of epistemic aims (vices)

5. Reliable and unreliable processes for achieving epistemic aims – what processes does a student hold as good for developing knowledge?

They note the increased focus in philosophical literature of 'testimony' as a source of knowledge (particularly, pragmatist in nature – see e.g. Craig (1999), and Fricker's work, e.g. (2012)), which has been largely ignored (or, rejected) by epistemic cognition literature. This shift to readmit the notion of testimonial knowledge is also reflected in some recent psychological literature which notes the importance of 'believing what you're told' in many contexts, including educational (see e.g Harris (2012)). These considerations are important in that they shift the focus to student's beliefs (or behaviours) of 'coming to know', rather than their beliefs regarding the structure of knowledge. This shift is perhaps also one from a focus on omission (a failure to see knowledge as complex, for example), to commission (an active act to ignore certain sources). This may recast some of the empirical work in the area, for example to refocus on what students do when they encounter problems they cannot solve[19] or their dealings with good informants, who they 'downgrade' epistemically (epistemic injustice – see Fricker (2009)). Such a focus accords well with their consideration of epistemic 'virtues' and 'vices', although they make an important caveat that: " judgements of whether dispositions such as open-mindedness should be regarded as an epistemic virtue or vice can vary according to the context" (C. Chinn A. et al., 2011, p. 156).

They thus propose a shift from questionnaires to interviews and observations, suggesting that:

> *A very important area for future EC research is to explore social aspects of epistemic cognition, several of which we have considered in this article. We have discussed beliefs about testimony, a pervasive social source of knowledge. We have also discussed beliefs about reliable social processes (E.g., argumentation, peer review, media processes) for achieving epistemic aims. Future research could also examine cognitions related to social features of the other components of the framework. (C. Chinn A. et al., 2011, p. 163)*

Of particular interest is their suggestion that:

> *What we have called tacit epistemic beliefs might better be called epistemic commitments (C. A. Chinn & Brewer, 1993)[20]. Some theorists may be uncomfortable with the idea that one can have a tacit 'belief' that cannot be expressed, and the term epistemic commitment avoids reference to such beliefs. An epistemic commitment reflects a tendency to act in specified ways, such as a proclivity to provide justifications based on personal experience (C. Chinn A. et al., 2011, p. 146).*

[19] The 'epistemology of silence' – see e.g. http://people.kmi.open.ac.uk/knight/2013/05/when-no-answer-is-answer-enough-2/

[20] Despite this earlier reference to epistemic commitments (or, epistemological commitments at that point), and a longer history of this research little reference to it is made in most current epistemic cognition research – including Tsai's work, which is explicitly about epistemic commitments. These two areas of research appear to have somewhat separated, with (broadly) the former focussing on epistemic commitments in the context of scientific theories – so, commitments as beliefs regarding what a model should look like – and the relationship of that to conceptual change, while epistemic cognition research has tended more towards exploring finer grained judgements of credibility and relevance by students. However, in some cases the two are used interchangeably (see e.g. the recent Zeinddin and Abd-El-Khalick (2010)) reflecting the shared history of some of the epistemic beliefs and commitments work. To be clear, where I refer to 'commitments' henceforth, I make a fairly strong distinction between that notion, and epistemic cognition or beliefs research – as shall be described.

Thus Tsai (2004) outlined: “epistemological commitments involve an individual’s explanatory ideals, that is, his or her specific views about what counts as a successful explanation in the field (e.g. science) and his or her general views about the character of valid knowledge or information” (C.-C. Tsai, 2004, p. 105). Tsai proposes a framework with a range of commitments for evaluative standards from ‘authority’ to ‘multiple sources’ and “a range of views from ‘functional’ (such as the ease of retrieving or search information) to ‘content’ (the relevancy to the intended search) is used for assessing the usefulness of Web-based materials. The framework also reflects an information-searching strategy ranging from ‘match’ to ‘elaboration and exploration’ (by metacognitive and purposeful thinking.”)” (C.-C. Tsai, 2004, p. 105).

In this early study, two experts in web-based instructions and 10 students were interviewed regarding assessing accuracy, judging usefulness and describing search strategies. In accord with other work they found that experts tend toward multiple sources – which they tried to integrate – while students emphasise ‘authority’ and ‘matching’ information to needs precisely. Despite this, experts placed more emphasis on defining search purpose and being “metacognitive”, while students tended to look for ease of access. Based on this small study, Tsai thus proposed a framework for information commitments as in Figure 6.

Information Commitments	Possible Orientations
Standards for correctness	Multiple sources ⟷ Authority
Standards for usefulness	Content ⟷ Functional
Searching strategy	Elaboration & exploration ⟷ Match

Figure 6 - A Framework for Information Commitments (C.-C. Tsai, 2004, p. 109)

While this development is an interesting one for this learning analytics project, and within a philosophical framework, it is still problematic not least because as Wu and Tsai (2005) highlight, students may utilize both the information commitments, ‘multiple sources’ and ‘authority’, at the same time when evaluating the accuracy of the materials on the Web, and they also probably utilized both the information commitments, ‘content’ and ‘technical issues’, to evaluate the usefulness of the materials on the Web.

7.2 An overview of methods in epistemic constructs research

From this work a number of methods can be identified. Methodologically the developmental models have tended towards interviews and laboratory tasks, while multidimensional models have emphasised paper and pencil self-report measures (DeBacker, Crowson, Beesley, Thoma, & Hestevold, 2008). Both of these approaches reflect the fixed perspective on beliefs from which theory they stem. Importantly, although three major survey instruments have been developed and deployed, – including in information seeking tasks (C. Lin & Tsai, 2008; Schommer, 1990) – they are criticised for their psychometric properties (DeBacker et al., 2008). Furthermore, while some studies have used interview (Barzilai & Zohar, 2012; Mason et al., 2009), think-aloud protocols (Barzilai & Zohar, 2012; L. E. Ferguson, Bråten, & Strømsø, 2012) or systematic observation (Scherr & Hammer, 2009) such methods may be limited in their insights, particularly where self-report data is to be used and interpreted by researchers. Importantly, they are also not appropriate for the study of online, collaborative, or geographically and temporally spread activities – in particular, online information

seeking, or information processing more broadly. These approaches reflect the epistemology of current assessment regimes, as indicated in Section 2, and seem to implicate the view of 'fixed' psychological constructs – whether intelligence, or epistemic beliefs, as further discussed throughout Section 3.

Table 10 - Overview of Methods Used in Epistemic Cognition Research

Method	Assumptions	Example studies	Advantages	Issues	Issues linked to use in Information Seeking
Think-aloud (Ericsson & Simon, 1980).	Epistemic beliefs are explicitly, consciously, brought to bear on information tasks.	Probing of epistemic beliefs in dealing with conflicting documents (L. E. Ferguson et al., 2012) Combined with retrospective interviews in exploration of evaluation and integration of sources (Barzilai & Zohar, 2012)	Access to 'stream of consciousness' which is *relevant* to the task, some evidence these reports are accurate (Duell & Schommer-Aikins, 2001).	Reports often incomplete. "…may consume limited attentional resources…may enhance metacognitive awareness by calling attention to the demands that the task requires (Schraw, 2000)" (B. K. Hofer, 2004a, pp. 50–51) Experimentation may give better cognitive access (Nielsen, Clemmensen, & Yssing, 2002).	In particular novice users may struggle with concurrent think-aloud and IR (Branch, 2001). May be related to metacognitive level & self-regulation – both of which also have concerns with use of think-aloud and its demanding nature (see, e.g. Schraw & Impara, 2000). Inappropriate for (co-located) collaborative work.
Self-report questionnaires (for a review of self-report measures and the theories behind them, see Duell & Schommer-Aikins, 2001)	Epistemic beliefs are something one can directly access, and report on in decontextualised settings and in ways that can – a priori – be categorized.	Analysis of relationship between self-report measure and ways of treating online science information (C. Lin & Tsai, 2008) Analysis of relationship between self-report measure and ways of summarising information and subsequent test score (Schommer, 1990)	Quick, easy, scalable and reliable. Can be used with other methods to provide access to various quantitative analysis types.	Three major surveys have been developed but are heavily criticised for their psychometric properties (DeBacker et al., 2008). Survey instruments tend to deliver decontextuaised, limited results lacking external validity.	Fail to account for the co-construction of knowledge, how epistemic beliefs are brought to bear on IR tasks is a complex relationship between the user(s), the systems, and the activity within which they are embedded.
Interviews (for a review of self-report measures and the theories behind them, see Duell & Schommer-Aikins, 2001)	Epistemic beliefs are something one can directly access, and report on in decontextualised settings	Retrospective interviews to probe epistemic metacognition regarding internet IR about a scientifically controversial topic (Mason et al., 2009) Combined with think-aloud protocols in exploration of evaluation and integration of sources (Barzilai & Zohar, 2012)	Allow more (unstructured) or less (structured) further probing of answers than questionnaires to ensure greater understanding. Can be used to provide coding schemes, and can be coded with these.	Interviewer bias and effects (on the individual, and potentially on wider environment).	Fail to account for the co-construction of knowledge, how epistemic beliefs are brought to bear on IR tasks is a complex relationship between the user(s), the systems, and the activity within which they are embedded.

Method	Assumptions	Example studies	Advantages	Issues	Issues linked to use in Information Seeking
Trace data (see particularly Greene et al., 2010) (note, 'think aloud' is sometimes thought of as trace data, although they are presented separately here)	Epistemic beliefs will be brought to bear on knowledge tasks in ways that can be meaningfully captured, in particular using technology systems (e.g. the way agents represent knowledge in mind maps)	Stadtler and Bromme (2007) analysed the ways participants found, extracted, and moved information – which could be used to reveal information about their beliefs (e.g. visiting few websites indicates trust in those (Greene et al., 2010).	Direct access to real behaviours in unobtrusive ways – high external validity. Improving technology makes these methodologies more robust, extensible, scalable and useful.	While trace data is unobtrusive, it may give an incomplete picture, in particular agents may have particular reasons for behaviour in particular ways that cannot be probed using such data.	Easier to track using online systems than offline. Behaviours may (at least partially) represent artefacts in the systems, as opposed to underlying cognitive constructs.
Collaborative dialogue	Epistemic beliefs are most interestingly studied in naturalistic settings – including collaborative ones, in which the beliefs may be thought of as co-constructed to a greater or lesser degree.	Rarely directly studied (see below). Tillema and Orland-Barak (2006) used various methods including observation (with dialogue analysis) to explore how "professionals' views on knowledge/knowing relate to the understandings gained through collaborative knowledge construction" (p.593) – however, the target of analysis is beliefs regarding the status of the collaborative group in the construction of knowledge.	Talk is part of the process (not a demanding 'add-on', as in think-aloud studies). Allows analysis of epistemic beliefs in a situated, naturalistic context in which they may be co-constructed and brought to bear on a particular problem.	Some epistemologically interesting facets may be salient to collaboration – for example, 'given' or assumed expertise – which may be difficult to capture or ascertain via in-situ dialogue alone.	Not all IR will involve collaboration (although there is a high incidence). Complex relationship between context – including the group, the wider setting (e.g. classroom), and the specific tools being used (e.g. worksheet and search engine) – the individuals, and epistemic beliefs.
Systematic observation	See trace data Observation allows direct analysis of physical behaviours, and the capturing of dialogue-based data (in quantitative form).	From observations of collaborative groups highlighted five behaviour clusters identified as 'epistemic frames' (Scherr & Hammer, 2009) – note the target of analysis here is the student's beliefs regarding the status of the collaborative group in the construction of knowledge.	Allows naturalistic data collection with (broadly) lower inference levels than other methods require.	Limited data available and in particular, only external behavioural indicators may be recorded. Potential for observer bias.	Trace data may be more interesting for IR tasks given access to log data.

See also Hofer (2004a, pp. 49–51) for a review of methods

One recent development (Bråten, Strømsø, & Samuelstuen, 2005) which is of particular interest to the information seeking element of this work is the development of the 36-item Internet-specific epistemological questionnaire (ISEQ), which has been associated with self-reports of internet-search and communication activities. Subsequent work (Helge I. Strømsø & Bråten, 2010) using this tool has found that:

1. Students who believe internet information is a source of detailed factual information are less likely to report problems with information seeking on the internet, and
2. Students who thought that the wealth of information available on the internet was an advantage, were more likely to report seeking expert help in their information seeking.
3. Similarly, those considering internet information to be detailed and concrete engaged in more self-regulatory activities.
4. Interestingly, those believing facts needed checking (and reasoning) were more likely to report engaging in self-regulatory strategies like planning.

Further ISEQ work in the context of a medical issue with conflicting information on the web (Kammerer, Bråten, Gerjets, & Strømsø, 2013), which analysed the ISEQ results in the context of log files, eye tracking, and verbal protocols found that:

5. Students with beliefs in the internet as a source of reliable, accurate, and detailed facts were less likely to reflect on the credibility of sources and URLs while maintaining more certainty in their search-decisions.
6. Correspondingly, those who had doubts about the need to check sources were more likely to have a one-sided representation.

While the use of both self-report and trace data studies in this highly relevant area is of interest, the applicability to domain general activities is still of concern. Furthermore, it is not clear how to interpret these results. For example, finding (6) may explain the claim in (1) that students report fewer information seeking problems, although it appears to support claim (4). Similarly, finding (5) may also explain the claim in (1), while appearing to contradict the claim in (3). Thus some self-report measures here may have probed self-efficacy (see Tsai for internet self-efficacy scale (2004)) motivating some claims (1) which may be explained by subsequent work, however that subsequent work may contradict some of the more substantive – and epistemic – claims made.

Similar work on developing a scale for epistemic-commitments (Y. Wu & Tsai, 2005; Y.-T. Wu & Tsai, 2007) shows some strong preliminary results (including indicating that students with more sophisticated evaluative standards also have more sophisticated scientific views (C. Lin & Tsai, 2008)) however it is not clear that the approach – for example, seeking agreement on the following item – is appropriate for dynamic internet search (where differing strategies may be more or less appropriate in varying contexts).

Multiple sources as accuracy (Multiple Sources) (Internet Commitment Scale ICS example item):

1. *I will discuss with teachers or peers, and then judge whether the information is correct*
2. *I will explore relevant content from books (or print materials), and then evaluate whether the information is correct*
3. *I will try to find more websites to validate whether the information is correct*

Other recent work in developing and deploying search strategy based scales also indicates that: 1) explicit strategies, i.e. actual behaviour tracked through trace, is better related to search outcome

than implicit (measured by the Online Information Searching Strategy Inventory, OISSI), and; 2) that there was little correlation between explicit strategies and OISSI scores (M.-J. Tsai, Hsu, & Tsai, 2012). Fundamentally, while the specific measures given here may provide some insight into student epistemic behaviour, each of them suffers from particular issues, and all of them can be critiqued on the grounds outlined in Table 10. Therefore if we wish to understand students' information actions in context, we may have to look elsewhere for methods to do so, notably, naturalistic user activity traces recorded in software logs, examined next.

7.3 Section Summary

This section has particularly sought to challenge the first claim in Schraw's summary of the salient literature in this area (as below), in particular by drawing on the second, third and fourth points regarding the context sensitivity and contextual elements of epistemic activity. Given that the literature highlights some important implications of epistemic beliefs – and their changes – for educational outcomes, this is an important area to address. Schraw notes that:

1. *Beliefs and world views can be measured using different types of measurement strategies*
2. *Beliefs and world views are complex and change over time*
3. *Beliefs and world views affect teaching practice – but there are inconsistencies between stated beliefs and practices*
4. *Beliefs and world views are context bound*
5. *Interventions and instruction can have effect on epistemological beliefs and classroom practice*
6. *Teachers' beliefs affect students' beliefs*

(Schraw, 2013, pp. 26–28)

In light of the challenges above, I now move on to discuss some other methodological approaches. These have particularly significance for educational settings, since the possibility of measuring epistemic cognition in-situ in authentic, messy learning environments would permit the exploration of localised, and co-constructed, belief-in-action. If we think that the way students treat information, and knowledge (their epistemologies in action) matter – which there is good empirical and theoretical reason to suppose – then we should seek to cultivate these as dynamic and context sensitive traits through our use of formative assessment. However, traditional approaches – including the oft used questionnaire – are likely to be inadequate for this purpose,

> *therefore some researchers (B. K. Hofer, 2004a; Maggioni & Fox, 2009; Mason et al., 2011, 2010; Mason & Boldrin, 2008) have further contextualised the study of epistemic cognition by moving beyond self-report inventories and using online think-aloud methodology (Ericsson & Simon, 1980). (L. E. Ferguson et al., 2012, p. 106).*

This is a continuing trend, with (Barzilai & Zohar, 2012) also utilising a thinking aloud. Yet even this method does not account for the nature of epistemic-action, which – insofar as it relates to "knowledge" – is by its very nature, social, collaborative, and normative (Craig, 1999).

Given the characterisation of epistemic cognition as "a lens for a learner's views on what it is to be learnt" (Bromme et al., 2009, p. 8) there are grounds for concern here – epistemic cognition research makes explicit the role of "action", and perspectives or beliefs in their context as facets of "doing", yet it does not draw from enough of the philosophical literature in this area, nor from methodological approaches which seek to explore the actions in context – actions as ways of doing, and ways of being.

The interest is thus particularly in information seeking, within the context that:

> *most concerns raised by teachers and educators – who complain about the lack of students' ability to orient themselves on the web – are not questions of technology but rather questions of beliefs about the nature of knowledge and knowing, which may facilitate or constrain searching and evaluating sources of information on the internet (Mason et al., 2011, p. 139)*

8. Epistemic ~~cognition/beliefs~~ Commitments

8.1 LA – Tools for Trace?

As noted above, a recent development in the epistemic cognition literature is that of 'epistemic commitments' . In conceptualising the salient construct as centered on 'commitments' as observed in actions, there is a shift from more generalised and developmental models of epistemic belief which often use self-report measures, to methods which explore the ways commitments are enacted in the context of particular information or sets of information. This has often – but not always – involved the use of trace, and as noted above explicit strategies are a better measure of performance, and are not well related to self-report of, search activities (M.-J. Tsai et al., 2012). Thus it is desirable to explore approaches which attempt to track and highlight the salient features of web navigation, such as the Meta-Analyzer environment through which students may conduct their information seeking, and teachers subsequently view their behaviours, a facility both found useful and desirable (Hwang, Tsai, Tsai, & Tseng, 2008; Tseng, Hwang, Tsai, & Tsai, 2009). This may be particularly promising given that students with less sophisticated epistemic beliefs are more likely to simply select and bookmark results from the top of search engine pages, in contrast to those with more sophisticated views – who hold that knowledge is constructed from multiple sources and expertise – who select a more diverse array of search results (Salmerón & Kammerer, 2012). Interface decisions and foregrounding may be particularly important in educational contexts given the interesting finding that, when 'sensitised' by being asked how they would proceed to confirm knowledge, students engage with more sources and are more evaluative of them (although this finding was more pronounced for those with prior sophisticated epistemic beliefs) (Porsch & Bromme, 2010).

One means to foreground such epistemic commitments visually is that of the "navigation flow map (NFM)" – a graphical display of the "fluid and multilayered relationships between Web navigation and information retrieval that students use while navigating the Web" (C.-C. Lin & Tsai, 2007, p. 689)[21]. These maps depicted the sequences of actions in search and retrieval visually, alongside quantitative metrics for: number of keywords (as a measure of search diversity); maximum depth exploration (how many pages consulted for each task); web page adoptions (how many pages used for each task); total depth of web page adoptions (depth of pages used); revisited web pages; additional web pages used for refinement (the method is not reported, but this could be measured

[21] Unfortunately the exact method of production for these NFM's was not reported in the study, or on a link (http://www.cctsai.net/NFM) which is now defunct, and appears not to be cached or redirected.

by tracking when questions were first answered, and classifying each subsequent page as a 'refinement' source).

This research indicated that the search strategies of six graduate volunteers on a socio-scientific task involving both recall and exploratory search, and a 'social-related' task involving mostly recall could be classified in to two types: match or exploration. They suggest that the exploration group "usually used richer keywords to find relevant pages, browsed and revisited more pages deeply, selected multiple sources to complete tasks, and refined previous answers with more conscious reflection" (C.-C. Lin & Tsai, 2007, p. 691). They thus conclude that exploration students tended to "compare, filter, and integrate information when searching on the Internet; by contrast, members in the Match group showed more simplistic searching strategies when seeking materials for a specific task" (C.-C. Lin & Tsai, 2007, p. 692).

This method is of particular interest given its attempt to assess how students' commitments to treating information are made explicit. However, the use of two tasks (the 'social-related' task of which – although not stated in the paper – was largely recall based), may be problematic. In particular, such tasks are unlikely to fulfil the sorts of educational aims discussed in section 4.3.1 in relation to 'exploratory search', nor are they likely to involve the kind of complex information processing related to epistemic activity around conflicting information in multiple document processing. Furthermore, analysis of the NFMs appears to have both been a source to derive the scoring mechanism from (number of pages/keywords/etc.) and a way of classifying students, to then make claims regarding their scores on those metrics. That is, there is a circularity in the assumptions which may be problematic – analysis of the NFM appears to have been used both to derive groups (match v. exploration) and metrics to assess those groups by, which were then used to support the existence of those groups. Given that the metrics used were embodied in the NFM this is a concern, although the general approach – both of visualisation of navigation flow, and of deriving metrics from search log behaviour – is certainly interesting and will inform my own approach to foregrounding information seeking commitments.

The NFM approach is also problematic given that metrics around keyword numbers, websites visited, and depth of navigation might all be interpreted as signs of 'lostness' (difficulty in navigating to useful information). Thus further work should be conducted – while some such work may seek to operationalize these activities in relation to concepts (such as 'lostness'), my own will focus instead on the ways in which users interpret and use these metrics in action (as shall be discussed further below). Indeed this point is further reinforced by one subsequent study which contradicts these earlier claims suggesting:

> *Two distinct groups of students could be discerned. The first consisted of more competent students, who during their navigation visited fewer relevant pages, however of higher credibility and more specialized content. The second group consists of weaker students, who visited more pages, mainly of lower credibility and rather popularized content.* (Dimopoulos & Asimakopoulos, 2010).

One possible explanation for these contrasts, is that Dimopoulos and Asimakopoulos not only tracked metrics of user behaviour, but also websites visited – including measures of site text quality (reading level, etc.), structural indicators of quality (number of colours in images, for example), etc.

which allowed them to assess the *quality* of navigation in a deeper – although not unproblematic – way (Dimopoulos & Asimakopoulos, 2010).

This issue highlights the importance of attempting to understand the sensemaking significance behind particular semantic moves in a given environment. Thus, Greene et al. (2010) described one method of trace analysis for epistemic beliefs built on information moves. Other examples of such trace capture could also be structured such as to gather student data in particular ways – some of which may be quite naturalistic (capturing search queries, or Facebook posts to explore 'problems' encountered, or interactions made (S. Knight, Buckingham Shum, & Littleton, 2013b)), and others of which might push students into information structuring activity in which they would not otherwise engage, such as argument mapping (De Liddo, Buckingham Shum, Quinto, Bachler, & Cannavacciuolo, 2011). The benefits of such structuring are further discussed in section 10.2.

In encouraging explicit conceptual structuring by learners, and claiming that this captures information about what they are doing, some may argue that we are simply reifying the constructs we have set out to explore. That is, if we are interested in epistemic beliefs, and set up a system to push students to make epistemic beliefs explicit, it does not matter whether those students *have* underlying epistemic beliefs because the system forces them into making some (it makes them reify). While for psychologists who wish to uncover underlying beliefs this is problematic, I do not see this as a concern for this project, because in the discursive, sociocultural, pragmatic approach presented the interest is in beliefs as "theory-in-action". In this view, the claim is not that the measurement of beliefs is not possible, but rather that when we take measurements, the discursive context is fundamental to the practices being observed, and the ways that the beliefs are instantiated in action. Thus, LA provides a means to tackle the static, decontextualized view of epistemic beliefs instantiated by questionnaire methods, offering a more authentic perspective on epistemic action than experimental contexts.

8.2 Epistemic Action and Epistemic Games

One approach which has explored this action-oriented perspective from an epistemic (although not 'epistemic beliefs/cognition'[22]) framework is that of Shaffer and his work on Epistemic Games as models of 21st Century assessments (Shaffer, 2008). Shaffer (2006) argues that epistemic frames – as ways of viewing what should be known, as normatively defined and achievable – and the games built around these frames, are important new ways of assessing a student's understanding of what it is to be engaged in some particular (domain related) knowledge activity. Thus, in epistemic games, students are asked to engage in some computer based activity, through which various sorts of pre-operationalised decisions are captured and mapped such that a map of their activity on various epistemic-facets can be created which offers a picture of the distance between their own activity, and that of an expert in the particular field they are working on.

However, while this is more akin to the action-oriented suggestion above, Shaffer's focus on *episteme* as a relationship between discursive practices and knowledge structures places less focus on the action, and more on the context in which that action occurs (that context of course

[22] To my knowledge, only one conference paper (M. D. Johnson, Reimann, Bull, & Fujita, 2011) has made this link between the two related areas of research previously.

representing the *episteme,* or community in which the action occurs)[23]. While this is no doubt important, it still glosses the choices that epistemic agents make in varied contexts – which may or may not be related to the occupation in which they are engaged or employed. This perspective also approaches epistemic-actions as contextually bound, but not context-creating – that is, the theory around epistemic games is not interested in the implications of particular epistemic moves outside of their appropriateness in the particular context of the game. Furthermore, while there may well be particular epistemic practices in particular occupations, the role of epistemic-action in everyday life goes beyond these particular occupational practices, which may well be instantiations of more general (non-epistemic) traits (and indeed, may well be good ways to learn to apply such general traits). In addition, the approach is also – at least arguably – agnostic as to the particular types of epistemic practices we would wish our students to adopt. That is, while it does hold – and assess – a set of normative practices, adopted from 'experts' in the community, to be important it does not assess these practices except in the context of the communities of practice from which they are drawn. This uncritical adoption of occupational norms is problematic methodologically for at least two reasons: Firstly, inter, and intra-group variability makes it difficult to make judgements regarding the appropriate community against which one's epistemic behaviours ought to be assessed; secondly, it raises an is/ought concern – something *being* the case does not necessarily imply that it *should* be thus.

What Shaffer offers, is a method through which one can imagine a set of normative (dynamic) principles being created which, through some appropriate assessment task, allow students to explore their own discursive practices and knowledge map in a supportive environment which could offer them formative advice for developing their epistemic practices. The point here is two-fold: first that the particular assessment system – as a dynamic, formative, epistemic-practice (rather than product) based device – may be a better epistemic practice in itself; second, that even if epistemic practices cannot be 'transmitted', this sort of system may aid in the formative guidance towards some set of normative practices which are noted and co-constructed as 'epistemically valuable'.

