Open Research Online

The Open University's repository of research publications
and other research outputs

Developing Learning Analytics for Epistemic Commitments in a Collaborative Information Seeking Environment

Thesis

oro.open.ac.uk

Developing Learning Analytics for Epistemic Commitments in a Collaborative Information Seeking Environment

Simon James Goodwin Knight (BSc, PGCE, MA, MPhil)

KNOWLEDGE MEDIA

KMi

INSTITUTE

Acknowledgements

Permissions

I retain non-exclusive copyright over most of the work herein, subject to the individual reuse permissions as below. Particular thanks to Elias (2011) for releasing the table listed as Table 1:2 in this paper under a Creative Commons BY-NC-ND license, and the Journal of Learning Analytics for publishing Knight, Buckingham Shum and Littleton (from which Figure 1:1 is taken, and preliminary sections adapted) under the same license (with authors retaining copyright), in both cases enabling easy reuse of material.

Individual thanks go to:

- Ivar Bråten for permission to reprint items from the psychometric reported in Bråten and Weinstein (2004)
- Table 2:5 is Reprinted from Knight and Littleton 'Learning from Collaborative Information Seeking' in Hansen.P., Shah.C., Klas C.P (Forthcoming) Collaborative Information Seeking: Best Practices, New Domains and New Thoughts. London, Springer International Publishing, p. 6, Copyright (2015), with permission from Springer International Publishing
- Figure 2:1 is Reprinted from Computers and Education, 53(4), Brand-Gruwel, Wopereis, and Walraven: A descriptive model of information problem solving while using internet, p.1209, Copyright (2009), with permission from Elsevier
- Figure 2:3 is Reprinted from Information Processing and Management, 46(6), Evans and Chi: An elaborated model of social search, p.661, Copyright (2010), with permission from Elsevier
- Figure 2:4 is Reprinted from Innovations in Education & Teaching International, 41(1), Tsai: Information commitments in Web-based learning environments, p.109, Copyright (2004); Taylor & Francis is pleased to offer reuses of its content for a thesis or dissertation free of charge contingent on resubmission of permission request if work is published.

Declaration

Much of the conceptual work presented in this thesis draws on my published work. Appropriate

citations are given to such work throughout. No publications are given in their entirety, or in place

of specific chapters/sections, in this thesis with the exception of sections 1-2, much of which are

adapted from Knight, Buckingham Shum and Littleton (2013a, 2014) (published under a CC-By-NC-

ND license, with authors retaining copyright), with the introduction to section 1 drawing on

http://www.edfutures.net/Learning_Analytics (licensed under a CC-By license), for which I am the

only contributor to the sections drawn on.

In addition, drawn from one of my MPhil assignments are: (1) 'Table 1: Dimensions of epistemic

cognition' (separately published in S. Knight & Mercer, Forthcoming, p. 3); (2) 'Table 3 – An

overview of approaches to the study of Information Seeking (adapted from Sundin & Johannisson,

2005b)' (S. Knight, 2012a, p. 3); (3) 'Table 10 - Overview of Methods Used in Epistemic Cognition

Research' (S. Knight, 2012b, pp. 18–19). Adapted from a table in my MPhil thesis is: (4) 'Table 9 –

An overview of models of epistemic cognition' (S. Knight, 2012b, p. 15); all are given appropriate

in text citation. Thematic contributions of published literature can be identified as follows.

The particular stance on epistemic cognition presented in section 2:4 has drawn on analysis in

earlier work (S. Knight, Arastoopour, Williamson Shaffer, Buckingham Shum, & Littleton, 2014; S.

Knight & Mercer, Forthcoming), and continued theorising around a social account of epistemic

cognition (S. Knight & Littleton, in submissionb).

A number of publications are drawn on in sections regarding the educational value of

collaborative information seeking (S. Knight & Littleton, in press); the need to understand the

educational context in CIS research (Newman, Knight, Elbeshausen, & Hansen, in press); the

collaborative benefits of CIS over and above individual information seeking (S. Knight & Littleton,

2015b); and the role of exploratory dialogue in collaborative information seeking (see empirical

evidence in, S. Knight, 2012b; S. Knight & Mercer, 2015); and its utility as a lens on epistemic

cognition (S. Knight, 2012b; S. Knight et al., 2013a; S. Knight, Buckingham Shum, & Littleton,

2013b; S. Knight & Mercer, Forthcoming).

This work has informed the development of my empirical work, the model for which was

presented at the LAK15 conference (S. Knight & Littleton, 2015a) and ongoing conceptual work

around discourse centric learning analytics (S. Knight & Littleton, in submissiona) and a socialised

account of epistemic cognition (S. Knight & Littleton, in submissionb).

1:1.1.1 References to relevant prior work

Knight, S. (2012a). Methodological justification: A framework for the analysis of students' collaborative use of search engines for information retrieval in the classroom. University of Cambridge.

Knight, S. (2012b, July). *Finding Knowledge - The Role of Dialogue in Collaborative Information Retrieval in the Classroom* (Master's). University of Cambridge, Cambridge.

Knight, S., Arastoopour, G., Williamson Shaffer, D., Buckingham Shum, S., & Littleton, K. (2014). Epistemic Networks for Epistemic Commitments. In *International Conference of the Learning Sciences*. Boulder, CO: International Society of the Learning Sciences. Retrieved from http://oro.open.ac.uk/39254/

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. http://doi.org/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., Buckingham Shum, S., & Littleton, K. (2014). Epistemology, assessment, pedagogy: where learning meets analytics in the middle space. *Journal of Learning Analytics*, *1*(2). Retrieved from http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3538

Knight, S., & Littleton, K. (in submissiona). Dialogue as data.

Knight, S., & Littleton, K. (in submissionb). Epistemic Cognition Socialised.

Knight, S., & Littleton, K. (in press). Learning through Collaborative Information Seeking. In P. Hansen, C. Shah, & C.-P. Klas (Eds.), *Collaborative Information Seeking: Best practices, New Domains, New Thoughts*. Springer. Retrieved from http://collab.infoseeking.org/cisbook2015/

Knight, S., & Littleton, K. (2015a). Developing a multiple-document-processing performance assessment for epistemic literacy. Presented at the The 5th International Learning Analytics & Knowledge Conference (LAK15): Scaling Up: Big Data to Big Impact, Poughkeepsie, NY, USA (Forthcoming). Retrieved from http://lak15.solaresearch.org/

Knight, S., & Littleton, K. (2015b). Thinking, Interthinking, and Technological Tools. In R. Wegerif, L. Li, & J. C. Kaufman (Eds.), *The Routledge International Handbook of Research on Teaching Thinking* (p. section 7(al)). Routledge. Retrieved from http://www.routledge.com/books/details/9780415747493/

Knight, S., & Mercer, N. (Forthcoming). The role of collaborative, epistemic discourse in classroom information seeking tasks. *Technology, Pedagogy and Education*.

Knight, S., & Mercer, N. (2015). The role of exploratory talk in classroom search engine tasks. *Technology, Pedagogy and Education*, *24*(3), 303–319. http://doi.org/10.1 08 0/14 75 939 X .2014.931884

Newman, K., Knight, S., Elbeshausen, S., & Hansen, P. (in press). Situating CIS – The importance of Context in Collaborative Information Seeking. In P. Hansen, C. Shah, & C.-P. Klas (Eds.), *Collaborative Information Seeking: Best practices, New Domains, New Thoughts*. Springer. Retrieved from http://collab.infoseeking.org/cisbook2015/

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abstract

Learning analytics sits at the confluence of learning, information, and computer sciences. Using a distinctive account of learning analytics as a form of assessment, I first argue for its potential in pedagogically motivated learning design, suggesting a particular construct – epistemic cognition in literacy contexts – to probe using learning analytics. I argue for a recasting of epistemic cognition as 'epistemic commitments' in collaborative information tasks drawing a novel alignment between information seeking and multiple document processing (MDP) models, with empirical and theoretical grounding given for a focus on collaboration and dialogue in such activities. Thus, epistemic commitments are seen in the ways students seek, select, and integrate claims from multiple sources, and the ways in which their collaborative dialogue is brought to bear in this activity. Accordingly, the empirical element of the thesis develops two pedagogically grounded literacy based tasks: a MDP task, in which pre-selected documents were provided to students; and a collaborative information seeking task (CIS), in which students could search the web. These tasks were deployed at scale (n > 500) and involved writing an evaluative review, followed by a pedagogically supported peer assessment task. Assessment outcomes were analysed in the context of a new epistemic commitments-oriented set of trace data, and psychometric data regarding the participants' epistemic cognition. Demonstrating the value of the methodological and conceptual approach taken, qualitative analyses indicate clear epistemic activity, and stark differences in behaviour between groups, the complexity of which is challenging to model computationally. Despite this complexity, quantitative analyses indicate that up to 30% of variance in output scores can be modelled using behavioural indicators. The explanatory potential of behaviourally-oriented models of epistemic commitments grounded in tool-interaction and collaborative dialogue is demonstrated. The thesis provides an exemplification of theoretically positioned analytic development, drawing on interdisciplinary literatures in addressing complex learning contexts.

Chapter 1:   INTRODUCTION

With the rise of online learning, at increasing scale, there has been a growing interest in the use of data from learning platforms to support that learning. Learning analytics and educational data mining are emerging fields, aimed at the analysis of this learning data for the purposes of insight into, support of, and intervention in learning. In the introductory sections (1:1-1:3) to this thesis, I introduce my theoretical account of learning analytics as an assessment technology. This account focuses learning analytic attention on developing models of the pedagogic, epistemological, and assessment purposes for which analytic devices are deployed. As such, the ability to access behaviours that were previously challenging to track offers opportunity to investigate key learning constructs, for the support (or assessment) of learning activities.

One such construct is literacy, which is a core component of learning including, but not limited to, the online context. I take literacy to be not just a mapping of phonemes (qua spoken words) to graphemes (qua written words) but the ability to *comprehend* in rich multi-media environments, to evaluate resources (and their perspectives), and integrate information from across resources. Bound up with this ability to effectively find and evaluate information in given texts (or multimedia resources) is students' beliefs regarding what knowledge is, how claims are justified and inter-related, what role students themselves play in constructing knowledge (versus 'taking it' from the page) – their epistemic cognitions. In sections 2:1-2:2 I develop a literacy based account of information seeking (the identification of needs, and finding information to address those needs) and multiple document processing which (in section 2:3) I associate with epistemic cognition. This connection between information seeking, literacy, and epistemic cognition is important in a variety of contexts. Imagine, for example, situations in which a parent is attempting to understand information around childhood vaccinations; a voter wants to investigate the plausibility of a politician's climate change denial; or someone seeking to lose weight wishes to investigate the merits of diet versus regular foodstuffs or supplements. In each case, the information seeker requires more than just the ability to read content; the information seeker must make decisions about where to look for information, which sources to select (and

corroborate), and how to synthesise (sometimes competing) claims from across sources. These information skills are key literacy skills, and they are actioned through the lens of epistemic cognition.

I note that much information seeking and epistemic cognition research has been conducted in individual contexts, and (in the epistemic cognition case) through self-report methods (sections 2:4.1 and 2:4.2). However, as section 2:2 illustrates, information seeking can be seen as a collaborative activity, with potential for learning gains through effective collaboration (see section 2:2.6). Moreover, I argue (sections 2:4.3-2:4.4), the ways in which individuals interact with information seeking tools and resources, and the dialogue they engage in while doing so, provide a novel lens onto epistemic cognition; through methodological approaches oriented to these observed behaviours – including collaborative behaviours and dialogue – researchers can gain insight into student epistemic *commitments.* The promise of learning analytics (as in section 2:5), then, is the development of analytic methods for constructs arising from learning science research, associated with learning outcomes.

My thesis work, then, engages in the theoretically grounded design of a large-scale study to investigate student epistemic commitments. Two primary tasks are designed, one in which multiple documents are provided to students with varying source qualities and claims, the other in which students search online for resources. Both tasks require students to write an epistemically salient text outlining the 'best supported claims' around a scientific issue (a herbicide and a food supplement). In both cases, the tasks are collaborative, and make use of an innovative browser addon to facilitate the collaboration and task completion, and to track behavioural data. Analysis of the behavioural data is driven by the stance on epistemic commitments that is argued for, making use of established metrics and a novel approach to treatment of the collaborative dialogue. Results provide insight into the explanatory power of behavioural trace for outcomes on the written task, with exemplifications of the epistemic nature of the tasks drawn particularly from the dialogue data, from which excerpts are presented alongside commentary. This analysis indicates reasonable explanatory power from the

behavioural data (up to 30% of the variance explained), with clear epistemic indicators in the dialogue data (as section 7:1.1.1 highlights).

The substantive contribution of this thesis is to provide an example learning analytic development, bringing together analytic approaches and a learning construct in a theoretically grounded study design. The particular study, involving collaborative information seeking and multiple document processing is, to my knowledge, the largest empirical study in either collaborative information seeking or epistemic cognition research and the first drawing alignment between the two domains. This design necessitated a theorised recasting of epistemic cognition in light of the theoretical approach and empirical evidence evaluated in the literature review to focus on dialogue, and behavioural trace – characterised as epistemic commitments. This novel behaviour-oriented approach to epistemic cognition represents a further contribution of the thesis. The analysis foregrounds the potential of a focus on 'epistemic commitments' in action, through behavioural trace (including dialogue), reinforcing the strength of the approach to learning analytics research argued for in the introductory section. Moreover, the analysis and discussion sections foreground areas of challenge for learning analytics research, indicating the complexities of addressing learning data in less-controlled environments and the compromises necessary in such analysis. This thesis thus presents the potential and challenges of learning analytics through the development of a learning analytic approach – comprising pedagogically grounded task design and trace indicators – for epistemic cognition in an information seeking context.

1:1 Introducing Learning Analytics[1]

There is an increasing interest in the development of learning analytics that are grounded in learning theory, building on established strong conceptual and empirical work in the fields of education, learning science, psychology and the information sciences (see, for example, Pardo & Teasley, 2014; Sharples et al., 2012). The first Learning Analytics and Knowledge conference

---

[1] Section 1.1 (introduction) is adapted from http://www.edfutures.net/Learning_Analytics, under a CC-By license, to which I am the primary author (see
http://www.edfutures.net/index.php?title=Learning_Analytics&diff=3363&oldid=3356#Where_has_learning_analytics
_come_from.3F). Sections 1.1 (onwards) and 2 are adapted from Knight, Buckingham Shum and Littleton (2014)

4

(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 context of formal education's use of virtual learning environments and learning management systems (VLEs/LMSs) . 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. The two most recent conferences (LAK14 and LAK15) have moved to further examine the "Intersection of learning analytics research, theory and practice" (LAK14; Pardo & Teasley, 2014; Pistilli et al., 2014), and most recently a shift in focus from data, to impact, with use of existing technologies for scalable impact on learning (Baron et al., 2015, p. 15).

Charting this progression, in "The State of Learning Analytics in 2012: A Review and Future Challenges" Ferguson (2012) tracks the broad developmental course of analytics for learning. Ferguson describes how 'big data' in business intelligence has received increasing interest, for example through targeted advertising, and collaborative filtering. The opportunity to apply these business intelligence techniques in learning contexts has arisen from an increased use of Virtual Learning Environments (VLEs), Content Management Systems (CMSs), and Management Information Systems (MIS) for education, with a commensurate increase in digitally available data regarding students in the form of background data (often held in the MIS) and learning log data (from VLEs). However, while there is interest in applying business intelligence techniques to digitally available educational data, it is not clear how to apply these systems in order to support learning contexts. This knowledge-gap is particularly pressing given the potential of analytic tools to provide evidence of 'progress', support good professional practice, and evidence high quality

learning design and pedagogy, and the commensurate potential that analytic tools (well designed or otherwise) may be used by management systems for use in both internal and external accountability structures. As a result of this tension between the potential of analytics to the upper-levels of the stakeholder hierarchy (governments and institutions) in accountability structures and the application of these accountability tools to teachers – who may also wish to use high quality analytics for learning – Ferguson notes that analytics is at an important stage. This stage has, therefore, involved increasing interest in the pedagogic affordances of learning analytics. This claim is supported by a systematic review of empirical articles (Papamitsiou & Economides, 2014, p. 59) in learning analytics, resulting in 40 key studies, that suggests four major 'axes' in the emerging research: pedagogy-oriented; contextualisation of learning; networked learning; and educational resources handling. This increasing interest in the pedagogic affordances of learning analytics has come alongside an increasing pressure to engage with analytics due to economic concerns, for example around the increasing costs of education, and the desire to scale learning in Massive Open Online Courses (MOOCs) (Sharples et al., 2013). 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:

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

However, while these issues relate to the purpose of analytics insofar as they consider the audience of those analytics, there is a wider concern, illustrated above. This concern regards the

kinds of learning activity we measure, and why, from measuring attendance and drop-out metrics, to understanding learning processes and practices. One of the contributions of this thesis is to provide theoretical discussion of learning analytics, relating them to epistemology, assessment and pedagogy in particular. This theoretically grounded approach motivates a focus on a particular learning construct – epistemic cognition – around which analytic techniques could be developed. The third Learning Analytics conference (LAK13, 2013) sought to bring 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, 2013) 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 context – how our tools and techniques might address learning, and how we can characterise learning. A contribution of this thesis is to provide an exemplification of this approach, bringing to bear the perspectives of psychology, learning science and information sciences to the construct of epistemic cognition. The use of a particular, collaborative information seeking (CIS) context – which will be described further throughout – provides a productive applied context within which to focus the empirical and theoretical evidence drawn from the different disciplines. It thus serves as an example of the kinds of multivocal analytic approach learning analytics researchers should draw on. The next section (adapted from my paper at that conference (S. Knight et al., 2013a) and subsequent journal article (S. Knight, Buckingham Shum, et al., 2014)), 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.

1:1.1 The Relevance of Epistemology

A fundamental question in education is: How do we, as educators, researchers, and assessors, know when a student knows something? That is, how we understand what it means to 'know' something, and what that knowledge is – how it manifests, what its structure is, what its warrants

or justifications are. 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 holistic inter-connections), and types of knowledge, for example knowing *how* (skills), versus 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 debate in relation to learning analytics 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 thesis work, the epistemological stance 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. That is, 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 and educators – to know, and what it means for our students to know.

1:1.2   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., 2014; S. Knight 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 high stakes exams suggest they are the fairest way to reliably, and fairly, assess competence. 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 (see, for example, 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, regarding internal validity, is that, while assessments clearly measure something, their outcomes (grades) do not necessarily reflect subject knowledge or mastery (Davis, 1999). These fundamental issues are highlighted in a significant body of research in the philosophy and psychology of education (see, for example, Davis, 1999; Gardner, 2011; Hopmann et al., 2007).

It should be noted that, while there may well be an empirical concern here – regarding reliability in measurement – 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 interrelated concepts, an understanding of links and of the *use* of knowledge, a 'holist' perspective on epistemology rejects the idea that individual tokens of knowledge may be separately distinguished – an ideal which high stakes testing strives for in a drive for reliability.

1:1.2.1 The epistemological concern

The epistemological concern raised above is strongly related to the wider educational curricula,

the nature of teaching (pedagogy), and assessment as articulated here by Greene, Azevedo and

Torney-Purta (2008, pp. 142–143) 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)*

(Greene, Azevedo, & Torney-Purta, 2008, pp. 142–143)

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 1:1[2].
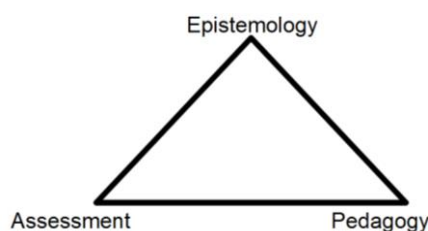
Figure 1:1: The Epistemology–Assessment–Pedagogy triad

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* 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 and ways in which, our methods for assessment offer 'access' to students' knowledge

---

[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'.

states. The triadic-relations are,, therefore, complex – 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. Throughout the thesis I draw out some of these with respect to learning analytics – which may be conceptualised as a component of assessment. Furthermore, I suggest that assessment regimes do implicate particular epistemological stances.

1:1.2.2 Our Learning Analytics Are Our Pedagogy

Buckingham Shum (2012b) used the shorthand "our learning analytics are our pedagogy" to draw attention to the arguments set out in more detail above: that the types of analytic we chose to deploy, and the ways in which we deploy them implicate particular approaches to learning and assessment. This is particularly important given that any use of analytics, will occur in the context of a wider educational ecosystem, as has been noted (Crook & Lewthwaite, 2010) in the context of more general educational technologies. It is this relationship between the types of analytic we deploy and our pedagogies which I now consider.

The relationship between learning analytics as a form of assessment 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, given in turn below, and learning analytics. These are not intended as comprehensive reviews, but rather as brief overviews of how the relationship between pedagogy and learning analytics might be conceptualised. The following section expands on some key ideas here, before moving on to explicate the core topic of this section – a sociocultural learning analytic – and one proposed instantiation of a learning analytic based on this approach.

1:1.2.2.1 Transactional or instructionalist approach

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

in many curricula, and the technological ease with which these may be automatically assessed underlies their current dominance of learning analytics and automated assessment (formative and summative).

Analytics Implications: learning analytics based on transactional approaches will tend to focus on metrics such as test scores, not requiring deeper analysis of more complex artefacts, or the processes by which they were derived (many standardised assessments, such as the Sats, might be said to fall into this category of analytic device).

1:1.2.2.2 Constructivist approach

Constructivist models focus on those forms of learning which occur in the learner's guided exploration of, and experimentation on, the world, typically in classrooms or online environments. Constructivist models are likely to measure learning in terms of quality of construction, with learners experimenting with their environment, and being capable of using tools which are appropriate for their given age.

Analytics Implications: learning analytics with a constructivist focus will focus on *progress,* particularly through tracking and judging the modifications made to a set of materials, resources or tools selected and arranged by the educator. An example of analytics in this tradition would be tracking the evolution of digital artefacts within the Scratch visual programming environment and community (Maloney, Resnick, Rusk, Silverman, & Eastmond, 2010).

1:1.2.2.3 Subjectivist or affect based approach

Subjectivist perspectives can be characterised as de-emphasising learning qua academia, with more attention to personal affect. While individual affect is a concern for educationalists, it is rarely if ever the *overarching* concern in the consideration of learning. One context in which affect is important is learning in complex socio-technical challenges: while there are certainly better and worse answers, there is too much information and no known best solution. Information seeking in such contexts can draw on subjectivist approaches that measure whether the user is 'satisfied' with the information they have found. Another context in which self-report is an important proxy for learning is research into dispositions (Deakin Crick, Broadfoot, & Claxton, 2004) and 'mindsets'

(Dweck, 2006) – learners' willingness to engage with opportunities that will challenge them, or stretch other transferable competencies such as questioning or collaborating.

Analytics Implications: In tandem with other approaches, learning analytics based on subjectivist approaches are likely to provide motivation assessments for understanding *why* someone is (or is not) engaging in particular actions. Such analytics may focus on self-report through survey tools (see for example, conceptually, Buckingham Shum & Deakin Crick, 2012; and empirical evidence Tempelaar, Niculescu, Rienties, Gijselaers, & Giesbers, 2012) or affect-based semantic mark-up such as blog tagging (R. Ferguson, Buckingham Shum, & Deakin Crick, 2011), alongside automated approaches such as textual sentiment analysis.

1:1.2.2.4 Apprenticeship approach

Apprenticeship approaches are sometimes used in learning analytics with an interest in whether the learner has become part of a community of practice or enquiry. 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. Epistemic Network Analysis of user data from gaming environments is designed to quantify the degree to which learners demonstrate behaviours valued in a professional community (Shaffer et al., 2009). The creation of social capital might be considered another proxy for community membership, overlapping with the next category.

1:1.2.2.5 Connectivist approach

Connectivism (Downes, 2007; Siemens, 2006) claims to highlight a perspective on epistemology which translates into a learning analytics 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 would look at how networks' size, quality and changes over time can serve as proxies for effective learning (Dawson, 2010; Haythornthwaite & de Laat, 2010).

1:1.2.2.6   Pragmatic, sociocultural approach

Pragmatic approaches (building on, for example, Dewey, 1938a) hold that learning occurs in the development and negotiation of a mutually shared perspective between learners targeted at some ends. Pragmatists suggest that, as human knowers, our conception of some given thing is bound up in our understanding of its practical application; 'truth' should thus be thought of in terms of pragmatic, activity-oriented use. 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. I follow Greeno, Collins, and Resnick (1992) in drawing a parallel between pragmatic, and sociocultural perspectives on learning (for example, the work of Vygotsky, 1987).

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 and purposes. Analytics tools in pragmatic-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. Various approaches to understand learner's social interactions might be characterised in this way (see, for example, Buckingham Shum & Ferguson, 2012). Analytics within this approach might attend particularly to quality of discourse for learning, development of critical literacy skills, and capabilities around 'learning to learn'. 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 learning analytics fits into the context of formative assessment or assessment for learning (AfL) and pedagogy, to which I now turn.

1:1.3   Epistemology and LA

Having summarised the sorts of relationships that might be seen between pedagogical approaches and learning analytics, I now turn to epistemology. The stance we take with regard to the relationship between epistemology, assessment and learning analytics relates to the issue of whether we place analytics in the role of summative diagnosis, or a kind of biofeedback tool to reflect on processes and practices of learning. Is learning analytics (and assessment) serving as the end point of, or an integrated component of pedagogy? As a diagnostic, we seek to accredit learning through defining behavioural proxies taken as evidence of knowledge and competencies. As biofeedback, learning analytics is seen as part of a wider system, 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 – necessarily – 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, for example, Black & Wiliam, 2001; Gardner, 2011). If evidencing process-centric competencies is defined as part of the summative assessment criteria, then the two categories may converge. For example, the process competencies of evidencing sound argumentation in discourse, higher resilience when stretched with new challenges, or the creation of social capital within a community of practice, might conceivably be assessed summatively through analytics.

The relationships highlighted in 1:1.2.2.1-1:1.2.2.6 serve as general pointers to the sorts of relationships one might see between pedagogy and learning analytics. There I also highlighted 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, accreditation implicates epistemological stances regarding when knowledge may be claimed (or

not). The preceding analysis suggests roles for learning analytics in accrediting mastery in three senses:

- **Mastering curriculum content:** this approach to learning analytics seeks behavioural markers of content knowledge using e-assessment technologies of varying sophistication, in order to generate summaries at varying granularities, for both individuals and cohorts. (see: transactional and constructivist approaches based on curricula mastery over a developmental course)

- **Evidencing membership and processes:** this approach to learning analytics 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. (see: affect-based, apprenticeship, and connectivist approaches based on connections to professional knowledges)

- **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. One characterisation of this family of approaches has been as Social Learning Analytics (Buckingham Shum & Ferguson, 2012) (see: connectivist and pragmatist approaches).

These three broad conceptualisations of learning analytics relate to the issue of whether or not we are deemed to *consume, discover,* or *co-construct* knowledge – is it 'out there' to be transmitted and acquired; unearthed through investigation; or emergent from activity-oriented contexts. This is not only a claim about learning or pedagogy, but a related claim about the status of knowledge, and its assessment. This thesis takes the third of these approaches – a pragmatic approach – focussed on use of knowledge, and the explanatory and supportive role of learning analytics in that area, rather than focussing on accreditation of learned curricula content.

1:1.4 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 high stakes standardised assessments 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, in the 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.

A broad pragmatic-sociocultural approach (see, Vygotsky, 1987; Wells, 1999; Wells & Claxton, 2002; Wertsch, 1993) takes it that:

1. Human activity (and thought) is best thought of in social terms, that we are by our very nature social beings, and that we come into being through social interaction;
2. That this social interaction is mediated by tools, most prominently language, which simultaneously shape and are shaped by our interaction with them;
3. Thus that cognition and learning – insofar as they are seen as separate from action – arise through, and are a part of, a process of tool mediated interaction within our (social) environment

Thus prompting a focus on humans as social, fundamentally entwined with their (social) environment, and the tools therein. As such, pragmatic-sociocultural approaches focus on dynamic interaction with both material and social resources, for example webpages and browsers and the social (collaborative, and cultural) environment. To be clear, these approaches should not be taken to be postmodern in a relativist sense (rejecting notions of fixed truth) or relativist-normative ('truth' as defined by the dominant theme of the time), but rather an approach focuses on the purposeful-activity and sensemaking over mapping of facts to states of the world. In their analysis of theories of cognition and learning, Greeno et al., (1992, pp. 27–28) indicate a set of design principles for the kind of broad pragmatic-sociocultural approach described here in which:

1. Design of learning environments should:
    a. Involve practices of social participation in enquiry learning – including use of social and material (such as textual) resources;
    b. Target development of epistemic identity, and support for that development (i.e., that environments should support students in making claims about what they know and how they know, in interaction with other knowers);
2. In design of curricula consideration should be given to:
    a. A focus on developing disciplinary discourse and representation practices – such that learners learn to engage in disciplinary dialogue communities;
    b. The use of realistic problems to be solved (and formulated) – authentic tasks that involve using salient constructs;
3. And in constructing assessments there should be focus on:
    a. Assessing participation in the enquiry practices (and social environment) of the community;
    b. Students as active participants in their own assessment; and
    c. Understand the sociocultural context of assessments as both shaped by, and shaping of, learning contexts and their wider societal setting in order to develop

> high quality assessment which respects the various ways in which learners and educators find value in particular learning practices and outcomes.

Thus, as described briefly above, pragmatic approaches have traditionally focused less on assessing the products of learning, and more on the process. Learning analytics in these approaches might encourage learners to reflect on their own contextualised activity, respond to formative feedback regarding their processes and practices, and engage in co-constructive behaviours with their peers and educators. This thesis is grounded in a sociocultural pragmatist approach. I develop a theoretically grounded information seeking task, making use of authentic resources in a naturalistic – social – context. In this task, students will engage in collaborative activity, to co-construct a shared resource. In line with this approach, analysis will draw on students' active processes of sharing and building knowledge. These will be aligned with assessment on a collaboratively written text, making use of a pedagogically motivated peer and self-assessment model.

1:1.5 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 policy-based 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 text they have not previously studied, making use of that text, a contemporary text, a short biography of the author and perhaps a media file from the publication period. They may be given unfamiliar resources, and permitted to source information for themselves from the internet. Thus, while Danish students in the pilot programme were expected to evidence mastery of curriculum knowledge – knowledge of facts – they must also exhibit a higher level of 'knowing-how' – success in use of resources – for example around information

---

[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). See also http://sjgknight.com/finding-knowledge/2013/07/danish-use-of-internet-in-exams-epistemology-pedagogy-assessment/

processing, synthesis, and metacognitive abilities, abilities that are less focussed on in countries restricting access to external resources that might otherwise enhance the student's capability. While this is of course simply one other (controlled) assessment 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, 1998, 2005; Davis & Williams, 2002; Katz, 2000).

We can contrast such assessments with high stakes testing regimes whose construct validity and external validity have been questioned. For instance, 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) in which 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 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's claim here is an explicitly sociocultural and pragmatic epistemological one: that there are important literacies and practices to be mastered in learning; that those should themselves be objects of assessment; and that 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. Thus attention should shift from assessment of repetition of facts within examination contexts to understanding the variety of processes of meaning-making and their situated context. 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 such examinations provide some wider access to information, this access alone 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 seeking and processing 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). An ethical concern

strengthens the motivation to explore these issues in order to understand why it is the case that the potential benefits of being online cannot be reaped through access alone (Hargittai, 2008). This need is particularly strong (Hargittai, 2010) given the internationally consistent finding that income, race, and educational level are associated with poorer ability to find and access information on the web (DiMaggio, Hargittai, Celeste, & Shafer, 2001)[4]. This finding indicates that there is a "second level digital divide" (Hargittai, 2002) between skilled and less skilled users of the internet, such that internet resources are 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).

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. In both cases, the understanding of internet use in the context of our understanding of knowledge (and, as discussed further below, literacy skills) is key.

It is important to draw a distinction here between various types of knowledge present in the context of information seeking. In particular, we can distinguish between a kind of propositional, fact based 'knowledge that', and a more procedural skills based 'knowledge-how'. These kinds of knowledge can be roughly mapped to Scardamalia and Bereiter's (2006) 'knowledge-about' and 'knowledge-of' respectively, where a rich conception of the latter is marked as crucial in developing 'understanding'. In the information seeking context we can take it that the type of knowledge retrieved in, for example search engine use, is propositional in nature. Such knowledge can be analysed in terms of suitability for answering any particular question – was the 'correct answer' retrieved. The accessing itself can be characterised as the 'knowing-how'. While certainly the former is important for understanding the latter, many educational contexts – and

---

[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.

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, and the literacy pre-conditions for developing understanding of sought knowledge. Research should, therefore, explore these literacy skills in the educational information seeking context. This is particularly important to aim for:

> *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 issues this divide raises. The importance of research into the digital divide is particularly salient given that access to the internet in and of itself may not confer its benefits in the absence of other pre-requisite conditions for meaningful use (Hacker & Mason, 2003). Thus, internet use is not simply about access to a sort of propositional knowledge. It is also about knowing *how* to access that knowledge as a shared cultural artefact, and develop understanding:

> *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 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 – there is a clear ethical duty to study these relationships. This claim is independent of, although strengthened by, concerns regarding the digital divide.

1:1.5.1 Critical Skills for Information Seeking

Crucially, the provision of pre-moderated reading lists and curated library resources are no longer

as core a focus for students or educators, who instead expect, and want, to find and use

information from the web in addition to those more traditional materials. 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. The broad concern regarding people's

ability to find and process information from the internet has been noted by a range of studies

across age ranges, as I discuss below, suggesting that across age ranges a deeper understanding of

internet based information seeking and processing is an important target of research.

Although the use of search engines is common by among young children and teenagers, many

report some issues with finding information (Livingstone, Bober, & Helsper, 2005, p. 10). A recent

review (Bartlett & Miller, 2011) of UK 'digital fluency' research between 2005-2010 (17 articles)

combined with a survey of 509 teachers in England and Wales paints a bleak picture of

information literacy skills across a range of ages. Supporting the reviewed research, the survey

data from that study indicated that teachers claim that: no information quality checks at all are

made by a quarter of 12-15 year olds; that their resource selection tends to a focus on resource

aesthetics over quality; and that the inclusion of websites within search engine results pages is

taken to be an indicator of the veracity of those resources. Worryingly, two thirds of respondents

from a broader age range (9-19 year olds) claimed never to have been taught how to judge the

reliability of the information they find, while concerns were expressed that students do not

understand how to conduct searches by over half of teachers searches (Bartlett & Miller, 2011). Indeed, across ages there is an overreliance on search engine ranking as a proxy for credibility, with many users attending little to cues of web page trustworthiness or reliability (Kammerer & Gerjets, 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 [since the introduction of web technology]" (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 (59 articles), 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. This is particularly concerning given

that "Little of the work on online credibility assessment has considered how the information-

seeking process figures into the final evaluation of content people encounter." (Hargittai,

Fullerton, Menchen-Trevino, & Thomas, 2010, pt. abstract).

1:1.6 Section Summary

To summarise, this section has argued that a consideration of epistemology is important to

learning analytics 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. Learning analytics are not objective or neutral: data does not "speak for itself" but arises from and is analysed using interaction with technological tools , which implicitly or explicitly, perpetuate the pedagogical and epistemological assumptions that come with any assessment instrument (Williamson, 2015).
- 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 general educational claim, that the quality

of a student's enquiry processes are 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 the context of internet-inclusive assessments 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 have not been

well explored. This is particularly true of collaborative contexts – that, in socioculturally grounded

pedagogic environments, can be educationally productive – this work addresses that gap. It is

thus that this thesis 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 cognition in the context of information seeking
2. Explain the relationship of epistemic cognition to the broader notion of epistemology
3. Describe the significance of epistemic cognition in the context of the triad outlined above

1:2 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 in order to answer any particular question – their *epistemic cognition* regarding the nature of the question, and how it may be answered (Black & Wiliam, 2009). 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). Epistemic cognition has thus been conceptualised as "internal conditions of learning" (Bromme et al., 2009, p. 1) embedded into self-regulation as facets of metacognition (Bromme et al., 2009)[5]. Indeed, there is evidence that more sophisticated epistemic cognition is generally associated with more productive learning strategies (Bromme et al., 2009; Schreiber & Shinn, 2003). For example, with more complex perspectives on knowledge in high school students associated with better adaptation to varying task-requirement complexities (Pieschl, Stallmann, & Bromme, 2014). Wider yet, when information is sought, the ways in which it is sought and processed matter; as Snow notes "Reading does not take place in a vacuum. It is done for a purpose, to achieve some end" (Snow, 2002, p. 15), and as Sundin and Johannisson put it in the information seeking context, "…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).

---

[5] See also Barzilai and Zohar's (2014) recent discussion of the relationship between metacognition and epistemic cognition

As shall be further expanded in section 2:3 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 reflection about knowledge, and knowing, in online information searching (Mason, Ariasi, & Boldrin, 2011). 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, 2014). 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 (Strømsø, Bråten, & Britt, 2011); indeed across students there is greater trust in textbooks than news sources, with a focus on content over data of publication in making judgements regarding trustworthiness (Bråten, Strømsø, & Salmerón, 2011). In addition, better learners engage in more sense-making on reliable sites than unreliable and by a larger margin than poorer learners (S. R. Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012).

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 section thus considers 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 Cognition:** Which I now introduce, and relates to the intrapersonal, psychological conceptualisations that individuals hold regarding knowledge

There is an increasing body of research exploring these epistemic cognitions – the ways in which learners conceptualise knowledge and coming to know (for an early review, see Schraw, 2001, and more recently, 2013). This body of research has used an array of labels, variously:

epistemological beliefs (Schommer, 1990); epistemic beliefs (Schraw, Bendixen, & Dunkle, 2002); epistemological understanding (Kuhn, Cheney, & Weinstock, 2000) or thinking (Kuhn & Weinstock, 2002); personal epistemology (B. K. Hofer & Pintrich, 2002); epistemic commitments (C.-C. Tsai, 2001; Zeineddin & Abd-El-Khalick, 2010) or in related work, information commitments (C.-C. Tsai, 2004; Y. Wu & Tsai, 2005); and epistemic cognition (Kitchener, 2002) which Greene, Azevedo, & Torney-Purta (2008) suggest is adopted as the overarching term for the constructs of interest in this thesis (a suggestion I follow).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 thesis I have generally opted to refer to 'epistemic cognition' when talking about epistemic beliefs or cognition and epistemological beliefs research (although quotations retain their original terminology); epistemic commitments, however, are discussed as such throughout the thesis.

Table 1:1 indicates four dimensions of epistemic cognition, for which there is general agreement across the various models[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, providing a lens onto students' understanding of their own learning, task demands, and how to meet those demands.

Table 1:1: Dimensions of epistemic cognition*

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"

*Tabulated from Mason, Boldrin, & Ariasi, 2009, p. 69; also presented in tabulated form in Knight & Mercer, Forthcoming, p. 3

---

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

1:2.1 The need for learning analytics focussed on epistemic cognition

Literacy is fundamental to learning, and ability to deal critically with dynamic texts is core to this. Building on work such as that of Rouet and Britt, the 2009-2015 PISA definition of reading indicates that: "Reading literacy is understanding, using, reflecting on and engaging with written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society." (OECD, 2013, p. 9). In the preceding sections I have introduced the construct of epistemic cognition, as a construct of interest to learning analytics, related to a particular epistemology, and salient in the context of the important and current skill of information seeking. This claim of the salience of epistemic cognition in the context of information seeking is true across a range of ages and contexts. This focus is on a subset of what might be considered broader notions of 'digital literacies' (see, for example, Littlejohn, Beetham, & McGill, 2012), and well aligned with the call for authentic tasks in the assessment of such broader literacies (Beetham, McGill, & Littlejohn, 2009). In the following section ('*2:1. Information Seeking; Defining Needs*') I outline some theories of information seeking, before indicating an issue with some of these theories, an issue which also raises a complication for epistemic cognition research – that of collaboration information seeking.

In section 1:1.5 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 many high-stakes assessment systems – 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 cognitions – 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 – is a good candidate for study under the sort of epistemology I put forward in the introduction.

It is thus that there have been recent calls for a focus on such notions of literacy in assessments, through performance based assessment and evidence centred design (S. R. Goldman, Lawless,

Pellegrino, & Gomez, 2012; Lawless, Goldman, Gomez, Manning, & Braasch, 2012). As Benjamin

notes with regard to performance assessments (Darling-Hammond & Adamson, 2010; Linn, Baker,

& Dunbar, 1991; Stecher, 2010):

> *Performance assessments are authentic problems or simulations of real world issues that*
> *are comprised of an assortment of documents, such as tables, figures, graphs, newspaper*
> *reports and photographs. The Internet permits richer versions of the tasks to be developed*
> *and on-line delivery permits the performance assessments to be administered, scored,*
> *analyzed and reported to the students and their institutions more quickly, accurately, and*
> *inexpensively. The scoring of the performance tasks provide rich diagnostic information*
> *about the students' writing, analysis, and critical thinking skills* (Benjamin et al., 2009, p.
> 3)

Such an approach should consider the elements of 'evidence centered design' (Mislevy, Behrens,

Dicerbo, & Levy, 2012) which moves through an evidence-based analysis of: (1) the constructs the

assessment aims to probe , (2) the behaviours (in this thesis, including tool use and interaction)

indicative of the constructs, and (3) a contextually grounded task design to elicit the salient

indicators identified in (2); to develop performance assessments (Darling-Hammond & Adamson,

2010; Linn et al., 1991; Stecher, 2010), which as Pellegrino notes, "do not offer a direct pipeline

into a student's mind. […] an [performance] assessment is a tool designed to observe students'

behavio[u]r and produce data that can be used to draw reasonable inferences about what

students know." (Pellegrino, 2013, p. 261). Aligned with the sociocultural account outlined above,

evidence centred design (Mislevy et al., 2012), rejects highly constrained assessment models,

favouring a model which can incorporate traditional assessments alongside data from informal

assessment, open ended tasks, collaborative simulations, and game based assessments. The aim

here is to focus on higher order skills, such that "unlike standardized multiple choice tests,

performance assessments are tests worth teaching to" (Benjamin et al., 2009, p. 3).

Along with the increase in internet use has come an increasing prevalence of ICTs such as Virtual

Learning Environments, bringing a growing interest in learning analytics: the use of trace-data

from such systems to make claims about learning (R. Ferguson, 2012). However, presently even

within a survey of 186 computer supported collaborative learning articles (from the International

Journal of Computer Supported Collaborative Learning, ijCSCL), only a minority of measures (12%

of 340 measures identified) assess process data including dialogue data, with most relying on self-report measures (33%) or assessment of output products (19%) (Gress, Fior, Hadwin, & Winne, 2010). Little research in epistemic cognition has taken a learning analytic approach, taking trace data as a data source for analysis (for related exceptions, discussed in section 8, see for example, Dimopoulos & Asimakopoulos, 2010; Greene, Muis, & Pieschl, 2010; Hsu, Tsai, Hou, & Tsai, 2013; Hwang, Tsai, Tsai, & Tseng, 2008; C. Lin & Tsai, 2008; Tseng, Hwang, Tsai, & Tsai, 2009). There is untapped potential here for the development of new learning analytics; as Winne notes:

> *trace data operationalize what learners do as they do it. Trace data avoid shortcomings of (a) asking leaners what they believe they do and (b) asking learners to perform mental calculations of unknown kinds (c) using sample fractions of past or possible future experiences that have unknown size and biases. When traces are faithful operational definitions of theoretical cognitive and metacognitive operations, they provide sturdy grounds for testing theories about when, whether, and how [self regulated learning] processes affect learning* (Winne, 2010, p. 275)

Epistemic cognition is thus one example of the type of construct which a pragmatically motivated learning analytic might probe. However, it is also a particularly *good* example given epistemic cognition's relationship to our everyday dealings with the world of information, and their relationship to pedagogy, assessment, and classroom practices (B. K. Hofer, 2001). It is a contribution of the thesis to investigate this learning science construct (epistemic cognition) using an innovative browser addon in the context of a pedagogically grounded information seeking task.

1:3 Research Aims and Contribution

The PhD sits firmly in the middle space – bridging the learning and information sciences, the psychology of epistemic cognition, and analytics approaches. It provides an exemplification of the interpretive flexibility in the use of educational technologies (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., 2014). In developing my approach to epistemic cognition, I provide an account which is socially-oriented, and describes the mediating role of the task design and tools.

Reflecting this, this PhD produces 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 (CIS). 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 epistemic cognition in CIS contexts are understudied.

This PhD takes the developed learning analytics lifecycle (Clow, 2012) which moves through: learners, data from or about learners, processing of data into metrics, and intervention. 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 1:2. In particular, I 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. Thus, in this PhD I start with a task context not uncommon to many students designing a well theorised pedagogically grounded task (the first of Clow's stages), and consider the types of data that may be captured about this context (Clow's second stage). A theorised account is given to describe the processing of this data into appropriate metrics – how to interpret the data (the third stage). The fourth stage of Clow's cycle (intervention) is, though, outside the scope of this developmental work, as it requires initial work (such as this) to provide a grounding from which to develop intervention strategies. As such, 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. It is also important to note theoretical considerations in task

design, selection and operationalization of constructs and the trace indicative of those constructs. I note that this accords well with my 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 1:2 - 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

Common to these models of research, analytics, and assessment is the need for clearly defined outcomes. Similarly across these models we see a mapping of trace indicators and behaviours, to theorised constructs.

A starting question for evidence centered design, built on in this thesis, is, "What are students supposed to do when they study multiple documents? And what kind of mental representation of such materials do they form?" (Rouet, 2006, p. 65); in answering questions like these Rouet, later built on by Rouet and Britt (2011), developed a literacy model: the Multiple Documents—Task-based Relevance Assessment and Content Extraction (MD-TRACE) model. The focus of this thesis is not on mental representation qua individual cognition, but rather on the ways in which epistemic literacy practices are represented in tool-mediated activity, student's dialogue around this activity, and the task-contextualised written outputs produced.

In this model, there are five main steps: task model construction; information need assessment; document processing; task product creation; task product assessment. At each stage external resources – including task specifications; search device; sources; text organisers; and reader generated products – and internal resources – including prior knowledge, reading/search skills,

and self-regulation skills – are brought to bear. These steps unfold interactively (i.e. they are not linear), and represent a more complex view on text processing than simple views of meaning-extraction. Of particular interest is the third stage of the process, in which: (a) decisions are made regarding whether information is relevant; (b) the text content is processed and (c) a document model is created/updated, perhaps feeding into (and drawing from) some task product. As Rouet (2006, p. 177) (citing Britt & Gabrys, 2001) notes, crucial to developing such literacy – and mature internet use – students should be taught:

1. *Skill of integration: the ability to connect prior and new information, including across documents, and including where claims are inconsistent or contradictory*
2. *Skill of sourcing: the ability to identify parameters that characterize the author and conditions of production of the information*
3. *Skill of corroboration: the ability to check information against multiple sources for its accuracy*
   (Rouet, 2006, p. 177)

Aula (2005) suggests that "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). Indeed, at first glance this is true, and the particular skills and processes involved in search are of key interest to this thesis (as will be discussed in the following section). However, as the model described above indicates, even within a set of pre-defined documents, the process of information seeking – deciding whether to probe a text for promising-looking links, building a document model incorporating inter-textual ties, corroborating information, and so on – is in itself complex, and parallels many tasks we might be more inclined to associate with search-engine use.

This research explores these issues, with a research aim to:

Investigate patterns of information seeking as epistemic processes

The work carries out this investigation through the theorised design of tasks to probe epistemic constructs in the context of a literacy task based on a multiple document processing and

information seeking task. Building on the theoretically grounded account of learning analytics in section 1:1, and empirical research outlined in subsequent sections, the works has a pragmatic focus on literacy inclusive of dialogue and interaction with a rich set of internet-based resources. As such, a rich collaborative context is seen as having pedagogic value in its own right, and providing a productive research lens onto student epistemic-activity in information seeking tasks.

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. In the following sections, I first (section 2:1) outline theories of information seeking, relating them to the particular epistemology and research context briefly outlined above. I then (section 2:2) elaborate a developing area of research – Collaborative Information Seeking – describing its incidence and relevance to my educational context. In section 2:3 this exposition is brought to bear specifically on information seeking as an epistemic process, before (section 2:4) I offer a full account of the particular stance taken in this work towards epistemic cognition as a construct. In the concluding section (section 2:5) I reiterate the potential of developing learning analytics, specifically relating information seeking analytic techniques, and peer assessment techniques to my own research. The following chapters thus outline the empirical work conducted for this thesis research, describing in turn the methods, and results before concluding with an extended discussion.

Chapter 2:   LITERATURE REVIEW

2:1 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. Importantly, as Cole notes, "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 my 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 the reader to models of information seeking, noting the increasing focus

on searchers as a crucial – in my view, epistemic – component of the information seeking system,

before, ending with a summary of useful facets explored across models.

2:1.1   Information Seeking Models

2:1.1.1 The Classical and Cognitive Models of Information Seeking

As various historic reviews of the field have described (see for example, Hearst, 2009b, Chapter

3), the classic model of information retrieval involves:

1. The Identification of an information need by an information seeker
2. Querying a system
3. Examining the results retrieved by that system
4. If needed, modifying the query and repeating the process

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 techno-centric because the process

through the model is seen as linear, and the model's description ends at the point when results

---

[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).

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).

However, more cognitively oriented frameworks develop this notion further, for example 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
5. Examining the results returned by the system
6. Reformulating the problem and its expression, and finally
7. Using the results

Thus rejecting a purely linear model of search (in '6') and emphasising human interaction with the

information sought – particularly at stages 5 and 6 – of key interest in consideration of

information seeker's decisions regarding selection, evaluation, and connecting of information.

Subsequent models thus 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. Importantly though, we see how models inclusive of the human agent offer

insight into the ways and stages at which epistemic cognition might impact on information

seeking behaviour. However, while these more cognitive models move to elucidate some of the

user-activities in interaction with the system, they are still rather linear, and system-oriented,

models.

2:1.1.2 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: first 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; and

second 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 are found along the way. Thus,

while classic models describe the search process as a matching of needs to retrieved information,

the berry picking model recognises that learning during the course of searching may change the

course of the information seeking episode; in sociocultural terms, search tools as mediational

tools in the search process.

2:1.1.3 Information Needs

Models such as the 'berry picking' model mark a shift from more techno-centric 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).

Cole's perspective thus emphasises the situated context within which information needs arise,

and are in mediation – it has a pragmatic focus on "a user's innate mechanism to generate

knowledge formation while seeking, finding, and using information during information search"

(Cole, 2011, p. 1220), indicating a clear epistemic factor in information seeking, and the

significance of theoretical accounts in development of models (and tasks) to understand

information seeking behaviour.