This point is also important because it places a different sort of burden on the individual student. The expectation is not that all students should be equally good at *acquiring* knowledge, but rather that they: "*learn* whom and what to trust. And [they] do so, in part, by learning about [their] own and others' cognitive selves. While we may not be able to choose to improve many of our faculties, we can choose to be sceptical about them, to override them, to ignore them." (DePaul & Zagzebski, 2003, pp. 250–251)

8.3 Discourse-centric Trace – A Path to Epistemic Commitments

A tool for such analysis may come through the use of trace data, which is more or less implicitly created by the student. For example, Stadtler and Bromme (2007) analysed the ways participants found, extracted, and moved information, while Bråten and Samuelstuen (2007) tracked highlighting and similar behaviours in document processing – which could be used to explore information about their beliefs (e.g. visiting few websites indicates trust in those sites visited (Greene et al., 2010)). Importantly in this study, users were either given evaluation prompts regarding multiple documents in the medical domain, or not, and those who received such prompts subsequently recalled more

[23] This focus then is on agents engaging in behavior fitting to "contexts" (qua, community practices), rather than qua community aims and purposes.

facts and were better able to evaluate sources. To return to the points made earlier in this work around epistemology and assessment – we should give serious consideration to the activities we ask students to engage in and their relationship to learning and knowledge, and how best to assess. Certainly if systems of prompts were to promote worse learning or relate poorly to an epistemological stance, we should be concerned, where however they can support high quality learning such methods should be explored.

Furthermore, Greene et al. (2010) point out that many behaviours which would ordinarily be difficult to observe can be explicitly elicited in the context of Computer Based Learning Environments (CBLEs), for example:

> …participants who report belief in objective truth and omniscient authority may self-regulate quite differently than participants with a desire to evaluate multiple forms of justification. Likewise, participants who believe in the inherent subjectivity of all knowledge may, on average, select more representations than those who look for an objective truth. (Greene et al., 2010, p. 254)

The claim is thus that epistemic beliefs will be brought to bear on knowledge tasks in ways that can be meaningfully captured, in particular using technology systems (e.g. the way people represent knowledge in mind maps). Trace data thus offers direct access to real-time behaviours in unobtrusive ways, and is thus high in external validity, although it is of course within the context of the system which is set up to capture such information. Furthermore, while trace data is unobtrusive, it may give an incomplete picture. In particular, people may have reasons for some behaviours which cannot be probed using such data; these reasons may range from epistemic (it is not epistemically 'sophisticated' to hold a multiplist view of knowledge regarding the nature of the earth as flat), to practical (ICT failures), to pragmatic (the demands of the task place a short time restriction on the activity), and so on. Thus, it is important to remember that while analytics regarding epistemic beliefs may be – at best – a dirty lens onto those beliefs, when analytics are considered *in action* as a tool for sensemaking, they may provide an insightful tool for learners to dissect their own metacognitive and self-regulatory – or, co-regulatory – behaviours.

8.4 Making it Explicit – Epistemic Commitments in Exploratory Dialogue

An approach based on a rather different epistemological, theoretical background, and as such appropriate for consideration in the analysis of language as a tool 'to do' rather than 'to represent' is that of Mercer and colleagues. That work has focussed on the ways in which language is used "as a social mode of thinking – a tool for teaching-and-learning, constructing knowledge, creating joint understanding and tackling problems collaboratively" (Mercer, 2004, p. 137).

Evidence indicates that collaboration – and high quality discourse – are strongly related to positive educational outcomes – but only if they are mediated by the kind of reasoned discussion which is known as Exploratory Talk (Mercer & Littleton, 2007; see also the collection edited by Littleton and Howe (2010)). "Wherever education is taking place, commonality – a shared perspective – is key, and dialogue[24] is the tool used to create such a perspective (D. Edwards & Mercer, 1987)" (S. Knight, 2013a). This shared perspective has been termed "common knowledge" (D. Edwards & Mercer, 1987), the body of shared contextual knowledge which is built up through discourse and joint action,

[24] Dialogue and discourse are used interchangeably in this paper

and forms the basis for further communication. Thus, in this perspective, "common knowledge" forms a key constitutive part of context for speakers in a discourse, as well as being a fundamental aspect of education – in which a mutuality of understanding is crucial. Recently Littleton and Mercer (2013) consider the complexity of common knowledge context as both historical and dynamic :

> *Successful interthinking requires partners to have, and to develop, a foundation of common knowledge to underpin their discussions. We have distinguished two types of common knowledge, both of which can be important. The first of these is accumulated through the activities of a group, as members develop a shared history. They have knowledge in common because it has been generated by their joint activities and associated conversations. It is the kind of common knowledge which allows a teacher to refer only briefly to the content of a previous lesson and expect students to have some recollection what it had been about. We have called this dynamic common knowledge, because it is produced by the dynamics of the group's own extended activity. The second type, which we call background common knowledge, is that which any established member of a community of practice can take for granted as being shared with other members and does not therefore need to be explained from first principles. It is the kind of knowledge which enables any two physicists, Beatles fans or people who have grown up in the same town to take certain kinds of understanding for granted, even if they have never met before.*

Indeed, the strong consensus among researchers is that in a variety of contexts, high quality dialogue is associated with learning (see the collection edited by Littleton and Howe (2010)). That research shows that, "Engaging children in extended talk which encourages them to 'interthink' and explain themselves…stimulates both their subject learning, and general reasoning skills (Mercer, Dawes, Wegerif, & Sams, 2004; Mercer, Wegerif, & Dawes, 1999; Mercer & Sams, 2006; Rojas-Drummond, Littleton, Hernández, & Zúñiga, 2010), as well as their social and language skills (Wegerif, Littleton, Dawes, Mercer, & Rowe, 2004)" (S. Knight, 2013a).

Mercer and colleagues have extensively researched such dialogue, developed an intervention strategy called 'Thinking Together', and highlighted a particular form of productive dialogue which, adapting the term from Douglas Barnes' (Barnes & Todd, 1977) original broadly individualistic description, they have termed 'exploratory'. They contrast this with two other types of, typically less productive, talk – disputational, and cumulative, as in Table 11.

Table 11 - Summary of Typology of Talk

Type of Talk	Characteristics	Analysis
Disputational	"Characterised by disagreement and individualised decision making. There are few attempts to pool resources, to offer constructive criticism or make suggestions."	"short exchanges, consisting of assertions and challenges or counter-assertions ('Yes it is.' 'No it's not!')."
Cumulative	"Speakers build positively but uncritically on what the others have said. Partners use talk to construct 'common knowledge' by accumulation."	"Cumulative discourse is characterized by repetitions, confirmations and elaborations."
Exploratory	"Partners engage critically but constructively with each other's ideas. Statements and suggestions are offered for joint consideration. These may be challenged and counter-challenged, but challenges are justified and alternative hypotheses are offered. Partners all actively participate, and opinions are sought and considered before decisions are jointly made. Compared with the other two types, in exploratory talk knowledge is made more publicly accountable and reasoning is more visible in the talk."	Explanatory terms and phrases more common – for example, 'I think' 'because/'cause', 'if', 'for example', 'also'

(Adapted from Mercer and Littleton 2007, pp. 58–59)

Similar characterisations of effective dialogue have emerged from the work of other researchers across a range of ages (Michaels, O'Connor, Hall, & Resnick, 2002; Resnick, 2001). In this research, Accountable Talk is described as encompassing three broad dimensions:

1. accountability to the learning community, in which participants listen to and build their contributions in response to those of others;
2. accountability to accepted standards of reasoning, talk that emphasizes logical connections and the drawing of reasonable conclusions; and,
3. accountability to knowledge, talk that is based explicitly on facts, written texts, or other public information. (Michaels, O'Connor, & Resnick, 2008, p. 283)

As with the typology of talk developed by Mercer and colleagues, the emphasis of Accountable Talk is not on learning particular subject or topic knowledge and language, but rather on learning to engage with other's ideas, and in doing so use skills of explanation and reasoning, learning to use language as a tool for thinking and – in the terms of Mercer and Littleton – *interthinking*.

Educational researchers within the sociocultural tradition would highlight the importance of dialogue as not only constitutive *of* context (that is, representing context), but constituted *in* context (that is, creating context). This distinction highlights the need to understand that context should not only be assumed from the state of the dialogue at any particular point (assuming dialogue represents context), but rather, we should also explore the ways in which the context changes over time as a feature of the dialogue (assuming dialogue involves the co-construction of context).

It is in part due to this consideration that sociocultural researchers have emphasized the use of both qualitative and quantitative methods, in which – in contrast to some other qualitative methods – the quantitative data is taken to aid the understanding of the qualitative, as opposed to the converse. It is thus that such researchers often include excerpts of talk, concordance analysis, and other contextual markers such as cohesive ties in their reporting. This technique is drawn from 'systemic functional linguistics', which takes it that types of text have contexts by being members of a particular genre, which is revealed through the way such texts are written[25] – thus, context is imbued into texts at the time of writing, for example through the use of contextualising key words (e.g. describing a book as a 'textbook' compared to a 'course book' raises different expectations), headings (e.g. the traditional article headings from abstract to conclusion), and positioning markers (e.g. positioning with respect to existing literature, such as citations and phrases such as 'we agree with'). In sociocultural discourse analysis, this assumption is adapted from that of 'texts' to the co-construction of context through dialogue in which "'context' is created anew in every interaction between a speaker and listener or writer and reader. From this perspective, we must take account of listeners and readers as well as speakers and writers, who [co]create meanings together" (Mercer, 2000, p. 21). It is thus that sociocultural researchers may seek to understand the temporal aspects of context, as involving continuity across talk, by looking for repetition of words, synonyms and ways of approaching problems, to understand how "speakers can jointly, co-operatively create cohesion in…their speech" (Mercer, 2000, p. 62). Such analysis has commonly be conducted using concordance software, which facilitates the exploration of 'Key Words In Context' (KWIC) by

[25] See Halliday, Hasan and Christie (1989)

displaying words searched for in their original context (typically, showing a sub-portion or whole sentence in which the keyword is located).

A related perspective, which arose from some common work (including Edwards and Mercer's 'Common Knowledge' (1987)) is Discursive Psychology – which has a particular interest in "the kinds of naturally occurring interactional talk through which people live their lives and conduct their daily business" (D. Edwards, 2005, p. 258). This approach is explicitly non-cognitivist in nature, with less interest in the ways people reproduce cognitive constructs via questionnaires, interviews, experiments and so on. Discursive Psychology thus describes cognitive psychology as treating discourse as "an abstract logical and referential system – language – rather than a locally managed, action oriented, co-constructed resource" (Potter & Edwards, 2003, p. 95). It is thus explicitly motivated by the type of pragmatic theoretical perspective discussed in the earlier sections of this report. This perspective explicitly argues that language is a tool to represent the world, where talk may be seen as "a window (a dirty window, perhaps) on the mind" (D. Edwards, 1993, p. 208).

Discursive psychology is of particular interest because, while sociocultural discourse analysis is interested in language as a cultural tool for learning, Discursive Psychology has focussed more on the respecification of commonly held psychological constructs in terms of their linguistic, situated, co-creation. That (respecification) is, it is interested in the emergence and fluidity of psychological constructs "in action" as co-constructed in, and mediated by, language, as opposed to their status as neural or cognitive entities. Thus, for a specific analysis of a construct – epistemic beliefs – this approach may be highly appropriate.

The importance of context is also a familiar topic in Natural Language Processing (NLP). In accord with their psycho-social analogues in the educational world (sociocultural theory, discursive psychology, many varieties of humanistic psychology) much emerging work has eschewed the focus on cognitivist models which seek to understand the *beliefs* and *intentions* of agents, instead focusing on attempts to understand the *contextual* and *action-based* nature of talk, as a thing "to do" rather than its role in the abstracted expression of underlying beliefs. This is the shift "from the view of language as a tool of representing the world to its view as a means of interacting with the world." (Peregrin, 2005, p. 39)[26].

In the context of epistemic beliefs, Discursive Psychology posits that we should not see beliefs and communication as "two separate 'objects' that can affect each other, but as more integrated aspects of cognition and/or behaviour" (Österholm, 2010, p. 242). This perspective describes "the activity, the discourse, as the site where epistemological beliefs come to existence, through explicit or implicit references to prior experiences (epistemological resources)" (Österholm, 2009, p. 262).

Österholm's argument is that this perspective can be combined with Hammer and Elby's 'resources' model, in which epistemic beliefs are viewed neither as fixed, nor developing cognitive models ranging over one or more domains, but are rather seen as dependent upon the resources available to the cognizer at any time – in our case, discourse being key to this. The emphasis of this perspective as "theory-in-action" – in which context, domain, culture, and task conditions interact – makes an important point, that:

[26] See also the rest of this Special Issue of *Pragmatics and Cognition* devoted to discussing the work of Robert Brandom

> *A sophisticated epistemology entails context-sensitive judgements. Thus they point out that it is not very sophisticated to view the idea that the earth is round rather than flat as 'tentative' whereas theories of dinosaur extinction do require a more tentative stance (Barzilai & Zohar, 2012, p. 42).*

Importantly, Discursive Psychology is also not interested in the socio-political or phenomenological elements of language. Instead, its focus is on the use of language as a tool – language, *in use.* This set of approaches recognise that consideration of the usefulness of knowledge and language 'in action' at work in the world, is preferable to trying to get at the 'real world'. As such, their focus is not on verification of correspondences between linguistic labels and 'things in the world', but on the ways in which knowledge and language acts on and in the world. The implication of such approaches is that information needs should be considered as they relate to communities of justification, and the purposes for which knowledge is deployed (e.g. practical v. academic nursing knowledge). Thus, the interest is not "what does it [language] represent? But, what is going on?" (D. Edwards, 1993, p. 218).

Recently, Sandoval (2012) has made similar claims, calling for epistemic cognition researchers to take seriously a 'situated' approach, building on similar theoretical (post-Vyogotskian) foundations to this work. In that work the point is made that:

> *One important way to understand the epistemic ideas that people bring to bear is to examine their participation in practices of knowledge evaluation and construction. Changes in the form of participation are indicators of changes in the meaning that individuals make of the activity in which they are engaged...Change in participation can indicate a shift in epistemic perspective, but it is the shift itself that suggests what particular epistemic ideas are brought to bear in the first place* (Sandoval, 2012, p. 350)

Thus, the focus is the emergence of information needs within groups, and the use of implicit and explicit criteria to assess the suitability of information for meeting information needs – both those arising from the groups themselves and those arising from the task, setting, and so on. Importantly, "…information seeking is not carried out for its own sake but to achieve an objective that lies beyond the practice of information seeking itself." (Sundin & Johannisson, 2005b, p. 107). Within this pragmatist socio-cultural epistemology:

> *…judging the truth of an idea becomes a question of whether the idea makes any difference to practice or not, whether the idea provides us with a useful tool or not. (Sundin & Johannisson, 2005a, p. 27).*

That is, analysis cannot focus solely on whether some clearly defined need which reflects a deficit in the 'real' world, is satisfied because that is not the nature of knowledge construction or use. Information needs arise from, and are addressed through activities in which knowledge is distributed, bi-directional, and in constant negotiation – it is through this process that our information needs are defined, and addressed; but that is not to say that for each information need there is some 'correct' abstracted answer, rather – answers are situated

8.4.1 Situated, Collaborative Epistemic Activity[27]

Indeed, this perspective – of right answerism – is exactly that discussed in the preliminary sections to this work. Student's framing of activities as the production of answers for the teacher or test, as opposed to gaining understanding, implies a particular epistemic stance towards their education (Hutchison & Hammer, 2010). Indeed, such perspectives may be observable in the behaviours of collaborative groups (Scherr & Hammer, 2009), and the use of collaborative knowledge building tools (such as Knowledge Forum) may not only encourage higher levels of engagement, but also greater collaboration, reflection, and a shift to more constructivist epistemological beliefs (see Hong and Lin (2010) for evidence in teacher trainees).

It is of interest to consider how language mediates and represents learner's views on their learning. Hutchinson and Hammer (2010) provide a science classroom case study, illustrated by excerpts of the type seen in sociocultural discourse analysis, in which framing of a broadly sensemaking perspective (and in my terms 'exploratory' or dialogic) is given in contrast to a more absolutist perspective. For example, at one point a student (Bekah) offers and explains an equation to illustrate her understanding – this is taken up and referred to collectively as "Bekah's Law", illustrating – in the terms described above – a cohesive tie (the repetition of terms through a text) demonstrating a type of common knowledge built up in that classroom (Hutchison & Hammer, 2010).

This claim – of a relationship to exploratory talk – is further supported in Rosenberg, Hammer and Phelan's work (2006). In that study, a case study was presented of a 15 minute discussion of the 'rock cycle' by a group of 8th graders – again, making use of dialogue excerpts to illustrate. Rosenberg et al., note that in the initial stages students were engaged in largely unproductive talk (there was some accretion of knowledge, with little explanation or evidence of understanding – it was largely *cumulative* in nature), suggesting this was because: "They [were] treating knowledge as comprised of isolated, simple pieces of information expressed with specific vocabulary and provided by authority" (Rosenberg et al., 2006, p. 270). After a brief intervention by the teacher, suggesting the students might build on their own knowledge, this talk instead shifts to more productive dialogue, seeking coherence and understanding in trying to create a theory and use terms they understand – the description, and excerpts provided here suggest this talk might be characterised as more 'exploratory' in nature.

8.4.2 Exploratory Dialogue – Epistemically Salient Dialogue

Furthermore, the epistemic nature of the sort of talk described by Mercer and colleagues, has recently been described by Reznitskaya and Gregory (2013). In that article, they point out that the sort of dialogic talk related to exploratory talk (described in Wegerif, 2006) stands in stark contrast to the kinds of 'monologism' described by Bakhtin (1984) in which dialogue portends to readymade and singular truth. In doing so, they elaborate theory which is in strong accord with that described above. They point out (p.118-119) that dialogic learning contexts are:

1) About recognising expertise and its limits

[27] This section comprises a review of articles which cited Rosenberg et al., (2006) and involved analysis of collaboration as a data source (not just as an interventional strategy, pre/post interviews, etc.), for analysis of epistemic cognition, and articles citing those – this numbered 5 articles, one of which (Sandoval, 2012) was discussed above. Rosenberg et al., (2006) do not cite any work explicitly using a similar method to their own in the epistemic cognition context.

2) Centred on divergent questions – this is key for us, talking points, exploratory talk, ill-structured problems
3) Metacognitive in nature, involving both products and processes, awareness of others

And indeed, they agree with the (pragmatic) Dewey (1938) that "inquiry, understood as the search for reasonable belief, has the general structure of generating hypotheses in response to well-formed questions and testing those hypotheses with evidence and arguments in order to arrive at the most reasonable conclusions" (Reznitskaya & Gregory, 2013, p. 119) – an accord which fits well with a focus on ill-structured exploratory search problems.

8.5 A Proposed Model of Epistemic Commitments

Reznitskaya and Gregory (2013) note that, more sophisticated epistemic cognition of the 'evaluativist' variety, is closely associated with the kind of exploratory talk which – as described above – is associated with educational gains. However, while elaborating a theory of dialogic talk in the context of epistemic cognition, Reznitskaya and Gregory's analysis focuses on the developmental classificatory system of Kuhn (1991) in which learners develop from absolutists, to multiplists, to evaluativists. While the theme of epistemic development is related here, concerns with this approach – and its corresponding methodological implications – were raised above (and indeed, by Reznitskaya and Gregory, pp.125-6). This type of dialogue is closely associated with a component of my approach to epistemic commitments around openness to ideas, and justification for them. The inclusion of this approach in studying epistemic commitments receives further support from the fact that the 'superordinate' (overarching non-domain specific) codes used in epistemic games include many of the keywords we would expect to see in exploratory dialogue episodes – *because, so, I think*, etc. – as described above (although in the epistemic game platform, computational analysis extends beyond simple keyword matching).

In addition to exploratory dialogue, other components of epistemic action are highlighted in the literature. Earlier I noted Mason et al.'s (2009, p. 69) claim that across models of epistemic cognition, there was a focus on the certainty, simplicity, source and justification for knowledge. I then noted Tsai's (2004, p. 109) framework for information commitments, comprised of: standards for correctness; standards for usefulness; and searching strategy. We can imagine recasting these two positions such that our focus is on:

1. Which sources (of testimony) are used – comprised of credibility decisions (from corroboration of sources, to trust in authoritative sources)
2. How they are used (in action – to justify claims, to make decisions) – comprised of justifications and source use (from dialogic approaches using talk of an exploratory nature, to attempts to match information to answers and use basic factual information)
3. How links between them are created (or not) – comprised of claims, (explicitly in language and through structured environments, as well as implicitly through search patterns) made around connectedness of concepts (from a holistic perspective of knowledge to a piecemeal)

This recasting aligns well with the specific context being studied here – that of online collaborative information seeking. It also provides three conceptually distinct (although probably empirically associated) constructs for study. In the two other models highlighted – Tsai's information commitments, and the general model described in Mason's analysis of the literature – it is not clear to me that each component can be conceptually distinguished. Specifically, 'certainty' in the general

model seems likely to be a function of justification and simplicity (the purpose for which the information is being deployed, and the other information to which it is being associated – and indeed, whether one holds a complex enough view of knowledge to recognise the instability of certain information). Similarly, it is not clear to me that 'searching strategy' is a useful conceptualisation of an information commitment given its strong relation to the tools at hand, and (as in the other components) the type of task set and justificatory framework required for that task. In contrast, building on the philosophical work discussed above, the proposed focus on who we believe, what we do with information, and how we relate information still provides a lens onto the sorts of beliefs explored in other models, from a recast model of epistemic commitments. For example, certainty is recast in light of our standards for credibility, explanation, and relating components of information such as new and old, or geographically located information; simplicity is most clearly related to the third focus on connectivity; source to the first; and justification to the second. Furthermore, the rhetorical shift both in the foci, and in the notion of 'commitments' over 'cognition' motivates an operationalization centred on:

1. Source selection, the number of links opened (corroboration of sources, operationalisable as 'recall') and the types of links repeatedly visited (e.g. is the 'BBC' used an authoritative source).
2. The type of justificatory framework used, the assertion of information (perhaps closely related to 'cumulative' talk and the style of search emphasising 'precision' to 'match' information to with little consideration to its wider impact) versus reasoning and understanding activities (closely related to dialogic and exploratory dialogue)
3. The sorts of connections made between concepts both in structured environments (As shall be discussed below in section 10.2), and in the ways that users build links between information in their search patterns (building on search terms by rephrasing and appending new query terms, following internal links, and using terms from opened sources to find new ones all imply some commitment to holistic perspectives on knowledge).

This model thus describes both a conceptual, and practical means to explore epistemic commitments in information seeking environments, and will be the model adopted in this work.

8.6 Section Summary – bringing epistemic commitments and information seeking together; consolidating the research

The preceding sections have sought to outline:

1. The nature of learning analytics (section 1) introducing the notion that they might be viewed through the lens of 'Assessment, Epistemology, Pedagogy' (which I described as LÆP-ing into the middle space at LAK13 (S. Knight, Buckingham Shum, et al., 2013a)) (section 2.1), expanding that relationship further in section 2.2 with respect to specific types of analytics.
2. Section 4 also started to introduce a topic of particular interest – epistemic cognition. This was highlighted as an interesting topic for assessment (and policy thereof, section 2.2.4) with respect to our epistemological stance. It was also noted that information seeking – related to epistemic cognition – was a challenge for many young people, including those in countries which assess (indirectly) such skills (section 2.2.4 and 3).
3. I then introduced some models of information seeking (section 4), highlighting lessons to be learnt from this work, and in particular noted that 'exploratory search' (section 4.3.1) is explicitly learning-oriented. I also noted that the measurement of 'success' in such contexts be necessity moves beyond a simple accord between facts in the world and retrieved tokens of information (section 4.3.2) – relating this claim to the epistemological stance advanced earlier.

4. However, I also noted the increasing interest in collaborative information seeking – with a growing body of evidence indicating the importance of such activity (section 5). This section (5) thus outlined some additional factors of significance in CIS, including awareness and communication. CIS was thus introduced as an interesting notion because: evidence indicates it is quite common, including in education; some of the factors related to CIS are of educational interest – including communication and awareness tools (of the CSCL variety – as shall be discussed in section 10); learning analytics may also offer insight into, and support for, this type of collaborative activity.
5. I then presented some research indicating the importance of epistemic cognition in information seeking (section 6), noting though, that in terms of the epistemology presented earlier, and evidence regarding dialogue in CIS (section 5.3) much of this research could be critiqued for excluding the notion of information commitments (section 7) in active use of information in collaborative contexts. Section 8 thus presented some work towards this end, and made some suggestions for further developments to explore CIS in the context of language in-action around information commitments.

The claim of this thesis is not just that we should be interested in systems to mediate help-seeking behaviours of information seekers (see e.g. Puustinen and Rouet (2009)) but that we should see the collaborative interaction as a fundamentally important way to improve information access – in a broad, epistemically complex sense – and that search provides a good context in which to create and observe such interactions. Two key related – yet theoretically distinct – elements of interest in this research paradigm are: topics around exploratory search in which information needs are ill structured and searches involve learning and investigation; ill structured issues on which exploratory search is based, such as socio-scientific issues involving epistemic judgements; the collaborative interaction around such tasks, which can be characterised as dialogic and thus involving exploration of multiple perspectives. As Gagnière, Betrancourt, and Détienne (2012) summarise:

> *Ge and Land (2004), in proposing peer interactions as a scaffold of ill-structured problem solving process, pointed out its potential to improve different phases of this process. In the problem representation phase, peer interactions may direct each other's attention to particular features of the problem they do not understand, leading to a more complete problem representation. In addition, it seems that peer interactions are useful for developing solutions, in exposing students to different perspectives. In the argument construction phase, peer interactions provide a context for constructing arguments and making justifications (A. King, Staffieri, & Adelgais, 1998). Lastly, in making the thinking process visible and available for examination, peer interactions have a potential to improve the monitoring and evaluation phase (Ge & Land, 2003)* (Gagnière et al., 2012, p. 75)

We can start to sketch a diagram relating these aspects as below in Figure 7. Of course, a key aspect of such a research paradigm which has not yet been discussed is the tasks they are based on, and it is to this issue that I now turn (section 9), before discussing some specific tools (section 10) to support learners, and provide meaningful data for analysis.

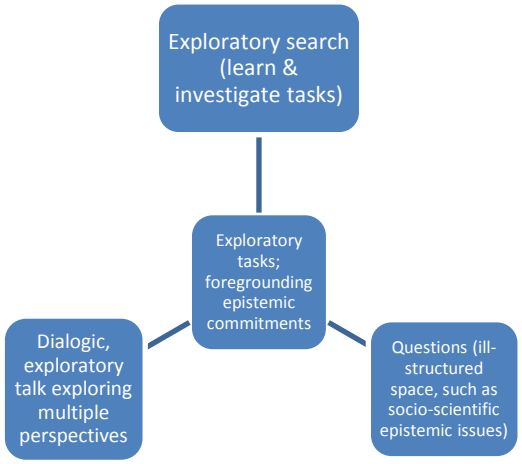

Figure 7 - Relating Components of this Thesis

9. Experimental Tasks To Probe Epistemic Commitments

Having discussed epistemic commitments in the context of collaborative information seeking (of an exploratory nature), I now relate these to particular types of task to probe such activities. Specifically I discuss the sorts of task we might use to: probe epistemic commitments in online search (section 9.1); prompt collaborative communication (section 9.2); and to frame collaborative information seeking (section 9.3). I have discussed two key approaches: multiple document processing; and information seeking tasks. However, in both of these instructions must be given to participants to provide context to their activity – and it is to this that this section now turns.