2:1.1.4 Information Seeking Process (ISP)

Related to this focus on users as contextualised, task- and information-needs oriented, 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 (in contrast to models that take

the information need as a given). Kuhlthau developed her model from field and case studies with

student library users – including a field study of 385 students. Kuhlthau'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 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 (derived from interview data) impact on the

various stages of ISP. In particular, Whitmire reports that indicators of epistemological belief developmental levels derived from interview data suggest 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 students were deemed not to reach this stage). However, this was a small scale study, using self-report interview data to analyse a (separate) learning process. In addition, the study operationalized stages of the search process in alignment with the epistemic constructs being studied, as such associations between the two are to be expected. 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.

2:1.1.5 Information Problem Solving 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 first year undergraduates and PhD students, trainee teachers, and secondary school students). They suggest (as in Figure 2:1) 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)
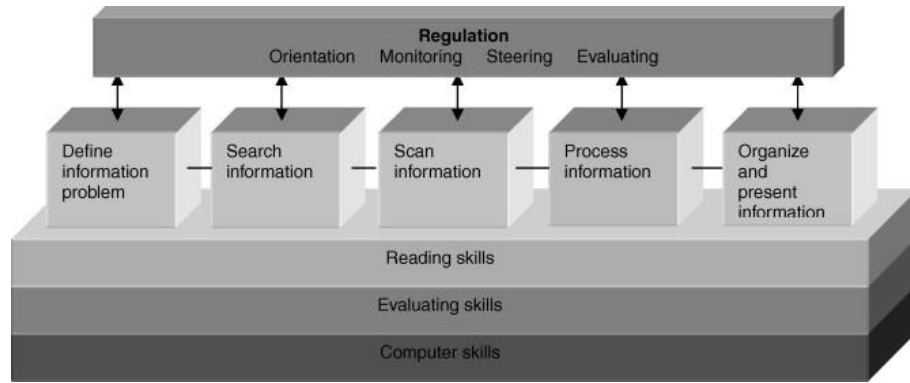
Figure 2:1 - The Information Problem Solving using the Internet Process (IPS-I) (Brand-Gruwel et al., 2009, p. 1209)

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:

1. *Define information problem*
2. *Search information*
3. *Scan information*
4. *Elaborate information*
5. *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; the identification of such epistemic processes cannot be captured by purely techno-centric models of search.

2:1.2   Information Seeking – Connections to the Epistemic

2:1.2.1 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).

Of particular interest here is Marchionini's association of differing levels of search activity with

levels on Bloom's taxonomy of educational objectives (Bloom, 1956) – which is a tool commonly

used in educational contexts for conceptualising a hierarchy of complexity in questions.

Marchionini notes a relationship between relatively lower level factual questions (for example,

"When did WWI end?") and precision oriented 'lookup' search tasks in which the aim of the

searching is to retrieve a single correct response. He contrasts these with more complex

investigations oriented to learning (for example, "What were the causes of WWI?"), which he

relates to 'exploratory search'.[8].

Marchionini thus notes that 'learning' searches, involve the kinds of literacy practices of interest

to this thesis: iteration, managing multiple resources, evaluating and comparing documents, and

synthesis:

> *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).

---

[8] See http://sjgknight.com/edusearch-tips/ or an abridged published version (S. Knight, 2013a) for some example questions and search tasks organised around Bloom's taxonomy, alongside some further tips for using search engines in school classrooms.

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).

2:1.2.2 Information Seeking –Epistemic Assumptions in Search 'Success'

Marchionini's exposition of 'exploratory search' in information seeking highlights the complexity of 'success' in search activities. Assessing the 'success' of question-answer routines, or fact retrieval searches, involves ascertaining whether the retrieved information matches some external feature of the world (it is 'true'). Ascertaining 'success' in the context of exploratory search may be more complex. Recalling the introductory sections of this thesis, there are both epistemic issues (regarding how student's conceptualise information problems and needs), and epistemological ones (regarding how task and system design imply particular perspectives on knowledge) at play here.

With regard to the latter, 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 2:1), 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 – which, as should be clear from the earlier sections of this thesis 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. In subsequent work, Sundin and Francke (2009) elaborate this theory for a Vygotskian, socio-cultural perspective on information literacy, stressing that "information literacy is in fact a socio-technical practice, incorporating knowledge of the epistemological aspects of the information sources as well as of the technology and systems that make up their material dimension" (Sundin & Francke, 2009, para. 12). As such, an

alignment should be drawn between sociocultural models of information seeking and of learning, highlighting the ways in which – in both cases – agents come to know through tool-mediated interaction with their (social) environment.

Table 2:1 – An overview of approaches to the study of Information Seeking (summarised from Sundin & Johannisson, 2005) (table adapted from S. Knight, 2012a, p. 3)

Approach	Epistemology	Methodological interest	Methodology	Methods	Implication for Information Seeking	Limitations
Structure Approach	Realist (aligned with transactional/ instructionalist approaches)	Emphasises structural features (socio-economic status, gender, etc.) in understanding behavior	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. 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	Excessive focus on systems with less focus on individuals. Assumptions regarding objectivity of information needs and access to these.
Individual Approach	User-centred, closer to constructivism	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' for individual ability, and affect (motivation & satisfaction) to address. Needs develop over time.	Excessive focus on individuals rather than their sociocultural context. Could be considered 'idealist'.
Affective/ motivational approaches	Subjectivist	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.	Relativisitic focus on affect over other aspects of cognition, system design.
Communicative Approach	Pragmatist/neo-pragmatist	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 its purpose and practices around it (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 of the search practices of users, or evaluation of search systems, cannot focus solely on whether some clearly defined need which reflects a deficit in the 'real' world, is plugged by retrieved information. This is because exploratory information seeking extends beyond the verification of facts into deeper epistemic and literacy processes. Indeed, as outlined in the introduction to this section, information needs are iterative, bound up with our learning and cultural context, and constantly renegotiated – through our information seeking we both define and address our information needs.

2:1.3 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* (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 as an epistemically salient information seeking process

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) – Knight and Spink's fourth element – and in understanding tasks. Moreover, by explicating the exploratory information seeking process as a socioculturally mediated cognitive process, we can draw parallels between models of information seeking and the more focussed literacy-based model of multiple document processing introduced in section 1:3; this broad alignment is depicted in Figure 2:2. This figure shows an alignment between various stages of information seeking (including in the context of multiple document processing), with horizontal arrows indicating alignment at each stage (or stages), and vertical arrows indicating the process of stages.
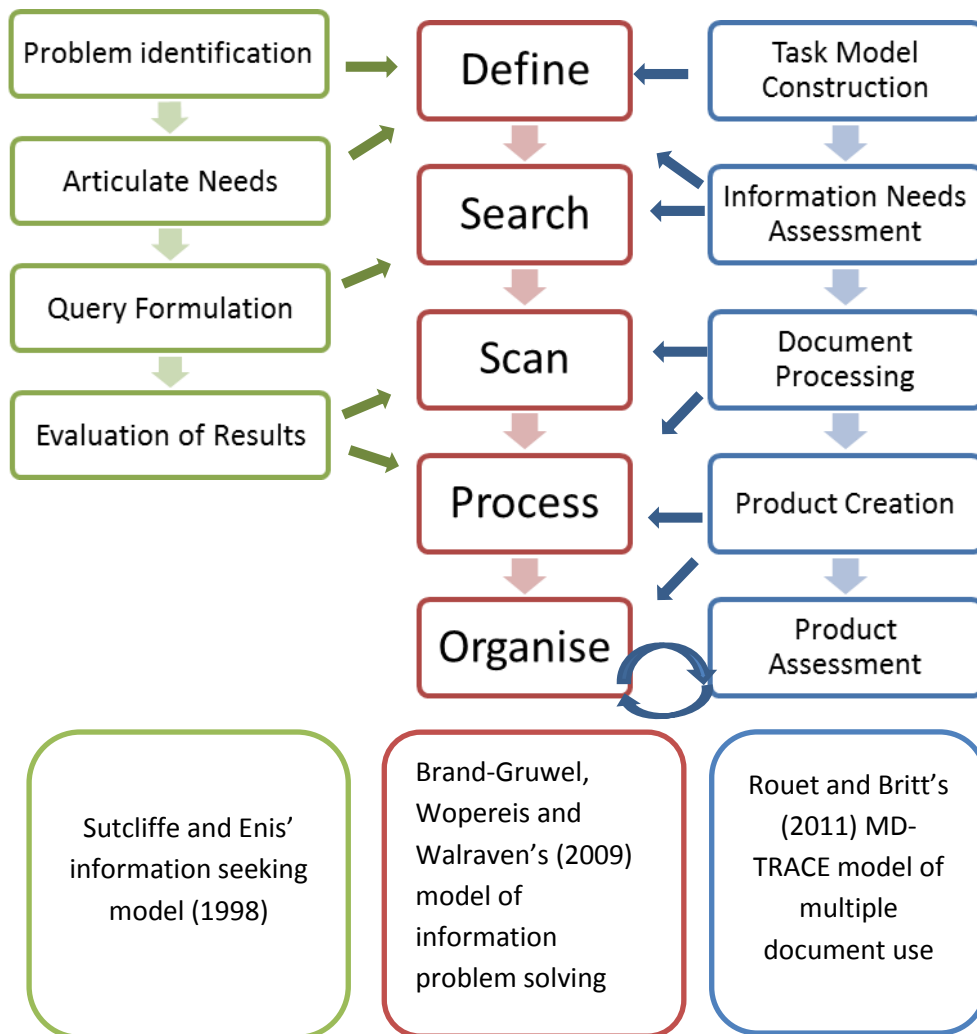
Figure 2:2 – Identifying common themes in models of: information seeking; information problem solving; and multiple document processing

Examining Figure 2:2 we see parallels between a sample of information seeking models, and the MD-Trace model, and limitations in taking any one of these models alone. Overall, the MD-Trace model provides the most appropriate model for the work in this thesis, as it incorporates most of the information seeking stages, alongside an iterative model of product creation and assessment. This model, understood in the context of search activity, provides a lens onto the decisions people make about defining their information needs, selecting, assessing, and collating information, and developing a document or topic-model within an output product, of the found information. Importantly, all three models in Figure 2:2 involve (to varying extents) the human, cognitive, element in information seeking and processing as well as the iterative nature of addressing information needs – both sites for consideration in epistemic cognition research.

In addition to the model as described, as briefly noted in the introduction and motivated further below, the role of dialogue and collaboration in literacy is important, but underexplored. Indeed, search is often collaborative too, and this collaboration forms an important component of the search context. Thus each stage of the model should be considered as iterative, and collaborative – involving a dialogue between partners, and both collective and individual activity. I will now turn to introducing collaborative search, its incidence, and models to represent it. In section 2:3 then I more fully explicate the relationship between information seeking and epistemic cognition, before (in section 2:4) describing a proposed model for epistemic cognition (or 'commitments') which resonates well with models of collaborative information seeking.

2:2 Collaborative Information Seeking

While in many cases information seeking is an individual task, this is not always so. This is particularly interesting given, as I discuss below (section 2:2.6), the potential educational benefits of collaboration, and a desire for contextually sensitive models of assessment and learning (as in chapter 1:1). Indeed, this is reflected by recent growing research attention on the issue of 'collaborative and social information seeking' with a section of the edited book "Web Search Engine Research" devoted to the various forms of "searching together" (Lewandowski, 2012), and a recent book length treatment from Shah (2012b) and separate edited volume (Hansen et al., Forthcoming).

This attention is warranted by the empirical evidence. Morris' early Microsoft study indicated that just over half of participants said they had ever collaborated on search activities, with analysis of that subset indicating that, within the last month, a substantial percentage (87.7%) had been 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 this 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 the kinds of collaborative search described in the preceding paragraph. 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 broader definition of collaboration, that included more indirect (for example, information shared via documentation) and explicit (the kinds of collaboration described above) interaction, Hansen and Järvelin (2005) found collaboration to be remarkably common at all stages of the information seeking process in a patent office context – 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 the notion of search. While certainly this lends support to my 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 concerning that the longstanding issue that most modern browsers do not facilitate this activity (see, for example, Twidale, Nichols, & Paice, 1997) stands, with several issues with "seamless collaborative information retrieval/seeking/behavior" systems remaining (Shah, 2009, p. 1). 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. 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 in CIS, to which I now turn.

2:2.1 What is the Collaboration in CIS?

In his review of CIS theory, Shah (2012b, 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.*
- *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.*

(Shah, 2012b, pp. 62–63)

Noting elsewhere (Shah, 2012b, p. 23) that, for effective CIS, systems should facilitate

communication and the coordination of contributions, including mechanisms for participants to

explore their differences and negotiate roles or responsibilities. Participants, in turn, should use

such systems to engage in constructive dialogue and negotiate their activities.

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, 2012b, 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, 2012b, p. 76).

2:2.2   What is the Information Seeking in CIS?

A question within CIS research has been the extent to which established models of information
seeking (as presented in section 2:1.1) apply to the CIS context. Research 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 case
study, (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, recording their
activity in a diary, and engaging in three interviews with reference to their diary reports. Their
reports suggested that although the broad ISP process was followed, aligning the activities
engaged in over the task process to particular stages was more challenging. This was because
group members would often work on separate activities and have different affective states
(towards a shared task). 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. Using survey, diary and interview analysis they report 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.

Later 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). Shah and González-Ibáñez's (2010) analysis of this data accorded with earlier work (Hyldegård, 2006, 2009), indicating that affect is not 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, might reflect the iterative nature of information seeking, rather than tightly constraining activities to individual stages of a process. 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.

2:2.3 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, 2012b, p. 67), to which we might later stages at which collaborative learning might include processes such as peer assessment of created products. During the process of CIS, as Table 2:2 below indicates, a number of potential learning benefits have been highlighted.

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. 679).
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) In the defining of information needs, where information such as important keywords or URLs might be exchanged and parameters for searching

(2) During the searching itself, where exchanges might involve attempts to understand the information found ,

(3) And in the latter stages of evaluation and use of information, where exchanges might involve organising information together, creating shared artefacts and perhaps externally sharing these resources with others

And proposing a model (Figure 2:3) of such search, indicating stages at which collaborative

defining and sharing of knowledge occur in the information seeking process.
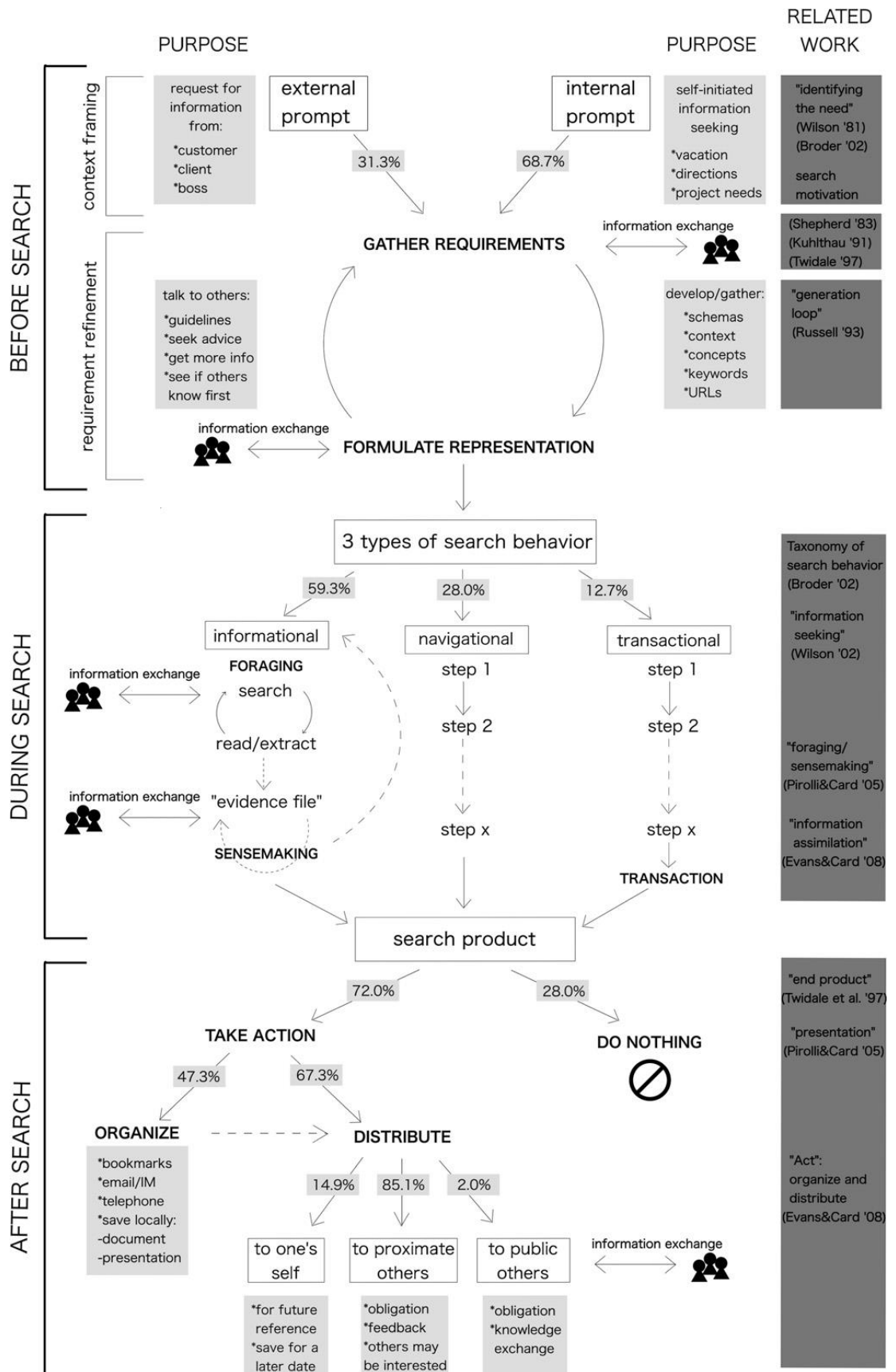
Figure 2:3 – Canonical social model of user activities before, during, and after a search act, with occurrence (%) indicated, including citations from related work in information seeking and sensemaking behavio[u]r (Evans & Chi, 2010, p. 661). (Examples of situations are given by asterisks)

2:2.4 Collaborative Information Seeking Models

In order to foreground 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, 2012b, p. 51), suggesting that these elements (as summarised in Table 2:3) be selected dependent on the particular research interests and task-contexts, for example by varying the roles given to users of the system.

Table 2:3 – Summary of Shah's elements of CIS (summarised from Shah, 2012b, 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

2:2.5   Further Factors in CIS – Awareness and Communication

2:2.5.1 Awareness[9]

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)

Relatively little CIS work has explored this issue, despite its importance for understanding the

nature of, and conditions of success in, CIS contexts. One study (Crescenzi & Capra, 2013) using

survey-data from 307 participants recruited via Amazon Mechanical Turk between 18 and 68

years old, 64% female and 32% students, asked respondents to consider "a situation in which you

have done a Web search in cooperation with another person (or several people)" (Crescenzi &

Capra, 2013, p. 2) using follow up questions to explore the particular context. This survey

explored the combinations of search computers used and shared display variables, reporting that:

- 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 share information. Thus, at times searchers

collaborate without sharing technology, by communicating about their searches and results.

---

[9] See also section 4.4.3 of Shah (2012b) for a review of awareness in CIS systems.

Building on established work, (Liechti & Sumi, 2002) Shah (2012b) thus proposes a taxonomy addressing four kinds of awareness: Group; workspace; contextual; and peripheral. The first (group) of which regards awareness of the group members' activity and status; the second (workspace) a space to share and create a common product; the third (contextual) regards the task-context and the needs imposed by that context; and the final (peripheral) regards awareness of the individual and collective's information history – what they have viewed and done previously.

Indicating the need for understanding coordination, and communications efforts in relation to awareness tools in CIS and the potential of well designed technology to support a variety of information seeking and collaborative aims (for example, as in Table 2:3). I now turn to discuss this combined aspect, of communication and awareness.

2:2.5.2 Communication and Awareness

The need for collaborator-awareness tools, including communication, 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 information-seeking 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 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[10]. Their trial of the tool, using groups of 3 participants 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 (see, for example, Clark & Brennan, 1991; Pickering & Garrod, 2004). 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 Capra et al., (2010), 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' (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).

Similar results were found in subsequent lab-based research with eleven asyhncronous collaborators (Capra, Chen, McArthur, & Davis, 2013). In this study, collaborators engaged in

---

[10] 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/

'think aloud' while working on a 30 minute task to write a report using a set of documents in a search system which included a tool to allow them to view their collaborators' previous work as they wished. This research indicated that collaborator's work was often leveraged at the query formulation and results examination stages, as well as when new ideas were being sought for research avenues. 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 (short extracts from the pages) 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 2:4 indicates.

Table 2:4 - 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.

2:2.5.2.1   Evidence for Dialogue in CIS

Thus CIS studies have noted the ways in which chat tools are made use of, 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).

This highlights the need to further explore the particular functions of the dialogue. Iin this last study (Shah & González-Ibáñez, 2010), 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 although affective considerations were relevant while participants found and shared information, the affective element in their chat disappeared in their transitions from one task to another and in the final stage of 'Presentation'. They also note 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); that is, that individuals would find information, and then share it for mutual consideration.

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, [participants] 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).

In both cases we see a literacy process similar to that described in the MD-TRACE model described above, making use of epistemic judgements – of evaluating and reflecting on information – within a collaborative information seeking context.

2:2.5.2.2 What Role Does Dialogue Play 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)[11][12].

Hertzum (2008) discusses the role of this shared motivation towards knowledge accumulation alongside the types of dialogue which facilitate such information seeking activities in the context of uneven distribution of information. Hertzum notes that communication is significant because collaborators: "may interpret the information known to them in different ways or be unable to make coherent sense of it" (Hertzum, 2008, p. 958), suggesting this is both the benefit of communication in CIS, and a cost to CIS activities in terms of effort and constraints. 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[13]. This accords well with the perspective taken in this work, and highlighted in Table 2:1 under the communicative, or pragmatic approach (Sundin & Johannisson (2005a).

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)

---

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

[12] 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.

[13] 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. 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, which has been associated with 'exploratory dialogue' in which there is a focus on understanding others ideas, as contrasted with cumulative and disputational dialogue in which interaction is characterised by statement and summary or disagreement and dispute respectively (Littleton & Mercer, 2013; Mercer & Littleton, 2007). This triadic typology (discussed further in section 2:4.4) can be compared in turn to Ellis et al.'s (2002) study, in which dialogue was analysed for: monochromatic active (language focuses on bigger picture); monochromatic passive (language focuses on approval of ideas); and polychromatic (language focuses on non-relevance of ideas) Ellis et al., (2002, p. 890). In Ellis et al.'s (2002) study the focus is on remotely located expert intermediaries who support the information seeking process, and the full range of dialogue is not analysed as a data form separately. However, the striking similarities in language types in this comparison indicates the potential of educational theory for analysis of talk in CIS contexts, and the potential of bringing together disciplines for the middle space of learning analytics research.

Similarly, in an educational context Foster (2009) analysed discourse using a classroom investigation task, although his work was on undergraduate students studying information management (n = 10) – 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).

2:2.6   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. Self-report measures of collaborative use certainly provide analytic insight. However, they are also open to sampling bias, interviewer effects and concerns regarding subjectivity. Moreover, self-report measures do not capture the specific ways in which dialogue mediates collaborators' contact with information. In order to maximise the potential of high quality collaboration in information seeking, the focus of our analysis should be on collaboration *in action.*

Indeed, there is support for this epistemic component of information seeking in the context of collaborative educational tasks from a between-subjects observational comparison of individual and collaborative students (n = 25, mean age 20) (Lazonder, 2005). Lazonder's proposal was 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. His suggestion is that this is important, because he claims that teenagers 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). Based on coding 'planning', 'monitoring', and 'evaluation' incidence, Lazonder suggests that the paired participants: performed significantly better; needed significantly less time to complete the tasks; engaged in a richer set of search strategies; evaluated more and tended to get an incorrect answer less (although neither of these results was statistically significant). The implication here is that, by encouraging effective collaboration through creation of common ground or knowledge, we may facilitate better information seeking processes through making shared language explicit. Indeed, some evidence suggests that in terms of query diversity (although not success), this is true of collaborative search tasks – that more diverse query language is related to collaborative processes (Yue, Han, & He, 2013). The suggestion may be further supported by Jucks and Paus' research (Jucks & Paus, 2013; Paus & Jucks, 2012) in which – in the latter study – giving dyads documents with different words for the same concept encouraged participants to engage in explicitly elaborated dialogue and led to better knowledge acquisition. However, Lazonder's research was a small scale study based on fact-finding (rather than exploratory) search processes 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, although with much younger participants, (Large, Beheshti, & Rahman, 2002), with a focus on gender, analysed 53 male and female 12 year olds engaging in collaborative 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. They report relatively higher rate of 'natural language' (as opposed to keyword) searches in girls, and the relatively faster speed at which boys navigated pages; however, these results are hard to interpret because most results were not statistically significant. Qualitative analysis of dialogue in use may thus be productive in this context, providing insight into the types of language used in

cases in which fewer keywords were entered in searches, versus dialogue use around more 'natural language search', and the association of both to outcome 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).