9.1 Overview of Tasks

Much of the CIS and epistemic cognition gives participants pre-assigned tasks, and requires them to cover particular aspects of topics (or particular types of response) in meeting the task requirements. However, the purpose for which participants seek information may matter to how they perform, and what sorts of activity they engage in. In particular, in the epistemic cognition research a distinction has been made between 'summarising' and 'argument construction' conditions in dealing with multiple documents. That is, how do participants behave if they are asked to construct an argument in multiple document comprehension tasks, versus simply being asked to summarise the content of those documents?

However, the current evidence indicates conflicting evidence with respect to summarising, versus argument construction in task design to probe epistemic cognition. One study found that readers who are asked to construct arguments and summarise information build deeper and more integrated perspectives than those asked to produce general overviews, with those with more sophisticated epistemic beliefs gaining more benefit from the argument task (Bråten & Strømsø, 2009), and another study suggests that summarisation tasks are superior to argument tasks (Gil, Bråten, Vidal-Abarca, & Strømsø, 2010). This may relate to context, and instruction, specific factors, for example students were more likely to engage in note taking and intertextual links to construct arguments than to summarise – and subsequently performed better (Hagen, Braasch, & Bråten, 2012). Significantly for my purposes, trace data appears to be superior to task-specific self-report; for example, tracking students' highlighting of key terms is more predictive of student's reading task

performance (which is predictive of PISA literacy tests) than self-report on practices (Bråten & Samuelstuen, 2007).

Another factor which is relevant to task design is the type of support or prompt given to participants in their search process. For example, Wu and Tsai (2010) discuss research comparing two groups – an unguided group who were told to find and summarise information, and a guided group who were reminded to search relevant information from multiple perspectives (science-oriented, ecology-oriented, etc.). They report that students in the guided task condition outperformed the unguided students in terms of richness and extent of understanding, and use of 'comparing', 'inferring' and 'explaining' skills, but that these students were no better at applying their reasoning after the task – a flaw in such support which might be addressed by task design to support high quality collaborative dialogue.

9.2 Task and Communication

Thus a key component of task design may be both the purpose of the activity, and the collaborative framing provided. One study explored both communication and epistemic cognition in the context of the co-construction of arguments or summaries by triads in a wiki-environment through reciprocal peer questioning. This study found that students with lower epistemological beliefs gained more from argumentation than summarisation (while there was no difference for more sophisticated students) (Cho, Lee, & Jonassen, 2011). That study also reports overall benefits to collaborating on argumentation, suggesting that such tasks (and CSCL tools) hold benefits for all students (Cho et al., 2011). In particular, those authors noted that particular forms of questions – e.g. deep-reasoning questions, comprehension questions – elicited particular responses – knowledge-integration, and knowledge-relating responses, respectively (Cho et al., 2011) – indicating the importance of dialogue quality in such environments.

Indeed, Park and Lim suggest that more sophisticated epistemic cognition is related to more advanced use of online communication tools, and – as Tsai (C. Tsai, 2004) suggested of web-based learning – such tools also support learners in developing perspectives of learning. Interestingly, although this study did not directly explore epistemic-cognition, it coded how students perceived online communication for learning, for example coding 'learning as acquisition' as related to knowledge from authority.

This factor of epistemic cognition in online learning is important, given the literature discussed in the context of collaborative information seeking (section 5), and the importance of dialogue for learning (section 8.4). Furthermore, one study reports that, on a Likert measure (with a 5 point scale ranging from "not at all typical of me", to "very typical of me") there is at least some support (2.51 average rating, 1.03 SD) for use of communication strategies in online learning – importantly, with those:

> *who conceived of knowledge as given and stable were less likely to take advantage of the opportunity for Internet-mediated communication…The reason for this might be that students who believe in given and stable knowledge do not see the point of participating in mutual negotiations, often involving multiple conflicting interpretations, about the meaning of subject content* (Bråten & Strømsø, 2006, p. 1038)

which is in accord with other work on online communities of practice sharing practices (e.g. (Teng, 2010).

9.3 Collaborative Information Seeking Tasks

9.3.1 Exploratory Search

In a systematic review of the literature to make recommendations for eliciting exploratory search, Wildemuth and Freund (2012) note:

> *a set of task characteristics associated with exploratory search tasks are identified: exploratory search tasks focus on learning and investigative search goals; they are general (rather than specific), open-ended, and often target multiple items/documents; they involve uncertainty and are motivated by ill-defined or ill-structured problems; they are dynamic and evolve over time; they are multi-faceted and may be procedurally complex; and they are often accompanied by other information or cognitive behaviors, such as sensemaking. (Wildemuth & Freund, 2012, p. 1)*

Building on literature describing these characteristics, they suggest some key lessons for exploratory search task design:

1. Tasks should be focussed on learning and investigation
2. Context and situation should be specified but the topic or request may introduce enough ambiguity and open-endedness to produce exploratory behaviours
3. Multiple facets should be included in the task and search topic
4. Possibility for eliciting dynamic and multi-stage search should be considered; in some cases tasks can be written to provoke this, but this will not always be the most appropriate approach
5. Data collection and evaluation should be aligned with the goals of the task

9.3.2 Imposed and Self-Generated Tasks

Noting a somewhat different distinction, Sams and Seligson (2011) briefly review the literature on task design in CIS, pointing out that while many queries are self generated, "imposed queries may result from classroom assignments, tasks given by a supervisor, or even requests by a friend" (Capra et al., 2011, p. 1). Their own research indicated that users were more engaged in self-generated tasks than in imposed tasks. This aligns with prior work indicating that imposed tasks result in more analytic approaches and direct paths to answers (Thatcher, 2006) and that – in children – self-generated tasks are preferred and performed better than assigned or semi-assigned tasks (Bilal, 2002). This trichotomy (Bilal, 2002) is an interesting extension though, in which a distinction is made between imposed, fully self-generated, and semi-assigned tasks in which the latter provides a main topic for the participant who must then pursue any aspect of the topic that interests them.

In other work (Capra et al., 2012) the development of an asynchronous collaborative task scenario to elicit naturalistic behaviours is described. That scenario asked students to search a news database for writing a paper, rating articles as they went, in advance of (a hypothetical) meeting with their team members. In addition, they 'seeded' a system so it appeared the participants were collaborating with a team of searchers who had conducted some searches and saved some articles previously. In this case the task design was based on a TREC task[28], chosen to ensure the task was engaging, and that not too many documents would be relevant (too easy) or too few (too hard). Participants were asked to research a paper on curbing population growth, and used collaborators'

[28] Text REtrieval Conference (TREC), see http://trec.nist.gov/. "…to encourage research in information retrieval from large text collections" – provides a corpus and tasks for experimental research.

previous queries to understand how far the collaborator had got and what articles they rated highly, and to start their own search process (or get ideas when they were frustrated) (Capra et al., 2012). This suggests another aspect of task design – the way that the tools placed at participants; disposal are brought to bear on task completion, and collaboration.

Further analysis in this area comes from Gagnière, et al., (2012) who found some support for the hypothesis that metacognitive prompts positively impact collaborative information problem solving performance, reporting that at the needs definition stage – describing the qualities of sources required, and the keywords to find these – such prompts increased metacognitive language, and that overall those groups performed better. This indicates the role a system to support collaboration might have in supporting epistemic dialogue in CIS contexts, but in this case the task design itself was not particularly 'exploratory' in nature requiring participants to select three relevant resources on performing a press analysis. These last two studies, together, indicate an interesting area in exploring the ways that collaboration can be brought to bear on exploratory search tasks in the context of particular tool use – and it is to this that I now turn.

10. Software Tools: Introduction

Some work has been conducted on analysing or supporting information seeking, and of course collaboration, in various disciplines. However, while many systems support individual elements – in particular, either collaboration *or* information seeking – of this problem domain, most have limited scope for use in my research. For example, above (section 8.1) I mentioned Meta-analyzer (Hwang et al., 2008) – which generates dashboards of students' information seeking behaviour on pre-assigned tasks for teachers to view. However, this is not a suitable tool for collaborative information seeking (although some of its features may be). Similarly, there may be useful features (indeed, see Co-sense discussion below) in Bateman, Teevan and White's (2012) dashboard which takes individual's searches, and aggregates them to be represented to the user and compared to the search profile of an archetypal expert. They report that over time such a dashboard supported users to improve their search skills even where the expert comparison element was removed (Bateman et al., 2012). Indeed there is preliminary work to create a search engine based on these principles, to share "community search logs" – query and browsing histories pooled from collaborators – in order to assist in sharing knowledge and experiences among participants who view similar web documents (by providing each document a set of related documents and queries) (Flanagan et al., 2013).

Another tool, designed to support learners in making epistemic judgements in online information seeking around health resources – met.a.ware (Stadtler & Bromme, 2008) – aims to use ontological classification firstly to enhance the structure of notes, and secondly to assist a focus on relevant categories in searching, finding that prompts to attribute and rate sources were effective in raising metacognitive awareness and content learning. However, while this tool builds on Suthers' work on representational guidance (Suthers & Hundhausen, 2003; Suthers, 2001) which suggests that externalisation of information can both constrain and structure – for instance, through the use of ontological categories in this case – Suthers' focus was primarily in collaborative learning contexts, while this work's focus is on individual learning. Furthermore, the sites assessed in this context were pre-selected (15 sites), which were to be classified according to 6 content-specific labels (e.g. function of cholesterol, causes of high cholesterol, etc.). Thus, while the reference to computer

supported cooperative learning (CSCL) work is of interest, I would hope to draw from this literature in a *collaborative* context – and it is to this that I now turn, first with respect to CIS tools, and then CSCL tools.

10.1 Collaborative Information Seeking Tools

A number of tools exist in the CIS literature to facilitate the CIS process, including a number which facilitate collaborative reordering of search results through algorithmic mediation (e.g. FXPAL's Cerchiamo (Golovchinsky, Adcock, Pickens, Qvarfordt, & Back, 2008), and Querium (Golovchinsky & Diriye, 2011)). While building in such approaches may be of interest, in order to analyse explicit, discourse-mediated collaboration further facilities will be needed. Building on early work to visualise and provide facility to annotate search processes (Twidale, Nichols, Smith, & Trevor, 1995), SearchTogether (Morris & Horvitz, 2007) (and its extension, CoSense (Paul & Morris, 2009)) offered facilities to view collaborator query histories, page views, and comments/ratings on those pages viewed. This tool allowed users to divided searches such that: 1) for any search engine results page (SERP) each user sees only half of the results; 2) each user searches a different engine (e.g. Google, and Bing); 3) or finally, instant messaging (IM) could be used to manually divide up search tasks. Those researchers found that generally users preferred using the IM feature to automated division tools. A similar tool from a different research group – Coagmento (Shah, 2010) – has also integrated IM, shared query and page history, and annotations into a browser add-on, along with a shared document space (Etherpad) in which users may engage in collaborative writing around their CIS topic. Work on this tool reports positively on user experiences, and of interest to my research, reports that remotely located participants were more effective at finding diverse information than co-located (although, they also preferred audio-support to reduce cognitive load and negative affect) (González-Ibáñez, Haseki, & Shah, In press).

Other work has developed a tool to facilitate awareness in classrooms by displaying queries being made (ClassSearch (Moraveji, Morris, Morris, Czerwinski, & Henry Riche, 2011) and SearchParty (Gubbels, Rose, Russell, & Bederson, 2012)), displaying query-centric search recommendations and bookmarks to collaborators making similar queries to colleagues in a programming environment (Bateman, Gutwin, & McCalla, 2013) and co-located collaborative search (e.g. CoSearch (Amershi & Morris, 2008), discussed in section 5.3).

Paul and Reddy (2010) note that "studies of CIS have focused on how people find and retrieve information together, with little attention paid to how people work together to synthesize and understand the different pieces of information that are shared during a CIS activity" (Paul & Reddy, 2010, p. 321), suggesting that – in the context of an emergency department – awareness tools which display not only actions (e.g., searches made) but activity (the order in which those searches were made, embedded in context) provides superior support.

In work on online CIS Paul and Morris (2011) note again that most sensemaking research has focussed on individuals, rather than collaborators who – in a formative study – indicated that in CIS tasks:

1. Participants highlighted the importance of the temporality of search process – the chronological orderings of content were desired to better understand path of navigation,

along with a persistence of sensemaking products and the ability to make notes not only on pages found, but on the task itself
2. Participants highlighted the need for 'awareness' of collaborators' actions – in particular notifications for chats, page views and summary creations were desirable
3. These factors were particularly important for collaborators who joined a search task later (asynchronous searching) who found it hard to see what others were doing and distinguishing old from new information (Paul & Morris, 2011)

Paul and Morris (2011) thus designed Co-Sense, which provides an interface for four displays:

1. A search strategy tab containing: URLs visited by the group, and by individual group-members; keywords aggregated across the group, and for each group-member; the number of pages visited and keywords used; and a query timeline (broken down by member using colour coded lists)
2. A timeline tab containing: a chronological list of: queries issued; web pages viewed; chat messages; comments; and page ratings
3. A workspace tab containing: summaries for web pages saved (comments and ratings and who has visited it); and a 'notes' space for general comments
4. A chat tab containing: group-chat which was colour coded by group member; clicking on chat messages showed the webpage associated with that message.

In this research, they found that: at the start of the task most activity was devoted to sensemaking using these tools; group members tended to use the chat rather than commenting on individual pages; and the search tab and chat tab were the most viewed features, with the 'tag cloud' of queries being particularly useful (Paul & Morris, 2011). Given the exploratory nature of the task in this study, and the high levels of sensemaking supported by the tool, such design features give insight into potentially useful features for a tool to explore explicitly *epistemic* sensemaking in CIS.

10.2 Foregrounding Epistemic Commitments – CSCL Literature

As noted above, the CSCL literature has a rich history of supporting collaboration, and foregrounding particular assumptions in systems to support students in their reasoning and collaboration. Section 8.3 highlighted the use of discourse-centric trace as a path to epistemic cognition. Building on this work, and CSCL work I first introduce some CSCL research[29], before proposing a model for epistemic-commitments based on this research.

10.2.1 Salient CSCL Literature

Building on earlier work (Suthers & Hundhausen, 2003; Suthers, 2006), Suthers (2008) reports on three possible influences of representations on collaborative processes:

1. Negotiation potentials – because the representation is shared, participants feel obliged to negotiate over changes to it
2. Referential Resource – because the representation has shared history, it becomes imbued with meanings
3. Mutual Awareness – Because the representation is external, it is a shared resource which creates a shared frame for activity

[29] For a review of CSCL literature related particularly to computer supported argumentation see Schneuer, Loll, Pinkwart and McLaren (2010), and for a discussion of the sorts of constraints and benefits of conceptually explicit notations see Suthers (2008).

Recall, in section 5.3.1 I briefly discussed the notion of 'awareness' in the CIS context, with four key facets: awareness of other's activities; awareness of shared products; awareness of shared goals; awareness of shared historic activities. The sorts of representations of interest in CSCL could – with some constraints – fulfil such roles, particularly where they include facilities for IM or asynchronous chat.

In a review of the literature on awareness in CSCL, Janssen and Bodemer (2013) note the importance of both content and social (or relational) group awareness – with the former relating to aspects such as awareness of knowledge states while the latter relates to the quality of collaborative processes. This is of particular interest given the concern raised that too much CSCL research has focussed on information sharing, at the cost of analysing the interactional conditions for learning, despite the fact that informational sharing is not a good predictor of collaborative performance (Suthers, Medina, Vatrapu, & Dwyer, 2007; Suthers, Vatrapu, Medina, & Dwyer, 2007, 2007; Suthers, Vatrapu, Medina, Joseph, & Dwyer, 2007). This is of course particularly interesting given that this is a concern that could be raised of CIS research too, with its lack of analysis of dialogue as a form of interaction. In those studies, Suthers et al., note that despite one group outperforming another on knowledge construction involving the integration of multiple sources (of obvious interest, see section 6.1.1), those groups appeared not to share any more information (as indicated by individual referencing in an essay) and that their performance was best associated with 'interaction', as characterised by 'round trips' of information. These 'round trips' describe the reuse of information previously stated, the building of ideas between collaborators – perhaps the *interthinking* of collaborators on shared artefacts. Given the characterisation of 'exploratory dialogue', and associated talk given in section 8.4, this characterisation of interaction, and systems which can monitor it is of interest to analysis of epistemic commitments in CIS contexts. We can for example imagine a 'round trip' in which: 1) a collaborator (A) expresses an idea, saves a page, or enters a query; 2) the object becomes available to their partner (B) and (B) sees it; 3) (B) expresses some related idea – runs a similar query, opens a result from (A)'s results page, uses terminology from (A)'s snippet; 4) (B)'s response becomes available to (A), and (A) sees it.

Such interactions – and a richer analysis similar to the 'Navigation Flow Maps' discussed in section 8.1 – can be inferred via analysis of log data created 'contingency graphs' (e.g. (Medina & Suthers, 2009). In these graphs, contingencies represent the ways participants' actions are built up from (contingent upon) prior actions and artefacts (including dialogue). Such an approach recognises that acts do not occur in isolation, but are built up – and should be analysed as such. Vertices can thus be described as events, including creation of objects, manipulations of them (e.g. edits), and perception of them (e.g. opening messages), while arcs (directed edges) are contingencies in which one event enables a subsequent event. Contingencies are easy to detect in some cases such as where they use the same shared object, but in other cases are harder to detect (for example, where an idea is semantically related to an earlier one but not explicitly linked to it).

10.2.2 The Temporal Aspect

This raises the interesting issue of 'temporality' in learning design. As Mercer notes[30], temporality is an understudied facet of learning dialogue, yet "as learning is a process that happens over time, and

[30] Further discussion of Mercer and colleagues' perspective on temporality in language for learning will be given in the introduction to the first component of preliminary work.

learning is mediated through dialogue, we need to study dialogue over time to understand how learning happens and why certain learning outcomes result" (Mercer, 2008, p. 35). Building on Mercer's work, similar claims have been made in CSCL, noting "Independently of the context of the learning—on the level of the individual, the group, the situation, or in the interaction of these—the main object of analysis in CSCL is a process, something that unfolds over time" (Reimann, 2009, p. 239). Thus the claim is particularly that we should move from an experimentalist focus on generalizable variance in constructs – in which variance in tightly defined independent variables can be mapped to variance in tightly defined dependent variables – to refocus on processes, in which variables are in mediation with each other (thus rejecting the input-output notion of causality of experimentalists).

This notion has also be raised by Suthers and Medina (2010) , who note that under various models of learning – particularly citing Gerry Stahl's notion of 'group cognition' (Stahl, 2010) – understanding group learning as geo-spatially and temporally distributed is crucial to understanding the process as beyond the inner working of one individual's mind. Interestingly, they also note that situated cognition – in its rejection of cognitive models with their underlying mental representations and notions of mutual belief – can have a tendency to over emphasise aspects of the physical environment (including cultural artifacts) and local-social interactions, leading to a bias to visible elements of cognition and an "uncritical acceptance of reliance on short transcripts" (Suthers & Medina, 2010, p. 2). They thus suggest a method of microanalysis which addends a global analysis to local features, by:

1. Identifying episodes such as segments of chat in which a problem is solved
2. Looking to prior episodes to explain (1) by:
    a) Engaging a 'global level of analysis' to look at prior episodes for similar functions, solutions, inscriptions, etc.
    b) Identifying the point at which the development of episodes in (2.a) is completed - at this point the episodes are "available" for future use
3. Repeating 1-2 of this process, to look for chains of contingencies (actions dependent on prior actions to occur) and uptake (notions which are functionally utilised in subsequent episodes)
4. Finally a *local level of analysis* works forward within episodes to look at meaning making in artifacts and chat at that level, including the use of globally identified constructs (or common knowledge)

These features of language use, particularly in CSCL and computer environments generally, are important. This section has introduced the significance of time in analysis of dialogue – in the following worked example, temporality will be highlighted with illustrative examples, and a proposal for an analytic technique for my own work.

10.2.3 A Worked Example for a CSCL Epistemic-Commitments Platform

One example of how knowledge-building CSCL tools (Scardamalia & Bereiter, 2006) are relevant to this thesis (as presented in Knight et al., 2013a, 2013b), is built around the Cohere argument mapping tool (Buckingham Shum, 2008) and previous work on sociocultural discourse-centric LA (De Liddo et al., 2011). Cohere is a web application for mapping ideas, concepts and arguments, which can be annotated directly onto source websites. Users enter ideas – nodes with meaningful classifications – and are then invited to "make the connection" with meaningfully labelled edges, to create a conceptual graph. Both ideas and connections may also be tagged, to add a further level of

semantic data. Cohere is designed as a tool to enable users to build their own structures, but also to share these, and integrate the nodes and connections of other users, thus building up communities of enquiry around particular disciplinary topics. More advanced use of Cohere is to directly annotate web materials, such that a node representing an idea, concept or claim (e.g. "Many pupils with high exam grades fail at university") may have multiple websites associated with it (news stories; journal papers; etc). This facility enables it to be used a qualitative analysis tool to analyse a literature (Jelfs, Buckingham Shum, & De L, 2011) or online student discourse mediated via other platforms (De Liddo & Alevizou, 2010).

Cohere facilitates analysis of the ways that users create and build on ideas, and the epistemic implications of such creation. At a basic level, this could simply be an analysis of the number of idea nodes, and connection types, used. A more advanced analysis might compare individuals' Cohere use on the same task, and provide analytics based on such comparison; these notions are discussed further below. However, neither of these explores the semantic qualities of ideas and connections. Using the broad epistemic 'dimensions' described above (Table 1) some correspondences between those descriptors, and possible trace can be identified as in

Table 12[31] which also gives 'suggested guidance', intended to be indicative of the sorts of challenges which might be posed to students to extend their epistemic cognition and probe their learning processes.

However, within the approach described above it should be understood that while the trace data given here is theoretically tied to the constructs, both the constructs and the trace should be seen in their situated context – as components of a sociocultural environment, interacting with the relevant agents (students, teachers, designers, etc.), and the wider cultures and subcultures. Thus, the possible trace markers and guidance are conceptually related to the work discussed above but these should be dynamic tools, and empirical work will be needed to explore the relationship between feedback given, representations allowed, student responses to feedback and the impact of this on learning.

Table 12 thus proposes one set of traces from which meaningful data could be captured, and from which simple computations could be implemented to feed back the number of ideas, and connection types used for self-reflection and to provoke meaningful dialogue regarding what these other types

[31] Following previous work (De Liddo, Buckingham Shum, Quinto, Bachler, & Cannavacciuolo, 2011) the basic analytic statistic is constructed as a percentage representation of the target type, over the total types created by the user. For example, the number of 'opinion' nodes created, as a percentage of the total number of nodes created by that user.

might be used for, or why they have not thus far been used. Similarly, constructive discourse might occur around the reasons why one student's map is more connected (but perhaps not appropriately so) than another's.

Table 12: Trace & Guidance for Epistemic Beliefs

	Trace	Guidance/Challenge
Certainty	Presence of competing claims (e.g. supports/ challenges). Presence of stability markers – e.g. current references, geographic repetition.	Are there two sides to this idea? Could you explore XY contrasting example? Is this idea consistent across time/place? Have you looked at XY map?
Simplicity	Number of connections between nodes.	Are any of these ideas connected? Have you considered how WX and YZ might be connected?
Source	Presence of 'I think' or restatement of fact, few additional nodes made other than those created as quotations.	What do you think of these ideas? or How does the evidence relate to your view?
Justification	Judgments of relevance, and supporting or explanatory notes ('this evidences/ explains x'). Ties to method 'ideas'.	What evidence do we have for this idea? Is it 'good' evidence? Why/why not?

Indeed, further metrics can be imagined as in the preliminary example given in Table 13 which indicates how Tsai's framework for information commitments might be operationalized within a CSCL environment.

Table 13 - Information Commitments in the Cohere Platform

Information Commitments	***Possible Orientations & Indicators***	
Standards for Correctness	**Multiple Sources** Claim nodes are associated with multiple sources. Multiple websites are visited.	**Authority** Use of single source to support claims, unsubstantiated claims or evidence used in place of claim nodes. Repeatedly visiting the same few websites
Standards for Usefulness	**Content** Claims involve integration of evidence from multiple sources.	**Functional** Use of single sources as direct reply to single issue-level nodes

Search Strategy	Elaboration & Exploration Builds 'web' of nodes including using other people's nodes. If tracking queries – modifies query terms, builds on those of collaborators, opens more than one link and goes beyond first SERP.	Match Builds linear nodes, mostly based on own nodes (rather than collaborators). If tracking queries – tends to only conduct individual searches.

Table built from Tsai's framework (C.-C. Tsai, 2004, p. 109)

10.2.4 Pitfalls of Collaborative Tools and Tasks[32]

It should be noted that, despite CIS and CSCL research indicating the positive potential of awareness and structuring support, some research indicates potential pitfalls with such systems. The ways groups are constructed is likely to impact on the nature and success of the collaboration, including by impacting on the discourse used. To give a concrete example, where students work together in front of a shared computer, the results of their inputs and outputs on the screen become common ground and therefore, perhaps arguably, necessarily implicit. This may reduce rather than enhance the need and possibility of individual's articulating knowledge explicitly through talk (Clark & Brennan, 1991; Pickering & Garrod, 2004). Indeed, the evidence discussed in section 5 provides some discussion of such issues in the context of CIS.

In CSCL environments though, it should be noted that

> *In spite of these positive effects of CSCL, many studies have also identified possible pitfalls when using CSCL (Kreijns, Kirschner, & Jochems, 2003). Examples of these problems are escalating conflicts among group members (e.g., Hobman, Bordia, Irmer, & Chang, 2002); free riding behavior and unequal participation (e.g., Lipponen, Rahikainen, Lallimo, & Hakkarainen, 2003; Savicki, Kelley, & Ammon, 2002); and discussions that lack depth, high-quality reasoning, and argumentation (Munneke, Andriessen, Kanselaar, & Kirschner, 2007, p. e.g.,). Although these pitfalls are not unique to CSCL (they also occur during face-to-face collaboration), some problems that learners may encounter in CSCL environments seem to be enhanced in these environments, for example, due a lack of social presence or limited nonverbal cues such as gestures and facial expressions (Daft & Lengel, 1986; Kreijns et al., 2003; Short, Williams, & Christie, 1976)* (Janssen & Bodemer, 2013, p. 40).

Any design should thus a) bear in mind the prior work in this area, and b) explore their implications for CIS and epistemic commitments. It is also important that we bear in mind the type of dialogue we are seeking, and the role of both tools and tasks in making explicit that dialogue both for the collaborators, and the analysts. Furthermore, structured environments are unlikely to provide the flexibility necessary to facilitate dynamic CIS processes, nor permit the analyst access to the collaborative dialogue around search and processing decisions made *during* the processes, as opposed to at the later 'structuring' stage of organising retrieved information. It is to these aspects that I now turn.

10.3 Articulating Reasons – Beyond Structured Dialogue

[32] I remain grateful to Charles Crook, with whom I first discussed this aspect of CIS, and who first drew my attention to initial relevant references.

While the analysis of moves in a structured CSCL environment holds benefits (as highlighted in section 10.2.1) and provides a target for some analysis, it is unlikely to provide a sufficient picture of epistemic commitments, nor to fully support students in their CIS activities – and indeed, many CIS tools have provided one or both of IM, or asynchronous messaging tools to facilitate collaboration. Indeed, in one study using common everyday (non-digital) materials (office supplies) in a controlled face-to-face context in which verbal and facial communication was not permitted (through use of a separating divider), while materials were used to fulfil many of the roles highlighted by CSCL and awareness literature (grouping, semantic ties, 'ownership' of ideas, and so on) so too were non-verbal communicative strategies to raise real time awareness (such as tapping an area of the workspace) or regulating the workflow – including through written messages (Dwyer & Suthers, 2006). It is thus important to consider both CSCL tools, and more discursive CMC devices to support activity giving due attention to:

1. How such tools could be integrated into a CSCL environment for CIS
2. What benefits this would hold for collaborators
3. What analytic benefit we would gain from integration of such tools

The potential for combing of representational and discursive tools is interesting given that artefact-centered discourse is very common in both professional meaning-making (Sumner & Buckingham Shum, 1998) and more formal education settings; yet many online environments do not support it (Suthers, 2003). Suthers (2003) argues that such representations can lead to new negotiations of meaning, foregrounding collaborator's decisions around additions or modifications to representations, and encouraging consensus and clarity around the meaning of representations – the most effective discussions around which are likely to be 'exploratory' in nature (R. Ferguson, Whitelock, & Littleton, 2010), and discussion anchored on which leads to learning gains (Eryilmaz, Pol, Ryan, Clark, & Mary, 2013). In designing a system to support such 'attached dialogue' (R. Ferguson et al., 2010) or 'artifact-centered discourse' (Suthers & Xu, 2002) a set of desiderata were identified for online discussion environments:

1. They should allow inclusion of visual artifacts such as graphs, videos, knowledge maps, etc.
2. These artifacts should not just be 'attachments' but should exist outside of the messages, and remain visible during the conversation.
3. Artifacts should be changeable in natural ways
4. Participants should be able to refer to individual artifacts (preferably, parts of artifacts) ion their contributions. They note that this preference for parts of artifacts is important given artifacts may be composed of many parts (e.g. paragraphs) potentially contributed by many actors, at many stages, on various topics. (Suthers & Xu, 2002)

These desiderata appear to have been embedded in the design of 'nStudy' (Beaudoin & Winne, 2009) – a tool to support learners in individual and collaborative online research, which allows learners to read web content and link selected content to new or previously created learning objects. Those learning objects made up of:

- Bookmarks – for URLs
- Notes – about learning objects, webpages, etc.
- Forms – which provide a structured space for re-useable notes and extraction of metadata
- Tags – specifically learning related tags, which are not dissimilar to the default node types in Cohere (including 'pro/con' style tags, descriptive tags such as 'method', 'evidence', etc.)

- Documents – including webpages and user created documents such as essays and reports
- Chats – IM style chat among collaborators
- Terms – terms define key concepts (and an ontology), when a term is used it is automatically linked to all other uses of the term
- A term net – which provides a network of term use
- The library - displays metadata tables for learning objects according to user's preference
- Maps - display relationships among information objects (nodes in the map)

(Beaudoin & Winne, 2009)

The features described above thus align well with the desiderata in CIS environments, and other CSCL research. Crucially, awareness of salient factors of the task is key – chat, domain features (such as 'terms'), and shared artefacts. In related work, Wegerif (2010) argues that, when teaching thinking with technology we should think about:

1. Opening dialogic spaces (e.g. by adding comments to blogs), but also teaching to do this (e.g. through the use of ground rules for talk, and philosophy for children)
2. Widening dialogic spaces – understanding more points of view, and the background behind them, for example through WebQuest activities in which different perspectives – and their assumptions – are explored
3. Deepening dialogic spaces – increasing reflection on assumptions made in arguments by students and others, shared awareness tools to make explicit the arguments being made (and their structures) can support such deepening.
4. Teaching content through induction into fields of dialogue – Wegerif notes "interactivity makes it easy for software to simulate multiple points of view in a dialogue, thus allowing leaners to be inducted into a field of dialogue rather than into fixed 'truths'" (Wegerif, 2010, p. 350) noting that, the internet can be a cacophony of voices, rather than a dialogue, but through designed spaces – such as WebQuests, and the emailing of links between geographically distant groups – presence and dialogue can be mediated to encourage reflection and learning.

This final point is fundamental to the interest in epistemic commitments, search tasks, and collaboration – collaboration, and well designed search tasks can provide a means through which to expose students to a "field of dialogue", and to explore their shared commitments around such multiple perspectives. This is not only a claim about collaborative dialogue, but one about the very nature – the unstructured, messy nature – of the internet, and its use for developing space to explore multiple viewpoints, and – using learning analytic techniques – making claims about students' commitments in such activities. Crucially, as noted in the preliminary sections of this work, if our targets are higher level reflection and conceptual understanding, such space must be created in contrast to many current educational systems, reiterating the point that collaborative task context is as important as collaborative tool design (Rick & Guzdial, 2006). These considerations will form a core component of this work, as will now be discussed in reporting the first year's practical work (section 11), and research proposal (section 12).

11. Report on Preliminary Practical Work

The simplistic addition of communication or collaboration functionality to knowledge management tools will not guarantee higher level epistemic commitments. Rather, these features – such as instant messaging in search interfaces, the highlighting of specialist (or, just different) language

through foregrounding search terms, or the inclusion of saved search results – open up the possibility of such activity: it foregrounds those possibilities, and allows their exploration through the use of appropriate tasks. By analogy, just as Mercer et al's interest should not properly be thought of as in the use of 'because', and explanations, but rather in a deeper notion of interthinking (see Littleton & Mercer 2013), so my interest is not in improved search, or better saved results – but in a deeper notion of evaluative epistemic commitments (which may involve interthinking).

This can be thought of in two respects: firstly, interthinking (as discussed in section 8.4) is an epistemic mode of activity; secondly, interthinking opens up a dialogic space around which further epistemic commitments may be made explicit – through tasks which encourage reflection on many points of view (point 2 above), the shared use of a tool (point 3 above), and in particular the shared use of such tools and dialogue around information found (point 4 above) a prerequisite for finding and working with documents on the web.

Above I noted Wegerif's claims for technological support in dialogic education, around the opening, widening, and deepening of dialogic spaces, and teaching of content via induction into fields of dialogue. Information seeking tasks are a prime example of such considerations in play:

1. Dialogic spaces may be opened both through exposure to (as in, the presence of) multiple perspectives in search results (and indeed, in the entering of search queries), and through the sharing and supported discussion of such searches and results where 'commenting' on these is enabled
2. Dialogic spaces may be widened through facilitating students through exposure to (as in, the exploration of) multiple perspectives in search results (and queries), and through the use of tools to support structuring multiple results in relation to each other and claims – requiring an understanding of perspectives, highlighting the possibility of alternative viewpoints, etc.
3. Dialogic spaces may be deepened by the shared nature of such search and processing, sharing both processes and claims around the structure of arguments and relationships between results.
4. Finally, the internet and collaborative information seeking both provide means through which students are exposed to fields of dialogue, both with respect to listening to collaborator understandings, and via the use of search tools to support understanding (from provision of dictionary definitions, to facilities to search for particular vocabulary, to query suggestions, and so on). As Wegerif notes, the emailing of links can create space to encourage reflection and learning; as noted in section 5, there are other forms and methods of collaborative information seeking, and the claim of this work is that these too create an environment conducive to dialogic education.

The preliminary work discussed here explored developments so far in creating a classifier for exploratory dialogue, here at the OU – the Exploratory Discourse Detection Module EDDM. The preliminary work provides a review of relevant literature in machine learning techniques particularly as deployed on educational dialogue, making suggestions for how the EDDM – or Exploratory Discourse Analytics Module (EDAM) – may be further developed to be appropriately deployed on my own data.

11.1 Overview of work to date

In this section I briefly introduce some core activities from the year, and how they relate to my research and development agenda. This section is a version of a post on my blog (as are some other elements of this report) and this is characteristic of my year's activities – having often drafted thoughts on the blog, where they are searchable and organised into posts. I intend to continue such activity throughout the PhD, although my main additional plan for next year is to move the blog to my own domain name for longevity's sake. In relation to my primary work, this section is organised as follows: I first discuss some conferences and meetings I have engaged in over the year; I then discuss some specific projects and work-products.

11.1.1 Conferences and Meetings

I've been lucky enough to attend a few conferences this year, with a couple of other events lined up for the future; I briefly discuss these here, mentioning some key outcomes and relationships to my research.

11.1.1.1 Wikimedia visits (including San Francisco and Lincoln AGM)

Over the course of the year I have engaged in editing Wikipedia (including the article on Learning Analytics), and thinking about how the platform – mediawiki – might be used for learning analytics[33]. Wikipedia is particularly interesting to me given that it is perhaps the best known collective intelligence tool. It also has a number of features (including an article feedback tool, and measures of editor reputation) that might offer insight into epistemic commitments[34], and may also support CIS[35] and collaborative information sharing (see below).

More broadly I am also interested in the potential of collective intelligence tools such as Wikipedia for developing learning environments based around OER (and talked about this at the WMUK AGM/Conference in Lincoln earlier in the year (S. Knight, 2013b), and at the WikiMedia Foundation office in San Francisco[36]), and have actively encouraged people to edit relevant Wikipedia articles (including the Learning Analytics article[37], and Massive Open Online Course article[38]). Prior to this year I had only ever made one minor correction to a Wikipedia article; my work on ORBIT involved using the platform which helped develop the skills to edit (and think about mediawiki for OER), but working on writing *for* Wikipedia has been a useful challenge as has thinking about its potential for learning analytics – something I hope to continue to work on. In particular I have one project (discussed below) on going, which may inform further work around tracking epistemic commitments in CIS on mediawiki.

11.1.1.2 CSCW13 (both workshops) - San Antonio, Texas

In terms of collaborative tools, I have also familiarised myself with some literature in the computer supported cooperative work (CSCW) area – the conference for which includes a large number of Wiki based research. In February I attended the 16th CSCW in San Antonio, Texas[39] with two

[33] See http://people.kmi.open.ac.uk/knight/2013/01/mediawiki-for-learning-analytics/ and http://people.kmi.open.ac.uk/knight/2013/07/badging-wikipedia-contributions/

[34] See http://people.kmi.open.ac.uk/knight/2013/01/wikipedia-feedback-ratings-as-an-epistemic-tool/

[35] See http://people.kmi.open.ac.uk/knight/2013/05/collaborative-information-seeking-on-wikipedia-talk-pages/

[36] See http://people.kmi.open.ac.uk/knight/2013/07/contributions-to-the-sum-of-knowledge-wikimedia-foundation-meeting/

[37] See http://people.kmi.open.ac.uk/knight/2013/01/wikipedia-learning-analytics-editathon/

[38] See http://people.kmi.open.ac.uk/knight/2013/05/an-invititation-to-the-massive-online-open-course-mooc-wikipedia-page/

[39] See http://people.kmi.open.ac.uk/knight/2013/03/cscw2013-2-workshop-papers-texan-fun/

workshop papers. One of those was on the use of a toolset (including Cohere) to support collaborative sensemaking in collaborative information seeking environments; in an updated form this is included as part of my first year report. In addition, that workshop has led to a collaboration with two other attendees and co-authoring a paper on the nature of 'context' in CIS. The other workshop - on the relationship between CSCW and Education - was also a useful networking event, from which I have maintained contact (and met up with again) at least one other attendee.

11.1.1.3 Learning Analytics and Knowledge – Papers and Online Course

Almost immediately upon starting the PhD we (my supervisors and I) set about writing a submission for the 3rd Learning Analytics and Knowledge conference[40] held in Leuven, Belgium; that paper was nominated for 'Best Paper' award, and will be revised and updated for submission in the first issue of the Journal of Learning Analytics (as well as forming a significant chunk of the earlier parts of my first year report). Much of the work on this paper also informed my and Simon Buckingham Shum's work on the 6th week of the Learning Analytics Open Course this year[41] – which was on epistemology and learning analytics, and included talks from me, George Siemens (on connectivism), and David Williamson Shaffer (on epistemic games – see earlier sections, and below).

In addition to our LAK13 conference submission, Karen Littleton and I wrote a paper for the Discourse Centric Learning Analytics workshop on the importance of context for educational discourse, and some challenges to DCLA of context. This paper informed my subsequent analysis of DCLA techniques and a paper (in draft, included as a Work in Progress in the report below) on the multiple levels of context in the analysis of exploratory dialogue, some challenges for machine learning techniques, and a proposed method.

11.1.1.4 LASI - Palo Alto, California

In early July, I was at Stanford University at the Learning Analytics Summer Institute. The week was very useful in terms of concentrated time to think about learning analytics and network. In particular: there are a number of hopeful collaborations (e.g. a group of researchers looking at information seeking/knowledge management); I had useful conversations with someone at Google who is intending to build a tool to explore student's sourcing from multiple documents (which I may be able to use); I had informative talks with Carolyn Rosé about machine learning for processing educational dialogue.

11.1.1.5 Society of the Query Conference - Amsterdam

One of my continued interests is in how we conceptualise knowledge, particularly in the context of tools such as Google, and Wikipedia. Paul Matthews wrote a piece on this in the context of social epistemology (and a) was at CSCW, and b) I hope to write something with in the not too distant future), and I also have great hopes for the Extended Knowledge Project[42] (which I hope to be able to contribute to). In this area, I've been invited to contribute to the 2nd Society of the Query Conference in Amsterdam this November on the subject of Education and the role of context (see below), and submit a piece to their reader. The particular panel description is 'Search in Context':

[40] See http://people.kmi.open.ac.uk/knight/2013/04/lak13/

[41] See http://people.kmi.open.ac.uk/knight/2013/03/lak13-mooc-week-on-epistemology-assessment-and-pedagogy/

[42] See http://people.kmi.open.ac.uk/knight/2013/03/the-extended-knowledge-project/

There is a long-term cultural shift in trust happening, away from the library, the book store, even the school towards Google's algorithms. What does that mean? How are search engines used in today's classrooms and do teachers have enough critical understanding of what it means to hand over authority? We think we find more and in a faster way, while we might actually find less or useless information. The way we search is related to the way we see the world – how do we learn to operate in this context?

11.1.2 Specific projects

Over the year I have also been working on specific projects, particularly developing skills and practical work around the literature I've been reading and writing about.

11.1.2.1 Developing a CIS environment for Epistemic Commitments

A core deliverable for my work is the development of an environment on which to conduct my research. Over the year I've been exploring a range of internal (Cohere and Evidence Hub particularly) and external (mediawiki, and the existing CIS environments in particular) tools which could be used for my research.

As a part of this I learnt how to use WAMP server to mirror an existing Wiki (the Schome project at the OU) and download and install extensions to that Wiki. I have also explored the use of Google Analytics for tracking user behaviours on websites (not appropriate due to constraints on identifiable information), and the potential to use external feeds (RSS) to seed another environment (Cohere). From these explorations and my reading I designed a specification for a tool in collaboration with my supervisors and Michelle Bachler (a developer in KMi) which will be a Firefox addon for the EvidenceHub tool, developed by Michelle (as was discussed above).

11.1.2.2 EDAM

Another core piece of my work is around educationally productive dialogue. In particular, given my interest in epistemic commitments a core part of my project is to identify when commitments are 'accountable' in the group (i.e. are within the scope of exploratory/accountable dialogue). To some extent keyword spotting may be enough here particularly within the constrained environment of the EvidenceHub (and indeed, that is the finding of the Epistemic Games group). However, given existing work in the department to further develop from bag of words approaches, and my own paper with Karen Littleton discussing the role of context in exploratory dialogue, it was of interest to explore how machine learning techniques might be used for such classification.

Therefore, over the course of the year I have learnt to deploy the existing Exploratory Discourse Detection Module (see section on Maturing EDAM in the first year report) which is built on the MALLET command line tool. My checking of outputs from this tool has been conducted in Excel (the tool essentially produces a .csv format), and I would not anticipate delving further into work with MALLET. I have, however, further explored the use of GATE, and to some extent WEKA (and I am pleased to have met one of its founders - Ian Witten - at LASI). I have also had useful discussions with Elijah Mayfield who developed the LightSIDE tool, and has been kind enough to share the code he used to detect 'authoritative talk' using an 'Integer Linear Programming' approach. I would hope over the course of the PhD to be able to utilise GUI tools such as LightSIDE and GATE in appropriate contexts, while also working with machine learning specialists to develop custom tools. To that end I have had helpful technical conversations (for which I am very grateful) with Carolyn Rosé and Elijah Mayfield, Zhongyu Wei (who conducted much of the original work on the EDDM tool), and Yulan He (who also worked on the EDDM tool, and who we hope to continue to work with). The joint paper

with Karen Littleton on Maturing EDAM includes a technical proposal for continued work which we hope provides a specification for the next generation tool. This paper is provided as a separate document to this report.

11.1.2.3 So.cl

In 2012 Microsoft released a dataset from the social search tool 'so.cl' to researchers. So.cl is an experimental social network in which when one searches, a post is created based around that search, to which interesting results from the search may be pinned. It is multimedia intensive, and visually quite attractive. The original intention was that the tool be used particularly in universities, although that focus appears to have shifted. However, one interesting new development along more learning-oriented purposes may be in the use of so.cl TEDActive in which: "conference-goers can assemble images, research links, videos, and text into collages that express their reactions and associations around the TED Talks."[43]

In addition to the literature review which offers a justification for the interest in dialogue around CIS, some public blog posts around interesting so.cl discussions (e.g. on whether Aliens built the pyramids[44]) indicated it might potentially hold some interesting exploratory dialogue. Thus, in order to attempt to investigate the exploratory properties of dialogue around CIS the dataset was requested, ethical clearance granted, and the dataset opened in R. R was used to: 1) subset rows to language based rows based on their action ID, 2) classify these using a language detection package,3) subset to English language only results, 4) export these into .csv for processing using EDDM (see above), 5) reimport and reintegrate the data into the whole dataset. No results are reported here because: 1) as this work was being conducted, I was also addressing the suitability of EDDM for such purposes as described in the draft paper, and it was decided that it was likely to be unsuitable, 2) few results were returned from the whole dataset as being likely to be 'exploratory' in nature, 3) the method of deploying R for such purposes is a poor use of R's power; a more suitable approach would start with data in a database from which R processes could be run, and within which more complex – database based – processes could be conducted. Such an analysis may be conducted at a later date.

11.1.2.4 CIS on Wikipedia (R)

Wikipedia Talk pages are a place in which editors can make sense of, and share, information - I was hypothesised that we could see these two distinct types of behaviour in link patterns; first moving from articles to talk pages (sensemaking), second from talk pages to articles[45]. With some generous help from Aaron Halfaker (who scraped the edit histories of Wikipedia for each edit on each page, storing every link added or removed), I set about trying to use R to process Wikipedia LinkFlow data[46] – this is another case using a large dataset where the data would be more appropriately stored in MySQL than R.

While at Stanford hosted by the Lytics Lab René Kizilcec and I worked further on this data to – without first loading it into R – work it into a readable form for uploading to MySQL, and then

[43] http://www.bing.com/blogs/site_blogs/b/search/archive/2013/02/26/bing-microsoft-research-bring-the-power-of-social-search-and-whimsy-to-tedactive-with-ted-so-cl.aspx

[44] http://blog.fuselabs.org/post/29422205211/some-argue-that-aliens-built-the-pyramids-a-heated

[45] See http://people.kmi.open.ac.uk/knight/2013/05/collaborative-information-seeking-on-wikipedia-talk-pages/

[46] See http://people.kmi.open.ac.uk/knight/2013/06/using-r-to-process-wikipedia-link-flow-data/

reshape it (or, just count) as per the discussion in that blog post in which I discuss looking for strings of ATDR (Inserted on Article; Inserted on Talk; Deleted from Article; Removed from Talk [essentially the same thing but we need to distinguish the two]). By counting such strings we can get an overview of the number of times any particular link has moved from A to T or vice versa (as well as the other doubles) and we could insert in S/N - same user, not same user – on each double to explore that aspect too. One could further extend this work by indexing each link such that if the same link appears on multiple pages it has the same ID – that would allow us to start exploring SNA potential too. Once this work is finished I am hoping to write it up in a 'note' style paper for conference submission.

11.1.2.5 ENA

One of the outcomes of networking at (and before) LAK13 was an invite to the University of Wisconsin-Madison Epistemic Games group as a visiting scholar. One of the hopes with the so.cl dataset (above) was that it might provide some interesting coded data on which their Epistemic Network Analysis tool could be used in order that I could learn how to use this method (described briefly below). However, due to concerns with the number of appropriate length posts and interactions from that dataset, I instead recoded a familiar dataset (my MPhil data) and conducted analysis on that.

ENA is based on the theory of Epistemic Frames which posits that the important component of 'knowledge' is not facts and skills in isolation, but understanding how those are connected. For example, in the case of information seeking, seeing that users' seek 'authority' is, in isolation, not particularly informative (because standards of 'authority' for knowledge may be inappropriate or appropriate depending on other contextual factors). However, understanding that a searcher's 'authority seeking' talk is connected to other talk related to community practices (perhaps around who we assume authorities to be, such as scientists – a 'value' of that community), or seeing that searchers engage in what Shaffer would call 'epistemic' talk and what I have called accountable or exploratory talk (to justify their selection of authorities) is interesting. So, we see in this example a case where simply exploring one component in isolation provides relatively little information, while looking for combinations offers more insight. For example, we could explore not only a reliance on authority/corroboration in sourcing, but also instances in which they are more/less likely to be connected to particular types of justification (attempts to understand the material v. simple matching information to plug answers in).

The theory of epistemic frames that ENA is based on is founded on an assumption that epistemic communities can be characterised by their: skills; knowledge; identity; values; and epistemology – their epistemic frame. While this theory has a very specific theoretical and operationalized notion of what these concepts mean, we can briefly easily imagine how – at face value – the concepts can be related to CIS activities. Specifically, we can conceptualise: Skills as those related to searching; knowledge as utterances related to the question at hand; identity as related to the group-work terms used (e.g. collective versus individualistic terms); values as instructions or imperatives ('we have to'); and epistemology as related to matching or understanding as described in the epistemic commitments section above. Indeed, in this case keyword identification – as opposed to a deeper analysis of exploratory dialogue – may be precise enough to garner interesting results. Furthermore we can imagine hypothetical work leveraging semantic web technologies (which KMi has expertise in) that would identify where key content terms were being used by making use of our prior

knowledge of that content. For example, the for a given question the 'knowledge' category could be populated by taking either the query or – in classroom tasks – the set question, and using a source such as DBPedia to gather relevant content key terms related to that query which could be identified in dialogue episodes among CIS teams. While this description is a hypothetical long term possibility, preliminary work has been conducted on my MPhil data to explore the informativeness of ENA for a known data-set. That work has required me to learn how to use the ENA tool, and again I have a paper in draft which I hope to finish by mid-November.

12. Research Proposal

12.1 Introduction

Seeking information in online environments is an increasingly important activity in a world in which students are no longer directed to pre-selected course books and materials. Yet, searching is a skill with which many – across age ranges – struggle. While support for technical aspects of searching may be of some assistance, we agree with Mason, Ariasi and Boldrin's claim that complaints regarding students' abilities at navigating the web are not technological, but rather epistemic issues around "the nature of knowledge and knowing, which may facilitate or constrain searching and evaluating sources of information on the internet" (Mason et al., 2011, p. 139). This work takes as its focus the seeking of information, claiming that this activity – particularly as mediated by and conducted through search engines – provides an *epistemic lens* through which researchers may explore the commitments learners make explicitly and implicitly about knowledge. These commitments are implicated in the ways students *select* sources, *use* them, and *make connections* between them in any information seeking task. Analysis to explore issues at this level, and understand how to support learners to engage more effectively in their search practices is important. Both professional and academic learning contexts require high levels of information literacy; as such, an activity-oriented perspective on developing skills to support such literacy is an important contribution. This work builds on recent contributions calling for a shift from psychometric assessments for epistemic cognition, towards an exploration of the situated contexts in which epistemic practices are brought to bear. Specifically, we argue for a new approach: epistemic *commitments* – action-oriented ways of working – rather than *beliefs*, and the analysis of such commitments through exploration of connections between epistemic modes of information seeking. We suggest a focus on trace indicators of behaviors, and the connections between particular types of behavior offers a productive new approach to the investigation of epistemic practices.

12.1.1 Seeking Information as an Epistemic Lens

In describing the established epistemic cognition literature, Mason (2009, p. 69) highlights broad agreement across models on the importance of two main facets – *what knowledge is, and how one comes to know*. Within the first area, two dimensions are noted: the *certainty* of knowledge (how stable or tentative knowledge is); and the *simplicity* of knowledge (how holistic a perspective of interrelated concepts, or simplistic a perspective of compartmentalized facts is taken). Similarly Mason identifies two dimensions of the second area: the *source* of knowledge (from transmission to constructivism); and the *justification* for knowing (what warrants a knowledge claim – from authority

to rules of inquiry). These models have informed analysis of the comprehension of multiple online sources – which may vary radically in the nature of their sourcing and justifications – in the understanding that students who regard knowledge as simple and finite may conduct brief and perfunctory searches with little recourse to integration or multiple sourcing (Barzilai & Zohar, 2009; Bråten & Strømsø, 2006). We thus agree that, "exploring students' thought processes during online searching allows examination of personal epistemology not as a decontextualized set of beliefs, but as an activated, situated aspect of cognition that influences the knowledge construction process" (B. K. Hofer, 2004a, p. 43).

Research in this area indicates that students with more sophisticated inquiry stances are more likely to evaluate websites, and to do so meaningfully, while those with more sophisticated perspectives on the multiplicity of knowledge (or multiple perspectives) are more likely to integrate and critically evaluate multiple online sources (Barzilai & Zohar, 2009; Bråten & Strømsø, 2006). While epistemic cognition is not a significant factor in understanding converging perspectives in online sources, in conflicting sources those with evaluativist beliefs (who critique claims) perform significantly better in their comprehension (Barzilai & Eshet-Alkalai, 2013). A growing body of work associates search and sourcing patterns with particular patterns of epistemic metacognition (Mason et al., 2009), with think-aloud research indicating that students engaged in web-based learning spontaneously engage in epistemic reflection, particularly around source selection and credibility (Mason et al., 2011, 2010), where students who verbalised about source credibility and information veracity significantly outperformed those who evaluated only sources (Mason et al., 2011). It should be noted, however, that the use of think-aloud protocols may increase such practices (Schraw & Impara, 2000; Schraw, 2000). We return later to the possibility that the *collaborative* search context may have higher external validity, and methodologically as a means of providing insight into a group's epistemic practices.