Indeed, in earlier work (S. Knight, 2012b; S. Knight & Mercer, Forthcoming, 2015) exploring collaborative classroom dialogue and fact-retrieval-based search tasks it was precisely the kinds of language used, while working together, that were related to the success of the children (11 year olds). In that work, despite generally similar academic attainment, the success of the small number of groups appeared to be directly related to their ability to use language to share and build ideas (their effective use of 'exploratory dialogue', as I shall introduce below in section 2:4.4). In that work we noted that the least successful group also made fewer attempts to explain their ideas to each other and build knowledge together (exploratory talk), as well as reflecting very little on the relationship between the information they found and the purpose for which they were seeking it. Indeed, the weakest group appeared to be primarily concerned with the quantity, ease of access, and aesthetic value of information. In contrast the other two groups focussed on the 'importance' of information and particularly that information was 'explained'; and the detail and novelty of information, respectively.

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 in doing so may ignore means to support higher quality collaboration. Thus, while the classic view of information seeking

– even collaboratively - implicates individual benefit over collective, a broader view of the benefits of CIS should be taken. In the context of my interest in epistemic cognition, there are at least two ways in which information seeking goes beyond individuals: First, it often involves complex information needs and complex practices beyond simple query-answer processes; and second:

> *when we seek information, particularly on the web, we engage with a network of linked documents with a rich set of intertextual ties; in a very real sense, reading much of the web involves an interaction with the thoughts of many people.* (S. Knight & Littleton, 2015b, p. 5)

I therefore see CIS as of educational interest in its own right, but further work is required to understand CIS 'outcomes' and the role of effective collaborative dialogue in those outcomes (S. Knight & Littleton, in press), the kind of dialogue we know to be associated with improved educational outcomes in other contexts (see section 7.4, and the collection edited by Littleton & Howe, 2010). Such work should foreground the ways discourse helps to shape the epistemic properties of particular tasks, including information seeking.

Consideration of the motivations and outcomes of CIS is important to gain full educational benefit (and indeed, to understand varying CIS contexts, Newman et al., in press). As Spence and Reddy (2012) note, CIS activities are highly contextualised, at the individual, team, organisational, and technological level. It is important, then, to relate the evidence of CIS in educational contexts to the wider literature around CIS and its contextual factors. To return to the list provided by Shah (Table 2:3),

Table 2:5 gives an indication of some of the focal points for CIS, and their relationships to educational interests and contexts – for example, understanding what sorts of collaboration are likely to be of interest to, and/or of high incidence in, educational settings.

Table 2:5 - Focal Points for CIS Research (S. Knight & Littleton, in press, p. 6)

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	Either synchronous (e.g. in class) or asynchronous (e.g. homework) within a constrained timeframe

	not?	
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 they are collaborating
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.

2:2.7 Section Summary

CIS is complex, and multi-faceted – and this complexity is carried over into assessment of its

success. As Shah notes, measures of effectiveness (such as finding more results) or efficiency

(finding results faster) may not capture the deeper benefits of CIS, such as a deeper

understanding of the information. In addition, "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, 2012b, pp.

115–116). However, as Shah indicates, the potential of CIS contexts for participants to develop

greater understanding, view a more diverse range of perspectives, raise engagement, and –

central to this work – probe 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

will now explore in more detail its relationship to epistemic cognition, before describing: some

methods for probing that construct, particularly in the online context (section 2:3) and some

possible advances in tracking epistemic cognition (or, as I will introduce – epistemic

commitments) (sections 2:4 and 2:5).

2:3 Information Seeking as an Epistemic Process

2:3.1 The role of epistemic cognition in information seeking[14]

2:3.1.1 Multiple Document Processing

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)[15]. 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 cognition, 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.

Building on the epistemic cognition literature, Bråten, et al., (2011) outline the empirical evidence

linking epistemic-cognition to the MD-TRACE model (presented in particularly in section 2:1.3) as

indicated in Table 2:6 which shows a summary of the hypothesized relationships between MD-

TRACE and epistemic cognition facets (which could be extended to information seeking models

more broadly, using the alignment in Figure 2:2).

Table 2:6 Summarised relationships between MD-TRACE and epistemic cognition*

---

[14] In literature search here I have particularly limited my focus to epistemic (or epistemological) cognition (or beliefs) 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 the relationship between information seeking and epistemic cognition research.

[15] Bråten (2008) reviews the relevant literature (to 2008) in epistemic cognition and multiple document processing in the context of learning within internet technologies, while Ferguson (2014) provides a more recent discussion of that research domain. A selection of this literature, and the literature since, is discussed throughout this thesis.

Facet of cognition	Less adaptive	More adaptive
Simplicity	Characterised by accumulation of facts, and a preference for simple sources	Characterised by integration of facts/claims (inter and intra-textual ties) and a downplaying of simple sources
Certainty	Characterised by use of single documents for sourcing claims	Characterised by corroboration of claims across multiple sources, and the tendency to represent a diversity of views in a complex perspective
Source	Characterised by an emphasis on one's own opinion as a justification, and a lack of differentiation between sources of varying qualities	Characterised by an emphasis on source characteristics, and (commensurately) a tendency to differentiate between sources of varying trustworthiness
Justification	Characterised by an emphasis on authority as justification, and a downplaying of corroboration	Characterised by an emphasis on argument schemas, and the use of corroboration and authority seeking in combination

*(relationships summarised from Bråten, Britt, et al., 2011; table adapted from S. Knight & Littleton, 2015a, p. 2)

This research indicates a range of relationships between epistemic cognition and document processing behaviours. Ferguson and Bråten (2013) report that from a sample of 65 10th grade students (mean age 14.9) those with less belief in personal justification, and stronger beliefs in justification by multiple sources on a self-report measure perform best on text-comprehension measured through assessment of three short open written answers designed to probe students integration and weighing of competing claims (Bråten, Ferguson, Strømsø, & Anmarkrud, 2013; L. E. Ferguson & Bråten, 2013).

Bråten, Ferguson, Strømsø, and Anmarkrud (2014) further report that from a sample of 51 undergraduates engaged in talk-aloud while reading six documents with conflicting claims regarding cell phone radiation those who justified knowledge claims by corroboration were more likely to include explicit referencing and had better argument structures in a subsequent written essay output (Bråten, Ferguson, et al., 2014).

In addition, Strømsø and Bråten (2009) indicate that, from a sample of 51 undergraduates engaged in reading six documents containing conflicting claims, there is a relationship between think-aloud indicating attempts to corroborate, and explicit source citations and inter/intra-textual ties in subsequent written output (Strømsø & Bråten, 2009). Moreover, a set of studies (Bråten & Strømsø, 2006a, 2010; Karimi & Atai, 2014) indicate that in a range of students (10 student-teachers; 49 law students; 64 midwifery students, respectively) more sophisticated

epistemic beliefs (on survey items) are associated with better synthesis and intra/inter-text integration as indicated by their written texts (Bråten & Strømsø, 2006a) or performance on an inference and comprehension task (Bråten & Strømsø, 2010; Karimi & Atai, 2014).

Further evidence from 204 participants indicates that students asked to integrate conflicting arguments are less likely to rebut arguments where they believe knowledge is passed down by authorities, and more likely to produce stronger arguments where they believe in a complex and tentative perspective on knowledge (Ku, Lai, & Hau, 2014). Recent modelling suggests that these individual difference variables have an indirect impact, via student effort and adaptive processing strategies, on multiple document comprehension (Bråten, Anmarkrud, Brandmo, & Strømsø, 2014) and that belief in scientific justification (over personal justification) is predictive of science achievement (Bråten & Ferguson, 2014).

2:3.1.2 Epistemic Cognition Online

Extending the multiple document paradigm, recent work has explored epistemic cognition in the comprehension of multiple *online* sources – which, in a naturalistic context, may vary radically in their justifications and source qualities – on the basis that students who perceive knowledge as simple and finite may have little recourse to synthesis, integration and corroboration, instead tending towards brief and perfunctory searches (Barzilai & Zohar, 2009; Bråten & Strømsø, 2006b). 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 meaningfully evaluate websites, with integration and critical evaluation of multiple online sources more likely of those with more sophisticated perspectives on the "multiplicity of knowledge" (Barzilai & Zohar, 2009; Bråten & Strømsø, 2006b). Further preliminary work suggests an association between "evaluativist" beliefs and comprehension of multiple conflicting online sources, but not multiple converging perspectives in online sources(Barzilai & Eshet-Alkalai, 2013).

Survey studies (Bråten & Strømsø, 2006b; Bråten, Strømsø, & Samuelstuen, 2005; Mokhtari, 2014) indicate relationships between aspects of epistemic cognition, and internet information seeking behaviours; although as indicated in Table **2:8** while self-report measures give productive insight and useful indictors of variance (for example, in beliefs), they can be challenging to interpret in the context of behavioural processes or apply across disciplines and task types.

A growing body of work associates search and sourcing patterns with particular patterns of epistemic cognition (Mason et al., 2009). Think-aloud studies in this line of research, indicate 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). This research finds 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 may – as an artefact of the method – increase practices such as credibility judgements (Schraw, 2000; Schraw & Impara, 2000). 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 Section2:4.

2:3.2   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 for our approach to learning analytics.

In the following section (section 2:4.1) I give an overview of theories of epistemic cognition, followed by (section 2:4.2) a discussion of methods to assess this construct.. In light of theoretical and methodological limitations, recent approaches to epistemic cognition are then discussed

(section 2:4.3). I use this discussion to motivate a novel methodological and theoretical approach to epistemic cognition (or commitments) (section 2:4.4 onwards) particularly in the context of collaborative, dialogue mediated information seeking. These sections thus (1) provide the theorised account of a broadly sociocultural account of epistemic cognition as oriented to behavioural trace in interaction with technological tools, and collaborative dialogue; an account which is applied to epistemic cognition in a CIS context, and (2) provide a methodological account with reference to prior work, discussing the analysis of behavioural trace and particular kinds of salient dialogue. Grounded in this novel theoretical stance, sections 2:4.5 and 2:5 then put forward an original approach to analysis of epistemic commitments through behavioural trace, offering a new methodological tact in epistemic cognition research.

2:4 Epistemic Cognition, Commitments, and Dialogue

2:4.1 An overview of theories related to epistemic constructs

Within the epistemic cognition literature, three broad models of epistemic cognition have been proposed (as Table 2:7 illustrates) namely cognitive developmental models, multidimensional perspectives, and the resources view. Firstly, cognitive developmental models (P. M. King & Kitchener, 2004; Kuhn & Weinstock, 2002) suggest that there is a sequence of increasingly sophisticated epistemic beliefs that individuals progress through. Similarly, multidimensional perspectives (B. K. Hofer, 2001; Schommer, 1990) suggest a separation of epistemic beliefs into dimensions, within which levels of sophistication can be identified (Greene et al., 2010, p. 248). Across both developmental and multidimensional models there is an assumption of uni-directional, fixed trajectories in which epistemic beliefs are applied uniformly across (and within) domains. Contrastingly, the resources view, emphasises the ways in which resources mediate, and are interacted with by the believer, highlighting the flexibility of resource-use available to a cognizer across stages in a task (Hammer & Elby, 2003). Thus, across the models, there is disagreement regarding the psychological nature and scope of epistemic cognition as a construct and whether it itself is stable – developmentally, and across domains – or shaped in some way by

resources or beliefs. This concern is similarly reflected in methods used for analysis of epistemic

cognition, as Table 2:7 indicates.

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)." (Greene et al., 2010, p. 248)	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.	Glosses 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)." (Greene et al., 2010, p. 248)	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.	Glosses 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." (Greene et al., 2010, p. 248)	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.

*Material adapted from Greene et al., (2010, p. 248). Multiple theoretical frameworks have been developed and reviewed (B. K. Hofer, 2004b; B. Hofer K. . & Pintrich, 1997; B. K. Hofer & Pintrich, 2002; Muis, 2007; Schraw, 2013). This table is adapted from Knight (2012b, p. 15).

Insofar as each model attempts to understand different facets of epistemic cognition, my own research may take an agnostic stance on the particular cognitive model of epistemic cognition, although a conceptual perspective will be offered in section 2:4 and a methodological one in section 2:4.5. 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 section2:4.4). Thus the models will instead inform the ways in which the data is understood – the ways that epistemic cognitions are brought to bear through activity, and the particular methods that have probed this. Importantly, as introduced in section 1:2, across the three broad models, there is agreement on two main areas – what knowledge is, and how one comes to know, as Table 1 indicated.

2:4.2 An overview of methods in epistemic constructs research

From this work a number of methods can be identified. Methodologically, interviews and laboratory tasks have been favoured by cognitive developmental models,, while paper and pencil self-report measures tend to be used in multidimensional models (DeBacker, Crowson, Beesley, Thoma, & Hestevold, 2008). In both cases, the methodological choices reflect the fixed theoretical perspective on epistemic cognition of the models. Importantly, although three major survey instruments (Schommer, 1990; Schraw et al., 2002; Wood & Kardash, 2002) 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), with interview data suggesting existing models of epistemic cognition fail to capture nuance in epistemic-perspectives (Greene & Yu, 2014), and qualitative analysis of open-question survey items taken alongside one measure (Kuhn et al., 2000) indicating incompatibilities between written justifications and survey item responses (Ahola, 2009). Furthermore, alternative methods used have tended to self-report (interview (Barzilai & Zohar, 2012; Mason et al., 2009), and think-aloud protocols (Barzilai & Zohar, 2012; L. E. Ferguson, Bråten, & Strømsø, 2012)) and methods with potential for researcher

subjectivity (particularly interview, and systematic observation (Scherr & Hammer, 2009)

methods). Importantly, as these methods rely on self-report (and the ability to self-reflect), and

researcher interpretation, they offer only partial insight into online, and collaborative (especially

remote) activities – in particular, online information seeking, or information processing more

broadly. The methods described may reflect a broadly 'fixed' view of psychological constructs that

is aligned with the epistemology of some assessment regimes, as indicated in Section 2. In

exploring epistemic cognition, particularly in online contexts, there are therefore a number of

considerations, with a range of advantages and disadvantages, across the variety of analytic

methods, along with underlying assumptions and issues directly related to information seeking –

as Table 2**:8** indicates.

Table 2:8 - 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 cognition is explicitly, consciously, brought to bear on information tasks.	Probing of epistemic cognition 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). Very useful for individual tasks and understanding decision making processes/assumptions.	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).	Novice users especially may struggle with concurrent think-aloud and information seeking (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 cognition is 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 analyses. Useful to model relationships between multiple cognitive constructs and behaviours.	Three major surveys have been developed but are criticised for their psychometric properties (DeBacker et al., 2008). Concerns re: external validity of survey instruments.	Does not account for the co-construction of knowledge, how epistemic cognition is brought to bear on information seeking 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 cognition is something one can directly access, and report on in decontextualised settings	Retrospective interviews to probe epistemic metacognition regarding internet information seeking 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. Useful for in depth understanding.	Interviewer bias and effects (on the individual, and potentially on wider environment). Conducted separately from tasks, thus relationships between claims made and behaviours hard to establish.	

Method	Assumptions	Example studies	Advantages	Issues	Issues linked to use in Information Seeking
Trace data (see particularly Greene et al., 2010) ('think aloud', sometimes thought of as trace data, is presented separately here)	Epistemic cognition is brought to bear in ways that can be meaningfully captured, in particular using technology systems	Analysis of individuals finding, extracting, and moving information (Stadtler & Bromme, 2007) – which could be used to reveal information about their beliefs (e.g. visiting few websites indicates trust in those sites (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 cognition is 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 & dialogue analysis to explore how "professionals' views on knowledge/knowing relate to the understandings gained through collaborative knowledge construction" (p.593) – however, this targeted 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 information seeking 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 appropriate to information seeking tasks given access to log data.

*See also Hofer (2004a, pp. 49–51) for a review of methods. Adapted from (S. Knight, 2012b, pp. 18–19).

One recent development (Bråten et al., 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, in a sample of 157 undergraduate students, has been associated with questionnaire based self-reports of internet-search and communication activities. Subsequent work (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 experimental ISEQ work with 79 undergraduate students (plus one excluded for having high self-reported prior knowledge), used the context of a controlled information seeking task using pre-selected conflicting information (from the internet) regarding a medical issue (Kammerer, Bråten, Gerjets, & Strømsø, 2013). That research 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.

Both findings suggest a clear relationship between epistemic cognition and internet information, with the use of both self-report and trace data methods in this highly relevant study is of interest. However, 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, as discussed in sections 2:1-2:2).

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*

(Y. Wu & Tsai, 2005, p. 379)

Other recent work in developing and deploying search strategy based scales also indicates that: 1) explicit strategies, i.e. actual behaviour tracked through trace, are 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 **2:8**. 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, as I now examine.

2:4.3   Recent Developments – Philosophical Commitments

Perhaps because of the array of research applications, conceptual labels, and methodological approaches – including their various advantages and disadvantages – there has been some concern regarding the nature of epistemic cognition as a construct. For example, some have argued that greater attention should be given to the distinction between epistemic and ontological cognition (Greene et al., 2008; Schraw, 2013), and more generally calls for a philosophical approach to conceptualising epistemic cognition (Alexander, 2006; C. Chinn A.,

Buckland, & Samarapungavan, 2011; Greene et al., 2008; Muis, Bendixen, & Haerle, 2006). This recent work (C. Chinn A. et al., 2011; Greene et al., 2008; Muis et al., 2006) has sought to bring lessons from the philosophical literature on epistemology, into the psychological literature on epistemic cognition.

However, this recent interest has, perhaps in deference to existing models of epistemic cognition, and perhaps also because the literature in philosophical epistemology is expanding rapidly, tended to a narrower focus on philosophical literature than that literature offers. In particular, it has tended to emphasise: classical models of reasoning or evidentiary standards (for example rationalism, empiricism); analytic epistemology, emphasising justified true belief (JTB) and the commensurate need for an account of 'justification'; and individualised accounts of 'knowing' in which the agent is an individual knower, abstracted from their social context. For example, 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 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. 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 that it might draw well on.

In contrast to the models described in Table 2:7 , Chinn, Buckland and Samarapungavan (2011) propose a model with two key differences, "(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).

Thus, 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, pragmatic in nature - see for example, 1999; and Fricker's work, for example, 2012), which has been largely ignored (or, rejected) by epistemic cognition literature. There has, for example, been recent philosophical interest in the kinds of 'good' practices associated with the acquisition of knowledge – epistemic virtues (see, Axtell, 2000; DePaul & Zagzebski, 2003; Roberts & Wood, 2007); and more generally the relationship of 'knowledge' to our social nature and environment (see, Baehr, 2011; Goldberg, 2010; A. Goldman & Whitcomb, 2011; Haddock, Millar, & Pritchard, 2010), including a focus on testimony as a source of knowledge (Fricker, 2009;

Goldberg, 2010; Lackey, 2008; Lackey & Sosa, 2006) and the implications of these developments for education (see, Baehr, 2011; and in particular, Kotzee, 2013).

This interest is motivated by a desire to understand how social ways of 'knowing' can be normative – that is, bound up with the norms of our social group – while still retaining its value – that is, if it is normative, then how do claims that "I know x" still hold weight beyond 'mere' belief (see, A. Goldman, 2010; and, Greco & Turri, 2013). A particular focus of this work has been an interest in how one comes to 'know' through testimonial knowledge, that is, the circumstances under which I might claim knowledge of some thing, because you (an informant) have told me about it (see, Adler, 2014). There are also recent developments to apply such theorising to the use of technologies, prompting questions such as "What do I know when I have pervasive access to an encyclopaedia?" under various theorised accounts of cognition (see for example, Palermos & Pritchard, 2013; S. Knight, 2014).

For the purposes of this work, we may remain agnostic regarding the specific concepts and arguments around their scope. What is of interest here is the more general claim: that claims of 'knowledge' are normative, and that it is only by understanding the social context within which they are made that such claims can be understood. This marks a shift in the understanding of knowledge away from a delineation of its a priori constraints, such as the a priori conditions for 'justification' under a 'Justified Truth Belief' (JTB) model of knowledge. Instead, these approaches focus on a naturalised understanding of the 'function' of knowledge – as a socially deployed, mediated, and communicative construct (see particularly, Craig, 1999)[16]. In the case of virtue epistemology this has particularly focused on the types of intellectual characteristics associated with the reliable production of knowledge. A core interest of much of this broad work has been the analysis of how knowledge standards are maintained at both a micro level (what are the conditions under which I may claim 'knowledge' from your testimony or my experience) and macro level (what are the conditions of 'knowledge' within this epistemic-group). Of course, these

---

[16] Some readers may see similarities here between 'Communities of Practice' (Lave & Wenger, 1991; Wenger, 1998) and the kind of social epistemology of interest here. For one discussion of a Deweyan social epistemology in this light see Garrison (1995).

concerns are of interest in epistemic cognition research too, particularly in its consideration of the conditions under which: authority, personal-knowledge or experience, or corroborative weight, in the sourcing of knowledge; and/or argument structure, evidential kind, or methodological process, in the justification of knowledge, lead to knowledge.

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, for example, P. L. 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). Such a focus accords well with Chinn and Buckland's 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 future research should include a shift to "social aspects of epistemic cognition" such as "argumentation, peer review, media processes" (C. Chinn A. et al., 2011, p. 163) – all of interest to my information seeking literacy context. 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)[17]. 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).

---

[17] 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, for example, the recent paper by Zeineddin & 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 the 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 2:4. Subsequent work (Dong, Liang, Yu, Wu, & Tsai, 2014) indicates that graduate students tend to emphasise content and ease of access more than undergraduates – that they want to access relevant information easily; and that in undergraduate students belief in innate or fixed ability was associated with reliance on ease of access and source-authority, a relationship which did not hold for graduate students. This indicates again the complexity of indicating 'sophisticated' and 'naïve' strategies.

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

Figure 2:4 - A Framework for Information Commitments (C.-C. Tsai, 2004, p. 109)

This development is an interesting one for this learning analytics project, and within the philosophical framework. However, it is still problematic including because, as Wu and Tsai (2005) highlight the information commitments may not be best thought about as a continuum. That is, both 'multiple sources' and 'authority' may be utilised when evaluating the accuracy of the materials on the Web; commensurately both , 'content' and 'technical issues', may be used in the evaluation of the usefulness of Web materials.

2:4.4   Epistemic Commitments & Epistemic Dialogue

2:4.4.1 Dialogue in epistemic commitments

I use the term epistemic *commitment* here, following on from Chinn et al., (2011), Knight, Arastoopour, Williamson Shaffer, Buckingham Shum, & Littleton (2014; see also, S. Knight & Littleton, in submissionb), and Wu and Tsai's description of information commitments (C.-C. Tsai, 2004; Y. Wu & Tsai, 2005) in order to foreground my position that, through our actions – encompassing the discursive stances we take – we commit to particular ways of seeing, and dealing with, the world; our epistemic commitments. This perspective accords with Chinn et al.,'s (ibid) suggestion of 'epistemic commitments' as behaviouristic traits in contrast to tacit 'beliefs'; an approach well aligned with learning analytics. Core to my approach is a move to see epistemic cognition as a co-constructive enterprise – a shift from individual activity and cognition, to understanding epistemic commitments as a function of small group activity (S. Knight & Littleton, in submissionb).

I argue that dialogue is an important consideration both in terms of supporting literacy – encompassing collaborative information seeking and epistemic cognition (see section 6 and S. Knight & Littleton, in press, 2015b) – and in terms of probing those constructs, offering a lens onto students' understanding of their own learning. I base this claim on the relationship between literacy and dialogue (see, for example, Snow, 2002), and that document use may often involve spoken or written communication (see Rouet & Britt, 2011, p. 26 and the sections on dialogue in collaborative information seeking), dialogue is an important area of interest in both supporting and probing complex literacy practices in collaborative information seeking. For example,

although Rouet and Britt devote little space to the issue, they note that much document-based activity involves collaboration through communication of various forms, thus: "Human agents may provide information that complements the documents, and they may also participate through dialogue, question answering or coaching", (Rouet & Britt, 2011, p. 26). This is also reflected in a special issue on the topic (see S. R. Goldman & Scardamalia, 2013) in which only one paper (Jucks & Paus, 2013) considers the role of discussion in dealing with multiple documents. As Goldman and Scardamalia (2013) note, dialogue is key in collectively authoring written outputs, particularly around "constructive uses of authoritative sources," that engages students both in understanding what is being claimed, and how to contribute to developing new knowledge (S. R. Goldman & Scardamalia, 2013, p. 260). They argue that we need two foci:

1. Productive use of metadata and meta-discourse – By metadata they mean the use of credentials, dates, source locations, quote or paraphrasing, citations, commentary on primary/secondary sources and so on. By meta-discourse they mean to highlight the ways in which discourse around this metadata becomes an object for discourse itself – meta-discourse
2. Use of authoritative sources – By this they mean the stating of claims, and citing of sources, with a focus on discourse for knowledge building through improving and developing ideas

In this perspective, we can see that an epistemic commitment might be taken from treatment, or mention, of particular sources. But moreover, we see that "constructive" dialogue requires an openness to ideas, a willingness to co-construct common knowledge – and that openness may be taken as an epistemic commitment too. Thus the first of these foci is closely associated with the kind of exploratory dialogue introduced below. The co-occurrence of this kind of meta-discourse with epistemic actions – the discussion around and selection of sources, the evaluation of claims, the creation of inter-textual ties, and so on – is indicative of particular ways of seeing knowledge

Such considerations align well with a perspective of language "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 & Howe, 2010). "Wherever education is

taking place, commonality – a shared perspective – is key, and dialogue is the tool used to create

such a perspective (Edwards & Mercer, 1987)" (S. Knight, 2013b). 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, where:

- Disputational talk is "Characterised by disagreement and individualised decision making. There are few attempts to pool resources, to offer constructive criticism or make suggestions." And language is characterised by "short exchanges, consisting of assertions and challenges or counter-assertions ('Yes it is.' 'No it's not!')."
- Cumulative talk in which "Speakers build positively but uncritically on what the others have said. Partners use talk to construct 'common knowledge' by accumulation." (Mercer & Littleton, 2007, p. 58) And language is characterised by "repetitions, confirmations and elaborations." (Mercer & Littleton, 2007, p. 58)
- And exploratory talk in which "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) And language is charactised by instances of exploratory terms and phrases– for example, 'I think' 'because/'cause', 'if', 'for example', 'also'.

Similar characterisations of effective dialogue have emerged from the work of other researchers

across a range of ages of learners (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*.

The target of my interest, then, is multiple-document processing tasks in which students collaborate on the processing of a range of sources, in order to create an output document, and particularly tasks in which we – as researchers – have access to chat and document-logs. In line with this argument that collaborative dialogue is of key interest to us, we suggest that the connections between trace-indicators of epistemic cognition and in particular, the kind of linguistic expressions associated with taking an 'epistemic stance' (Kärkkäinen, 2006), some of which (e.g. 'because', 'I think', 'so') are also associated with the kind of educationally productive dialogue known as 'exploratory talk' (Mercer & Littleton, 2007) or accountable talk (Michaels et al., 2002; Resnick, 2001) are key. These terms include: 'I think', 'he' or 'she' said, 'I don't know' 'I guess', 'I thought', epistemic adverbs such as 'maybe', 'probably', 'apparently', 'of course', and epistemic modal auxiliaries such as 'would', 'must', 'might', 'could', 'will', 'may'. Such stances indicate a linguistic positioning of the speaker(s) with regard to their linguistic target. It is thus that the consideration of dialogue data (including via chat systems), may inform the understanding of other trace data to understand the ways in which people engage in: selecting multiple sources; claims around source authority; connecting pieces of information in complex ways; and so on.

That is, analysis of epistemic commitments should not solely focus on the number (or even type) of resources opened, but on references to source metadata (authorship, publication date, publisher, etc.), alongside identification of connections made between corroborative and authority-identification behaviours, particularly in the information-oriented dialogue data. Students whom rely on authorities without corroborating, or those who look for repetition of information primitively 'corroborating' both engage in less sophisticated behaviour than those who corroborate by using authoritative sources (see S. Knight, Arastoopour, et al., 2014 for a preliminary description of this potential). Thus, with an information seeking session, understanding the proportion of resources students visit or make use of may be of interest, but it

is particularly informed by how they use source metadata (authorship, publication, etc.) and the kinds of metadiscourse markers I associate with taking an epistemic stance.

2:4.4.2 Evidence for dialogue in epistemic commitments

Discursive Psychology is a psychological perspective sharing some history with the sociocultural perspective outlined above (including "Common Knowledge" Edwards & Mercer, 1987) – having a particular interest in "the kinds of naturally occurring interactional talk through which people live their lives and conduct their daily business" (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 thesis. 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" (Edwards, 1993, p. 208).

Alongside sociocultural discourse analysis (Mercer, 2004), which uses a similar theorising to approach language as a cultural tool for learning, discursive psychology is of particular interest because, it has focussed on the respecification of commonly held psychological constructs in terms of their linguistic, situated, co-creation. That is, in respecification, we see a shift from psychological constructs as neural or cognitive entities, to a focus on the emergence and fluidity of psychological constructs "in action" as co-constructed in, and mediated by, language. Thus, for a specific analysis of a construct – epistemic cognition – this approach may be highly appropriate.

In the context of epistemic cognition, discursive psychology posits that we should not see cognition and dialogue or language more broadly 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 discursive perspective – which shares a view of language as a tool 'to do' with sociocultural theory – can be combined with Hammer and Elby's 'resources' model. In such a combined model, epistemic cognition is should not be viewed as fixed, nor domain-specific, but rather as dependent upon the resources – the mediational tools, dialogue, and interaction with the (social) environment – available. This perspective thus emphasises "theory-in-action" – in which context, domain, culture, and task conditions interact. Indeed, such context sensitivity is key to the resources model, and understanding sophistication in epistemic cognition:

> *A sophisticated epistemology entails context-sensitive judgements. Thus they [(Hammer & Elby, 2003)] 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?" (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). Therefore in the context of exploratory information seeking tasks, in which knowledge is co-constructed and negotiated between resources (webpages, etc.) and information seekers, our focus should not just be on 'correctness' or truth but, "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). This stance accords well with the position on high stakes assessment, and a need for a more situated approach to assessment practices, in the introduction.

This need for understanding the situated context is true of both assessment structures broadly, and student's own framing of assessment exercises. 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). Fundamentally, student's talk and interaction while navigation resources on the internet during formal-education web-based inquiry tasks:

constitute a possible entrance for understanding how they actually make sense of and employ prompts as structuring resources in their learning processes. During interaction, participants constantly make meanings and interpretations of situations, events and actions visible and observable to other participants as well as for us as analysts (Linell, 1998; Mercer, 2004)

(Furberg, 2009, p. 400)*.*

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 & Lin, 2010 for evidence in teacher trainees). 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') is given in contrast to a more absolutist perspective. For example, they report a student (Bekah) offering an equation and explanation to illustrate her understanding of some concept. This explanation is taken up collectively and referred to in the dialogue as "Bekah's Law", illustrating – in the terms described above – a cohesive tie (the repetition of terms through a text), and demonstrating a type of common knowledge built up in that classroom dialogue (Hutchison & Hammer, 2010).

This claim – that exploratory and accountable types of dialogue are epistemic in nature – is further supported in Rosenberg, Hammer and Phelan's work (2006). In that study, a case study of a 15 minute discussion of the 'rock cycle' was presented from a group of 8th graders; this presentation againmade use of dialogue excerpts (as in sociocultural analyses) 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, in which they make the (epistemic) suggestion that the students might build on their own knowledge, they observed a shift in the dialogue to be more productive, emphasising coherence and understanding in their creation of theory using terms that they understand. The description and excerpts provided by the researchers, indicate that the dialogue they observed could be characterised as 'exploratory' in nature.

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
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

Moreover, they agree with the (pragmatic) Dewey (1938b) 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. As Acosta (2013) notes, a focus on attempting to reach a conclusion – as a performative goal involving more 'cumulative' styles of talk – versus attempts to consider all evidence and come to a shared decision (which we would expect to parallel use of exploratory dialogue) should be seen as epistemic stances in complex collaborative in inquiry problems.

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). Importantly, though, their identification of an association between exploratory dialogue and more sophisticated epistemic stances lends support to my position.

Furthermore, a study exploring associations between the ISEQ psychometric described above (section 2:4.2), and self-report internet learning behaviours 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 for use of communication strategies in online learning (across participant responses, M = 2.51 rating, SD = 1.03) – importantly, with less sophisticated epistemic cognition (belief in knowledge as given and stable):

> *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, 2006b, p. 1038)

This finding accords well with other self-report survey research (Teng, 2010) indicating an association between less sophisticated epistemic cognition and lower level contributions in an online community of practice platform. Indeed, as Tsai (2004) notes, the internet's affordances for discussion, alongside information seeking, may also provide a useful tool to help advance student's epistemic commitments (see also discussion of dialogue around "the given" in S. Knight & Littleton, 2015b). This role of dialogue for epistemic commitments in online learning is important, given the literature discussed in the context of collaborative information seeking (section 2:2), and the importance of dialogue for learning (section 2:4.4).

2:4.4.3 Trace data – a path to epistemic commitments

Earlier I noted Mason et al.'s (2009) claim (Table 1:1) that across models of epistemic cognition, there was a focus on the certainty, simplicity, source and justification for knowledge. Above, Tsai's (2004) framework for information commitments was introduced, comprised of: (1) standards for correctness; (2) standards for usefulness; and (3) searching strategy. We can compare this with Rouet's claim regarding the skills students need for mature internet use (noted in section 1:3):

1. *Skill of **integration**, ability to establish connections between prior knowledge and new information, including across documents, and including where claims are inconsistent or contradictory*
2. *Skill of **sourcing**, ability to identify parameters that characterise the author and conditions of production of the information*
3. *Skill of **corroboration**, or the activity of checking accuracy of information against other sources*

(Rouet, 2006, p. 177 emphasis added)

I am particularly interested in the class of constructs related to how students engage higher level literacy skills of information selection and evaluation, creation and identification of ties between and within documents, and development of outputs based on these activities which might be more or less elaborated in their form. Across the literature discussed, a number of key observable behavioural indicators of interest can be identified:

> *1. How students select information through corroboration and reference to source-authority, and how these strategies are used in isolation, or combination – for example by corroborating across multiple sources, whilst making reference to the qualities of those sources.*
>
> *2. How students connect claims across and within sources whether claims are considered and stated in isolation, or integrated and synthesized while seeking information, and creating output texts.*
>
> *3. How students take claims and use them in task oriented ways; whether claims are stated without evaluation, or are evaluated and elaborated.*

(S. Knight & Littleton, 2015a, p. 2)

These three elements can be seen in light of epistemic commitments. Therefore we take it that when sourcing information – through selecting individual or multiple sources – those selections should be taken to as commitments to authority and corroboration, analysed in connection with student's linguistic 'stance taking' towards these actions. Such sourcing does not stand alone; it is embedded in and connected to the continued seeking of information, extraction and synthesis of claims, and deploying of that information in task-specific contexts. For example, through trace indicators such as logs of document use, or identification of key markers linking claims to their sourcing documents, we might identify that a particular claim has been sourced from a document; in such cases, it is of interest to also identify whether or not sourcing metadata (dates, authorship, genre, etc.) has been discussed or not.

This situated perspective on behavioural indicators highlights the importance of attempting to understand the semantic significance behind particular moves in a given environment. As Winne notes:

> *trace data operationalize what learners do as they do it. Trace data avoid shortcomings of (a) asking leaners what they believe they do and (b) asking learners*

> *to perform mental calculations of unknown kinds (c) using sample fractions of past or possible future experiences that have unknown size and biases. When traces are faithful operational definitions of theoretical cognitive and metacognitive operations, they provide sturdy grounds for testing theories about when, whether, and how [self regulated learning] processes affect learning.* (Winne, 2010, p. 275)

In the context of epistemic commitments, we can conceptualise trace behaviours, within designed tasks, that can be understood as semantically salient. 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, Bråten and Samuelstuen (2007) tracked highlighting and similar behaviours in document processing – which could be used to explore selection of salient claims across documents – and as Greene, Muis and Pieschl note visiting fewer websites may indicate trust in those websites visited(2010). As Greene et al. (2010) point out, 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 in interaction with technologies, the ways in which epistemic cognition is brought to bear on knowledge tasks may be meaningfully captured, (for example, through the ways in which people represent knowledge in mind mapping tools). Trace data thus offers direct access to real-time behaviours in unobtrusive ways, and is thus high in external validity. In conceptualising the salient construct as centred on 'commitments' as observed in actions, there is a shift from more generalised and developmental models of epistemic cognition 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. In developing lines of research this shift has methodological often – but not always – involved the use of trace data; particularly important given that observed explicit strategies are both better related to learning-performance, and poorly related to self-report of such activities, in search tasks (M.-J. Tsai et al., 2012). Significantly for my purposes, trace data may offer insight beyond task-specific self-report; for example,

Bråten and Samuelstuen's (2007) study of tenth grade students (n = 177, mean age 15 years 10 months) indicated that tracking students' highlighting of key terms was more predictive of student's performance on a reading-task (which is predictive of PISA literacy tests) than their self-report of reading-task practices (Bråten & Samuelstuen, 2007).

The need for such analysis is especially pressing given that "Little of the work on online credibility assessment has considered how the information-seeking process figures into the final evaluation of content people encounter." (Hargittai et al., 2010, pt. abstract). In that study, a sample of 1060 18-19 year olds on a self-report pencil/paper survey was used to gain a gender-skill stratified random sample of 102 participants for an observed task asking participants to comment as they searched. Analysis of recordings indicated that "the process of information seeking is often as important as verifying the results when it comes to assessing the credibility of line content" (Hargittai et al., 2010, p. 479). In addition, only 10% of the participants made remarks about the author or their credentials, and none actually verified those credentials, while superficial cues such as search rank and domain names (for example, .edu, .gov) were seen as credibility indicators.

This is particularly interesting given that, as Kobayashi (2014) indicated through a controlled experimental design, while participants given two texts of varying quality are more likely to favour high quality sources, they make little reference to source features (on average only 1.85 out of 10 features); and rarely (<6% of the 154 participants) explicitly use source information for justifying their evaluation of the text's explanation, that is, they do not make connections between source metadata and their evaluative stance. Moreover, Barzilai and Eshet-Alkalai (2015), using a similar controlled experimental design, demonstrate that students with more sophisticated (evaluativist) epistemic-perspectives have higher comprehension of conflicting sources. Such behaviours may also be mediated by topic, and topic knowledge as indicated by Bromme, Scharrer, Stadtler, Hömberg, and Torspecken (2014) who found that students were more likely to attend to citations (although not methods cited) and to report these as justifications in their texts when reading contrasting claims regarding a more scientifically based unknown topic (cholesterol) than a socio-

scientific one (climate change). Furthermore, in a correlational study, Kammerer, Amann and Gerjets (2015) find that those scoring higher on the 'justification' ISEQ component also spend longer on more 'objective' websites when searching for information on a novel health based topic.

A related experimental study, (Salmerón, Kammerer, & García-Carrión, 2013) developed a 'fake' search engine results page with fewer (experimental manipulation) or more (control) topically relevant pages displayed in the top search results. In that study, dwell time, links opened, and bookmarked pages while reading for a task on a socio-scientific issue indicated that students tended to employ a 'top link' heuristic – trusting search engine results page ranking as an indicator of credibility – while searching web pages for a socio-scientific topic, and that their background knowledge only played a role in the normal SERP, but not when less relevant documents were artificially placed at the top. They did, however, assess topic relevance and trustworthiness of web pages when deciding which to bookmark for subsequent reading. Recent large-scale log-data analysis (White, 2013) exploring the ways search engine and searcher bias relate, also indicates that in health based search, participants favour positive (i.e., affirming a claim) over negative (i.e. denying a claim) results, and that search engines are more likely to display such positive results, despite the fact this meant around half of answers searchers settled on were incorrect.

Across this range of studies we see the emerging use of behavioural trace indicators to explore psychological constructs, including a small number of epistemic cognition and information seeking. However, the numbers of such studies are small, and none have focused on epistemic cognition in the context of collaborative information seeking and document processing. The following section (2:4.5) describes further relevant work in use of behavioural trace data drawn from digital activity-logs. Section 2:5 then builds on this work and the theoretical account in section 2:4.4 to offer a novel methodological approach for learning analytics around epistemic commitments.

2:4.5 Learning Analytics – Operationalising Constructs To Behavioural Trace

In addition to a body of work making claims relating epistemic cognition to observable behavioural differences, there is a body of research which has explored approaches to track and highlight the salient features of web navigation. For example, 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 et al., 2008; Tseng et al., 2009). This research approach 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 more sophisticated prior 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)[18]. 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 by tracking when questions were first answered, and classifying each subsequent page as a 'refinement' source).

---

[18] 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.

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 2:1.2.1 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 may be a circularity in the assumptions such that analysis of the NFM appears to have been used both to derive groups (match v. exploration) and for 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; 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, p. 246).

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 by using surface features (Dimopoulos & Asimakopoulos, 2010).

A final study (Hsu et al., 2013) sampled 42 undergraduate and graduate students for a laboratory search task from a pool of 240 students who had completed a Scientific Epistemic Beliefs (SEB) survey. These participants were first asked to read two competing articles on a scientific dispute, and then asked to justify which they trusted more, and if their position changed (and why) during searching for justificatory material. The participant log data was recorded with a coding scheme of online behaviours giving a code to various acts. Students with higher SEBs were more likely to: show bi-directional sequences of 'query-results browsing'; and results browsing involving viewing more than one page of search engine results, than those with low SEBs. High SEB students were also more likely to use the 'back' button to browse earlier information. They thus conclude that high SEB students display more advanced search behaviours.

In each case, insight is gained through analysis of log data. Of course, such insights are within the context of more or less involved systems designed to capture such information – from unobtrusive collection of log data, to visualisation techniques to foreground navigation patterns to information seekers as they search. Furthermore, while trace data is often unobtrusive in collection, it may give an incomplete picture. For example, the reasons for some behaviours may be challenging to probe using such data; "these reasons may range from epistemic (sometimes

accepting the 'top' hit is the most sophisticated action to take), to practical (ICT failures), to pragmatic (the demands of the task place a short time restriction on the activity)" (S. Knight, Buckingham Shum, et al., 2014, p. 36). Thus, it is important to remember that analytics regarding epistemic cognition may be – at best – a dirty lens onto that cognition.

2:5 The Promise of Learning Analytics

The preceding sections have foregrounded the relationships between literacy practices (multiple document processing and information seeking) and epistemic cognition. In the latter parts of the literature review (particularly sections 2:4.4.3 and 2:4.5 above) I have highlighted analytic techniques for our construct of interest. In doing so I have tried to make it apparent to the reader that there is potential in exploring behavioural traces in the study of epistemic cognition.

In engaging students in tasks to specifically encourage explicit conceptual structuring, discussion, or writing by learners, using the data from these tasks to claim insight into what they are doing, it might be argued that we simply reify the constructs we set out to explore. That is, if we wish to investigate epistemic cognition (or, commitments), and we design tools to push students to make these commitments explicit, then it may be the case the students do not, in fact, have underlying epistemic cognitions – the tool simply forces them to actions that we reify as constructs. While for psychologists who wish to uncover underlying cognitive-constructs this may be problematic, I do not see this as a concern for this project, because my stance is oriented around epistemic commitments as "theory-in-action". In this view (as explicated in the introductory sections), the claim is not that it is not possible to measure cognitions, but rather that (as the sociocultural account outlined in section 1:1.4 describes), the activity and its (socially, discursively) mediated context is fundamental to the practices being observed when we take measurements, and as such fundamental to the ways in which cognition is instantiated in action; thought and action are, in this view, fundamentally bound up in their situated, tool-mediated, and social environment. This point also places a different kind of burden on individual students, such that the expectation is not that all are equally equipped in the *acquisition* of knowledge, but rather that they understand their behaviour and interactions with their environment, 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).

Indeed, the kind of dialogue introduced in section 2:4.4 is associated with student's developing faculties around their own knowledge construction abilities. Thus, learning analytics provides an alternative method to approach epistemic *commitments*, without recourse to a decontextualized view of epistemic cognition instantiated by questionnaire methods.

In his 2013 review of the literature, Schraw summarises the salient literature in epistemic cognition, noting 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)

This review of the literature has chiefly sought to probe the first of these claims, 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. 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).

As a result, this review of the literature has suggested a focus on information seeking tasks in the context of the open web, or across multiple documents. I have drawn parallels between internet based information seeking and multiple document processing (see particularly Figure 2:2), and highlighted the potential of CIS both as a pedagogic tool, and a lens onto epistemic-commitments. In developing this alignment between pedagogy and assessment I have sought to address the concerns in the introduction (see particularly Figure 1:1) regarding the relationships between assessment epistemology, assessment, and pedagogy. The methods chapter will describe the particular choice of two tasks (an MDP and CIS task) for this thesis work, choices which are grounded in the theorised account described above.

This work will specifically make use of one of both CIS and MDP tasks. Analysis of both types of task allows both the flexibility that CIS tasks provide, with their focus on open-web search; and the potential for control in MDP tasks in which the researcher is aware of the topics, sources, and inter-textual relations among the documents provided. Given, as I argue in section 2:1.3 both MDP and information seeking models can be seen within the same general model – involving information seeking, selection, and integration – it is important to develop research around both kinds of task, to better understand their relationships. This is particularly true given, as highlighted in the introductory sections (particularly 1:1.5) there is increasing pressure to use internet based sources and 'search' in contrast to pre-selected resources.

In considering information seeking processes, section 2:4 foregrounded a particular stance on epistemic cognition, noting the prior literature around metrics in information seeking and epistemic cognition tasks. Drawing on that literature, the next section will describe the metrics to be used in this research. While a broader range of analytic devices are available to us as researchers, those described are grounded in prior work, and establishing their relationship to epistemic cognition is an important contribution. In the following section (2:5.2) I discuss another promising, and pedagogically motivated, means through which learning analytics research has

been conducted – peer and self-assessment – highlighting its usefulness to my own research agenda. In the final section (2:5.3) I introduce the research questions for this work, aligning them with the literature reviewed.

2:5.1 Epistemic Commitments in Information Seeking for Learning Analytics

Above (section 2:4.5) I have highlighted some of the trace data used in the kinds of task of interest to my research. In section 2:1.2.2 (and briefly 2:2.7) I noted the complexities of evaluating search success in information seeking contexts. This is particularly true for CIS tasks, where evaluation might focus on the 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 metrics 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 tasks.

Drawing on the research outlined in this literature review, a set of metrics can be drawn from the (smaller number) of studies explicitly making use of trace data in their analysis. Across a broad set of studies (for example, C.-C. Lin & Tsai, 2007; Salmerón & Kammerer, 2012; Shah & González-Ibáñez, 2011; Shah, Hendahewa, & González-Ibáñez, in press) the viewing of pages has been highlighted, often along with some metric of 'page use' indicated by explicit marking actions by a user (for example, bookmarking) or implicit actions (for example, citing the page, copying text from the page). In addition, the number of search queries made (Shah & González-Ibáñez, 2011; Shah et al., in press), the 'depth' of query probing (i.e. various measures around searchers going past the first few results into subsequent pages) (Hsu et al., 2013; C.-C. Lin & Tsai, 2007; Salmerón & Kammerer, 2012; Shah & González-Ibáñez, 2011; Shah et al., in press), and the number of keywords (C.-C. Lin & Tsai, 2007; C.-H. Wu, Hwang, & Kuo, 2014) and number of keywords in relation to number of queries (Yue, Jiang, Han, & He, 2012) have been reported as recent analytic devices. Other metrics reported have included: time spent on pages or search engine results pages (Salmerón & Kammerer, 2012; C.-H. Wu et al., 2014); sequence analysis (for example analysis of patterns of query->page, or page->page navigation) (Hsu et al., 2013); page revisiting

(C.-C. Lin & Tsai, 2007; C.-H. Wu et al., 2014); and analysis of pages visited but not used (C.-H. Wu et al., 2014). Building on some of these measures, a set of metrics has been composed (C.-H. Wu et al., 2014) around the 'symmetry' of collaborative user activity, indicating whether one partner in a pair is contributing more to any particular metric than another – for example, whether one is adding more bookmarks than the other.

Finally, In the context of evaluating search success, Shah (2012b, Chapter 7, and more recently, 2014b), notes that research in the area has explored and utilised measures such as: asking collaborators to rate each other's collaborative or information seeking capabilities; qualitative descriptions of information seeking processes (providing a narrative for a search episode); and other more traditional 'information retrieval' (IR) metrics. These IR measures of success include precision (a measure of 'noise' in a set of retrieved results, such that lower precision indicates some of the documents retrieved were not relevant), recall (a measure of coverage of retrieved results, such that lower recall indicates relevant documents were missed) and F-measures (which combine precision and recall to give an overview metric). 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, 2012b). Similar methods have been used to 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. However, 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, commonly used in search-research and noted by Shah, is to take queries as the object of inquiry, and to explore the 'diversity' of queries issued by computing their

'Levenshtein 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).

Across these studies then, a broad set of metrics have emerged as objects of inquiry in their own right (rather than as components of other metrics), as indicated in Table 2:9. While the indicators in Table 2:9 are drawn from prior research, their application to learning analytics for the construct of epistemic cognition in a CIS environment is novel. There is strong potential for the application of such techniques to analysis of connections across these metrics and the kinds of epistemic commitments characterised by claim selection, corroborative-evaluative sourcing and synthesising.

Table 2:9 - Metrics in epistemic information seeking tasks

Metric	1	2	3	4	5	6	Used to understand…
Page use metrics*	●	●	●	●			Breadth of inquiry
Bookmarks, snips, etc.*	●	●	●				Discriminatory power
Metrics of pages viewed but not used*	●	●					Discriminatory power
Page revisiting		●		●			Depth of reading
Temporal sequence	●					●	Inquiry processes
Durations (of page/query views)		●	●				Depth of reading/inquiry
Query depth*			●	●		●	Depth of inquiry
Query vocabulary indicators*				●	●		Depth and breadth of inquiry
Query counts*	●	●					Breadth of inquiry
N of modifications made in answer	●	●		●			Synthesis and editing processes
Collaborative symmetry metrics*		●					Division of labour
Coded chat*	●						Collaborative process

1 (González-Ibáñez, Haseki, & Shah, In press; Shah & González-Ibáñez, 2011; Shah et al., in press)
2 Wu, Hwang and Kuo (2014)
3 (Salmerón et al., 2013)
4 (C.-C. Lin & Tsai, 2007, p. 689)
5 Yue, Jiang, Han and He (2012));
6 (Hsu et al., 2013)

As discussed in section 2:2.5.2, some work has involved analysis of chat in CIS contexts, including (as in Table 2:9) that of González-Ibáñez, Haseki and Shah (In press). The focus of that work was division of labour and affective components of learning processes. However, the use of chat data as a data source for analysis is productive for my work, and can be adapted to the particular constructs of interest. Specifically chat terms can be identified which address the model of epistemic commitments highlighted above which has a focus on: Claims being made, the metadata associated with those claims (building on Goldman and Scardamalia 2013), and the

metadiscourse around them, specifically the kinds of exploratory dialogue used indicative of active processes of constructing common knowledge, and developing new knowledge.