12.1.2 Situating Epistemic Commitments

The context of search is thus an interesting one for our investigations. Recent work has rejected an analysis of *beliefs* in favor of an action-oriented view::

> *What we have called tacit epistemic beliefs might better be called epistemic commitments (C. A. Chinn & Brewer, 1993). Some theorists may be uncomfortable with the idea that one can have a tacit 'belief' that cannot be expressed, and the term epistemic commitment avoids reference to such beliefs. An epistemic commitment reflects a tendency to act in specified ways, such as a proclivity to provide justifications based on personal experience (C. Chinn A. et al., 2011, p. 146).*

Furthermore, Sandoval (2012) has made related claims, calling for epistemic cognition researchers to take seriously a 'situated' approach:

> *One important way to understand the epistemic ideas that people bring to bear is to examine their participation in practices of knowledge evaluation and construction. Changes in the form of participation are indicators of changes in the meaning that individuals make of the activity in which they are engaged. […] Change in participation can indicate a shift in epistemic perspective, but it is the shift itself that suggests what particular epistemic ideas are brought to bear in the first place (Sandoval, 2012, p. 350)*

In a similar vein, Tsai (2004) suggests that information commitments involve "specific views about what counts as a successful explanation in the field (e.g. science) and […] general views about the character of valid knowledge or information" (C.-C. Tsai, 2004, p. 105). Tsai (2004, p. 109) proposing a three dimensions:

- Standards for correctness: evaluative standards from 'authority' to 'multiple sources'
- Standards for usefulness: assessment of the usefulness of web-materials, from 'functional' (e.g. ease of retrieval), to 'content' (e.g. relevance of retrieved information)
- Searching strategy: information-search strategy from 'match' (of simple claims to questions) to 'elaboration and exploration'

While this turn from epistemic beliefs to commitments is an interesting one for those who wish to analyse users' behavioural traces as proxies for epistemic beliefs, it is still problematic, not least because as Wu and Tsai (2005) highlight, students may utilize both of the information commitments ('multiple sources' and 'authority') at the same time when evaluating the accuracy of the materials on the Web – a scenario which this framework does not have conceptual resource to explain. That is, while the orientations are displayed as dichotomous, or scalar, it is not clear that it is appropriate to think of them in such a way.

In our view, the action-oriented shifts described above are best characterized by the connections learners make between aspects of their sourcing behaviour and information use. Thus, the focus should be on the emergence of information needs, and the use of multiple implicit and explicit criteria to assess the suitability of information for meeting those needs is dictated by a complex combination of searcher's action, task context, and technical mediation. Importantly, "…information seeking is not carried out for its own sake but to achieve an objective that lies beyond the practice of information seeking itself." (Sundin & Johannisson, 2005b, p. 107). Therefore credibility assessments do not stand alone, but are connected to the continued seeking of information, and the ways in which information is used. Thus individual activities should not be considered in isolation: selecting multiple sources; claims around source authority; connecting pieces of information in complex ways; and so on, are not in themselves complex or simple. Context sensitivity is fundamental for a sophisticated epistemology; it is not very sophisticated to view the idea that the earth is round rather than flat as 'tentative' whereas theories of dinosaur extinction do require a more tentative stance (Barzilai & Zohar, 2012, p. 42).

12.1.3 Epistemic Frames

Epistemic Frame Theory (EFT) provides a means to conceptualize these connections between commitments. Epistemic Frames can be thought of in terms of the connections between elements usually described as: skills, knowledge, values, identities, and epistemological rules, from any particular domain. EFT is explicitly discourse oriented, and argues that ENA may give insight into the frames of experts and novices working in a domain (Shaffer & Graesser, 2010; Shaffer et al., 2009). ENA thus offers a way to model the relations among elements of epistemic frames – which are constituted in discourse: particular facets of the frame (e.g. keywords indicating particular ways of working) become nodes, while connections between those nodes represent the patterns of connections between frame facets (e.g. the co-occurrence of keywords).

The search context is a particularly interesting one in which to deploy ENA. The theory takes as its unit of analysis any chunk of dialogue (a session) broken into meaningful chunks (stanzas). In the case provided in this paper we chunk stanzas by task, but for other analyses it may be more

appropriate to chunk by search query. ENA allows us to examine various types of connections, and broadly examine whether or not particular ways of making sense of information – in the confines of answering questions, or attempts at deeper understanding – co-occur with particular types of sourcing or connections between knowledge. Moreover, such analysis may offer insight into the quality of frame elements (nodes) – for example, claims about the 'authority' of websites might be rather trivial (for example, "it looks good") or more sophisticated (for example, "they used a scientific method"); understanding how such justificatory elements of the frame are connected to sourcing elements may give insight into the pedigree of those sourcing decisions which would be missed by looking only for 'authority' claims. When we seek information we search for both in the sense that we search for information, and we search for a purpose; how users engage with those purposes is what matters, and how they connect those purposes to their epistemic commitments. Their sourcing decisions and the way they conceptualize the complexity of information is crucial. Thus, while search strategies matter, and an overreliance on individual (authoritative) websites or the consistent use of multiple websites (corroboration) might be of concern, their relationships to other epistemic assumptions are key.

In the work reported in this thesis we take a previously analysed dataset, and apply ENA to the epistemic discourse around searching for information to address a number of pre-assigned questions. We discuss the dataset further below, note though, that the use of this pre-existing dataset allows us to compare insights gained through close textual analysis, and those offered through ENA, thus supporting the development of a 'proof of concept' model for ENA around epistemic commitments.

12.1.4 The Collaborative Lens

A fundamental component of understanding the social context and role of language in learning is an analysis of how language mediates and represents learners' views on their learning. This component of learning is also fundamental to the theoretical and practical application of ENA, which takes as its data the discourse used in the course of students' learning practices. As noted above, it also avoids the methodological risk of artificially activating metacognitive strategies through the use of think-aloud techniques.

High quality collaboration also entails particular – epistemic – ways of working. In the context of epistemic commitments, take for example Hutchinson and Hammer's (2010) case study from a science classroom, in which framing by students which could be characterized as 'sensemaking' in nature (and, as we note below, accountable or exploratory) is contrasted with a more absolutist perspective. For example, at one point a student (Bekah) offers and explains an equation to illustrate her understanding – this is taken up and referred to collectively as "Bekah's Law", illustrating a cohesive tie (the repetition of terms through a text) demonstrating a type of common knowledge built up in that classroom (Hutchison & Hammer, 2010). This type of talk bears striking resemblance to exploratory or accountable talk, research on which focuses on the ways in which language is used "as a social mode of thinking – a tool for teaching-and-learning, constructing knowledge, creating joint understanding and tackling problems collaboratively" (Mercer, 2004, p. 137). In exploratory dialogue:

> *Partners engage critically but constructively with each other's ideas. Statements and suggestions are offered for joint consideration. These may be challenged and counter-*

challenged, but challenges are justified and alternative hypotheses are offered. Partners all actively participate, and opinions are sought and considered before decisions are jointly made. Compared with the other two types, in exploratory talk knowledge is made more publicly accountable and reasoning is more visible in the talk. (Mercer & Littleton, 2007, p. 59)

In such talk, explanatory terms and phrases are more common, for example: I think; because/'cos; if; for example; and also. Similar characterizations of effective dialogue have emerged from the work of other researchers across a range of ages (Michaels, O'Connor, Hall, & Resnick, 2002; Resnick, 2001). This talk is thus explicitly epistemic, in that it embodies consideration of "the other's" perspective. The significance of this type of dialogue for the study of epistemic commitments receives further support from Reznitskaya and Gregory (2013) who note that more sophisticated epistemic cognition of the 'evaluativist' variety, is closely associated with the kind of exploratory talk which is associated with educational gains. This claim – of an epistemic relationship to exploratory talk – is further supported in Rosenberg, Hammer and Phelan's work (2006). In that study, a case study was presented of a 15 minute discussion of the 'rock cycle' by a group of 8th graders – again, making use of dialogue excerpts to exemplify. Rosenberg et al., note that in the initial stages students were engaged in largely unproductive talk (there was some accretion of knowledge, with little explanation or evidence of understanding – it was largely *cumulative* in nature), suggesting this was because: "They [were] treating knowledge as comprised of isolated, simple pieces of information expressed with specific vocabulary and provided by authority" (Rosenberg et al., 2006, p. 270). After a brief intervention by the teacher, suggesting the students might build on their own knowledge, this talk instead shifts to more productive dialogue, seeking coherence and understanding in trying to create a theory and use terms they understand – the description, and excerpts provided here suggest this talk might be characterized as more 'exploratory' in nature. Exploratory dialogue is thus closely associated with a component of our approach to epistemic commitments around openness to ideas, and justification for them. This is particularly interesting given evidence that collaborative information seeking is a common activity (see Shah, 2012 for a review), and may have benefits for information seeking in classroom contexts (Lazonder, 2005) .

12.1.5 A Proposal for Epistemic Commitments

In addition to exploratory dialogue, other components of epistemic cognition are highlighted in the literature. Earlier we noted Mason et al.'s (2009, p. 69) claim that across models of epistemic cognition, there was a focus on the certainty, simplicity, source and justification for knowledge. We then noted Tsai's (2004, p. 109) framework for information commitments, comprised of: standards for correctness; standards for usefulness; and searching strategy. We thus recast these two positions such that our focus is on:

1. *Which sources of information are selected* – comprised of credibility decisions (from corroboration of information across sources, to trust in the authoritativeness of sources)
2. *How information is used* (in action – to justify claims, to make decisions) – comprised of justifications and source use (from dialogic approaches using talk of an exploratory nature, to attempts to directly approach questions by matching information to answers)
3. *How links between information are created (or not)* – comprised of claims, (explicitly in language and through structured environments, as well as implicitly through search patterns) made around connectedness of concepts (from a holistic to a piecemeal perspective of knowledge)

This recasting aligns well with the specific context being studied here – that of collaborative information seeking. It also provides three conceptually distinct (although probably empirically associated) constructs for study. In the two other models highlighted – Tsai's information commitments, and the general model described in Mason's analysis of the literature – it is not clear that each component can be conceptually distinguished. Specifically, 'certainty' in the general model seems likely to be a function of justification and simplicity. The sophistication of one's perspectives on 'certainty' depends on the purpose for which the information is being deployed, and the other information to which it is being associated – and indeed, whether one holds a complex enough view of knowledge to recognize the instability of certain information. Indeed, 'certainty' could be characterized as a connection between a facet of the information (publication date metadata, for example) and justification (recency, or information being well 'established' for example). Similarly, it is not clear that 'searching strategy' is a useful conceptualization of an information commitment given its strong relation to the tools at hand, and the type of task set and justificatory framework required for that task. The proposed model focuses on whom we believe, what we do with information; and how we connect information. However, although 'epistemic commitments' recasts the constructs of other models, it still provides a lens for them. For example, certainty is recast in light of our standards for credibility, explanation, and relating components of information such as new and old, or geographically located information; simplicity is most clearly related to the third focus on connectivity; source to the first; and justification to the second. Furthermore, the rhetorical shift both in the foci, and in the notion of 'commitments' over 'cognition' motivates an operationalization centered on:

1. *Source selection*, the corroboration of information across opened links, and the types of links repeatedly visited (e.g. use of authoritative sites such as 'BBC', repeated use of top links in search engine results pages, use of source metadata in justificatory framework below).
2. *The type of justificatory framework used*, the assertion of information (perhaps closely related to a style of search which emphasizes precision of information with little consideration to its wider impact) versus reasoning and understanding activities (closely related to exploratory dialogue)
3. *The sorts of connections made* by students between concepts in their dialogue and document creation, and in the ways that users build links between information in their search patterns (building on search terms by rephrasing and appending new query terms, following internal links, and using terms from opened sources to find new ones all imply some commitment to holistic perspectives on knowledge).

This model thus describes both a conceptual and practical means to explore epistemic commitments in information seeking environments, and will be the model adopted in this work.

12.2 Research Questions and Contribution

The PhD sits firmly in the middle space – bringing the learning sciences (including psychology around epistemic cognition, and the discursive approach we advance), and analytics (including tool design) together. It provides an exemplification of the interpretive flexibility (Hamilton & Feenberg, 2005) around intentional design issues in the middle space between the learning sciences/educational research, and the use of computational techniques to capture and analyse data (Suthers & Verbert, 2013), bounded by the triadic relationships between *epistemology* (the nature of knowledge), *pedagogy* (the nature of learning and teaching) and *assessment* (S. Knight, Buckingham Shum, et al., 2013a).

This PhD aims to produce a new type of trace-oriented approach to a psychological construct, as motivated above, through the use of a novel collaborative information seeking paradigm. A target construct has been selected, around which tools to track potentially useful data for manual and automated analysis are being developed. While there is a potentially wide range of interests in this construct, this PhD focuses on a particularly salient context to education which where appropriately constructed may exemplify pedagogy supported by educational research on collaboration and dialogue – namely collaborative information seeking. This is a well motivated context for attention both because its incidence provides an external validity, and because the educational outcomes or/and processes of CIS are understudied.

This PhD thus takes the developed learning analytics lifecycle (Clow, 2012) which moves through: learners, data from or about learners, processing of data into metrics, and intervention. In this PhD we start with a task context not uncommon to many students, and consider the types of data that may be captured about those contexts. Theorising is needed to consider how to process this data into appropriate metrics – how to interpret the data. Once metrics are developed, consideration of interventions in this PhD will not delve into the full range of possible tutorial interventions, but rather focus on ways in which data might be meaningfully represented either for students or for expert educators in support of learners (in line with Clow's point that interventions do not necessarily need to involve a return of data to students). Given Clow's call for us to focus on assessment (which is what students focus on) and this PhD's agreement with (and parallel arguing for) that call, this model is well suited to this endeavour.

Furthermore, the PhD is scoped within the bounds of existing broad models of analytic cycles an overview of which are given by Elias (2011) as summarised in Table 14. In particular, we note that data selection and capture involves defining goals, and in the case of this PhD a clear rationale is given for the selection of the target construct (epistemic commitments). A set of data to be captured is defined below, with a clear theorisation around how this should be aggregated and reported to give insight into student epistemic commitments. The study described here involves processes of validation and comparison across task types (CIS and MDP) which lend themselves to refining, and open scope for prediction of CIS behaviour from the constrained environment of MDP. A deeper analysis of the ways in which data might be visualised, or used in sophisticated pedagogic strategies (for example, how best to *teach* towards sophisticated epistemic commitments) is beyond the scope of this PhD, although some preliminary indications will be given. We note that processing data involves a range of irreducible elements involving theoretical models, target constructs, and data constraints. For example, the segmentation of discourse data into chunks for processing is both a pragmatic decision regarding the selection of statistical technique, and also a theoretical one around the units at which data may be meaningfully discussed (see S. Knight & Littleton, in draft). We note that this accords well with our claims around the middle space, and this PhD's position in it: Theory, in this case largely psychological in nature, gives insight into learning that can lend itself to development of analytic techniques, both with respect to a manual analysis and in tandem through provision of proof of concept tools for learning technologists.

Table 14 - Comparison of Analytics frameworks and models, (Elias, 2011, p. 10)

Knowledge Continuum	Five Steps of Analytics	Web Analytics Objectives	Collective Applications Model	Processes of Learning Analytics
		Define goals	Select	Select
Data	Capture	Measure	Capture	Capture
Information	Report		Aggregate	Aggregate & Report
Knowledge	Predict		Process	Predict
Wisdom	Act	Use	Display	Use
	Refine			Refine
		Share		Share

This narrative contribution, and the literature reviewed in the first year report and brief above motivate three research questions as shall now be discussed in turn:

1. **How can analytics for Epistemic Commitments be characterised?**
2. **Is there a relationship between Commitments in MDP and CIS tasks, and how can it be characterised?**
3. **What characterises the Epistemic Commitments of the most 'successful' groups?**

12.2.1 How can analytics for Epistemic Commitments be characterised?

The first of these questions is addressed via theoretical work (in part stated above), and use of empirical research around the ways in which information seekers treat multiple documents, about which researchers have prior knowledge. That is, data to be treated as a proxy for EC will be gathered from two sources:

1. An existing validated psychometric questionnaire for information commitments (as detailed below)
2. A new, purpose-designed, multiple document processing (MDP) experimental task involving the seeking of information within a closed set of documents and the creation of a summary document (as detailed below)

This empirical research will use sociocultural approaches to analysis (Mercer, 2004) in which researchers move from close analysis of individual words, to utterances, to whole documents in light of the activity in which speakers are engaged. This approach is well suited to this research data and treated as a source of validation for automated analytic approaches, such as ENA (see S. Knight, Arastoopour, Williamson Shaffer, Buckingham Shum, & Littleton, 2013 which is in submission as a conference paper). This analysis of process data (including sourcing decisions from documents we have prior knowledge on) will be compared to: task outcomes including quality judgements and textual qualities for epistemic facets, and outcomes on the survey instrument. Answering this first question offers validation for the conceptualisation of the construct of epistemic commitments. We anticipate that the type of analytic approach discussed in this proposal will have a stronger relationship to outcomes than simple metrics alone, and by extension manual analysis.

12.2.2 What is the relationship between EC in MDP and CIS?

The second of these questions is addressed via a comparison of behaviours in the controlled MDP task, and open-web CIS task using search challenges over the internet. The CIS challenge is detailed below in the design section.

I anticipate using the broad tools of sociocultural analysis to investigate the meaning making across these two task types; such analysis frequently includes use of descriptive statistics to give surface level insight into the types of activity being engaged. There are two objects of analysis:

1. Task Outcomes, comprising:

a. The written output from each task, and its epistemic properties with respect to their commitments (particularly on the MDP task in which we have prior knowledge of the document qualities).
b. An assessment of text 'quality' for academic quality such as a comparison of expert and novice texts, use of a natural language rhetorical parser (e.g. Simsek, Buckingham Shum, Sandor, De Liddo, & Ferguson, 2013) and a qualitative assessment of text quality. Clear claims should be stated and related where appropriate, and justifications given for why they were the best supported claims to provide.

2. Task trace data, as summarised in the table overleaf, comprising:

a. Manual analysis of chat-data, and navigational trace
b. Automated analysis of chat-data and navigational trace

12.2.3 What characterises the EC in most 'successful' groups?

The third of these questions will be addressed via a between-groups comparison of 'success' and its relationship to epistemic commitments (as characterised above). Analysis at this stage will involve:

1. Information Commitments Survey results
2. Self reported pre-task-knowledge (detailed below)
3. A 'warm-up task' performance on a task such as those given at www.agoogleaday.com (a proxy for search skill)
4. Analysis of data as in 2 above (both trace and outcome).

Of course, such analysis must recognise the epistemic constraints of the task and assessment scheme (which is not a between-groups variable).
At this stage a set of sub-claims partially derived from the first research question may be assessed, including speculatively that:

1. Pairs that make connections between 'understanding' and sourcing decisions (corroboration v authority) will perform better than those who do not.
2. Pairs that make connections between knowledge claims (task-specific pieces of information), sourcing decisions, and information seeking (search) will perform better than those who do not.
3. Pairs that make connections between sourcing decisions and general knowledge claims (around broad task requirements) will perform better than those who do not.
4. Pairs that make connections between modes of the same dimension (for example, corroboration and authority) will perform better than those who do not.

12.3 Procedures

12.3.1 Pilot

Prior to the Christmas break 2013 a pilot study will be conducted. The pilot is intended to inform aspects of the methods including software suitability, and analysis. Two pairs (four participants) will be asked to engage in a full trial run of the procedure as below. Participants in the pilot will be separated and asked to use the tool as below, however they will be in the same proximal area, with the experimenter at close proximity. Participants will be asked to record difficulties experienced as they work, and interviewed after task completion to assess.

12.3.2 Main Study

12.3.3 Participants

Ten pairs of PhD students or academic staff and ten pairs of undergraduate students (40 participants total) will be recruited to form two groups: experts in general academic literacy and novices. Participants will be recruited outside of the specialist knowledge area of the tasks.

12.3.4 Design

The study has a mixed design, with one between-groups factor (expertise), and within groups factors arising from analysis of trace data and psychometric assessment. Mixed method approaches will be used for analysis of data. The primary tasks (Search (S), MDP and CIS) and psychometric assessment (P) will be counterbalanced.

Experts (n=10): P S MDP CIS

Experts (n=10): S MDP CIS P

Novices (n=10): P S MDP CIS

Novices (n=10): S MDP CIS

MDP and CIS will not be counterbalanced as MDP is framed as a task to learn how to use the system, it is a naturally preliminary task (i.e. has face validity as a training task), and is not a between-groups variable.

12.3.5 Tasks

There are four task units, broken into two counterbalanced blocks. The first block consists of an assessment using the Information Commitments Survey measure.
The second block consists of:

1. A brief measure of pre-task knowledge of the target content, this will consist of a simple binary response (yes/no) regarding the target task (e.g. "I have heard of…"), and a Likert type response giving a confidence level of knowledge (e.g. "My knowledge of …. is: 1 2 3 4 5 where '1' is 'I've never heard of …' and 5 is 'I have expert knowledge of …'").
2. Performance on a warm-up task (which we take as a proxy for search skill) The warm up task is likely to be taken from the set of examples (or similar tasks) here: https://docs.google.com/file/d/0BxlpTzK9iG-2a3VvZi1nZWl0eEE/edit?usp=sharing&hl=en&forcehl=1 The task is a retrieval task, and as such has a correct answer (or set of answers); good performance is indicated by a retrieval of the required facts, in a short period of time.
3. A Collaborative Multiple Document Processing (MDP) task
4. A Collaborative Information Seeking (CIS) task

Both the MDP and CIS tasks will involve processing a number of documents on a particular topic – in the former case pre-selected provided documents, and in the latter documents retrieved by participants – and producing a summary of the best supported information and arguments therein. Specifically the participants will be given the following instructions:

> *Your task is to act as an advisor to an official within the science ministry. You are advising an official on the issues below. The official is not an expert in the area, but you can assume they are a generally informed reader. They are interested in the best supported claims in the documents.* ***Produce a summary of the best supported claims you find and explain why you think they are.*** *Note you are not being asked to "create your own argument" or "summarise*

everything you find" but rather, make a judgement about which claims have the strongest support.

In the first case, a colleague has already found a number of documents for you to process with your partner. In the second you and your partner should work together to find relevant materials.

You should:

1. *Read the questions/topic areas provided, these will require you to find information and arguments [on the internet/in documents] to present the best supported of these, you should decide with your partner which are best as you search*

2. *As you visit pages, share ones you want to discuss with your partner using the tool to do so*

3. *Use the chat box, or the snippet tool, or cut and paste extracts from the pages into the etherpad shared document to discuss specific information or arguments from the pages (you'll get an introduction to these tools)*

4. *Tag saved elements to bundle perspectives or information together*

5. *Focus your time on finding and selecting information and building a consensus with your partner about which bits of information are best supported. Use the etherpad to collate this information. Once you've agreed as a pair that you've found enough information, work together to use the etherpad to collate the best supported information*

6. *You should explain why the claims you've found are the best available*

7. *You should spend no more than 1 hour on the first task, and 2 hours on the second*

The aim of these instructions is to guide the participants in their task, encouraging them to explain their decision processes as they go, while not directing them in particular to either sourcing via corroboration or authority (and explanations thereof). In the first instance participants will be directed to a set of documents around a selected topic with three primary themes.

In the CIS tasks, the participants will be directed to a different topic, selected because:

1) The Wikipedia article is not particularly high quality (it is rated 'b-class' in the alternative medicine WikiProject quality scale, and an article for the topic under a scientific name does not receive a rating on any relevant scientific or medical WikiProject scales, but is a stub article).

2) Google searches for the colloquial topic bring up rather different content to those for its scientific name (the Wikipedia article for which is a stub) (These different queries can be tracked in participant search logs)

3) Good well sourced accessible material is moderately challenging to find

4) The controversy is largely around restrictions and side effects (i.e. it is uncontroversial that it has a medical effect, although risks and scope of those effects are disputed).

12.3.6 Performance

Performance is assessed primarily through the analysis of an output document. The tasks used in this study have relationships to similar tasks such as the Civil Service Fast Stream tasks and the Collegiate Learning Assessment which is a particular instance of a scenario-based writing task or performance based assessment. Although these instruments seek to measure other facets of participant's capabilities (including work skills and inter-personal skills), the sample exercises in these cases have similar outcome measures to those proposed for this task, specifically:

1. A range of key topics are covered (pre-defined core topics in the MDP task, and relevant to the task in the CIS task).
2. A range of sources is used
3. Specific claims are stated, with quotations or/and figures given where appropriate
4. Evidence is evaluated for credibility and reliability
5. There is synthesis of information from across sources (i.e. claims are integrated not repeated, and associated claims from different sources are synthesized in the text).

12.3.7 Documents

A detailed description of topic selection can be provided upon request. Two topics were selected, one for the MDP task and one for the CIS task. Both topics were chosen because: they provide a focussed topical research area which can be studied in isolation; they are not topics that are high profile controversies; they are not topics about which there is scientific consensus; they are topics about which a variety of sources exist.

12.3.8 Psychometrics

The appendix gives sample items from three of the most commonly used, and three other related, psychometric assessments. The Information Commitments Scale (ICS) is used in this study with a preliminary vignette given to contextualise answers and questions modified for English use (the ICS has been translated into English by its original authors for publication). The vignette will provide participants with a context to consider their answers in around finding and supporting the understanding of health information for a friend. This scale has been chosen as it most closely relates to the theoretic model proposed in this work, and has clear implications for user actions on search engines. In addition to three of the other psychometric assessments being criticised for the psychometric properties (as discussed earlier in the literature review), the dismissed assessments lack face validity for my research, focussing on different aspects of epistemic cognition or/and attempting to assess large grain epistemic beliefs rather than context sensitive ones. For example, even in the Connotative Aspects of Epistemic Beliefs (CAEB) scale which is used at a domain or sub-domain level, the shift between specific claims one might make within a domain and the epistemic beliefs associated with those claims may not be respected by the scale.

12.3.9 Interview and critical incident analysis

At this stage no interviews are planned, except informal feedback during debriefing. However, the pilot study will indicate whether such interviews should be included in the study as a further form of validation. If this is the case these interviews are likely to take the form of critical incident analysis, or scenario based interviews in which specific decision stages in a task are discussed in light of

epistemic commitments (for example, discussing what information one might seek to support a particular claim, or what one might do with information found).

12.4 Candidate Trace Data for Epistemic Commitments

Both CIS and MDP analysis will be informed by the discourse and trace analytics given below such that for each claim made, a set of trace and discourse markers can be associated with that claim. In the case of MDP tasks analysis will be conducted in light of our prior knowledge regarding the document corpus. As such, a general behaviour model can be constructed for the MDP and CIS task, in addition to which in the MDP task claims can be associated with particular documents, and claims regarding sourcing decisions (around use of poor or stronger sources, corroboration across particular *types* of sources and so on) can also be made. This data might be used for quantitative analysis, or to group particular types of claims (as a kind of metadata), or to inform qualitative analysis. A summary of the documents is given in the preceding section. Recall, in each instance of trace data, it is *connections* that mark key commitments. Thus, for any claims made, a fundamental component of analysis will involve attempts to 'segment' claims, and their related dependencies. That is, for any claim analysis will involve looking for connections between that claim, trace markers (including documents read), and discourse related to that claim.

Table - Candidate Trace for Epistemic Commitments

Commitment dimension	Mode of engagement	Manual analysis	Trace data
Choosing sources of information *Which claims are selected*	*Corroboration*	Analysis of talk for connections between claims and multiple sources.	Citation of multiple sites for claims Recall (no. of sites used as proportion of sites used by all groups) Connections between metadata from multiple sources
Choosing sources of information *Which claims are selected*	*Authority (search)* *Authority (source)*	Metadata from only single sources referred to, use of individual sites. Claims about website authority made.	Using first presented documents Frequent use of same source Source metadata referred to
Justification and use of information *How claims are used and justified*	*Matching*	Reference to questions and tasks made. Bias towards positive claims.	Precision (narrow range of source topics covered) Use of keywords taken from the question
Justification and use of information *How claims are used and justified*	*Understanding*	Exploratory dialogue and use of argument markers	Use of 'exploratory' keywords. Possible development of machine learning techniques for exploratory dialogue.
Connectedness/complexity of information *How claims are connected or not*	*Piecemeal approach*	Connections made (or not) between individual claims in dialogue	Listing claims There are no connections made between claims (no shared tags for example, written text consists of single unconnected claims)
Connectedness/complexity of information *How claims are connected or not*	*Holistic approach*	Connections made (or not) between individual claims in dialogue	Query editing Link following Using terms from page 'a' in query 'b' Dialogue involves connections between claims.