The potential of learning analytics, then, is in developing trace indicators for processes of tool mediated activity, including indicators for the kinds of practices of social participation in enquiry learning and the development of epistemic identity (considerations in socioculturally grounded design of learning environments; see section 1:1.4). The indicators described briefly above (and outlined further in section 3:4.3 in the Methods chapter) are a starting point for development, using established measures in a novel context. The novel context is a pedagogically grounded, collaborative, epistemic information seeking and processing task with a focus on particular kinds of dialogue practices, and authentic tasks (considerations in socioculturally grounded design of curricula content; see section 1:1.4). Through this analysis, a kind of assessment of enquiry-practice participation can be developed (the 1st assessment consideration; although of course, this offers only one lens onto the kinds of behaviours engaged in, the 3rd assessment consideration – see section 1:1.4). In addition, bringing students into active participation in their own assessment (see section 1:1.4) offers potential in terms of a pedagogically and theoretically motivated task design, and in particular holds potential benefits in the learning analytic context, as I now discuss.

2:5.2   Peer and Self-Assessment for Learning Analytics

Within the context of learning analytics, one area of significant potential for understanding student learning outcomes is in understanding how to develop effective peer assessment models for large scale contexts such as massive open online courses (MOOCs) (Dillenbourg, Fox, Kirchner, Mitchell, & Wirsing, 2014). Indeed, the potential to connect learning outcomes from such assessment to behavioural trace as proxies for learning is strong, particularly where outcome facets may be related to particular behavioural traces – an approach which will be adopted in this thesis, to connect specifically epistemic outcomes to trace indicators.

Moreover, there is potential for MOOC peer assessment to act not just as a method of assessment, but as a formative learning activity in its own right (R. Ferguson & Sharples, 2014). As

such, one means to deal with a large number of student-generated outputs, and gain further insight into their qualities (in this case, qualities related to epistemic commitments) is through peer and self-assessments. Indeed, peer and self-assessment is identified as a key design component in fulfilling the pedagogic potential of massive open online courses (See, for example, Admiraal, Huisman, & Van de Ven, 2014; Kulkarni et al., 2015; O'Toole, 2013; Suen, 2014), learning analytics (See, for example, Klamma, 2013; Redecker & Johannessen, 2013; Sharples et al., 2014), and e-learning contexts more broadly (See, for example, Whitelock, 2010). Indeed, practices of peer-assessment – the evaluation of a peer's work for formative or summative purposes – have been an area of interest for some time (for an early review, see K. Topping, 1998) with a recent review (Gielen, Dochy, & Onghena, 2011) describing an expanding focus of interest on peer-assessment methods and a special issue (Strijbos & Sluijsmans, 2010) focussing on understanding the methodological, functional, and conceptual issues in peer assessment.

Across this work we see broad benefits to the use of peer-assessment. However, conducting peer and self-assessment reliably, and for learning gain, is a challenge. There is a need for well-designed pedagogic approaches to such assessment, and multiple approaches to developing reliability of results. In a study of 708 students across 16 courses at 4 universities evidence suggests that an aggregation of at least 4 peer ratings can be as reliable and valid as instructor ratings (K. Cho, Schunn, & Wilson, 2006), suggesting the benefits of peer-assessment for gathering reliable assessment data, under some conditions. One method to (a) increase learning of both assessors and assessed and (b) increase accuracy of peer and self-assessments is 'calibrated peer review' (Carlson & Berry, 2003; Robinson, 2001), recently described at scale for large classes (Watson & Ishiyama, 2012) and mooc contexts (Balfour, 2013). This method uses a kind of diagnostic-training stage in the review process such that students:

1. Complete the written task they will be assessing
2. Assess three assignments using a rubric for which there are reference values (i.e., they have been previously assessed by the instructor)
3. Receive feedback on their assessment of the training/diagnostic assignments in (2), and are assigned a 'calibration' score (which indicates their quality as a rater, based on distance of their reviews from the reference values)
4. Assess peer assignments using the same rubric

5. (Optional) Self-assess their own assignment again using the rubric

This method thus has the advantage of facilitating peer and self-assessment with both training (to support student understanding of the rubric) and a diagnostic (again to support learning, and to identify high quality assessors). There is strong potential of such learning analytic techniques to develop effective models of peer assessment at scale, with pedagogic value (Piech et al., 2013); its application in this thesis contributes to understanding this potential and its connection to other trace indicators.

2:5.3 Developing Learning Analytics for Epistemic Commitments in a Collaborative Information Seeking Environment

The research described in this thesis tackles the research aim *'Investigate patterns of information seeking as epistemic processes'.* Through meeting this aim, the ways in which people interact with a particular technology, in a particular (collaborative) setting, on a designed epistemic task are investigated. The preceding literature has addressed this aim by outlining:

1. My theorised account of the nature of learning analytics introducing the notion that analytic devices might be viewed through the lens of 'Assessment, Epistemology, Pedagogy' (section 1:1)

2. I used this theoretical approach to argue – drawing on the middle space notion – that learning analytics can most productively be developed when the development of the analytic techniques makes use of multiple disciplinary perspectives, grounded in learning science constructs. Epistemic cognition associated with literacy skills was noted as a particularly important construct for such analysis (section 1:2).

3. Following an introduction to the research described in this thesis (section 1:3), I then outlined in section 2:1 the importance of information seeking, offering a novel alignment of information seeking to models of literacy around multiple document processing. In particular I highlighted that 'exploratory search' (section 2:1.2.1) 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 2:1.2.2) – relating this claim to the epistemological stance advanced in sections 1 and 2.

4. However, I also noted the increasing interest in collaborative information seeking – with a growing body of evidence indicating the incidence and educational potential of such activity (section 2:2). This section (2:2) thus outlined some additional factors of

significance in CIS, including awareness and communication. CIS was introduced as a salient context for epistemic cognition because: it naturally occurs in educational contexts; aspects of it, including dialogue and shared awareness tools, are related to features of educational contexts including improved educational outcomes; it provides a lens onto information seeking processes, giving insight into for example the type of language collaborators use in CIS tasks.

5. Section 2:3 then foreground the emerging literature connecting epistemic cognition in information seeking. I noted in that section, and section 2:4, that the epistemological approach advanced in the introductory sections could be drawn on to develop a new social account of epistemic commitments. This account drew a novel alignment between evidence regarding dialogue in CIS (section 2:2.5.2), the epistemic context of information seeking, and methodological and theoretical perspectives for analysis of epistemic cognition.. Section 2:5 thus presented some work towards this end, highlighting the potential of learning analytics for insight into epistemic-commitments in CIS and MDP processes. This potential is realised through: (1) the investigation of established indicators in information seeking tasks and an analysis of dialogue aligned with the theoretical perspective presented and (2) the connection of these behavioural traces to a peer and self-assessment model designed for pedagogic gain and insight into learning outcome proxies which may be aligned with the behavioural indicators.

The claim of this thesis is not just that we should be interested in systems to mediate help-seeking behaviours of information seekers (see for example, Puustinen & Rouet, 2009) but that collaborative interaction has potential to improve information access and that information seeking provides a good lens onto epistemic commitments in action. I see the potential of exploratory information seeking tasks thus:

> *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, Betrancourt, & Détienne, 2012, p. 75)

This thesis will address this potential, by developing novel information seeking tasks to probe epistemic commitments, using a collaborative context and behavioural trace from that context as

an analytic lens onto student's epistemic commitments. Development of this approach will make use of a range of analytic devices, including the behavioural trace data, including dialogue data, described above (section 2:5.1), and peer and self-assessment (section 2:5.2). Analysis will focus not only on the application of established metrics, but a closer discussion of the processes engaged in by participants in the development of a written output, including a discussion of their dialogue data. In addition, contextual information on the participants will be collected, including established methods for the exploration of epistemic cognition (in particular, a psychometric instrument). This information provides salient context to aid in understanding how different participants engage with the tasks, as well as giving a clear indication of the relative benefits and disadvantages of the use of survey data in contrast to analysis of behavioural data. Thus, the research context is explored through the following research questions:

1. Can epistemic markers of selection, evaluation, and integration of claims in a written output from a designed information seeking task be reliably identified?
2. What relationships can be identified between facets of epistemic cognition on a psychometric instrument, and written outputs from a designed information seeking task (as in '1')
3. What is the relationship between trace behaviours on an information seeking task, epistemic properties of written outputs (as in '1'), and epistemic cognition measured on a psychometric instrument (as in '2')?

The use of peer and self assessment with a diagnostic training provides a pedagogically grounded scalable approach to developing outcomes data in learning analytics research; as indicated in the first question. The second question provides, to the best of my knowledge, the first attempt to connect to epistemic outcomes in written outputs to structures on a psychometric instrument. The final question brings together the association between learning outcomes (as identified in students' written outputs), student characteristics (through the psychometric instrument) and behaviour traces, analysed within the frame of the novel theoretical and methodological account provided above.

The thesis addresses these questions through the use of two collaborative tasks – a multiple document processing task (MDP) with assigned documents; and an information seeking task (CIS) in which students search on the internet for resources. Alongside these two tasks a set

of survey instruments, including the ISEQ described in section 2:4.2, are implemented.

Learning outcomes are identified through a pedagogically motivated peer and self-assessment task targeted at epistemic properties in the student written outputs. Analysis focuses on the insights information seeking analytic devices give into student success and epistemic commitments. The following methods section will describe this work in detail.

Chapter 3:   RESEARCH DESIGN AND METHODS

3:1 Design and Context

The primary empirical research described in this thesis took place in the weeks 06/10/2014 to 18/10/2014; the first week consisting of a 'lab based' session, and the second (13-18th) involving an 'at home' task.

The study took place at the Maastricht University School of Business and Economics, during skills sessions for a first year Quantitative Economics class. This school is highly selective, with a strong international representation in the student body (over two thirds of the cohort from an international background, mostly European), and English as the primary language of instruction. It also employs a student-centred learning approach called "problem-based learning" (PBL). As PBL involves small-group collaborative learning on open-ended problems, these students are familiar with the use of collaborative learning activities such as those used in this research. This method of curriculum design has demonstrated outcomes, with student's appreciating the style of learning, and gaining improved inter-personal skills for such tasks (H. G. Schmidt, Molen, Winkel, & Wijnen, 2009).

There were a number of stages to the study as described in the proceeding sections. In a pre-lab task (described in section 3:6.1) participants completed a survey item (described in section 3:4.1.1). The primary tasks, comprising survey items (described in section 3:4.1.2) and collaborative task (described in section 3:4.2), took place in a computer lab within the School of Business and Economics – the lab-session (described 3:6.2). A final component involved an 'at home' task, comprised of an assessment task (described in section 3:4.4) and a post-task survey (described in section 3:4.1.3), which were to be completed at a location convenient for the individual participant.

A between-subjects design was used, with two conditions – the CIS and MDP tasks described below in section 3:4.2 – and variables as described largely in section 3:4.3.1.1. The following sections then, introduce the participants (section 3:2), ethical considerations in this thesis work

(section 3:3), the materials used across the work (section 3:4), a pilot procedure (section 3:5), and the procedure followed for this empirical work (section 3:6), before closing with a chapter summary (section 3:7) linking the research aims, empirical design, and data collected.

3:2 Participants

Participants were students attending computer skills sessions in Quantitative Economics at Maastricht University. Participants had the option of either attending the 75 minute lab-session to complete the study and a 60-90 minute task at home, or completing a separate task which would have taken slightly longer (4-6 hours). Almost all students opted to participate in the research study. Participation was in one of 38 sessions during a week (6 on Monday, 8 on Tuesday-Friday), with each session lasting 75 minutes in total and comprising between 25 and 35 students. Participants worked on separate PCs, and mostly in pairs (although due to uneven numbers and late arrivals a small number worked in trios or individually).

In total 1148 students participated in the study. This thesis work focuses on 578 students who took part, as indicated in Table 3:1. A number of students are excluded from this figure. Early in the week, software failure resulted in data loss and students working independently (n = 271). As such, while this cohort participated (and received feedback on their output), no data was collected on their processes (and comparison of demographic data for this cohort is not possible). In addition, further software problems resulted in partial data-loss for other sessions, data from which will not be used in this analysis (n = 270). Finally to balance numbers in classes a small set of students (n = 18) worked in trios (using a triadic-based login for the system), and a smaller group independently (n = 11); analyses from these groups is excluded from this thesis. Assignment to classes within the week is not based on ability, indicating reasonable confidence that this data loss did not resulted in bias in the data obtained. This was confirmed using t-test comparison of means and chi-square test of independence, reported for the target sample compared to the excluded groups in section 4:1 of the results. Where salient, additional comparisons are indicated in the results to: demonstrate homogeneity between the target group and the full set of participants who completed the ISEQ (section 6:1.1.1.1); between those who

completed the ISEQ and those whom did not; and those who completed the assessment stages and those whom did not. The earlier days in the week provided opportunity to ensure the class-based procedures were running effectively, and that all lab assistants understood the procedures.

Despite this data loss, to the best of my knowledge this is the first study to explore epistemic CIS and MDP tasks on such a large scale (see the literature review for indicative sample sizes) or with an innovative collaborative browser technology.

Table 3:1 – Demographic data for study participants

	n total	n females	n males	M age	SD age
CIS	308	125	183	19.01	1.32
MDP	270	100	170	19.00	1.24
Total	*578*	*225*	*353*	*19.00*	*1.28*

3:3 **Ethics**

The Open University Human Research Ethics Committee reviewed and approved the research described in this thesis (reference: HREC/2014/66836/Knight/2). The British Psychological Society (2014) Code of Human Research Ethics, and British Educational Research Association (2011) Ethical Guidelines for Educational Research were followed.

Participants in this study were in a cohort of students who have, in advance, consented to use of their educational data, in an anonymous format, for educational and research purposes, a process described in Tempelaar et al., (2012, sec. 3.2–3.3); specifically consenting to analysis of self-report questionnaires or psychometric instruments and educational outcomes. In addition, participants were: informed in advance a week prior of the study; given a briefing at the beginning of the study-session; and given both a paper description and full online consent form for further detail (reproduced in Appendix 1). Participants consented by ticking an on-screen box, logging into the system (using a login given to them) and submitting their unique student ID number. In doing so the participants consented to Maastricht University sharing data with the Open University team and the standard terms of use of Coagmento (http://www.coagmento.org/terms.php) including the collection of browser-data (this was verbally noted, separately included in the consent details, and a link to the Coagmento terms provided). All procedures and materials were negotiated with our collaborator and gatekeeper at Maastricht University.

Participants could opt to take part in the study, or to complete a separate course-task; this is akin to the option for course-credit (although it should be noted that neither option provide *extra* course credits). Participants could withdraw and complete the alternative task at any time, and the key criterion for 'completion' was simply signing up (that is, participants were not penalised for minimal effort or success). This was true of both the 'in lab' and 'at home' components of the study. This is in line with relevant ethics guidelines. It should also be noted that relevant guidance on ethical use of learning analytics exist (See, Slade & Prinsloo, 2013; The Open University, 2014)[19] although these largely relate to implemented tools to aggregate data across sources and classify students within real courses, rather than the kind of exploratory work being conducted here; the fact the research did not impact directly on student's grades or courses directly is, though, aligned with that guidance.

The study took place in a Maastricht University lab setting using university PCs. Firefox and Coagmento (the browser add-on used, see section 3:4.3.1) was installed in advance on all PCs. The PCs were logged in to a generic user account, and kept on that account for the duration of the study. Between each session the researcher and lab assistants cleared the browser cache (removing history, saved passwords, active logins, form and search history, and cookies). These steps: Minimise the risk of students exposing data about themselves to subsequent participants using the same machine; minimise the risk of student data being compromised, for example, by being able to see the history of successful searches in a prior participant's session; minimise the risk of students exposing data about themselves to the researchers via logging their activity (because the browser does not have their usual bookmarks, history or logins saved and is not associated with their usual login); minimises risk of students accessing any other inappropriate content (because their access to the internet is over the Maastricht University proxy), and

---

[19] Note that the OU framework for learning analytics (The Open University, 2014) was not applicable to this research (because the data is not from OU students, nor targeted at intervention). Note too that the BPS guidance on online research (British Psychological Society, 2013) is largely not directly applicable in this case as that guidance is largely intended to cover the use of online surveys and similar instruments where participants may recruited entirely online, and there is low researcher cost to reaching out to large numbers of participants at any one time.

commensurate need for researcher disclosure (see BERA guidance on disclosure in research contexts 2011).

No sensitive data (per the Data Protection Act Office of Public Sector Information, 1998) was collected (e.g. racial or ethnic origin, religious or similar beliefs, disability or other health matters, sexual life). Personal data arising from the research is kept on a secure passworded computer. Where data is shared between research-collaborators, student IDs are used to match records, and then removed in place of a researcher-assigned unique identifier. Data is shared between Dr Dirk Tempelaar (Senior Lecturer and module convenor Quantitative Methods) at Maastricht University, Dr Chirag Shah (Assistant Professor and director of the InfoSeeking Lab which develops Coagmento) at Rutgers University, and the Open University supervisory team for research, educational, and dissemination purposes only.

In processing data, care has been taken to ensure that no names or other personally identifying information is revealed. As such, raw log-file data will not be released, and pseudonyms have been used in reporting where necessary. The data will be kept in full for the duration of the primary investigator's PhD research or until January 2016 (whichever is later), after which it will be reviewed with a view to full anonymisation (i.e. any personal data still held will be deleted). The data will be reviewed after that point with a view to full deletion when appropriate. Data stored will be kept in a password protected file in accordance with the Data Protection Act.

No risk was identified in participating in the research (as identified to the participants), and the task was framed with the research interest that: "the researchers are interested in developing tools to support students in finding and evaluating information together". This claim is a true but partial disclosure given the research emphasis is on understanding differences in patterns of information seeking and evaluation, rather than in aspects of software design or human-computer-interaction. A full debrief, describing the study more completely, was given following completion of both parts of the study (see Appendix 3).

The study was specifically designed to be educationally beneficial to participants in its own right, in addition to analysis holding wider benefit to our understanding of epistemic-commitments and source-evaluation. While the main topics were not directly relevant to the students (although a warm-up task was economics-focussed), the skills used in such tasks – of seeking, evaluating, and writing up information in small groups – are important transferable skills, and as discussed in 2:2.6 there are reasons to believe that collaborative information seeking holds educational value. Similarly, the second part of the study (involving the 'at home' assessment task) was designed to be educationally valuable in its own right (as discussed in the following section 3:4.4 on this task).

3:4 Materials

This section describes the set of materials and tasks used in the thesis research. The section is split into a number of parts. First I discuss a set of demographic data collected and self-report measures used at different stages of the study, I then go on to describe the tasks developed and the materials used in those tasks, before describing the kinds of trace-data obtained through use of a browser add-on in completion of the tasks. The section finishes with a discussion of the assessment method and materials used for the text-outputs before going on to describe the procedure used (section 3:6).

3:4.1 Psychometrics and survey instruments

In terms of survey instruments, participant data was collected in three ways: self-report in advance of the main lab-session; self-report in a short survey following the main task and assessment task; and collected by the research-collaborators at Maastricht University.

3:4.1.1 Pre-lab: Internet Specific epistemological Questionnaire (ISEQ)

As noted in section 2:4.2 one psychometric instrument of particular relevance to information seeking and epistemic-commitments is the ISEQ (Bråten et al., 2005), which has been deployed in a small number of similar tasks to the one described here (Kammerer et al., 2015, 2013; Strømsø & Bråten, 2010). The ISEQ is a 36-item instrument with a 4-factor conceptual structure mirroring the epistemic-cognition literature structure described in Table 1:1: simplicity; certainty; justification; and source of knowledge. Empirical validation of its structure (Bråten et al., 2005) with 157 Norwegian political science undergraduates indicated a two factor structure, rather than

the four factor structure initially conceptualised: a justification factor (α .70) with 4 items scores on which range from a perspective that internet-based knowledge claims can be accepted without critical evaluation to a perspective that they should be corroborated and critiqued; and a general internet epistemology factor (α .90) with 14 items scores on which range from a perspective that the internet can give true, specific facts, to a perspective that the internet is not a good source of true facts. A 7 point scale version of the English ISEQ (Bråten & Weinstein, 2004) was deployed in this study, completed in advance of the lab-session using the same system as the students use to complete similar survey items as part of their course (as described in Ethics section).

3:4.1.2 In-lab: Demographic variables and Trustworthiness items

A short survey was given to participants at the end of the lab-session (reproduced in Appendix 6 for the CIS group and Appendix 7 for the MDP). This survey was comprised of the trustworthiness assessment items (discussed in the following subsection, 3:4.1.2.1), and three items asking students to rate on a 1-10 scale: how familiar they were with their partner prior to the lab-session; their level of agreement with their partner during the lab-session; their knowledge of the topic prior to the lab-session. A final element to the survey in this session then asked participants to respond to five technical questions, specifically: how often they conducted searches using a search engine in an average day (0-3, 4-6, 7-10, 10+ or 'occasionally'); how experienced they were at using search engines (1-5 scale); what their most used search engine was (bing, google, yahoo or 'other'); what browser their main browser was (chrome, firefox, ie, safari, or 'other'); their primary operating system (e.g. linux, mac, or windows). At this stage participants were also asked to indicate their gender, and age (in years).

3:4.1.2.1   Trustworthiness Assessments

Following the main task (during the lab-session), participants were also asked to rate the trustworthiness of the documents (n = 11) or general websites (3 questions) on a 1-10 scale. This follows earlier research (Anmarkrud et al., 2014; Bråten, Braasch, Strømsø, & Ferguson, 2014) in which students were asked to read multiple conflicting documents and, following writing a short

report, rank those documents according to their trustworthiness. Following ranking, they were then asked to give reasons for their decision. In that earlier work, students were given only the title and metadata (e.g. author, publisher, date of publication) rather than the complete content of the document.

In this research, two tasks (as described below) were deployed, one as per the research above involving reading multiple documents, the other involving finding resources on the internet. In the former students were asked to rate 11 documents on a 1-10 scale (least to most trustworthy), rather than to rank the documents.

- Please rate how trustworthy you thought Document X was (10 = This source was high quality, and the information was very credible, 1=This source was low quality and the information lacked credibility)

Rating rather than ranking permits: documents to be rated at the same level; consideration of more documents (where ranking becomes more challenging the more items are to be ranked); greater discrepancies in rating to be captured (while ranks are ordinal, ratings are interval); and the possibility of factor analysis on ratings. Students were not asked to justify their choices (on the basis that source selection and evaluation was part of their main task).

In the second (internet-search or CIS) group, students were asked to give three ratings (on a 1-10 scale), indicating trustworthiness scores for the: most, least, and average trustworthiness of the resources they had found. In addition they were asked to give a URL for the most and least trustworthy resource, and to give general comments on the type of resources found:

- Please rate how trustworthy the information you found in this task was on average (10 = The sources were high quality, and the information was very credible, 1=The sources were low quality and the information lacked credibility)

- Please give an estimate rating for the *least* trustworthy page you found where 1 is 'not at all trustworthy' and 10 is 'very trustworthy' (if you can remember it, please post the URL here)

- Please give an estimate rating for the *most* trustworthy page you found where 1 is 'not at all trustworthy' and 10 is 'very trustworthy' (if you can remember it, please post the URL here)

- Please give any other feedback or commentary on examples of types or sources of documents you found, and how trustworthy you found them in the space provided.

3:4.1.3 Post-Lab: Task-survey

At the end of the study (following the at-home assessment tasks), participants were invited to give general feedback (in a text field) on the study. At this stage participants were also asked to rate their experience of the study, specifically they were asked to indicate on a 1-10 scale (where 1 indicates 'not at all' and 10 indicates 'very'): how satisfying they had found the collaborative element of the tasks; how satisfying they had found the tasks; how intuitive they had found the browser add-on.

3:4.1.4 Academic Performance

In addition to this self-report data, a numeric achievement grade was obtained from Maastricht following the research, indicating achievement across 4 course-assignments. The delay in obtaining these grades gave time for the fourth assignment to be collected and graded. Grades in the Netherlands are given on a 1-10 scale, with a passing mark of 5.5. These grades were averaged to give a 'grade point average' score for each student on a scale of 1 (low) to 10 (high).

3:4.2   Task Designs

Section 2:3 above describes the literature on multiple document processing and epistemic cognition tasks arguing for the epistemic nature of such literacy tasks, and information seeking more generally (section 2:1.2), particularly collaborative, exploratory search (section 2:2). This section builds on and expands that literature to describe the process of task design for this research.

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

(Bråten & Strømsø, 2009). 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?

Thus, a key component of task design may be both the purpose of the activity, and the collaborative framing provided. One study (Y. H. Cho, Lee, & Jonassen, 2011) 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) (Y. H. Cho et al., 2011). That study also reports overall benefits to collaborating on argumentation, suggesting that such tasks (and CSCL tools) hold benefits for all students (Y. H. 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 (Y. H. Cho et al., 2011) – indicating the importance of dialogue quality in such environments.

However, current research 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). As described below, in the research conducted for this thesis a task was designed to elicit summaries and explicit judgements regarding the quality of other arguments, thus asking

students to select, evaluate, and contrast claims rather than construct their own opinion or argument.

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

Characteristics which are drawn on in the topic-selection and task statement described below.

3:4.2.1 Task topic

After a warm-up task to familiarise participants with the respective tools and their paired collaborator, in this research, two themes were selected. One using multiple pre-selected documents and the second involving searching for information on the internet. Following the work described above, topics with conflicting perspectives and a variety of source-qualities were sought to foreground participant's commitments to varying source-content qualities. A number of sources giving 'scientific controversies' (see Appendix 8) were explored to select an appropriate topic, and two topics were identified which:

1. provide a focussed topical research area which can be studied in isolation, within a 1 hour session;
2. are not topics that are high profile or/and large scale controversies (such as climate change, or genetically modified crops, both of which receive a lot of press coverage);

3. have a variety of source-types and qualities which refer to them, from varying perspectives.