12.5 A software tool for analysis of CIS

Coagmento will be used for data capture. Coagmento is an established research tool for CIS, with Firefox and Chrome addons, and an Android app available. It was designed with CSCL and CSCW literature in mind in addition to the CIS requirements that it support the logging and sharing of search queries. It thus comprises: a query logger; a bookmark and 'snippet' tool to clip and share short website excerpts; a ranking tool to rank queries and bookmarks; tagging for bookmarks; a chat tool. Each dataset is associated with a particular 'project', and users can join projects which then gives them access to the data for that project in addition to a project etherpad. In addition to the end-user data as described, Coagmento tracks page views during browsing (and thus this data can be mined too).

12.5.1 Analysis

For each task, manual analysis will involve movement from a focus on individual utterances, to sets of topically related utterances). This analysis will be conducted using the kind of sociocultural approaches to analysis described in Mercer (2004) drawing particular inspiration from the notion of a 'stanza' described by Gee (See e.g. Gee, 1989) in which expressions are analysed in context by segmenting sets of topically related utterances; grouping them into stanzas. A key component of this analysis will involve consideration of a number of possible methods for this delineation here including:

1. Grouping discourse and trace by the document, query, or question with which they are associated
2. A variant on (1), grouping discourse and trace by topical clusters of documents, queries, or questions with which they are associated
3. Grouping discourse and trace by the claims with which they are associated (that is, associate documents and queries with claims made in the output document, and include their trace data in a single stanza or/and taking claims made (factual statements) in discourse and associating documents/queries with these to create a stanza).
4. Grouping discourse and trace using temporal markers (for example, every 5 minutes)

Careful analysis will be needed here. Suthers has described methods for analysis of 'uptake' or 'contingencies' of ideas, in which idea 'b' is contingent on idea 'a' iff idea 'b' is temporally latter to 'a' and conceptually contingent on 'a' (for example it involves developing ideas from 'a'). However, as Introne and Dreschler note, while documents contain a relatively stable set of concepts within any individual document, conversations often use the same word to mean different things at different times. Thus clustering all uses of the same term – and the activity trace around that term – may not be an effective means of stanza delineation. Indeed, Furburg and Ludvigsen (2008) and Furberg (2009) note that understanding interaction trajectories – the ways in which, as Littleton and Mercer (2013) would put it, dialogue is both context contingent and context creating over time – is fundamental to understanding how students are meaning-making in socio-scientific issues such as the tasks in this research.

The general structure of each group's dialogue can be analysed for differences related to their final output and the processes leading to it. In addition, for each stanza, a comparison may be made between stanzas associated with claims made, and those associated with no claim (i.e., not used in the final output). Finally, a comparison can be made between the processes of experts and novices.

Furthermore, in the case of the MDP task claims can be associated with their documents directly, with claims made regarding the qualities of those documents and the repetition of information across them. Thus this metadata may also be included within any given claim stanza. In the case of CIS tasks, the use of metadata extraction tools[47] (to automatically extract metadata from websites visited), tf;idf algorithm[48] and semantic concept tools[49] such as the Wikipedia miner (which disambiguates terms within any given document using the structure of that document and a comparison with possible Wikipedia sources) also provide sourcing information – although not information that has had normative judgements made regarding quality (as in the MDP task).

12.6 Timeline for Completion

12.6.1 Timetable

Q	Period	Milestone	Work to be done
Q1	Oct 12- Dec 12	• Preliminary/familiarisation work conducted • First conference submissions	• 2 CSCW workshop papers submitted • 1 LAK full paper submitted
Q2	Jan 13- Mar 13	• Papers accepted • First conference attendance	• Revise & resubmit papers • DCLA workshop paper submitted
Q3	Apr 13- Jun 13	• Second conference attendance (paper nominated for 'best paper') • Lit review to supervisors • Report to Supervisors (June)	• Write paper & workshop presentations & deliver • Compile work into lit review and add where appropriate • Finalise report
Q4	Jul 13- Sep 13	• Study visits at Wisconsin-Madison and Stanford • First Year Report due (August) for viva in September	• Refine Formal Proposal • ICLS paper written from Madison trip & workshop proposal • First year viva
Q5	Oct 13- Dec 13	• Report to Research School (October) • Empirical work planning completed	• Attend EduWiki (1st/2nd Nov) • Write & give invited talk at Society of the Query conference, Amsterdam (Nov 7th/8th) • Apply for ethical clearance • Recruit participants & conduct usability testing
Q6	Jan 14- Mar 14	• Usability testing & design alterations completed and written up • Main empirical work starts • Internal seminar (if accepted, on ICLS paper)	• Analysis of usability testing, possible workshop/conference submission • Recruit participants, task design & start main empirical work • Look for (and apply to) summer internships • If ICLS workshop accepted, review

[47] E.g. http://meta-extractor.sourceforge.net/

[48] E.g. http://code.google.com/p/tfidf/

[49] See e.g. http://texlexan.sourceforge.net/, http://dbpedia-spotlight.github.io/demo/ and http://wikipedia-miner.cms.waikato.ac.nz

			submissions, continue to organise
Q7	Apr 14- Jun 14	• Main empirical work is completed • Analysis starts • ICLS paper & proposed workshop (subject to acceptance)	• Conduct empirical work • Start analysis process
Q8	Jul 14- Sep 14	• Possible summer internship? • Second Year Review (September) • Thesis Overview and Table of Contents (October) • Internal Seminar (October) • Report to Research School (October)	• Possible internship work • Analysis and write-up starts • Presentations on work in progress
Q9	Oct 14- Dec 14	• Journal submissions (see below) • Second Year Review (September) • Thesis Overview and Table of Contents (October) • Internal Seminar (October) • Report to Research School (October)	• Write for journal submissions, continue writing thesis • Start preparing updated paper for relevant conferences • Phase 3 begins in earnest
Q10	Jan 15- Mar 15	• First draft of thesis	• Thesis and journal paper writing
Q11	Apr 15- Jun 15	• Final completed draft of thesis • Internal Seminar (July)	• Finalising thesis draft • Prepare for practice thesis defence at internal seminar
Q12	Jul 15- Sep 15	• Third Year Review (September) • Thesis Submission and Viva (October) • Report to Research School (October)	• Complete thesis submission, prepare for Viva, finalise paperwork • Celebrate success!

In addition, I have a work plan for publications directly related to my PhD, given below in order of authoring period (including two possible conference submissions in the academic year 2013/14, and one in 2014/15); these publications mark the contribution of this research:

- Following LAK14 I anticipate the Journal of Learning Analytics making a call for a special issue on Discourse Centric Learning Analytics, to which I will respond with a modified version of my workshop paper.
- August 2013 – write up collaborative information seeking in Wikipedia research, consider submission as CHI note (I was assisted with data capture by Aaron Halfaker who I would anticipate asking to co-author on this)
- October 2014 I anticipate the deadline for submission to LAK15 – I intend to submit to this conference for final year attendance.

A number of publications are in review or final draft:

- Karen Littleton and I have written a paper (in final draft form) on computational techniques for detection of educationally constructive dialogue which we intend to submit to ijCSCL following an internal workshop on the topic in January
- Learning Analytics Journal, on Learning Analytics, Epistemology, Assessment and Pedagogy (building on my LAK13 paper);
- A collaborative work with colleagues from the CSCW13 workshop on Collaborative Information Seeking, we have submitted to a special issue of Computer on collaborative dialogue to mediate CIS in various disciplinary contexts.
- A publication from my MPhil, with Neil Mercer, submitted to Technology, Pedagogy and Education (revisions submitted)
- I submitted a chapter (accepted) to the Society of the Query reader, this is in the editing stage

I have considered a number of other contributions, some (but probably not all) of which I would anticipate writing as follows:

- Learning Culture and Social Interaction, on machine learning for sociocultural analysis – promises and pitfalls (likely mid-2014);
- Journal of Learning Sciences, on sociocultural computational linguistics perhaps focussed on "data reduction" methods in sociocultural discourse analysis and parallels in debates across manual and automated techniques and how they might learn from each other (likely mid-2014)
- Educational Psychologist, on Exploratory dialogue for Epistemic commitments (likely late 2014);
- Instructional Science, on Tracking Epistemic Commitments in Multiple Document Processing and Collaborative Information Seeking (likely early 2015);
- ijCSCL on a CSCL platform for Epistemic Commitments (possible early 2014, post usability testing, but likely to include subsequent preliminary empirical work too);

12.6.2 Literature Gaps

The literature review as it stands has a number of gaps which will need to be addressed in the final PhD thesis as follows:

- On knowledge forum and other relevant collaborative tools
- Around socio-scientific issues and potentially other topic/subject related literatures
- Around philosophical assumptions – specifically, there is a thread through epistemic commitments, epistemology of assessment, and discursive psychology (as method and context) which should be better drawn out particularly as related to pragmatism. Addressing David Williamson Shaffer's perspective on communities of practice and Dewey, and some anticipated new literature (Jan Derry's new book in November 2013) will be targets for this issue.
- There is a gap around socio-technical systems and the ways in which technology mediate our interactions with the world and society
- There is also a gap around what our expectations for new technology should be regarding their transformative power (such exploration could lead to a learning analytics publication,

and this may be an interesting outcome from our proposed ICLS workshop on learning analytics for practice)

- There is a gap on data reduction in socio-cultural approaches, we are working on this paper
- Around specific technical aspects of analysis regarding methods in information retrieval and their relevance to this research (such as tf;idf)

12.6.3 Feasibility and Main Risks

This section provides a key set of risks in the completion of the PhD work, and risk management to mitigate those potentials:

- Coagmento software tool not completed (to deployable standard)
    - The specification is currently being worked on, and should be ready for usability testing in December. At that point a break can be taken to iterate and develop further.
    - The tool is already developed to a useable standard, and is maintained by colleagues at Rutgers. We have talked to them and agreed mutual support in supporting the tool and sharing research.
- Exploratory Discourse Analytics Module not developed
    - This risk may come about through lack of staffing. While deploying EDAM in a particular context (and providing feedback on its success) is desirable for this work, it is not a necessary component and other semi-automated (bag of word) approaches, and manual coding are both suitable alternatives. In addition, Yulan He has expressed interest in continued collaboration on this tool (having worked on the current work).
- Papers rejected,
    - I have begun a publication track record, and am working on a number of other papers now in collaborations. I have a number of papers planned (as above) – rejections are mitigated insofar as there should hopefully be some balance of acceptances (as is currently the case). By PhD I should be able to clearly demonstrate a contribution, and writing of a publishable quality via my publication record.
- Failure to recruit participants and other practical empirical work issues
    - I have planned well in advance including preliminary considerations regarding participants; a university has been contacted and I will work with them to make use of their participant pool.
- Overrunning the PhD
    - I have a timetable of milestones to reach, including some lag. I have a strong record of meeting deadlines. I am aware of the need to limit research scope to ensure timely completion.
- Gaps in the literature review [as it stands]
    - I am aware of some gaps in my coverage of the literature here which will need filling for the PhD thesis including:
        - On knowledge forum and other relevant tools
        - Around socio-scientific issues and potentially other topic/subject related literatures

        - Around philosophical assumptions – specifically, there is a thread through epistemic commitments, epistemology of assessment, and discursive psychology (as method and context) which should be better drawn out particularly as related to pragmatism. Jan Derry's new book (Nov 2013) relating the sociocultural theory of Vygotsky and pragmatism of Brandom will greatly aid this.
        - There is a gap around socio-technical systems and the ways in which technology mediate our interactions with the world and socieity
        - There is also a gap around what our expectations for new technology should be regarding their transformative power (I can imagine a publication on this in relation to learning analytics)
- Lack Analytic skills
    - Taking projects on to learn some R and related tools already building my capabilities in this regard (see section 11)
    - In addition these projects, among others (including the specification for tool development) indicate a capability to think analytically regarding practical methods.

13. Concluding Remarks: The Epistemic Role of Search Engines

Educational systems should consider the epistemological assumptions on which they are grounded. Learning Analytics is one tool to provoke such discussion – but it alone will not make changes; the problem is not a new one, new tools (books, for example) have come, and become embedded in systems which are slow to change and may sometimes seem to be fixed of necessity. Learning Analytics, and a close analysis of alternative methods of assessment (such as that given in the Danish example of permitting internet access in exams) may open the space up for discussion about what we wish to assess, why, and how.

This work has suggested that, across models of assessment, one interesting facet of student working is the ways in which they deal with information, how they conceptualise information problems and seek to address these using various tools – specifically, search engines. This provides an interesting space for a research paradigm which analyses the in-action decisions students make around their search behaviour, looking for their epistemic commitments.

This work argued for a particular stance on epistemic commitments, building on – but distinct from – previous research in epistemic cognition and beliefs. This stance was aligned with the epistemological position described earlier particularly in the context of learning analytics. Aspects of epistemic commitments were discussed and some indications of operationalisations were given – the importance of exploratory dialogue for justification, the analysis of sites visited, and the potential to explore information behaviours with structured environments for further insights.

Of course one facet of this exploration relates to the role of the search engine as a socio-technic, epistemic tool. We should be mindful of evaluations of such tools as epistemic tools – tools which influence and mediate our epistemic commitments through the ways they present (and fail to present) information to us[50]. Indeed, in the case of CSCL environments it is precisely this power – to

[50] See for example Simpson (2012) and many of my blog posts at http://people.kmi.open.ac.uk/knight/tag/epistemic/

support students in their reasoning, learning, and collective action – which motivates our research. We can imagine various outcomes of an analysis of epistemic commitments (the psychological component) in the context of socio-technical analysis of search systems. For example, given the finding that users may not seek contrasting viewpoints by themselves, but the presenting of such perspectives (and credibility information) supports students in more advanced understanding (Vydiswaran, Zhai, Roth, & Pirolli, 2012) we can imagine systems to foreground such information in CIS contexts. Similarly, journal article recommender systems might offer more fine grained information such as "disagrees with" rather than "is related to"; learners could be presented with *diversity aware search* results, results personalised to the individual learner based on presenting viewpoints which the learner might not otherwise consider; and alternative sources, or corroborating sources might be highlighted in search results to support a range of sourcing commitments in action.

Of course, just as search engines are socio-technical systems, so too are learning analytics. As we argued in Knight, Buckingham Shum and Littleton (2013a), we should consider not only what data we gather (and in what environment) but in what ways our analytics are deployed, how they inform, what transformative power they might have. This issue is one to be worked on over the course of the PhD, with proposals for support mechanisms to be made in the PhD thesis, although one core possibility will be explored here – the use of more constrained tasks (multiple document processing) as a diagnostic assessment for open-web tasks. The primary contribution of my empirical work is to develop a theoretical base, supported by empirical findings, to develop analytics around information commitments. Specifically, my empirical work will:

- Develop a framework through which epistemic commitments may be analysed – using the theoretical developments proposed above – through the lens of CIS and its dialogue.
- A core practical output of this work will be the development of a software environment to support such CIS.
- Furthermore, the focus on *dialogue* motivates an interest in discourse centric learning analytics; and this too will be a core contribution of my PhD work
- A key focus will be on comparing multiple document processing, and exploratory search tasks – as a means to validate assumptions made in open web search (where assessment of commitments may be more challenging), and potentially as a means to provide a 'diagnostic assessment' space to support students in developing more advanced epistemic commitments

Each of these contributions thus feeds into the overarching aim of the PhD work to develop a framework for the capturing of analytics around epistemic commitments in CIS, in order to support students in more advanced practices.

14. References

Amershi, S., & Morris, M. R. (2008). CoSearch: a system for co-located collaborative web search. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (pp. 1647–1656). Retrieved from http://research.microsoft.com/en-us/um/people/merrie/papers/cosearch.pdf

Amershi, S., & Morris, M. R. (2009). Co-located collaborative web search (p. 3637). ACM Press. doi:10.1145/1520340.1520547

Anmarkrud, Ø., Bråten, I., & Strømsø, H. I. (In Press). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences*. doi:10.1016/j.lindif.2013.01.007

Aula, A. (2005). *Studying User Strategies and Characteristics for Developing Web Search Interfaces* (PhD). University of Tampere, Tampere, Finland. Retrieved from http://tampub.uta.fi/bitstream/handle/10024/67544/951-44-6488-5.pdf?

Ba, H., Tally, W., & Tsikalas, K. (2002). Investigating children's emerging digital literacies. *The Journal of Technology, Learning and Assessment*, *1*(4). Retrieved from http://www.edtechpolicy.org/ArchivedWebsites/JCTE/v1n4_jtla1.pdf

Bakhtin, M. M. (1984). *Problems of Dostoevsky's poetics*. Minnneapolis: University of Minnesota Press. Retrieved from http://www.getcited.org/pub/101523787

Barnes, D., & Todd, F. (1977). *Communication and learning in small groups.* Routledge & K. Paul.

Bartlett, J., & Miller, C. (2011). *Truth, lies and the internet a report into young people's digital fluency*. Demos. Retrieved from http://demos.co.uk/files/Truth_-_web.pdf?1317312220

Barzilai, S., & Eshet-Alkalai, Y. (2013). The Role of Epistemic Thinking in Comprehension of Multiple Online Source Perspectives. In Y. Eshet-Alkalai, A. Caspi, S. Eden, N. Geri ,, Y. Kalman, & Y. Yair (Eds.), *Proceedings of the Chais conference on instructional technologies research 2013: Learning in the technological era*. Raanana: The Open University of Israel. Retrieved from http://innovation.openu.ac.il/chais2013/download/a1_1.pdf

Barzilai, S., & Zohar, A. (2009). The Role of Epistemic Thinking in Online Learning. In *Proceedings of the Chais conference on instructional technologies research 2009: Learning in the technological era*. Raanana: The Open University of Israel. Retrieved from http://telem-pub.openu.ac.il/users/chais/2009/morning/1_2.pdf

Barzilai, S., & Zohar, A. (2012). Epistemic Thinking in Action: Evaluating and Integrating Online Sources. *Cognition and Instruction*, *30*(1), 39–85. doi:10.1080/07370008.2011.636495

Bateman, S., Gutwin, C. A., & McCalla, G. I. (2013). Social navigation for loosely-coupled information seeking in tightly-knit groups using webwear. In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 955–966). Retrieved from http://dl.acm.org/citation.cfm?id=2441885

Bateman, S., Teevan, J., & White, R. W. (2012). The search dashboard: how reflection and comparison impact search behavior. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (pp. 1785–1794). Retrieved from http://dl.acm.org/citation.cfm?id=2208311

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, *13*(5), 407–424.

Beaudoin, L. P., & Winne, P. (2009). nStudy: An Internet tool to support learning, collaboration and researching learning strategies. *CELC-2009 Vancouver, BC Http://learningkit. Sfu. Ca/lucb/celc-2009-Nstudy. Pdf*. Retrieved from http://learningkit.sfu.ca/lucb/celc-2009-nstudy.pdf

Bilal, D. (2002). Children's use of the yahooligans! web search engine. III. cognitive and physical behaviors on fully self-generated search tasks. *Journal of the American Society for Information Science and Technology*, *53*(13), 1170–1183. doi:10.1002/asi.10145

Black, P., & Wiliam, D. (2001). Inside the black box. BERA. Retrieved from http://www.collegenet.co.uk/admin/download/inside%20the%20black%20box_23_doc.pdf

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 5–31.

Bloom, B. S. (1956). *Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain* (2nd edition.). Addison Wesley Publishing Company.

Bonfadelli, H. (2002). The Internet and knowledge gaps. *European Journal of Communication*, *17*(1), 65.

Branch, J. L. (2001). Junior high students and Think Alouds: Generating information-seeking process data using concurrent verbal protocols. *Library & Information Science Research*, *23*(2), 107–122. doi:10.1016/S0740-8188(01)00065-2

Brand-Gruwel, S., & Stadtler, M. (2011). Solving information-based problems: Evaluating sources and information. *Learning and Instruction*, *21*(2), 175–179. doi:10.1016/j.learninstruc.2010.02.008

Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novices: analysis of a complex cognitive skill. *Computers in Human Behavior*, *21*(3), 487–508. doi:10.1016/j.chb.2004.10.005

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education*, *53*(4), 1207–1217. doi:10.1016/j.compedu.2009.06.004

Bråten, I. (2008). Personal Epistemology, Understanding of Multiple Texts, and Learning Within Internet Technologies. In M. S. Khine (Ed.), *Knowing, Knowledge and Beliefs* (pp. 351–376). Dordrecht: Springer Netherlands. Retrieved from http://www.springerlink.com/content/j664674514614405/

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist*, *46*(1), 48–70.

Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (in press). Justification beliefs and multiple-documents comprehension. *European Journal of Psychology of Education*, 1–24. doi:10.1007/s10212-012-0145-2

Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (2012). Students working with multiple conflicting documents on a scientific issue: Relations between epistemic cognition while reading and sourcing and argumentation in essays. *British Journal of Educational Psychology*, n/a–n/a. doi:10.1111/bjep.12005

Bråten, I., & Samuelstuen, M. S. (2007). Measuring strategic processing: comparing task-specific self-reports to traces. *Metacognition and Learning*, *2*(1), 1–20. doi:10.1007/s11409-007-9004-y

Braten, I., & Stromso, H. (2006). Constructing meaning from multiple information sources as a function of personal epistemology. *Information Design Journal*, *14*(1), 56–67. doi:10.1075/idj.14.1.07bra

Bråten, I., & Strømsø, H. I. (2006). Epistemological beliefs, interest, and gender as predictors of Internet-based learning activities. *Computers in Human Behavior*, *22*(6), 1027–1042. doi:10.1016/j.chb.2004.03.026

Bråten, I., & Strømsø, H. I. (2009). Effects of Task Instruction and Personal Epistemology on the Understanding of Multiple Texts About Climate Change. *Discourse Processes*, *47*(1), 1–31. doi:10.1080/01638530902959646

Bråten, I., Strømsø, H. I., & Samuelstuen, M. S. (2005). THE RELATIONSHIP BETWEEN INTERNET-SPECIFIC EPISTEMOLOGICAL BELIEFS AND LEARNING WITHIN INTERNET TECHNOLOGIES. *Journal of Educational Computing Research*, *33*(2), 141–171. doi:10.2190/E763-X0LN-6NMF-CB86

Broder, A. (2002). A taxonomy of web search. *SIGIR Forum*, *36*(2), 3–10. doi:10.1145/792550.792552

Bromme, R., Pieschl, S., & Stahl, E. (2009). Epistemological beliefs are standards for adaptive learning: a functional theory about epistemological beliefs and metacognition. *Metacognition and Learning*, *5*(1), 7–26. doi:10.1007/s11409-009-9053-5

Buckingham Shum, S. (2008). Cohere: Towards web 2.0 argumentation. In P. Besnard, S. Doutre, & A. Hunter (Eds.), *Frontiers in Artificial Intelligence and Applications* (Vol. 172, p. 97). Toulouse, France: IOS Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=U5rWx0Kh4vMC&oi=fnd&pg=PA97&dq=%22%2B44+(0)+1908+655723%3B+Fax:+%2B44+(0)+1908+653196%3B%22+%22The+Web+2.0%

22+%22lot+is+being+written+about+the+Web+2.0+paradigm,+a+term+first+dubbed+in+2004%22+&ots=KsIXPgvtz5&sig=hQ0iDjMiv0DvpQF0vrihDJXs1q4
Buckingham Shum, S. (2012). Our Learning Analytics Are Our Pedagogy. Presented at the Expanding Horizons 2012 Conference, Macquarie University. Retrieved from http://www.slideshare.net/sbs/our-learning-analytics-are-our-pedagogy
Buckingham Shum, S., & Ferguson, R. (2012). Social Learning Analytics. *Educational Technology & Society*, *15*(3), 3–26.
Burghardt, M., Heckner, M., & Wolff, C. (2012). Chapter 2: The Many Ways of Searching the Web Together: A Comparison of Social Search Engines. In D. Lewandowski (Ed.), *Web search engine research*. Emerald Group Publishing.
Capra, R., Chen, A. T., Hawthorne, K., Arguello, J., Shaw, L., & Marchionini, G. (2012). Design and evaluation of a system to support collaborative search. *Proceedings of the American Society for Information Science and Technology*, *49*(1), 1–10.
Capra, R., Marchionini, G., Velasco-Martin, J., & Muller, K. (2010). Tools-at-hand and learning in multi-session, collaborative search (p. 951). ACM Press. doi:10.1145/1753326.1753468
Capra, R., Sams, B., & Seligson, P. (2011). Self-Generated Versus Imposed Tasks in Collaborative Search. In *Collaborative Information Seeking: Bridging the Gap Between Theory and Practice (CIS)*. New Orleans, Louisiana. Retrieved from http://collab.infoseeking.org/events/asist2011workshop
Chi, E. H., Pirolli, P., & Lam, S. K. (2007). Aspects of augmented social cognition: Social information foraging and social search. In *Proceedings of the 2nd international conference on Online communities and social computing* (pp. 60–69). Retrieved from http://dl.acm.org/citation.cfm?id=1784305
Chinn, C. A., & Brewer, W. F. (1993). The Role of Anomalous Data in Knowledge Acquisition: A Theoretical Framework and Implications for Science Instruction. *Review of Educational Research*, *63*(1), 1–49. doi:10.3102/00346543063001001
Chinn, C., A., Buckland, L., A., & Samarapungavan, A. (2011). Expanding the Dimensions of Epistemic Cognition: Arguments From Philosophy and Psychology. *Educational Psychologist*, *46*(3), 141–167. doi:10.1080/00461520.2011.587722
Cho, Y. H., Lee, J., & Jonassen, D. H. (2011). The role of tasks and epistemological beliefs in online peer questioning. *Computers & Education*, *56*(1), 112–126. doi:10.1016/j.compedu.2010.08.014
Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on Socially Shared Cognition*, *13*(1991), 127–149.
Clow, D. (2012). The learning analytics cycle: closing the loop effectively. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 134–138). Retrieved from http://dl.acm.org/citation.cfm?id=2330636
Cole, C. (2011). A theory of information need for information retrieval that connects information to knowledge. *Journal of the American Society for Information Science and Technology*, *62*(7), 1216–1231. doi:10.1002/asi.21541
Cole, C. (2012). *Information need*. American Society for Information Science and Technology by Information Today, Incorporated.
Craig, E. (1999). *Knowledge and the state of nature : an essay in conceptual synthesis*. Oxford: Clarendon Press.
Crescenzi, A., & Capra, R. (2013). Collaboration in Search Processes and Results Sharing. In *Collaborative Information Seeking Workshop*. San Antonio, Texas, USA. Retrieved from http://collab.infoseeking.org/resources/papers/cis2013/CrescenziPaper.pdf
Cunnane, S. (2011, May 12). The Danish gambit: online access, even during exams. *Times Higher Education*. UK. Retrieved from http://www.timeshighereducation.co.uk/story.asp?sectioncode=26&storycode=416090&c=1