The two topics (for MDP and CIS) are now discussed in further detail.

3:4.2.1.1 MDP Task: Glyphosate

For the MDP task, a set of eleven documents was produced, with queries seeded from a Guardian blog tagged as 'controversies in science' (Moses, 2013) (although this article was not included in the document set itself). The article led directly to the first article, from which further articles were seeded (as in Table 3:2, which presents articles in the order they were presented to participants).

Table 3:2 - MDP Document Sourcing

	Authors	Type	Date	Title	Seed
1	GM Freeze and Friends of the Earth	Press Release	June 13 (2013)	Government Urged to Act After Weedkiller Traces Found in Britons	Guardian seed
2	Medizinisches Labor Bremen	Scientific report	(2013)	Determination of Glyphosate residues in human urine samples from 18 European countries	1
3	Farmers Weekly/Case	Trade blog	June 13 (2013)	Scientists dismiss "unreliable" glyphosate study	Query: "Glyphosate urine"
4	Science Media Centre	Science literacy blog	June 13 (2013)	Expert reaction to press release from FotE and GM Freeze about glyphosate in urine	3
5	Acquavella, et al.,	Peer reviewed research article	(2004)	Glyphosate biomonitoring for farmers and their families: results from the Farm Family Exposure Study	Scholar query: "Glyphosate urine"
6	Gillam	News article (Reuters)	April 25 (2013)	Heavy use of herbicide Roundup linked to health dangers-U.S. study.	Query: "Glyphosate health"
7	Samsel and Seneff	Peer reviewed theory article	(2013)	Glyphosate's Suppression of Cytochrome P450 Enzymes and Amino Acid Biosynthesis by the Gut Microbiome: Pathways to Modern Diseases	6
8	Raeburn	Science-literacy blog	April 26 (2013)	Discover blogger Keith Kloor stumbles into nest of questionable studies and reporting on GMOs and multiple ailments	Query titles of 6/7
9	Williams, Kroes and Munro	Peer reviewed review article	(2000)	Safety Evaluation and Risk Assessment of the Herbicide Roundup and Its Active Ingredient, Glyphosate, for Humans	Scholar query: "Glyphosate risk"
10	Johal and Huber	Peer reviewed position article	(2009)	Glyphosate effects on diseases of plants	Scholar query: "Glyphosate risk"
11	O'Duke and Powles	Peer reviewed review article	(2008)	Glyphosate: a once-in-a-century herbicide	Query: "Glyphosate risk"

A simplified document-model (building on Rouet's work 2006, Chapter 3) is given in Figure 3:1, depicting the three key themes identified in this document set (the presence of glyphosate markers in human urine; the risks to human health of glyphosate; and the agricultural risks of glyphosate use), the document stance (broadly negative – orange; critique/broadly positive – green; largely neutral/scientific – blue) and the relations among them (+ - support; - - critique; note document 11 relates to 2 primary themes).
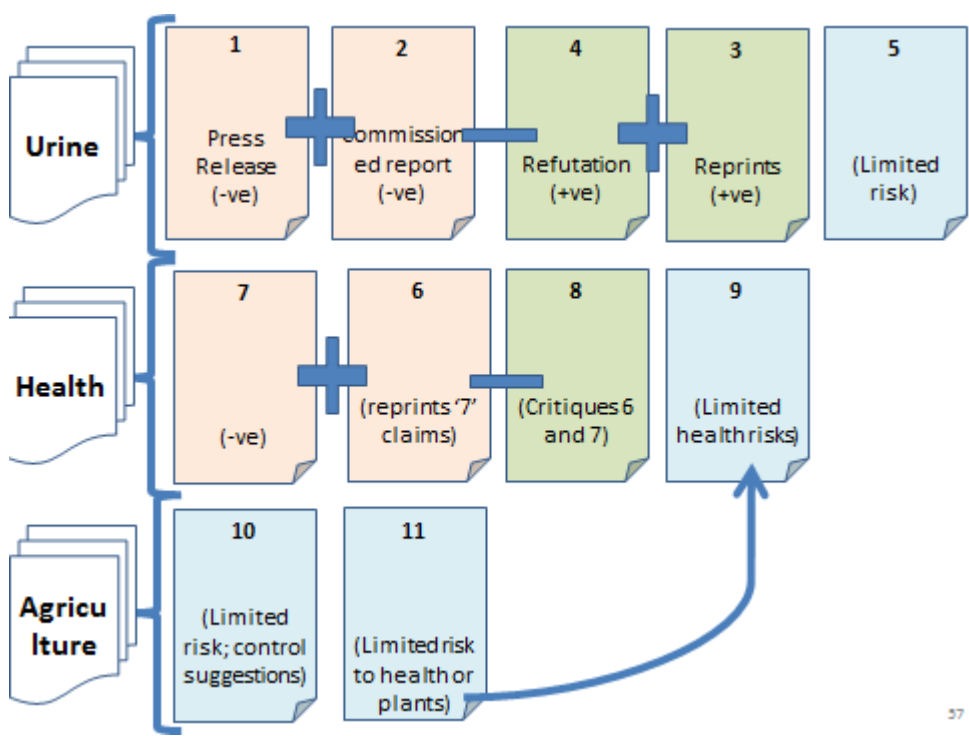
Figure 3:1 - Simplified MDP Document Module

For each document, the original source (HTML or PDF) was saved and formatted for presentation to students (to ensure it would load without scripts, etc., and would load in html without need to use a pdf viewer or other external reader). The documents were also cleaned, to remove extraneous detail and to reduce them to core claims around glyphosate. Only the abstracts of 5, 7 and 9 were provided, while 10 was reduced to the abstract and first section of the introduction; 11 was reproduced in an abridged form, it was also the most comprehensive document in terms of coverage of potential risks.

Note in particular that the set of documents selected is rather complex. For example, the author of 7 is criticised in 8. Ostensibly, 7 is more trustworthy because it is in a peer review journal and

republished by Reuters (6), while 8 is a blog. However, the critique provided in 8 (and the evidence referred to) is strong and the source features of the blog (also based at MIT) are also strong. Furthermore, the author of 7 has been criticised for publishing in an area they are not an expert in, (including praising the discredited Andrew Wakefield on autism), and while the journal is peer reviewed, it is primarily a physics journal not a health-sciences one. We also see in documents '3' and '4' a reprint on a trade website (Farmers Weekly; 4) of an independent critique (3) – something students might identify as raising concerns of bias in '4', although the content is identical. Thus the selection of documents provides a set of conflicting sources, of varying quality, with a range of sub-topics present. As such, the topic and selected documents provide good source material for probing students' abilities to extract, integrate and evaluate information from across sources.

3:4.2.1.2   CIS Task: Red Yeast Rice

For the CIS task, a theme was identified and explored by the author to ensure it was appropriate for use in the research. The topic of 'red yeast rice' was selected based on its presence under the sub-category of 'Medical controversies' (https://en.wikipedia.org/wiki/Category:Medical_controversies) in Wikipedia's Scientific controversies category (https://en.wikipedia.org/wiki/Category:Scientific_controversies). This case was identified as interesting because:

1. Using search engines to seek information on health issues, such as use of food supplements, is a common issue (See, for example: use of Wikipedia, Heilman et al., 2011; survey data, Horgan & Sweeney, 2012; and log data, C. W. Schmidt, 2012) and requires evaluation of claims from across various types of sources;
2. The Wikipedia article on 'red yeast rice' is not particularly high quality (it is rated 'b-class' in the 'alternative medicine WikiProject' quality scale). Monascus purpureus (its scientific name) does not receive a rating on any relevant scientific or medical WikiProject scales (but is a stub article, i.e. it is very short);
3. Search engine results pages show varying results for queries on 'red yeast rice' and monascus purpureus;
4. The controversy is largely around restrictions and side effects (i.e. it is uncontroversial that the substance has a medical effect, although risks and scope of those effects are disputed).

Further research indicated that the substance had received some public attention including from regulatory agencies in France and America (ANSES, 2014; FDA, 2007) based concerns regarding

the concentration of the active ingredient (Gordon, 2010), and concerns reported in the press

regarding its contamination with citrinin (Harding, 2008). Alongside this, there were also reports

in the popular press (Macrae, 2008) citing research (Lu et al., 2008) on its positive impact, and

medical advisory sites providing a science-literacy perspective on this (NHS Choices, 2008). In

addition, red yeast rice is widely (correctly) reported as containing the same active ingredient as

'statin' drugs, which have various known side-effects, and have had somewhat controversial

coverage in the press in their own right (See, for example, Boseley, 2014; Gallagher, 2014; Ridker

& Cook, 2013). Three key themes were identified in this search, first that red yeast rice should be

treated as a statin, second that the concentration levels of the active ingredient vary, and third

that some samples have been contaminated with citrinin. Again, the potential set of documents

participants might encounter in researching the assigned topic offers conflicting information from

sources of varying quality, with a range of sub-topics present. In this more open task, it is also

worth noting that the depth of information seeking participants engage in is of key interest as

covering the range of topics requires active information seeking beyond that of the MDP task

(where the range is present in the given documents). Indeed, given the connection of red yeast

rice to statins – which as noted above have received some controversial coverage – it is possible

some participants could engage in a very wide reaching information seeking activity. As such, the

topic provided offers a good seed for probing students' abilities to extract, integrate and evaluate

information from across sources.

3:4.2.2 Task Presentation

The tasks were presented using a website which required login provided to students at the start

of their session. The procedure is described further below, here I note that the main task involved

a set of instructions (below), a space to write their answer, on a separate webpage (using

etherpad, see section 3:4.3 discussing the software used) and – in the MDP case – a set of

documents to read. These documents were linked to from the task instructions, and presented

such that their title hyperlinked to the main article, and was followed by a short 'snippet' as given

in most modern search engines. The articles were presented in a fixed order (i.e. their appearance was not randomised), as in Figure 3:2.

Articles

Government Urged to Act After Weedkiller Traces Found in Britons *Government Urged to Act After Weedkiller Traces Found in Britons*. GM Freeze and Friends of the Earth. The Government is being urged to take…

Determination of Glyphosate residues in human urine samples from 18 European countries Document Title. Determination of Glyphosate residues in human urine samples from 18 European countries. Test Compound. Glyphosate and…

Scientists dismiss unreliable glyphosate study *Scientists* have *dismissed* suggestions that *glyphosate* could be unsafe after a report claimed traces of the weed killer is commonly found in…

Expert reaction to press release from FotE and GM Freeze *Expert reaction to press release from FotE and GM Freeze* about glyphosate in urine. Friends of the Earth and GM Freeze press released the…

Figure 3:2 - Article presentation

In both Anmarkrud, Bråten and Strømsø (2014) and Bråten, Braasch, Strømsø, & Ferguson (2014) students were given six texts to read (on the cancer-risks of mobile phones) with conflicting perspectives and varying source-feature trustworthiness, with the framing prompt to:

> *Imagine that a close friend has told you that she experiences discomfort when using her mobile phone. She has asked you for advice and you have searched the Internet for information about the topic. The search resulted in six results…* (Anmarkrud et al., 2014, p. 5; Bråten, Braasch, et al., 2014, p. 18).

The participants were instructed to read the six 'search results' over 40 minutes, in order to provide their friend with "well-grounded advice". They were then given an essay prompt, to address in 20 minutes, without access to the source-documents:

> *You are now going to write a brief report where you judge the health risk of cell phone use. Base your report on the texts that you just read and try to express yourself clearly and elaborate the information—preferably in your own words. Justify your conclusions by referring to the sources you have been working with. (Anmarkrud et al., 2014, p. 4; Bråten, Braasch, et al., 2014, p. 15)*

Building on the task design used in that research, the task prompt for the thesis research was written to foreground student's understanding of knowledge claims, and support for those claims. For example whether they corroborated, emphasised source features and source-credibility, or evaluated source-content and methods used in sources. Thus, students were not asked to "refer to the sources you have been working with", but instead asked to "Produce a summary of the

best supported claims you find and explain why you think they are." 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). The text below thus gives the full task instructions used in my

research (note differences in task are given '[ ]' thus: [CIS/MDP]):

> For this task you will be researching the safety of ['Red Yeast Rice'/'Glyphosate']
>
> 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.
>
> [You and your partner should work together to find relevant materials on the internet./ A colleague has already found a number of documents for you to process with your partner, you should use these to extract the best supported claims (without using the internet to find further material).]
>
> You should:
>
> **Read the questions/topic** areas provided, these will require you to find information and arguments to present the best supported claims, you should decide with your partner which are best as you read.
>
> **Group information together** by using headings in the Editor
>
> You should work with your partner to explain why the claims you've found are the best available
>
> You should spend about 45 minutes on this task
>
> *[The official has heard that French officials have raised some concerns about the safety of 'Red Yeast Rice' and potential contamination, and would like a briefing on its potential risk.*
>
> */*
>
> *A review is coming up for the license of Glyphosate, the official would like to know the best supported claims around its risks.*
>
> A colleague has collected some documents, available from [this page](.).]

3:4.3   Software and Website

The main tasks in this research involved collaborative information seeking processes. A number of

tools exist in the CIS literature to facilitate the CIS process, including several 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)). 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 2:2).

In the context of my research, the ability to explore communication between collaborators is key. 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.

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 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).

A similar tool, which is still being actively developed and supported, 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 (Shah, 2012a), 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 et al., In press). 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. As I describe further in the following section, Coagmento has been used extensively in research contexts, and provides a good lens onto the kinds of epistemic behaviours of key interest to this research, in collaboration with the Coagmento research group, it has thus been customised and deployed in this thesis work.

3:4.3.1 Coagmento: A software tool for analysis of CIS

Coagmento (Shah, Marchionini, & Kelly, 2009) was used for support of the collaborative process and data capture in my research. 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 requirement 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). Coagmento was customised (as depicted in Figure 3:3) in this case to:

1. Send users to google.co.uk in order to minimise the use of non-English resources
2. Remove some menus from the standard Coagmento interface (specifically the 'bookmark' 'rank' 'tagging' and 'project' elements) in order to reflect the needs of the study
3. Provide a default 'project' such that as students moved through the tasks the appropriate materials were displayed to them, with one project-etherpad created for the warmup task, and a second for the main task
4. Track participant's "copy" actions (using ctrl+c, or the edit->copy menu), to enable us to explore which websites were being used and what material drawn from them

Thus, as indicated in Figure 3:3.1, which shows Coagmento in use (with both the toolbar and sidebar displayed) and then each component, Coagmento provides:

1. A set of menus via a toolbar, as in Figure 3:3.2:
   a. 'Home' which takes participants back to the main introductory page (from which they can navigate forward)
   b. 'System Guide' which gives participants some instructions on using Coagmento
   c. 'Snip' which allows participants to create a shared copy of any "snipped" text from a webpage
   d. 'Task pad' which takes participants to the shared etherpad for the task-stage they are at
   e. 'Active Task' which opens a popup of the instructions for the stage they are currently on
2. A set of interaction interfaces via a sidebar, as in Figure 3:3.3 and Figure 3:3.4:
   a. A 'Chat' tab for the collaborators within any group, automatically using and storing their Coagmento login usernames

   b. A 'History' tab, which displays the searches and snippets of the user and their collaborator (note only 'Snippets' are displayed in Figure 3:3.3, for CIS users, the Sidebar would display 'Searches' as default, with 'Snippets on an additional tab next to it)
   c. A 'Submit' button, used to submit the answers at the end of the warmup and main task – the button asked participants to confirm with their partner (in the chat) and once confirmed, the page would reload with the subsequent task displayed.
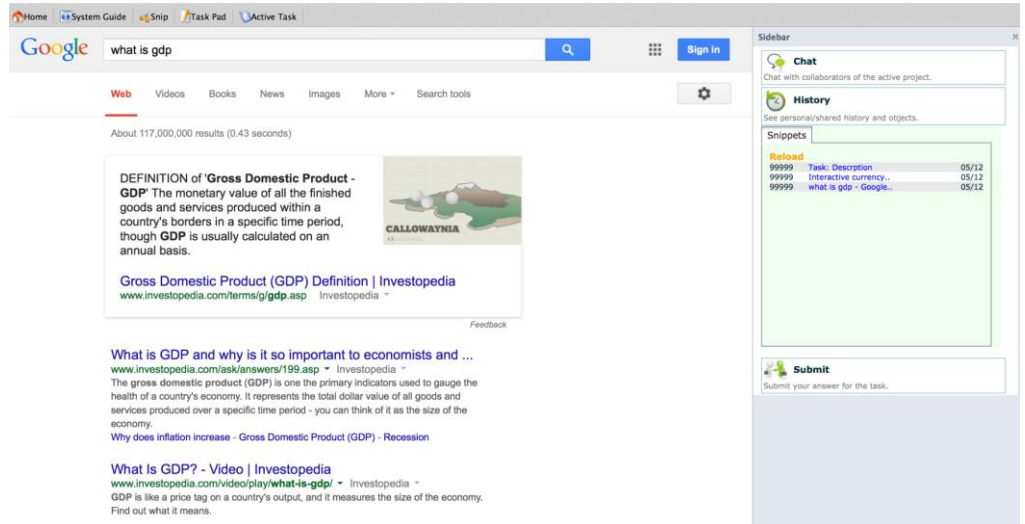
Figure 3:3.1:

Figure 3:3.2:
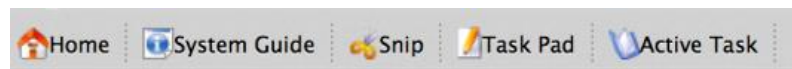
Figures 8.3 (left):

Figure 8.4 (right):
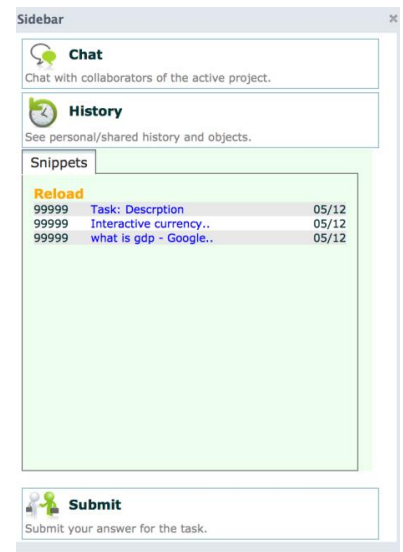
Figure 3:3: Coagmento Screenshots (from top: 8.1 A full screen display from a browser window; 8.2 The toolbar element; 8.3 Sidebar with Chat displayed; 8.4 Sidebar with Snippets displayed)

In order to deploy the study, a website was designed to guide participants through each stage – from logging in at the start of the lab-session, to completion of the at home task (see Figure 3:4 in section 3:6 for a visual timeline). For the lab-session, students were required to login to the website, and could then make use of a browser add-on (Coagmento) which was required for some

137

stages of the task. The add-on consisted of a toolbar (along the top of the browser screen), and a

sidebar (along the right-hand side of the browser screen). Functions in this add-on became

active/inactive depending on the task-stage. For the at home element students were not required

to use the browser add-on (and could use the browser of their preference), simply logging in to

the website using their student ID instead. The browser add-on was selected based on the

functions it provided, and ability to track particular information as now described.

3:4.3.1.1   Trace data foregrounded by Coagmento

The browser addon comprises three core elements: An etherpad; a chat tool (PhP free chat); and

a logger, which tracks website views and queries made. Data from these can be analysed using

the userIDs or projectID of the groups. We thus have free-text data from the etherpad and chat

elements, as well as textual data in the form of URLs and "copied" text (text copied from

webpages by participants), alongside various forms of metrics which can be computed from the

navigation-logs, as I now discuss.

3:4.3.1.1.1   Information Seeking Metrics

Drawing on the work described in section 8, this thesis research uses a mixture of approaches to

the log-data, from which a set of metrics are computed – including some simple counts (for

example, 'page views') as in Table 3:3 . Aligned with Table 2:9 in section 8.1, note that query-

vocabulary richness is used (as described by Yue, Jiang, et al., 2012) rather than simpler counts of

keywords, giving a richer indicator of vocabulary richness across queries. Similarly likelihood of

discovery is excluded from this analysis given that: there is no direct analogue for the MDP task; it

is not obvious how best to summarise the data (maxima, minima, average, etc.); nor is it clear

how likelihood relates to depth of search rather than lostness. As an alternative to likelihood of

discovery, the simpler metric of 'query depth' is computed; this indicates how many search

engine results pages beyond the 1st page of results are opened. 'Page use' in this study is classed

as any page where the URL is recorded in the chat, etherpad, or as a source of a 'copy' or 'snippet'

action – this parallels metrics around 'bookmarking' or 'snipping' from pages, bridging explicit and

implicit usage metrics. In order to capture information regarding pages visited but *not* used are

captured through the use of an f measure – which captures further data in addition. In addition, a

set of symmetry metrics were computed as discussed below in section 3:4.3.1.1.2, and chat-based

analysis in section 3:4.3.1.1.3. Duration and sequence measures are not reported in this thesis.

Table 3:3 – Information Seeking Metrics

Metric	Description
Page views	N of unique pages or urls viewed
Pages used	N of unique pages or urls referred to in the chat, entered into the etherpad, or copied/snipped from
F	An f measure, calculated as 2 * precision * recall /(precision + recall), where precision is the ratio of pages used to pages visited and recall is the ratio of the number of pages used by the individual group to the number of pages used by any group
Coverage	n of pages visited by group / n of total unique pages visited by any group
Queries*	Number of unique queries
Query terms*	Number of unique terms used in a group's queries
Query vocabulary richness*	Number of unique query terms / number of unique terms (Yue, Jiang, et al., 2012)
Query depth*	The sum of instances of unique pageviews of search engine results pages beyond the first page of results

*Note: Not computed for the MDP task

As noted in section 2:1.2.2, and earlier in the introductory section, there are theoretical

(epistemological) grounds for conceptualising learning outcomes – in educational, and

information seeking terms – as tied to use of information, and its communicative properties.

Indeed, information seeking processes should be understood in the context for which the

information is sought. Moreover, such communicative properties may be particularly

foregrounded by CIS contexts. Each of the methods above is problematic in its analysis of

'success' or 'performance' – although each may provide important insight into the ways in which

students engage in tool mediated interaction with the information they are seeking and

evaluating . 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. Given the

focus on educational contexts for CIS, educational aims form part of the context for task activity,

and judging task success.

3:4.3.1.1.2 *Contribution symmetry*

A set of metrics was designed to indicate the symmetry of contribution and interaction. These

metrics were designed to explore whether pairs were acting collaboratively, or whether one

partner was (seemingly) more active than the other. The first set of these is simply a difference measure of the measures in Table 3:3 (for example, the difference in page views between collaborators, where '3' indicates that one partner viewed 3 more pages than the other). On each of these we can also compute a symmetry score, taking the difference between each partner in any of the given metrics, and converting these to a 0-1 score where 0 indicates perfect symmetry and 1 indicates perfect asymmetry. In addition, two metrics were computed on the etherpad data. First, a contribution symmetry score indicating the percentage of the total text that can be attributed to a single author; closer to 0 indicating full symmetry (i.e. both authors contributed the same number of words), 1 indicating full asymmetry (i.e. one author contributed 0 words). Second, a count of 'touch points' – points in the text at which the author transitioned from one partner to the other, where 1 would indicate that there was a single authorial transition – or two chunks of text, each written by a single author. This metric was selected as an indicator of collaboration versus co-operation, where this distinction is operationalised as one of authorial-interaction, such that a low number of 'touch points' indicates co-operation (partners write collectively, but separately), while a higher number indicates collaboration (partners co-edit their work and each integrates their writing into the other's; following Southavilay, Yacef, Reimann, & Calvo, 2013) – two patterns of writing observed in earlier work (Hirsch, Hitt, Powell, Khalaf, & Balawi, 2013).

3:4.3.1.1.3   Trace data from chats

Finally, chat data was collected from all participants. As part of the preliminary analysis this data was analysed, using a term identification method such that each message received a binary code – present, or not present – for a typology of terms as in Table 3:4. A typology, in this research, provides a means to develop analytic categories that typify the kinds of dialogue engaged in by the participants, along dimensions derived from the research focus.

A process building on deductive content analysis (see, for examples, Elo & Kyngäs, 2008; Hsieh & Shannon, 2005; Mayring, 2000; and Hatch, 2002 for a deductive typology development) was used to draw out terms for a typology of chat messages. In this process, theoretic considerations are

used to identify key categories and concepts associated with them. Application of these categories can involve the identification of variables or definitions for the typology categories, and the development of new categories where themes emerge from the data which do not fit the a priori categories. In this research, three categories were selected based on analysis of the research literature and surface analysis of the chat data; topic, source quality, and exploratory chat. The first of these – labelled 'topic' – relates to the topic-content of the material being sought, 'topic' chat regards the theme (for example, mentioning 'red yeast rice'), the sub-themes (for example, mentioning 'urine'), and other content knowledge around the particular tasks (for example, references to 'health' impacts).. A second category 'source quality' relates to use of metadata and source citation – as Goldman and Scardamalia (2013) and researchers noted in section 2:4.4 highlight – terms were selected based on known metadata from the MDP documents and a sample of CIS documents (for example, the author name 'Gillam', or the domain name drugs.com), and terms relating to citation ('written by', 'published in', etc.). Finally, the category 'exploratory' includes terms which refer to the kind of meta-discourse Goldman and Scardamalia highlight, and which are related to exploratory dialogue (see section 7.4), and epistemic stance ('because', 'I think', 'therefore', etc.). The intention in analysis of this kind is not to develop an exhaustive coding scheme as such, but to highlight some structures underlying the chat data as indicated through the application of a typology of terms.