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science*, *32*(5), 554–571.
Davis, A. (1998). 3: Understanding and Holism. *Journal of Philosophy of Education*, *32*(1), 41–55.
Davis, A. (1999). *The Limits of Educational Assessment*. Wiley.
Davis, A. (2005). Learning and the Social Nature of Mental Powers. *Educational Philosophy and Theory*, *37*(5), 635–647. doi:10.1111/j.1469-5812.2005.00148.x
Davis, A. (2006). High Stakes Testing and the Structure of the Mind: A Reply to Randall Curren. *Journal of Philosophy of Education*, *40*(1), 1–16. doi:10.1111/j.1467-9752.2006.00492.x
Davis, A., & Williams, K. (2002). Epistemology and curriculum. In N. Blake, P. Smeyers, & R. Smith (Eds.), *The Blackwell guide to the philosophy of education*. Blackwell Reference Online. Retrieved from http://books.google.co.uk/books?hl=en&lr=&id=J_9WaYKxEygC&oi=fnd&pg=PA253&ots=DTsa4HxNxP&sig=DFNFobZy4E4227Up2K_CgqvE3DM
De Liddo, A., & Alevizou, P. (2010). A method and tool to support the analysis and enhance the understanding of peer-to-peer learning experiences. Presented at the OpenEd2010: Seventh Annual Open Education Conference, Barcelona, Spain. Retrieved from http://oro.open.ac.uk/23392
De Liddo, A., Buckingham Shum, S., Quinto, I., Bachler, M., & Cannavacciuolo, L. (2011). Discourse-centric learning analytics. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 23–33). New York, NY, USA: ACM. doi:10.1145/2090116.2090120
DeBacker, T. K., Crowson, H. M., Beesley, A. D., Thoma, S. J., & Hestevold, N. L. (2008). The Challenge of Measuring Epistemic Beliefs: An Analysis of Three Self-Report Instruments. *The Journal of Experimental Education*, *76*(3), 281–312. doi:10.3200/JEXE.76.3.281-314
Dede, C. (2008). A Seismic Shift in Epistemology. *EDUCASE Review*, *43*(3), 80–81.
DePaul, M., & Zagzebski, L. (2003). *Intellectual Virtue*. Oxford University Press. Retrieved from http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199252732.001.0001/acprof-9780199252732
Dewey, J. (1916). *Democracy and education: an introduction to the philosophy of education*. Macmillan.
Dewey, J. (1938). The theory of inquiry. *New York: Holt, Rinehart & Wiston*. Retrieved from http://www.secondroad.net/foundations/practical/01DeweyLogicOfInquiry.pdf
Dewey, J. (1998). *Experience and education*. Kappa Delta Pi. Retrieved from http://ruby.fgcu.edu/courses/ndemers/colloquium/ExperiencEducationDewey.pdf
DiMaggio, P., Hargittai, E., Celeste, C., & Shafer, S. (2001). From Unequal Access to Differentiated Use: A Literature Review and agenda for Research on Digital Inequality. *Princeton: Center for Arts and Cultural Policy Studies, Woodrow Wilson School, Princeton University*. Retrieved from http://www.eszter.com/research/pubs/dimaggio-etal-digitalinequality.pdf
Dimopoulos, K., & Asimakopoulos, A. (2010). Science on the Web: Secondary School Students' Navigation Patterns and Preferred Pages' Characteristics. *Journal of Science Education and Technology*, *19*(3), 246–265. doi:10.1007/s10956-009-9197-8
Dinet, J., Chevalier, A., & Tricot, A. (2012). Information search activity: An overview. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, *62*(2), 49–62. doi:10.1016/j.erap.2012.03.004
Duell, O. K., & Schommer-Aikins, M. (2001). Measures of people's beliefs about knowledge and learning. *Educational Psychology Review*, *13*(4), 419–449. doi:10.1023/A:1011969931594
Dwyer, N., & Suthers, D. D. (2006). Consistent practices in artifact-mediated collaboration. *International Journal of Computer-Supported Collaborative Learning*, *1*(4), 481–511. doi:10.1007/s11412-006-9001-1
Edwards, A. D., & Furlong, V. J. (1978). *The language of teaching: Meaning in classroom interaction*. Heinemann London. Retrieved from http://www.getcited.org/pub/102019252

Edwards, D. (1993). But what do children really think? Discourse analysis and conceptual content in children's talk. *Cognition and Instruction*, *11*(3-4), 207–225.
Edwards, D. (2005). Discursive Psychology. In K. L. Fitch & R. E. Sanders (Eds.), *Handbook of language and social interaction* (pp. 257–273). Erlbaum.
Edwards, D., & Mercer, N. (1987). *Common knowledge: the development of understanding in the classroom*. London, UK: Routledge.
Elias, T. (2011). Learning analytics: Definitions, processes and potential. Retrieved from http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf
Ellis, D., Wilson, T. D., Ford, N., Foster, A., Lam, H. M., Burton, R., & Spink, A. (2002). Information seeking and mediated searching. Part 5. User–intermediary interaction. *Journal of the American Society for Information Science and Technology*, *53*(11), 883–893. doi:10.1002/asi.10133
Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215. doi:10.1037/0033-295X.87.3.215
Eryilmaz, E., Pol, J. van der, Ryan, T., Clark, P. M., & Mary, J. (2013). Enhancing student knowledge acquisition from online learning conversations. *International Journal of Computer-Supported Collaborative Learning*, *8*(1), 113–144. doi:10.1007/s11412-012-9163-y
Evans, B. M., & Chi, E. H. (2010). An elaborated model of social search. *Information Processing & Management*, *46*(6), 656–678. doi:10.1016/j.ipm.2009.10.012
Evans, B. M., Kairam, S., & Pirolli, P. (2009). Do your friends make you smarter?: An analysis of social strategies in online information seeking. *Information Processing & Management*, *46*(6), 679–692. doi:10.1016/j.ipm.2009.12.001
Ferguson, L. E., & Bråten, I. (2013). Student profiles of knowledge and epistemic beliefs: Changes and relations to multiple-text comprehension. *Learning and Instruction*, *25*, 49–61. doi:10.1016/j.learninstruc.2012.11.003
Ferguson, L. E., Bråten, I., & Strømsø, H. I. (2012). Epistemic cognition when students read multiple documents containing conflicting scientific evidence: A think-aloud study. *Learning and Instruction*, *22*(2), 103–120. doi:10.1016/j.learninstruc.2011.08.002
Ferguson, R. (2012). *The State of Learning Analytics in 2012: A Review and Future Challenges* (Technical Report No. kmi-12-01). Knowledge Media Institute: The Open University, UK. Retrieved from http://kmi.open.ac.uk/publications/pdf/kmi-12-01.pdf
Ferguson, R., & Buckingham Shum, S. (2012). Social Learning Analytics: Five Approaches. Presented at the Proc. 2nd International Conference on Learning Analytics & Knowledge, Vancouver, BC: ACM Press: New York. Retrieved from http://projects.kmi.open.ac.uk/hyperdiscourse/docs/LAK2012-RF-SBS.pdf
Ferguson, R., Buckingham Shum, S., & Deakin Crick, R. (2011). EnquiryBlogger: Using widgets to support awareness and reflection in a PLE setting. In *1st Workshop on Awareness and Reflection in Personal Learning Environments*. Southampton, UK. Retrieved from http://oro.open.ac.uk/19576/1/Producing_Animations_to_educate_MSM.pdf
Ferguson, R., Whitelock, D., & Littleton, K. (2010). Improvable objects and attached dialogue: new literacy practices employed by learners to build knowledge together in asynchronous settings. *Digital Culture & Education*, *2*(1), 103–123.
Flanagan, B., Suzuki, T., Zeng, J., Yin, C., Sakai, T., Hashimoto, K., & Hirokawa, S. (2013). PEER KNOWLEDGE ASSISTED SEARCH USING COMMUNITY SEARCH LOGS. *International Journal of Digital Information and Wireless Communications (IJDIWC)*, *3*(2), 1–8.
Florida State. Florida Education Omnibus Bill. , Pub. L. No. H.B. 7087e3 (2006).
Foster, J. (2009). Understanding interaction in information seeking and use as a discourse: a dialogic approach. *Journal of Documentation*, *65*(1), 83–105. doi:10.1108/00220410910926130
Fricker, M. (2009). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, USA.

Fricker, M. (2012). Group Testimony? The Making of A Collective Good Informant. *Philosophy and Phenomenological Research*, *84*(2), 249–276. doi:10.1111/j.1933-1592.2011.00565.x
Furberg, A. (2009). Socio-cultural aspects of prompting student reflection in Web-based inquiry learning environments. *Journal of Computer Assisted Learning*, *25*(4), 397–409. doi:10.1111/j.1365-2729.2009.00320.x
Furberg, A., & Ludvigsen, S. (2008). Students' Meaning-making of Socio-scientific Issues in Computer Mediated Settings: Exploring learning through interaction trajectories. *International Journal of Science Education*, *30*(13), 1775–1799. doi:10.1080/09500690701543617
FXPAL. (2011). FXPAL - Collaborative Exploratory Search. Retrieved October 31, 2011, from http://www.fxpal.com/?p=CES
Gagnière, L., Betrancourt, M., & Détienne, F. (2012). When metacognitive prompts help information search in collaborative setting. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, *62*(2), 73–81. doi:10.1016/j.erap.2011.12.005
Gardner, J. (2011). *Assessment and learning* (2nd ed.). SAGE.
Ge, X., & Land, S. M. (2003). Scaffolding students' problem-solving processes in an ill-structured task using question prompts and peer interactions. *Educational Technology Research and Development*, *51*(1), 21–38.
Ge, X., & Land, S. M. (2004). A conceptual framework for scaffolding III-structured problem-solving processes using question prompts and peer interactions. *Educational Technology Research and Development*, *52*(2), 5–22.
Gee, J. P. (1989). Two styles of narrative construction and their linguistic and educational implications. *Discourse Processes*, *12*(3), 287–307.
Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010). Understanding and Integrating Multiple Science Texts: Summary Tasks are Sometimes Better Than Argument Tasks. *Reading Psychology*, *31*(1), 30–68. doi:10.1080/02702710902733600
Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. (2012). Comprehending and Learning From Internet Sources: Processing Patterns of Better and Poorer Learners. *Reading Research Quarterly*, *47*(4), 356–381. doi:10.1002/RRQ.027
Golovchinsky, G., Adcock, J., Pickens, J., Qvarfordt, P., & Back, M. (2008). Cerchiamo: a collaborative exploratory search tool. *Proceedings of Computer Supported Cooperative Work (CSCW)*. Retrieved from http://fxpal.fxpal.com/publications/FXPAL-PR-08-470.pdf
Golovchinsky, G., & Diriye, A. (2011). Session-based search with Querium. In *Proceedings of the HCIR 2011 Workshop, Mountain View, CA, USA*. Retrieved from http://fxpal.com/publications/FXPAL-PR-11-660.pdf
Golovchinsky, G., Pickens, J., & Back, M. (2009). A Taxonomy of Collaboration in Online Information Seeking. In *arXiv:0908.0704*. Retrieved from http://arxiv.org/abs/0908.0704
González-Ibáñez, R., Haseki, M., & Shah, C. (In press). Let's search together, but not too close! An analysis of communication and performance in collaborative information seeking. *Information Processing & Management*. doi:10.1016/j.ipm.2012.12.008
Greene, J. A., Azevedo, R. A., & Torney-Purta, J. (2008). Modeling Epistemic and Ontological Cognition: Philosophical Perspectives and Methodological Directions. *Educational Psychologist*, *43*(3), 142–160.
Greene, J. A., Muis, K. R., & Pieschl, S. (2010). The Role of Epistemic Beliefs in Students' Self-Regulated Learning With Computer-Based Learning Environments: Conceptual and Methodological Issues. *Educational Psychologist*, *45*(4), 245–257. doi:10.1080/00461520.2010.515932
Grice, H. P. (1975). Logic and conversation. In H. Geirsson & M. Losonsky (Eds.), *Readings in Language and Mind* (pp. 41–58). Oxford, UK: Blackwell.
Gubbels, M., Rose, A., Russell, D., & Bederson, B. B. (2012). *SearchParty: Real-time Support for Social Learning in Synchronous Environments* (No. HCIL-2012-21). Human-Computer Interaction

Lab, University of Maryland. Retrieved from http://hcil2.cs.umd.edu/trs/2012-21/2012-21.pdf
Hacker, K. L., & Mason, S. M. (2003). Ethical gaps in studies of the digital divide. *Ethics and Information Technology*, *5*(2), 99–115. doi:10.1023/A:1024968602974
Hagen, Å. M., Braasch, J. L. G., & Bråten, I. (2012). Relationships between spontaneous note-taking, self-reported strategies and comprehension when reading multiple texts in different task conditions. *Journal of Research in Reading*, n/a–n/a. doi:10.1111/j.1467-9817.2012.01536.x
Halliday, M. A. K., Hasan, R., & Christie, F. (1989). *Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective*. Oxford University Press, USA.
Hamilton, E., & Feenberg, A. (2005). The Technical Codes of Online Education. *Techné: Research in Philosophy and Technology*, *9*(1).
Hammer, D., & Elby, A. (2003). Tapping Epistemological Resources for Learning Physics. *Journal of the Learning Sciences*, *12*(1), 53–90. doi:10.1207/S15327809JLS1201_3
Hansen, P., & Järvelin, K. (2005). Collaborative Information Retrieval in an information-intensive domain. *Information Processing & Management*, *41*(5), 1101–1119. doi:10.1016/j.ipm.2004.04.016
Hargittai, E. (2002). Second-Level Digital Divide: Differences in People's Online Skills. *First Monday*, *7*(4). Retrieved from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/942/864
Hargittai, E. (2008). The Digital Reproduction of Inequality. In D. Grusky (Ed.), *Social Stratification* (pp. 936–944). Boulder, CO: Westview Press. Retrieved from http://webuse.org/pdf/Hargittai-DigitalReproduction2008.pdf
Hargittai, E. (2010). Digital Na(t)ives? Variation in Internet Skills and Uses among Members of the "Net Generation." *Sociological Inquiry*, *80*, 92–113. doi:10.1111/j.1475-682X.2009.00317.x
Harlen, W. (2007). *Assessment of learning*. London, UK: SAGE.
Harris, P. L. (2012). *Trusting What You're Told*. Harvard University Press.
Hearst, M. A. (2009a). Chapter 3: Models of the Information Seeking Process. In *Search User Interfaces* (1st ed.). Cambridge University Press.
Hearst, M. A. (2009b). *Search User Interfaces* (1st ed.). Cambridge University Press.
Hertzum, M. (2008). Collaborative information seeking: The combined activity of information seeking and collaborative grounding. *Information Processing & Management*, *44*(2), 957–962. doi:10.1016/j.ipm.2007.03.007
Hjørland, B. (1997). *Information seeking and subject representation: an activity-theoretical approach to information science*. Greenwood Press.
Hobman, E. V., Bordia, P., Irmer, B., & Chang, A. (2002). The expression of conflict in computer-mediated and face-to-face groups. *Small Group Research*, *33*(4), 439–465.
Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review*, *13*(4), 353–383.
Hofer, B. K. (2004a). Epistemological Understanding as a Metacognitive Process: Thinking Aloud During Online Searching. *Educational Psychologist*, *39*(1), 43–55. doi:10.1207/s15326985ep3901_5
Hofer, B. K. (2004b). Introduction: Paradigmatic approaches to personal epistemology. *Educational Psychologist*, *39*(1), 1–3.
Hofer, B., K. .., & Pintrich, P., R. .. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, *67*(1), 88. doi:10.2307/1170620
Hofer, B. K., & Pintrich, P. R. (2002). *Personal Epistemology: The Psychology of Beliefs About Knowledge and Knowing*. Routledge. Retrieved from http://books.google.co.uk/books?hl=en&lr=&id=JHcW74qBDLEC&oi=fnd&pg=PP1&dq=%22personal+epistemology:+the+psychology+of+beliefs+about+knowledge+and+knowing%22&ots=eAnf1b11Re&sig=0kYT0TMYmziVjwHfUnjfemzWFe0

Hong, H.-Y., & Lin, S.-P. (2010). Teacher-Education Students' Epistemological Belief Change through Collaborative Knowledge Building. *The Asia-Pacific Education Researcher*, *19*(1). Retrieved from http://www.ejournals.ph/index.php?journal=TAPER&page=article&op=view&path%5B%5D=129

Hopmann, S. T., Brinek, G., & Retzl, M. (Eds.). (2007). *PISA According to PISA: Does PISA Keep What It Promises?* Vienna, Austria: Wien Lit-Verlag. Retrieved from http://www.univie.ac.at/pisaaccordingtopisa/

Hutchison, P., & Hammer, D. (2010). Attending to student epistemological framing in a science classroom. *Science Education*, *94*(3), 506–524. doi:10.1002/sce.20373

Hwang, G.-J., Tsai, P.-S., Tsai, C.-C., & Tseng, J. C. R. (2008). A novel approach for assisting teachers in analyzing student web-searching behaviors. *Computers & Education*, *51*(2), 926–938. doi:10.1016/j.compedu.2007.09.011

Hyldegård, J. (2006). Collaborative information behaviour—exploring Kuhlthau's Information Search Process model in a group-based educational setting. *Information Processing & Management*, *42*(1), 276–298. doi:10.1016/j.ipm.2004.06.013

Hyldegård, J. (2009). Beyond the search process – Exploring group members' information behavior in context. *Information Processing & Management*, *45*(1), 142–158. doi:10.1016/j.ipm.2008.05.007

Janssen, J., & Bodemer, D. (2013). Coordinated computer-supported collaborative learning: Awareness and awareness tools. *Educational Psychologist*, *48*(1), 40–55.

Jelfs, H., Buckingham Shum, S., & De L. (2011). Mapping the Deeper Learning Literature using Cohere. Presented at the OLnet Project Seminar, Institute of Educational Technology, The Open University, UK. Retrieved from http://olnet.org/node/610

Johnson, J. D., & Meischke, H. (2006). A Comprehensive Model of Cancer-Related Information Seeking Applied to Magazines. *Human Communication Research*, *19*(3), 343–367. doi:10.1111/j.1468-2958.1993.tb00305.x

Johnson, M. D., Reimann, P., Bull, S., & Fujita, N. (2011). Epistemic Beliefs and Open Learner Models. In *Proceedings of the International Workshop on Personalization Approaches in Learning Environments* (Vol. 732, pp. 19–24). Retrieved from http://adenu.ia.uned.es/workshops/pale2011/PALE2011%20proceedings.pdf#page=25

Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2013). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during Web search on a medical issue. *Computers in Human Behavior*, *29*(3), 1193–1203. doi:10.1016/j.chb.2012.10.012

Karunakaran, A., Spence, P. R., & Reddy, M. C. (2010). Towards a Model of Collaborative Information Behavior. In *Workshop on Collaborative Information Retrieval*. Savannah, GA. Retrieved from http://www.google.co.uk/url?sa=t&rct=j&q=+Towards+a+Model+of+Collaborative+Information+Behavior&source=web&cd=1&ved=0CCgQFjAA&url=http%3A%2F%2Fworkshops.fxpal.com%2Fcscw2010cis%2Fsubmissions%2Ftmp10.pdf&ei=t8s7T-bWC4rH0QXizdRs&usg=AFQjCNFHjJUptg9R3LMdZTMmY6goknb85A&sig2=C9oJgmkmNdL-9ussTxVcyQ&cad=rja

Katz, S. (2000). Competency, epistemology and pedagogy: curriculum' 's holy trinity. *Curriculum Journal*, *11*(2), 133–144. doi:10.1080/09585170050045164

Kelly, G. J., Luke, A., & Green, J. (2008). What Counts as Knowledge in Educational Settings: Disciplinary Knowledge, Assessment, and Curriculum. *Review of Research in Education*, *32*(1), vii–x. doi:10.3102/0091732X07311063

King, A., Staffieri, A., & Adelgais, A. (1998). Mutual peer tutoring: Effects of structuring tutorial interaction to scaffold peer learning. *Journal of Educational Psychology*, *90*(1), 134.

King, P. M., & Kitchener, K. S. (2004). Reflective Judgment: Theory and Research on the Development of Epistemic Assumptions Through Adulthood. *Educational Psychologist*, *39*(1), 5–18. doi:10.1207/s15326985ep3901_2

Knight, S. (2011a). Why we should study...technology and situated cognition. *PsyPAG Quarterly*, (78), 23–26.

Knight, S. (2011b, August). *Knowing my extensions: THE IMPLICATIONS OF EXTENDED MIND FOR OUR CONCEPTION OF KNOWLEDGE AND ITS ASSESSMENT – DO I "KNOW" MY EXTENSIONS?* (Master's). Institute of Education, University of London, London. Retrieved from https://docs.google.com/file/d/0Bx_nJE-nmSkmNVpBTGprcWF4Tzg/edit

Knight, S. (2013a). Creating a supportive environment for classroom dialogue. In S. Hennessy, P. Warwick, L. Brown, D. Rawlins, & C. Neale (Eds.), *Developing interactive teaching and learning using the IWB*. Open University Press.

Knight, S. (2013b). *Mediawiki for OER and Learning Analytics*. Presented at the Wikimedia UK conference and AGM, Lincoln, UK. Retrieved from http://www.slideshare.net/sjgknight/media-wiki-for-oer-and-learning-analytics-understanding-learning-resources-and-learners/

Knight, S. A., & Spink, A. (2004). Toward a Web Search Information Behavior Model. In A. Spink & B. J. Jansen (Eds.), *Web search: public searching on the Web*. Springer.

Knight, S., Arastoopour, G., Williamson Shaffer, D., Buckingham Shum, S., & Littleton, K. (2013). Epistemic Networks for Epistemic Commitments. *Working Paper*. Retrieved from http://kmi.open.ac.uk/publications/pdf/kmi-13-03.pdf

Knight, S., Buckingham Shum, S., & Littleton, K. (2013a). Epistemology, Pedagogy, Assessment and Learning Analytics. Presented at the LAK 2013: 3rd International Conference on Learning Analytics & Knowledge, Leuven, Belgium: ACM Press. doi:10.1145/2460296.2460312

Knight, S., Buckingham Shum, S., & Littleton, K. (2013b). Tracking Epistemic Beliefs and Sensemaking in Collaborative Information Retrieval. In *Collaborative Information Seeking Workshop*. San Antonio, Texas, USA. Retrieved from http://collab.infoseeking.org/events/cscw2013workshop

Knight, S., & Littleton, K. (in draft). The importance of context in machine learning for tracking productive educational dialogue. *Working Paper*.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in Human Behavior*, *19*(3), 335–353.

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, *42*(5), 361–371. doi:10.1002/(SICI)1097-4571(199106)42:5<361::AID-ASI6>3.0.CO;2-#

Kuhn, D., & Weinstock, M. (2002). What is epistemological thinking and why does it matter?. In B. K. Hofer & P. Pintrich R. .. (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 121–144). Mahwah, NJ: Lawrence Erlbaum Associates.

Kuiper, E., Volman, M., & Terwel, J. (2009). Developing Web literacy in collaborative inquiry activities. *Computers & Education*, *52*(3), 668–680. doi:10.1016/j.compedu.2008.11.010

LAK11. (2011). 1st International Conference on Learning Analytics and Knowledge 2011 | Connecting the technical, pedagogical, and social dimensions of learning analytics. *LAK11*. Retrieved May 21, 2013, from https://tekri.athabascau.ca/analytics/about

LAK12. (2012). 2nd International Conference on Learning Analytics and Knowledge 2012 | Dialectics in Learning Analytics. Retrieved May 21, 2013, from http://lak12.sites.olt.ubc.ca/cfp/

LAK13. (2013). 3rd International Conference on Learning Analytics and Knowledge 2013 | Dialectics in Learning Analytics. Retrieved May 21, 2013, from http://lakconference2013.wordpress.com/

Large, A., Beheshti, J., & Rahman, T. (2002). Gender differences in collaborative Web searching behavior: an elementary school study. *Information Processing & Management*, *38*(3), 427–443. doi:10.1016/S0306-4573(01)00034-6

Lave, J., & Wenger, E. (1991). *Situated learning: legitimate peripheral participation*. Cambridge University Press.

Lazonder, A. W. (2005). Do two heads search better than one? Effects of student collaboration on web search behaviour and search outcomes. *British Journal of Educational Technology*, *36*(3), 465–475. doi:10.1111/j.1467-8535.2005.00478.x

Lewandowski, D. (Ed.). (2012). *Web search engine research*. Emerald Group Publishing.

Liechti, O., & Sumi, Y. (2002). Editorial: Awareness and the WWW. *International Journal of Human-Computer Studies*, *56*(1), 1–5.

Lin, C., & Tsai, C. (2008). Exploring the Structural Relationships between High School Students' Scientific Epistemological Views and their Utilization of Information Commitments toward Online Science Information. *International Journal of Science Education*, *30*(15), 2001–2022. doi:10.1080/09500690701613733

Lin, C.-C., & Tsai, C.-C. (2007). A Navigation Flow Map Method of Representing Students' Searching Behaviors and Strategies on the Web, with Relation to Searching Outcomes. *CyberPsychology & Behavior*, *10*(5), 689–695. doi:10.1089/cpb.2007.9969

Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and Instruction*, *13*(5), 487–509.

Littleton, K., & Howe, C. (2010). *Educational dialogues: understanding and promoting productive interaction*. Abingdon, Oxon: Routledge.

Littleton, K., & Mercer, N. (2013). *Interthinking: putting talk to work*. London: Routledge.

Livingstone, S., Bober, M., & Helsper, E. (2005). Internet Literacy among children and young people: findings from the UK Children Go Online project. London: LSE Research Online. Retrieved from http://eprints.lse.ac.uk/397/1/UKCGOonlineLiteracy.pdf

Livingstone, S., & Helsper, E. (2007). Gradations in digital inclusion: Children, young people and the digital divide. *New Media & Society*, *9*(4), 671. doi:10.1177/1461444807080335

Lloyd, A. (2007). Recasting information literacy as sociocultural practice: implications for library and information science researchers. *Information Research*, *12*(4), 12–4.

Maggioni, L., & Fox, E. (2009). Adolescents' reading of multiple history texts: An interdisciplinary investigation of historical thinking, intertextual reading, and domain-specific epistemic beliefs. In *annual meeting of the American Educational Research Association, San Diego, CA*.

Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, *49*(4), 41. doi:10.1145/1121949.1121979

Marchionini, G., & White, R. (2007). Find What You Need, Understand What You Find. *International Journal of Human-Computer Interaction*, *23*, 205–237. doi:10.1080/10447310701702352

Mason, L., Ariasi, N., & Boldrin, A. (2011). Epistemic beliefs in action: Spontaneous reflections about knowledge and knowing during online information searching and their influence on learning. *Learning and Instruction*, *21*(1), 137–151. doi:10.1016/j.learninstruc.2010.01.001

Mason, L., & Boldrin, A. (2008). Epistemic Metacognition in the Context of Information Searching on the Web. In M. S. Khine (Ed.), *Knowing, Knowledge and Beliefs* (pp. 377–404). Dordrecht: Springer Netherlands. Retrieved from http://www.springerlink.com/content/p8110v34340v3748/

Mason, L., Boldrin, A., & Ariasi, N. (2009). Epistemic metacognition in context: evaluating and learning online information. *Metacognition and Learning*, *5*(1), 67–90. doi:10.1007/s11409-009-9048-2

Mason, L., Boldrin, A., & Ariasi, N. (2010). Searching the Web to learn about a controversial topic: are students epistemically active? *Instructional Science*, *38*(6), 607–633. doi:10.1007/s11251-008-9089-y

McDowell, J. (1996). *Mind and World* (New edition.). Harvard University Press.

Medina, R., & Suthers, D. D. (2009). Using a contingency graph to discover representational practices in an online collaborative environment. *Research and Practice in Technology Enhanced Learning*, *4*(03), 281–305.

Mercer, N. (2000). *Words & Minds: How we use language to think together*. Oxon: Routledge.

Mercer, N. (2004). Sociocultural discourse analysis: analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, *1*(2), 137–168. doi:10.1558/japl.v1i2.137

Mercer, N. (2008). The seeds of time: why classroom dialogue needs a temporal analysis. *Journal of the Learning Sciences*, *17*(1), 33–59. doi:10.1080/10508400701793182

Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: ways of helping children to use language to learn science. *British Educational Research Journal*, *30*(3), 359–377.

Mercer, N., & Littleton, K. (2007). *Dialogue and the Development of Children's Thinking: A Sociocultural Approach* (New edition.). Routledge.

Mercer, N., & Sams, C. (2006). Teaching children how to use language to solve maths problems. *Language and Education*, *20*(6), 507–528.

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's Talk and the Development of Reasoning in the Classroom. *British Educational Research Journal*, *25*(1), 95–111. doi:10.1080/0141192990250107

Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education*, *27*(4), 283–297.

Michaels, S., O'Connor, M. C., Hall, M. W., & Resnick, L. (2002). Accountable talk: classroom conversation that works. *Pittsburg: University of Pittsburgh*.

Moraveji, N., Morris, M. R., Morris, D., Czerwinski, M., & Henry Riche, N. (2011). ClassSearch: facilitating the development of web search skills through social learning. In *Proceedings of the 2011 annual conference on Human factors in computing systems* (pp. 1797–1806).

Morris, M. R. (2008). A survey of collaborative web search practices (p. 1657). ACM Press. doi:10.1145/1357054.1357312

Morris, M. R. (2013). Collaborative Search Revisted. Presented at the CSCW, San Antonio, Texas, USA. Retrieved from http://research.microsoft.com/en-us/um/people/merrie/papers/morris_cscw2013_collab_search_revisited.pdf

Morris, M. R., & Horvitz, E. (2007). SearchTogether: an interface for collaborative web search. In *Proceedings of the 20th annual ACM symposium on User interface software and technology* (pp. 3–12). Retrieved from http://dl.acm.org/citation.cfm?id=1294215

Muis, K. R. (2007). The Role of Epistemic Beliefs in Self-Regulated Learning. *Educational Psychologist*, *42*(3), 173–190. doi:10.1080/00461520701416306

Muis, K. R., Bendixen, L. D., & Haerle, F. C. (2006). Domain-Generality and Domain-Specificity in Personal Epistemology Research: Philosophical and Empirical Reflections in the Development of a Theoretical Framework. *Educational Psychology Review*, *18*(1), 3–54. doi:10.1007/s10648-006-9003-6

Muis, K. R., & Franco, G. M. (2009). Epistemic beliefs: Setting the standards for self-regulated learning. *Contemporary Educational Psychology*, *34*(4), 306–318. doi:10.1016/j.cedpsych.2009.06.005

Munneke, L., Andriessen, J., Kanselaar, G., & Kirschner, P. (2007). Supporting interactive argumentation: Influence of representational tools on discussing a wicked problem. *Computers in Human Behavior*, *23*(3), 1072–1088.

Nielsen, J., Clemmensen, T., & Yssing, C. (2002). Getting access to what goes on in people's heads?: reflections on the think-aloud technique. In *Proceedings of the second Nordic conference on Human-computer interaction* (pp. 101–110). New York, NY, USA: ACM. doi:10.1145/572020.572033

Olson, D. R., & Bruner, J. S. (1996). Folk psychology and folk pedagogy. *The Handbook of Education and Human Development*, 9–27.

Österholm, M. (2009). Theories of epistemological beliefs and communication: A unifying attempt. In *Proceedings of the 33rd Conference of the International Group for the Psychology of*

Mathematics Education (Vol. 4, pp. 275–264). Retrieved from http://umu.diva-portal.org/smash/record.jsf?pid=diva2:228322

Österholm, M. (2010). Relationships Between Epistemological Beliefs and Properties of Discourse: Some Empirical Explorations. In C. Bergsten, E. Jablonka, & T. Wedege (Eds.), *Madif 7, the 7th Swedish Mathematics Education Research Seminar* (pp. 241–250). Stockholm, Sweden: Linköping, Sweden: SMDF. Retrieved from http://osterholm.pcriot.com/publ-madif7a.shtml

Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You. by Eli Pariser*. Viking.

Paul, S., A. .., & Morris, M. R. (2009). CoSense: enhancing sensemaking for collaborative web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1771–1780). New York, NY, USA: ACM. doi:10.1145/1518701.1518974

Paul, S., A. .., & Morris, M. R. (2011). Sensemaking in Collaborative Web Search. *Human–Computer Interaction*, *26*(1-2), 72–122. doi:10.1080/07370024.2011.559410

Paul, S., A. .., & Reddy, M. C. (2010). Understanding together: sensemaking in collaborative information seeking. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 321–330). New York, NY, USA: ACM. doi:10.1145/1718918.1718976

Peregrin, J. (2005). The nature of meaning: Brandom versus Chomsky. *Pragmatics & Cognition*, *13*(1), 39–57. doi:10.1075/pc.13.1.05per

Phillips, D. C. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher*, 5–12.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–189.

Pieschl, S., Stahl, E., & Bromme, R. (2013). Adaptation to Context as Core Component of Self-Regulated Learning: The Example of Complexity and Epistemic Beliefs. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 53–65). Springer New York. Retrieved from http://link.springer.com.libezproxy.open.ac.uk/chapter/10.1007/978-1-4419-5546-3_4

Porsch, T., & Bromme, R. (2010). Effects of epistemological sensitization on source choices. *Instructional Science*, *39*(6), 805–819. doi:10.1007/s11251-010-9155-0

Potter, J., & Edwards, D. (2003). Sociolinguistics, cognitivism and discursive psychology. *IJES, International Journal of English Studies*, *3*(1), 93–110.

Powell, S., & MacNeil, S. (2012). *Instituitional Readiness for Analytics* (A Briefing Paper No. Vol 1., No. 8). JISC CETIS. Retrieved from http://publications.cetis.ac.uk/wp-content/uploads/2012/12/Institutional-Readiness-for-Analytics-Vol1-No8.pdf

Puustinen, M., & Rouet, J.-F. (2009). Learning with new technologies: Help seeking and information searching revisited. *Computers & Education*, *53*(4), 1014–1019. doi:10.1016/j.compedu.2008.07.002

Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, *4*(3), 239–257. doi:10.1007/s11412-009-9070-z

Resnick, L. B. (2001). Making America smarter: The real goal of school reform. *Developing Minds: A Resource Book for Teaching Thinking*, 3–6.

Reznitskaya, A., & Gregory, M. (2013). Student Thought and Classroom Language: Examining the Mechanisms of Change in Dialogic Teaching. *Educational Psychologist*, *48*(2), 114–133. doi:10.1080/00461520.2013.775898

Rick, J., & Guzdial, M. (2006). Situating CoWeb: a scholarship of application. *International Journal of Computer-Supported Collaborative Learning*, *1*(1), 89–115. doi:10.1007/s11412-006-6842-6

Robson, A., & Robinson, L. (2013). Building on models of information behaviour: linking information seeking and communication. *Journal of Documentation*, *69*(2), 169–193. doi:10.1108/00220411311300039

Rojas-Drummond, S., Littleton, K., Hernández, F., & Zúñiga, M. (2010). Dialogical interactions among peers in collaborative writing contexts. In K. Littleton & C. Howe (Eds.), *Educational*

dialogues: Understanding and promoting productive interaction (pp. 128–148). Abingdon, Oxon: Routledge.
Rorty, R. (1981). *PHILOSOPHY AND THE MIRROR OF NATURE*. Princeton University Press.
Rosé, C. P., Dyke, G., Law, N., Lund, K., Suthers, D. D., & Teplovs, C. (2011). Leveraging Researcher Multivocality for Insights on Collaborative Learning. Retrieved from http://hal.archives-ouvertes.fr/hal-00722952/
Rosenberg, S., Hammer, D., & Phelan, J. (2006). Multiple Epistemological Coherences in an Eighth-Grade Discussion of the Rock Cycle. *Journal of the Learning Sciences*, *15*(2), 261–292. doi:10.1207/s15327809jls1502_4
Russell-Rose, T. (2013). *Designing the Search Experience: The Information Architecture of Discovery*. Morgan Kaufmann.
Säljö, R. (1999). Learning as the use of tools: A sociocultural perspective on the human-technology link. In K. Littleton & P. Light (Eds.), *Learning With Computers: Analysing Productive Interaction*. Psychology Press.
Säljö, R. (2012). Literacy, Digital Literacy and Epistemic Practices: The Co-Evolution of Hybrid Minds and External Memory Systems. *Nordic Journal of Digital Literacy*, (01), 5–19.
Salmerón, L., & Kammerer, Y. (2012). Selecting Pages from Google to Learn about a Controversial Topic: The Role of Epistemic Beliefs. *STAGING KNOWLEDGE AND EXPERIENCE: HOW TO TAKE ADVANTAGE OF REPRESENTATIONAL TECHNOLOGIES IN EDUCATION AND TRAINING?*, 181.
Salomon, G. (1996). *Distributed Cognitions: Psychological and Educational Considerations* (1st Pbk. Ed.). Cambridge University Press.
Sandoval, W. (2012). Situating epistemological development. Presented at the International Conference of the Learning Sciences (ICLS), Sydney. Retrieved from http://www.academia.edu/2385109/Situating_epistemological_development
Savicki, V., Kelley, M., & Ammon, B. (2002). Effects of training on computer-mediated communication in single or mixed gender small task groups. *Computers in Human Behavior*, *18*(3), 257–269.
Savolainen, R. (2012). Conceptualizing information need in context. Retrieved from http://informationr.net/ir/17-4/paper534.html
Scardamalia, M., & Bereiter, C. (2006). Knowledge Building: Theory, Pedagogy, and Technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97–118). New York: Cambridge University Press.
Scherr, R. E., & Hammer, D. (2009). Student Behavior and Epistemological Framing: Examples from Collaborative Active-Learning Activities in Physics. *Cognition and Instruction*, *27*(2), 147–174. doi:10.1080/07370000902797379
Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning*, *5*(1), 43–102.
Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, *82*(3), 498–504. doi:10.1037/0022-0663.82.3.498
Schraw, G. (2000). Assessing metacognition: implications of the Buros symposium. *Issues in the Measurement of Metacognition*, 297–321.
Schraw, G. (2013). Conceptual Integration and Measurement of Epistemological and Ontological Beliefs in Educational Research. *ISRN Education*. Retrieved from http://www.hindawi.com/isrn/education/aip/327680/
Schraw, G., & Impara, J. C. (Eds.). (2000). *Issues in the measurement of metacognition*. Buros Institute of Mental Measurements and Department of Educational Psychology, University of Nebraska-Lincoln.
Schreiber, J. B., & Shinn, D. (2003). Epistemological beliefs of community college students and their learning processes. *Community College Journal of Research &Practice*, *27*(8), 699–709.

Shaffer, D. W. (2006). Epistemic frames for epistemic games. *Computers & Education*, *46*(3), 223–234. doi:10.1016/j.compedu.2005.11.003

Shaffer, D. W. (2008). *How Computer Games Help Children Learn*. Palgrave Macmillan.

Shaffer, D. W., & Graesser, A. (2010). Using a quantitative model of participation in a community of practice to direct automated mentoring in an ill-formed domain. In *Proceedings of the 4th International Workshop on Intelligent Tutoring Systems and Ill-Defined Domains held at the 10th International Conference on Intelligent Tutoring Systems (ITS '10)* (pp. 61–68). Retrieved from http://edgaps.org/gaps/wp-content/uploads/ITS-workshop-shaffer-graesser-041510.pdf

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., … Mislevy, R. (2009). Epistemic network analysis: A prototype for 21st-century assessment of learning. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/ijlm.2009.0013

Shah, C. (2010). Coagmento-a collaborative information seeking, synthesis and sense-making framework. *Integrated Demo at CSCW*, 6–11.

Shah, C. (2012). *Information Retrieval: Collaborative Information Seeking: the Art and Science of Making the Whole Greater Than the Sum of All*. Springer.

Shah, C. (2013). Effects of awareness on coordination in collaborative information seeking. *Journal of the American Society for Information Science and Technology*, n/a–n/a. doi:10.1002/asi.22819

Shah, C., & González-Ibáñez, R. (2010). Exploring information seeking processes in collaborative search tasks. *Proceedings of the American Society for Information Science and Technology*, *47*(1), 1–7. doi:10.1002/meet.14504701211

Shah, C., & Marchionini, G. (2010). Awareness in collaborative information seeking. *Journal of the American Society for Information Science and Technology*, *61*(10), 1970–1986. doi:10.1002/asi.21379

Short, J., Williams, E., & Christie, B. (1976). The social psychology of telecommunications. Retrieved from http://www.citeulike.org/group/1662/article/327279

Siegel, H. (1998). KNOWLEDGE, TRUTH AND EDUCATION. In D. Carr (Ed.), *Education, knowledge, and truth: beyond the postmodern impasse* (pp. 19–36). London, UK: Routledge.

Siemens, G. (2006). *Knowing Knowledge*. Lulu.com.

Simpson, T. W. (2012). Evaluating Google as an Epistemic Tool. *Metaphilosophy*, *43*(4), 426–445. doi:10.1111/j.1467-9973.2012.01759.x

Simsek, D., Buckingham Shum, S., Sandor, A., De Liddo, A., & Ferguson, R. (2013). XIP Dashboard: visual analytics from automated rhetor-ical parsing of scientific metadiscourse. In *1st International Workshop on Discourse-Centric Learning Analytics*. Leu. Retrieved from http://libeprints.open.ac.uk/37391/1/LAK13-DCLA-Simsek.pdf

SQW. (2011, June). Evaluation of the Home Access programme - Final report. DCSF. Retrieved from https://www.education.gov.uk/publications/eOrderingDownload/DFE-RR132.pdf

Stadtler, M., & Bromme, R. (2007). Dealing with multiple documents on the WWW: The role of metacognition in the formation of documents models. *International Journal of Computer-Supported Collaborative Learning*, *2*(2), 191–210. doi:10.1007/s11412-007-9015-3

Stadtler, M., & Bromme, R. (2008). Effects of the metacognitive computer-tool met.a.ware on the web search of laypersons. *Computers in Human Behavior*, *24*(3), 716–737. doi:10.1016/j.chb.2007.01.023

Stahl, G. (2010). GROUP COGNITION AS A FOUNDATION FOR THE NEW SCIENCE OF LEARNING. In M. S. Khine & I. M. Saleh (Eds.), *New Science of Learning: Cognition, Computers and Collaboration in Education.* London: Springer.

Strømsø, H. I., & Bråten, I. (2009). Beliefs about knowledge and knowing and multiple-text comprehension among upper secondary students. *Educational Psychology*, *29*(4), 425–445. doi:10.1080/01443410903046864

Strømsø, H. I., & Bråten, I. (2010). The role of personal epistemology in the self-regulation of internet-based learning. *Metacognition and Learning*, *5*(1), 91–111. doi:10.1007/s11409-009-9043-7

Strømsø, H. I., Bråten, I., & Britt, M. A. (2011). Do students' beliefs about knowledge and knowing predict their judgement of texts' trustworthiness? *Educational Psychology*, *31*(2), 177–206. doi:10.1080/01443410.2010.538039

Sumner, T., & Buckingham Shum, S. (1998). From documents to discourse: Shifting conceptions of scholarly publishing. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 95–102). Los Angeles, California, USA: ACM Press/Addison-Wesley Publishing Co. Retrieved from http://dl.acm.org/citation.cfm?id=274659

Sundin, O., & Johannisson, J. (2005a). Pragmatism, neo-pragmatism and sociocultural theory: Communicative participation as a perspective in LIS. *Journal of Documentation*, *61*(1), 23–43. doi:10.1108/00220410510577998

Sundin, O., & Johannisson, J. (2005b). The Instrumentality of Information Needs and Relevance. In F. Crestani & I. Ruthven (Eds.), *Context: Nature, Impact, and Role* (Vol. 3507, pp. 107–118). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://www.springerlink.com/content/gb5gfmg8xn8u3gtx/

Sutcliffe, A., & Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers*, *10*(3), 321–351.

Suthers, D. D. (2001). Towards a Systematic Study of Representational Guidance for Collaborative Learing Discourse. *Journal of Universal Computer Science*, *7*(3), 254–277.

Suthers, D. D. (2003). The Influence of Knowledge Modeling on the Communication Process. Retrieved from http://scholarspace.manoa.hawaii.edu/handle/10125/22890

Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning*, *1*, 315–337. doi:10.1007/s11412-006-9660-y

Suthers, D. D. (2008). Empirical Studies of the Value of Conceptually Explicit Notations in Collaborative Learning. In T. Sherborne, S. J. B. Shum, & A. Okada (Eds.), *Knowledge Cartography* (pp. 1–23). Springer London. Retrieved from http://link.springer.com.libezproxy.open.ac.uk/chapter/10.1007/978-1-84800-149-7_1

Suthers, D. D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *The Journal of the Learning Sciences*, *12*(2), 183–218.

Suthers, D. D., Lund, K., Rosé, C. P., Teplovs, C., & Law, N. (Eds.). (Accepted). *Productive Multivocality in the Analysis of Group Interactions*. Springer.

Suthers, D. D., & Medina, R. (2010). The Temporal Development of Representational Practices: Implications for Theory and Analysis of Situated Learning. In *2010 43rd Hawaii International Conference on System Sciences (HICSS)* (pp. 1–10). doi:10.1109/HICSS.2010.397

Suthers, D. D., Medina, R., Vatrapu, R., & Dwyer, N. (2007). Information sharing is incongruous with collaborative convergence: The case for interaction. In *Proceedings of the 8th iternational conference on Computer supported collaborative learning* (pp. 715–717). Retrieved from http://dl.acm.org/citation.cfm?id=1599730

Suthers, D. D., Vatrapu, R., Medina, R., & Dwyer, N. (2007). Information sharing and interaction in collaborative convergence. *FRONTIERS IN ARTIFICIAL INTELLIGENCE AND APPLICATIONS*, *162*, 167.

Suthers, D. D., Vatrapu, R., Medina, R., Joseph, S., & Dwyer, N. (2007). Conceptual representations enhance knowledge construction in asynchronous collaboration. In *Proceedings of the 8th iternational conference on Computer supported collaborative learning* (pp. 705–714). Retrieved from http://dl.acm.org/citation.cfm?id=1599729

Suthers, D. D., & Verbert, K. (2013). Learning analytics as a “middle space.” In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 1–4). New York, NY, USA: ACM. doi:10.1145/2460296.2460298

Suthers, D. D., & Xu, J. (2002). Kukakuka: An online environment for artifact-centered discourse. In *Education Track of the Eleventh World Wide Web Conference (WWW 2002)* (pp. 7–11). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.17.8384&rep=rep1&type=pdf

Teng, Y.-T. (2010). *The relationship between epistemic beliefs and knowledge contribution to online communities of practice* (PhD). University of Illinois at Urbana-Champaign. Retrieved from https://www.ideals.illinois.edu/handle/2142/16126

Thatcher, A. (2006). Information-seeking behaviours and cognitive search strategies in different search tasks on the WWW. *International Journal of Industrial Ergonomics*, *36*(12), 1055–1068.

Tillema, H., & Orland-Barak, L. (2006). Constructing knowledge in professional conversations: The role of beliefs on knowledge and knowing. *Learning and Instruction*, *16*(6), 592–608. doi:http://dx.doi.org/10.1016/j.learninstruc.2006.10.006

Tsai, C. (2004). Beyond cognitive and metacognitive tools: the use of the Internet as an “epistemological” tool for instruction. *British Journal of Educational Technology*, *35*(5), 525–536. doi:10.1111/j.0007-1013.2004.00411.x

Tsai, C.-C. (2004). Information commitments in Web-based learning environments. *Innovations in Education and Teaching International*, *41*(1), 105–112. doi:10.1080/1470329032000172748a

Tsai, M.-J. (2004). Developing The Internet Self-Efficacy Scale (ISES). In *World Conference on Educational Multimedia, Hypermedia and Telecommunications* (Vol. 2004, pp. 4406–4408). Chesapeake, VA: AACE. Retrieved from http://editlib.org/p/11710

Tsai, M.-J., Hsu, C.-Y., & Tsai, C.-C. (2012). Investigation of High School Students’ Online Science Information Searching Performance: The Role of Implicit and Explicit Strategies. *Journal of Science Education and Technology*, *21*(2), 246–254. doi:10.1007/s10956-011-9307-2

Tsai, P.-S., Tsai, C.-C., & Hwang, G.-J. (2011). The correlates of Taiwan teachers’ epistemological beliefs concerning Internet environments, online search strategies, and search outcomes. *The Internet and Higher Education*, *14*(1), 54–63. doi:10.1016/j.iheduc.2010.03.003

Tseng, J. C. R., Hwang, G.-J., Tsai, P.-S., & Tsai, C.-C. (2009). META-ANALYZER: A WEB-BASED LEARNING ENVIRONMENT FOR ANALYZING STUDENT INFORMATION SEARCHING BEHAVIORS. *Computing, Information and Control*, *5*(3), 567–579.

Twidale, M. B., Nichols, D. M., & Paice, C. D. (1997). Browsing is a collaborative process. *Information Processing & Management*, *33*(6), 761–783.

Twidale, M. B., Nichols, D. M., Smith, G., & Trevor, J. (1995). Supporting collaborative learning during information searching. In *The first international conference on Computer support for collaborative learning* (pp. 367–370). Retrieved from http://dl.acm.org/citation.cfm?id=222831

Undervisningsministerie (Ministry of Education), & Afdelingen for Gymnasiale Uddannelser (Department of Secondary Education). (2010, September). Endelig rapport fra følgegruppen for forsøget med digitale prøver med adgang til internettet i udvalgte fag på stx og hhx (Final report of the Monitoring Group experiment with digital samples with access to Internet in selected subjects at stx and HHX). Undervisningsministerie (Ministry of Education). Retrieved from http://www.uvm.dk/~/media/Files/Udd/Gym/PDF10/101007_it_rapport_13%20september%202010.ashx

Van Strien, J., Brand-Gruwel, S., & Boshuizen, E. (2012). Do prior attitudes influence epistemic cognition while reading conflicting information? Retrieved from http://dspace.ou.nl/handle/1820/4390

Vansledright, B. A., & James, J. H. (2002). Constructing ideas about history in the classroom: The influence of competing forces on pedagogical decision making. *Advances in Research on Teaching*, *9*, 263–298. doi:10.1016/S1479-3687(02)80010-4

Vydiswaran, V. G., Zhai, C., Roth, D., & Pirolli, P. (2012). BiasTrust: Teaching biased users about controversial topics. In *Proceedings of the 21st ACM international conference on Information and knowledge management* (pp. 1905–1909). Retrieved from http://dl.acm.org/citation.cfm?id=2398541

Walraven, A., Brand-Gruwel, S., & Boshuizen, H. (2008). Information-problem solving: A review of problems students encounter and instructional solutions. *Computers in Human Behavior*, *24*(3), 623–648. doi:10.1016/j.chb.2007.01.030

Wegerif, R. (2006). A dialogic understanding of the relationship between CSCL and teaching thinking skills. *International Journal of Computer-Supported Collaborative Learning*, *1*, 143–157. doi:10.1007/s11412-006-6840-8

Wegerif, R. (2010). Dialogue and teaching thinking with technology. In K. Littleton & C. Howe (Eds.), *Educational dialogues: Understanding and promoting productive interaction* (p. 304). Oxford, UK: Routledge. Retrieved from http://books.google.co.uk/books?hl=en&lr=&id=cAXvoXtf-wwC&oi=fnd&pg=PA304&dq=rupert+wegerif+%22educational+dialogues%22&ots=2nNGRHkG3j&sig=l2sPQ4BR53tRReQPFzTcGVk-ZHU

Wegerif, R., Littleton, K., Dawes, L., Mercer, N., & Rowe, D. (2004). Widening access to educational opportunities through teaching children how to reason together. *Westminster Studies in Education*, *27*(2), 143. doi:10.1080/0140672040270205

Wells, G. (2002). The role of dialogue in activity theory. *Mind, Culture, and Activity*, *9*(1), 43–66. doi:10.1207/S15327884MCA0901_04

Whitmire, E. (2003). Epistemological beliefs and the information-seeking behavior of undergraduates. *Library & Information Science Research*, *25*(2), 127–142.

Wildemuth, B. M., & Freund, L. (2012). Assigning search tasks designed to elicit exploratory search behaviors. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval* (pp. 4:1–4:10). New York, NY, USA: ACM. doi:10.1145/2391224.2391228

Williams, K. (1998). Assessment and the challenge of scepticism. In D. Carr (Ed.), *Education, knowledge, and truth: beyond the postmodern impasse*. London, UK: Routledge.

Williams, P., & Rowlands, I. (2007, October 18). Information behaviour of the Researcher of the Future. JISC. Retrieved from http://www.ucl.ac.uk/infostudies/research/ciber/downloads/GG%20Work%20Package%20II.pdf

Wopereis, I. G. J. H., & van Merriënboer, J. J. G. (2011). Evaluating text-based information on the World Wide Web. *Learning and Instruction*, *21*(2), 232–237. doi:10.1016/j.learninstruc.2010.02.003

Wu, Y., & Tsai, C. (2005). Information commitments: evaluative standards and information searching strategies in web-based learning environments. *Journal of Computer Assisted Learning*, *21*(5), 374–385. doi:10.1111/j.1365-2729.2005.00144.x

Wu, Y.-T., & Tsai, C.-C. (2005). Developing the Information Commitment Survey for assessing college and graduate students' evaluative standards for Web information and their searching strategies in Web-based learning environments. In *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (Vol. 2005, pp. 2500–2512). Retrieved from http://www.editlib.org/INDEX.CFM?fuseaction=Reader.ViewPresentation&paper_id=21573&paperfile_id=34292

Wu, Y.-T., & Tsai, C.-C. (2007). Developing an Information Commitment Survey for assessing students' web information searching strategies and evaluative standards for web materials. *JOURNAL OF EDUCATIONAL TECHNOLOGYAND SOCIETY*, *10*(2), 120.

Wu, Y.-T., & Tsai, C.-C. (2010). The Effects of Different On-line Searching Activities on High School Students' Cognitive Structures and Informal Reasoning Regarding a Socio-scientific Issue. *Research in Science Education*, *41*(5), 771–785. doi:10.1007/s11165-010-9189-y

Yue, Z., Han, S., & He, D. (2012). Search Tactics in Collaborative Exploratory Web Search. In *Proceedings of HCIR 2012*. Cambridge, MA. Retrieved from http://crystal.exp.sis.pitt.edu:8080/ips/paper/hcir2012_cs.pdf

Zeineddin, A., & Abd-El-Khalick, F. (2010). Scientific reasoning and epistemological commitments: Coordination of theory and evidence among college science students. *Journal of Research in Science Teaching*, *47*(9), 1064–1093. doi:10.1002/tea.20368

Zillien, N., & Hargittai, E. (2009). Digital Distinction: Status-Specific Types of Internet Usage*. *Social Science Quarterly*, *90*(2), 274–291.