Thus, terms were selected through: a priori identification, using prior research and terms directly related to the typology themes, for example the 'exploratory' terms are derived from the terms used in the published literature, the use of 'written by', 'published in' in source quality chat is derived from prior knowledge of the typology theme; a deductive analysis of the materials, including the task instructions, documents provided (MDP task) and a sample of websites visited (CIS task), from which further topic and source terms were derived (for example, identification of metadata such as authorship in the documents); and finally a deductive analysis of the chat data (described further below), to identify further terms, and ensure the chat data aligned with the typology.

In order to conduct this deductive analysis of the chat data, messages containing each term were identified using a concordance-style analysis, and visually inspected. This analysis was conducted by selecting only those messages in which any given term occurred. Such analysis has commonly be conducted using concordance software, which facilitates the exploration of 'Key Words In Context' (KWIC) by displaying words searched for in their original context (typically, showing a sub-portion or whole sentence in which the term is located). In this research concordance analysis was not needed (given the short nature of chat messages) although the method is broadly the same. The approach described here extends my earlier (small scale) approach to epistemic language in information seeking (S. Knight & Mercer, Forthcoming) in which concordance analysis was used alongside identification of cohesive ties – recurring terms in a dialogue – to identify the ways in which epistemic stances were co-constructed through dialogue while seeking and evaluating information. This analysis is in line with sociocultural approaches to discourse which emphasise the use of both qualitative and quantitative methods (Mercer, 2004), and aids with understanding how terms are used, alongside a quantitative indication of their use (reported in the results section). In this sociocultural approach (Mercer, 2004) quantitative analyses (such as term counts), is made use of alongside the inclusion of excerpts of dialogue intended to provide exemplifications of the ways in which talk is being used as a tool for thinking together; thus, qualitative and quantitative analysis are taken together, to inform the research process. This approach is founded in the kind of sociocultural theory briefly described in the introductory sections of this thesis. Furthermore, it has been deployed in the analysis of dialogue "in action" described both with regard to language as a tool to learn, and (in the context of discursive psychology) a more general tool 'to do' and a lens onto psychological constructs (such as epistemic cognition).

In this analysis, terms aligned with the typology described, and as in Table 3:4, were identified and messages containing those terms marked. At the preliminary stage messages with these terms occurring in them were visually inspected, in order to explore how the particular term was being used, and its alignment with the conceptual theme (topic, source quality, exploratory). Those

terms that were not aligned with the theme were removed. In this way, false-positives are minimised. Not all of the terms identified were widely used, for example 'good page' was used on a single occasion, but broader constructions (for example 'page') were not specific enough in their use, or resulted in results based on quoting the task instructions (e.g. 'claim' was used in the task instructions, thus instances of 'claim' are often not to do with specific claims, while this is not true of the small number of instances of 'claiming'), thus selection was targeted at identifying those terms best related to the typology themes on both a theoretical basis, and derived from the chat data itself. In addition to the concordance-style analysis, the first 5000 of the 20913 raw messages (i.e., with blank messages, server log messages, etc. intact) were visually inspected for further terms using a top down deductive selection process to select terms associated with the developed themes. Thus, a broader set of terms could be identified than through a priori selection or identification of terms co-occurring with existing target terms alone, thus minimising false-negatives. Finally, the deductive analysis of the chat data indicated an additional category of chat message involving terms around writing activities, including: decisions to copy/paste; concerns around plagiarism; and suggestions to edit the etherpad. Terms identified in messages of this type were labelled 'synthesis', with a theoretically driven hypothesis that those participants who engaged in more discussion making use of these terms will have engaged in coordination of their writing, and that this coordination might be associated with synthesising behaviour, in particular that this discussion enables participants to integrate the information they each (separately) found.

Table 3:4 – Typology of Chat Terms

Type	Terms in typology (each separated by a bar '|', in some cases a 'stem' is given, e.g. 'conclu' to capture all affixes)
Exploratory[+] – related to sharing and co-positioning with relation to knowledge claims (see section 2:4.4)	cuz|because|cause|coz|cos|cus|caus|becuz|I think|I guess|for example|I dont know|I don't know|I guess|I thought|so |if |also |maybe|probably|apparently|aparently|of course|would|must|might|could|will|may|conclu|weigh|agree|therefore
sourQual[#] - related to identifying sources, source metadata, and source metadiscourse	written by|prof|publish|author|recent|source| site|good page|bias|scienti|blog|news|cite|sponsor|credible|samsel|senef|indep|reuter |kloor|gillam|duke|powles|gm|fote|foe|friends of the earth|laboratory|bremen|sanders|nutritional sciences|murphy|boobis|fda|method|they sa|\"| webmd| web md| medicinenet| medicine net| lexology| altmedicine| alt medicine| sciencebasedpharmacy| science based pharmacy| drugs.com| medicine.com| wikipedia| emedicinehealth| medicine health| mayo| mercola| maryland| pennstate| medicinenet| medicine net| reuters| ncbi| about.com| NCCAM| ANSES| FDA| efsa| clinicaltrials| dfg.de| consumerlab| umm.edu|webmd |web md |medicinenet |medicine net |lexology |altmedicine |alt medicine |sciencebasedpharmacy |science based pharmacy |drugs.com |medicine.com |wikipedia |emedicinehealth |medicine health |mayo |mercola |maryland |pennstate |medicinenet |medicine net |reuters |ncbi |about.com |NCCAM |ANSES |FDA |efsa |clinicaltrials |dfg.de |consumerlab |umm.edu |data|document|article|backed up|evidence|writer|medicine net|study|trust|claiming
Topic* - related to content terms, specific claims, and key-words related to the content theme	statin|ryr|red yeast|rice|citrin|health|heart|diseas|amount|monacolin|contamin|cholesterol |drug|liver|lovastatin|medicin|french|kidney|china|muscl|supplemenet|ingre di|toxi|prescri|diet|ferment glyphosate|roundup|monsato|urine|salt|2 microgram|2.5 litres|residues|farmer|weed|environ|cytochrome|obes|xenobiotic|gastro|heart|diab|depres|autis|infert|cancer|alzh|entropy|west|inflam|agri|health|diseas|stati|herbicid|volunteers|44 per cent|44%|44 percent|18 countries|samples|10%|90%|plant|crop|weedkil|pesticid
Synthesis – related to coordination regarding writing[α]	structure|phrase|formulate|plag|conclus|intro|heading|write|own words|summary|task pad|copy|paste

[+]Note: 'best support' is used in quoting the task instructions, and is thus excluded from this analysis; agree is mostly used e.g. "I agree" to partner (rather than, e.g. "these sources agree"); weigh is used as in 'weighing up the arguments'; conclu[de/sion] is used as in "we can conclude that" as is 'therefore' (rather than "the paper has a conclusion", etc.).

[#]Note: Author names extracted from the MDP documents are included here, although most were not directly referred to; 'Dr' rarely used, 'doctor' more so but often in context of the content rather than source; 'by' used frequently for purposes other than sourcing (e.g. "make headings by..."); similarly 'written' is used variably, and 'written in' (rare) used to refer to the etherpad, 'written by' is used to refer to sources (but is rare); 'old' was not used, nor was outdated or historic; 'new' used to refer e.g. to new tab, etc., while 'recent' was used to refer to sources; 'good' is too widely used (e.g. 'good health' 'good question', 'good job') to use globally; 'good paper', 'good web[site]', etc. not used, although 'good source' is used this is captured by use of 'source' as a keyterm, 'good article' and 'good site' are used rarely but included here; other sentiment terms (e.g. 'great', 'fantastic', 'excellent', 'wonderful', 'brilliant', 'splendid') are either not used, or not used re: sourcing ('great'); 'bad' is used only in the context of topics, not the sources; 'method' is used where participants are evaluating the material in the sources; note the \"" is to catch all instances of quotation. Finally, note url references are not coded here (because they are aggregate as 'link use') although they are related to sourcing discussion.

Data used to refer to "there isn't enough data", etc. 'Claim' is used to refer to source instructions, while 'claiming' is used e.g. "since the people who are tested are the people who are claiming that it's bad". Keyterms from the unique 'most trustworthy' and 'least trustworthy' URL domains were extracted and identified in the chat data (as below). No further content analysis or metadata extraction on these texts has been conducted.

```
View(chat[grep("webmd|web md|medicinenet|medicine net|lexology|altmedicine|alt
medicine|sciencebasedpharmacy|science based
pharmacy|drugs.com|medicine.com|wikipedia|emedicinehealth|medicine
health|mayo|mercola|maryland|pennstate|medicinenet|medicine
net|reuters|ncbi|about.com|NCCAM|ANSES|FDA|efsa|clinicaltrials|dfg.de|consumerlab|umm.e
du",chat$content, perl=T,ignore.case=T), ])
```

In line with the work on use of source metadata (e.g. author, title, etc.) discussed in section 7 a full analysis of metadiscourse (for example building on analysis by Hyland, 1998, in the context of written texts) is not conducted, as my interest is restricted to use of 'evidentials' (in Hyland's framework), i.e. in whether students refer to sources or their qualities directly.

* Note: These topic terms are taken from visual inspection of a tfidf output from the CIS and MDP chat corpora, and (particularly in the MDP case) prior knowledge of the topic materials.

[α] Synthesis key terms were extracted from analysis of messages. Note 'plag' is to capture plagiarism, plagiarised, etc. (and misspellings) – this has been manually checked for false positives with satisfactory results.

3:4.4   Assessment Structure

3:4.4.1 Self and Peer-Assessment Introduction

As noted in the literature review, one means through which to approach large amounts of

assessment data, with a pedagogically supported method, is through the use of peer and self-

assessment. Two key considerations in developing peer or self-assessments are the students'

motivation to assess, and their level of skill to do so reliably and accurately. These are particularly

important given the extra cognitive, social and emotional demands peer or self-assessment places

on students, and the desire (including, ethical) that this demand bring educational benefit to the

students. The benefit of peer and self-assessment should, ideally, come both through constructive

feedback (i.e., that the assessment fulfils the needs of high quality assessment), and through the

process itself (i.e. that the process of peer or self-assessment is in itself educationally productive).

As noted earlier (section 2:5.2), across a broad body of work peer and self-assessment is seen as

having beneficial educational outcomes. However, understanding the specific contexts of use, and

materials or style of assessment, is important. The following section outlines some key

considerations in this case.

In this thesis, an assessment model was constructed such that each participant marked:

1. Three 'training' exemplars, presented side-by-side (diagnostic-assessment)
2. Two peer-outputs (presented sequentially; peer-assessment)
3. His/her own output (self-assessment)

Across all assessments, the assessment materials were identical, involving the presentation of six rubric elements to be scored on a 9 point scale, followed by three short-text box answers, as described below. In addition, the researcher marked all outputs, with a second assessor marking a subset. This was to allow assessment of expert inter-rater reliability, and to engage in a deductive process similar to that used for the chat data, to address the written outputs in the context of the theoretically derived rubric facets (which are described below). This process entailed aligning the rubric facets with observed features in the written outputs, engaging in a confirmatory process for the a priori model (defined by the rubric) and, if necessary, deriving new elements for a typology as indicated by the data (i.e., developing new rubric facets). The background and materials for this process are described further below.

3:4.4.2 Diagnostic assessment and training

In university settings, first year students are often less exposed to peer and self-assessment, suggesting a need to provide further – consistent – exposure to such assessments to support learning outcomes (Nulty, 2011), with recognised potential for use of technology to support such a process (Davies, 2000; Mostert & Snowball, 2013). Generally, a review of the literature (van Zundert, Sluijsmans, & van Merriënboer, 2010) indicates that training improves peer assessment outcomes, and attitudes towards peer assessment (L. R. Harris & Brown, 2013; van Zundert et al., 2010). However in one survey with 233 students, conducted after peer and self-assessment, students indicated that although they experienced challenges and training is likely a better option, even without training the assessment was educationally productive and motivational (Hanrahan & Isaacs, 2001).

Various means through which to support such training have been used, for example through the provision of exemplar text-assessment pairs (Kean, 2012) and particularly such exemplar-pairs in cases where a teacher guides students through the assessment standards (Hendry, Armstrong, & Bromberger, 2012; Orsmond, Merry, & Reiling, 1996).

In the research described in this thesis, building on the calibrated peer review system described in section 2:5.2, a set of three exemplars were written for both CIS (see Appendix 11)and MDP (see

Appendix 12) tasks, with the exemplars varying along key dimensions – as described by the rubric discussed below. These exemplars were presented alongside each other to facilitate comparison, with each rubric element appearing one at a time. The exemplars fulfilled three key roles, as now described.

First, marking the three exemplars was intended to provide training for the participants with regard to understanding the assessment-criteria. This element was facilitated first by asking students to assign 'high', 'medium', and 'low' scores across the three texts for each criterion, and second by feeding back the 'reference values' to participants following their marking of the texts, along with some explanation for those reference values.

Second, marking the exemplars provided participants with an opportunity to read three texts, all of which were well written in some ways, while lacking in others. The association with the rubric and – following participant marking – feedback regarding the qualities of the texts was intended to provide educational benefit. That is, marking the exemplars was intended to provide a learning experience related to the context and style of the texts.

Finally, the exemplar marking provided a further data source for the research. Because the qualities of the texts are a known quantity, 'reference values' can be defined, with acceptable deviance from those values given, and 'quality' participant-assessors identified with respect to their deviance from the reference values. This affords two opportunities. First, it allows the identification of potentially more accurate assessors in the peer and self-assessment stages, with scope to remove those who do not fall into this category from the assessor data to increase peer-assessment reliability. Second, there is scope for analysis of such deviance in light of performance measures, process data, and the psychometric qualities of the individual assessors. Of particular interest to our context is that epistemological views (based on a somewhat different framework to the epistemic commitment one used in this study) are related to peer-assessment behaviour, with 'constructivist' perspectives on knowledge (as opposed to 'positivist' perspectives)

associated with better outcomes on the task, and higher quality feedback (C.-C. Tsai & Liang, 2009).

3:4.4.3 Rubric use

In order to facilitate the peer and self-assessment, a rubric was developed. Rubrics can be beneficial for formative assessment in many cases (Panadero & Jonsson, 2013), although unlike in this research, the rubric is often given to students in advance to facilitate structuring their work. In this research giving the rubric in advance would have been inappropriate as it would likely influence students' epistemic commitments instead of leaving an open question for students to consider how they understand what is meant by 'the best supported claims'. Despite this key difference, Panadero and Jonsson's review of the evidence suggests that rubric use can be effective in supporting students to improve their performance – a strong educational motivate for their use, and likely to also relate to student performance as peer-assessors. In addition, while there is some concern that students may over-score in peer and self-assessment exercises, rubric use is related to more valid results, although a bias towards over-scoring among friendship groups suggests anonymous grading is the most valid method (Panadero, Romero, & Strijbos, 2013).

A final consideration in the design of the assessment-system, and in particular the rubric, is the number of options presented for marking the rubric elements against. The design of scales involves consideration of both internal reliability, and test re-test reliability. In order to improve reliability, particularly test-retest, 7 point scales have been recommended, alongside the labelling of all points on the scale rather than just the end points (Weijters, Cabooter, & Schillewaert, 2010; Weng, 2004). In a 'mini-clinical evaluation exercise' scale, while 9 point scales only marginally improved inter-rater reliability over 5 point scales, 9 point scales did provide more accurate scores than 5 point scales on average (Cook & Beckman, 2009). This is in line with other work finding a 9 point scale superior in validity (or accuracy) over a 5 point scale in an observation task (Hancock & Klockars, 1991). In a rating exercise (Preston & Colman, 2000) – more closely mirroring our rubric use – participants were asked to rate aspects of a store or restaurant's service. Participants were given a rating exercise in which the only variant between participants

was the number of options available – 2-11, and a '100 point' continuous scale (from which a number needed to be indicated, rather than an option selected). The finer grain (10, 9, and 7) scales were preferred, with higher reliability from these scales too (Preston & Colman, 2000). Furthermore, simulations of scale length indicate that those with fewer points may be less reliable than those with more (typically 7) points (Jenkins & Taber, 1977; Lozano, García-Cueto, & Muñiz, 2008; Nishisato & Torii, 1971). Similar calls for a focus on longer (up to 11 point) scales have been made in marketing research (Darbyshire & McDonald, 2004).

In this work, a 9 point scale is used, with 3 directly labelled points on the rubric (at 1, 5, and 9) and a point range (1-3, 4-6, 7-9 or 'low', 'mid', 'high') given. While labelling all points on the scale has been recommended (Weijters et al., 2010; Weng, 2004), those studies compared end-point labelling only against all points, without consideration of intermediate options. Moreover, in this case labelling all 9 points might have overburdened students with information, given that the labels provided are not simple "strongly agree, agree, neutral...", etc. labels but descriptive sentences providing more information for the student assessment exercise. A 9 point scale was used to enhance the quality of the ratings, and provide an easy gradated set of ranges (1-3,4-6,7-9) for both the diagnostic or training assessment, and the peer and self-assessments. These scores can be treated either in their raw (1-9), or scaled form (1,2,3 or 'low', 'mid', 'high' respectively). On this basis, an assessment instrument was designed. For peer and self-assessment a 6 facet rubric with a 9-point scale was designed; this was reduced to a 4 facet rubric on the reduced 3-point scale for expert-assessment.

In recent related work, Anmarkrud, Bråten, and Strømsø (2014) and Bråten, Braasch, Strømsø, and Ferguson (2014) scored student essays for: explicit or implicit sourcing (i.e. explicit reference to the source, or indirect reference such as "one article spoke of [specific detail]" but without direct use of source information); and reference to trustworthiness of the source or information from that source (coding separately for negative and positive evaluations), and whether connections were made between content-source trustworthiness. In this work, as in that described in section 2:4.5, students did not routinely make explicit sourcing references

(approximately half, with the other half implicit) and did not make reference to the full list of sources (approximately 3 of 6 references). In other work, also on multiple document processing tasks, Goldman, Lawless, Pellegrino and Gomez (2012) identify three clusters of students from their written outputs: satisficers, who selected few sources; selectors who selected many sources but did not connect them; and synthesisers who selected sources and integrated them.

Throughout this prior work the use of key-content, implicit and explicit citation, evaluation of those citations and (separately) their content, and the synthesis of information are foregrounded. Building on the particular task design, and this prior work, the rubric designed (reproduced in Appendix 10) consists of:

1. Topic coverage – The text covers a range of different topics and relates them to the question (the risks of the substance)
2. Range of sources coverage – The text uses a range of sources
3. Quality of sources – The text evaluates the quality of sources cited
4. Clarity of the claims – The text states claims clearly. It gives quotations, numeric indications, and figures where appropriate
5. Quality of evaluation – The text evaluates evidence
6. Synthesis of information – The text synthesizes information from across sources

Where '1' indicates content coverage, '2' the sourcing of that content, '3' evaluation of the source features, and '5' the evaluation of the content of those sources. '6' is then about the ways intertextual ties are identified and made in the text, and '4' about how explicit and specific evidence is given, or quotation made. Note in the final rubric version (Appendix 13, used for expert evaluation), '4' is removed (as part of '1' or '6') as is '5' (subsumed in '3'), following the deductive process described in section 3:4.4.1).

In addition to the rubric element, students were also asked to write 3 improvements that could be made to the outputs they assessed (described in the following section), tagging these with the rubric element labels (or 'other'). Participants were asked to email the researcher or course leader if they had any problems, either with the assessment, or with the free-text feedback received from their peers.

3:4.4.4 Addition of feedback element

While rubrics support assessment, they offer limited scope for the assessing students to engage in qualitative feedback, and the commensurate lack of such feedback to the students being assessed. Xiao and Lucking (2008) contrasted two types of peer-assessment: quantitative only, based on a rubric, versus the same quantitative rubric alongside qualitative feedback. In a quasi-experimental study – with 232 university students – those students in the combined qualitative and quantitative condition were more satisfied with the peer-assessment, and the assessments of peer-assessors in that group were more reliable and valid. In that study, written feedback involved structure - two good things, two improvements, and 150 words of qualitative feedback. This qualitative element is particularly important given that providing comments as a reviewer is educationally productive (Y. H. Cho & Cho, 2011). In addition, in work on a CSCL environment, students receiving feedback expressed satisfaction with a qualitative peer-assessment process, again suggesting the importance of this component of peer-assessment (Prins, Sluijsmans, Kirschner, & Strijbos, 2005).

In the research described in this thesis, in addition to a rubric, the participants were asked to give three ways the text could be improved, and to 'tag' these (using a drop-down menu) with the rubric-elements (or 'other') as labels. Thus for each assessment there is qualitative, labelled, feedback alongside the rubric scores. This method has the advantages of giving qualitative feedback – both to the assessor and assesse – maintains a link to the rubric, and provides another potential source of data for considering the qualities of particular texts and raters (although it is not directly reported in this thesis work).

3:4.4.5 Section Conclusion

In order to assess the written-outputs a marking scheme was developed, giving a numeric indicator for various epistemic-components of writing. The intent of the marking scheme was to provide an outcome variable related to epistemic-commitments, rather than to assess any measure of 'correctness' or scientific accuracy (the final marking scheme appears in Appendix 13). From the literature described above, a number of key lessons have been identified, and implemented in the design of the peer and self-assessment:

1. Peer and self-assessment can be pedagogically useful for both the assessor and assesse particularly where training is provided

2. To avoid bias in assessment (and maintain confidentiality) anonymisation of results should be used
3. Rubrics can be an effective way to (a) give and (b) receive assessment, from peers or experts
4. The inclusion of qualitative feedback in assessment exercises is preferred
5. Longer scales (especially up to '7') are preferred to shorter scales
6. In use of marking-scales, more points than just the end-points should be labelled

Research in a culturally diverse settings like those in our context confirms that when students receive appropriate training, peer- and self-assessment can lead to important insights for students and teachers (Rienties, Willis, Alcott, & Medland, 2013). However, it is worth noting that, as highlighted above, while students can manage self and peer assessment the scope for practice and training in this research was limited, and this may limit reliability of results. For example, a qualitative study of 14 students by Andrade and Du (2007) indicates that students need extended practice to gain the full benefits of self-assessment. Both Reddy and Andrade (2010) and Topping (2010) note while rubrics can be well-liked by students and instructors, and provide valid and reliable feedback, more robust work is needed to establish validity and reliability, particularly across settings and in relation to learning. In addition, it should be highlighted that output content is not assessed by the assessment design, nor the 'correctness' of judgements regarding the sources – thus inaccuracies in content or inaccurate evaluations of the source qualities will not be penalised directly by the rubric, which has not been validated against expert judgements or student understanding. Finally, the text reference values in the diagnostic are not validated against expert judgements (although modifications were made to them in light of external-marking), and are not designed to 'match' student outputs, thus their validity as a training tool may be questioned. Nonetheless, the instruments provide an educationally productive tool for aligning learning outcome indicators with the behavioural indicators described in section 3:4.3.1.1, for exploring the potential of learning analytics for epistemic commitments in a collaborative information seeking environment.

3:5 Initial Piloting

Prior to the main study, three single-pair pilot studies of the lab-based session were conducted, one co-located with the researcher, and two remotely.

The first pair (one male one female) were British PhD students in geology and educational technology respectively. This pair worked co-located, based in the same room, in the presence of the researcher. They were free to ask the researcher questions, or point out issues as they arose, and were then given a short interview afterwards. From that session a number of amendments were made to the task instructions, and to the design of the browser add-on used to enable the collaborative element. These users confirmed that ~45 minutes was an acceptable duration for the main task. However, they noted three key issues: The 'chat' functions at times had a 'lag' in sending/receiving messages; the toolbar appeared to 'lock' the browser at times, preventing any other activity while locked; and some of the documents required further abridging. These participants also confirmed that the readability (formatting and subject-level) of the documents provided, and the sample materials they found in their searching, was appropriate for non-experts (note though that the pilot participants were PhD students).

The remote participants completed the tasks in their own time (collaboratively). The first group (two educated German males, one with a PhD) were then interviewed, while the second (two educated US female library specialists) sent feedback via email. Amendments made to the text instructions in the earlier pilot were effective, and these two remote provided some further small adjustments to those instructions. This remote session also highlighted that conducting the session co-located was preferable, to ensure participants were satisfied they had understood the instructions, and worked together without distraction.

Common to the first two sets of pilot participants were concerns regarding the functionality of the browser add-on, which appeared to "lock" the browser at times, and that chat function for which was sometimes slow. Two rounds of improvements were made to address these concerns, with updates to the software; following these updates the second piloters continued their work and were satisfied that the software was functional. No such concerns were raised by the third piloters. As described in section 3:4.3.1, the Coagmento software has been extensively used in research studies including usability assessment (Shah, 2012a).

In all three cases, the types of chat (in the browser add-on), search, and output created were aligned with researcher expectations, demonstrating clear epistemic activity. Each pair found the task satisfying and interesting. The work described in the rest of this thesis is the main study, with a large sample size, used for exploratory purposes to develop learning analytic techniques in the area.

3:6 **Procedure**

As noted above, the procedure followed a number of steps: a pre-lab; in-lab; and post-lab elements, as indicated in Figure 3:4 which gives an overview of the procedures followed. The figure includes a number of arrows indicating group-splits, such that all tasks in the middle lines were taken by all participants, while splits indicate that the MDP and CIS task differed and those tasks were thus defined by topic.

Time	Weel 1: Pre-lab	Week 2: Lab				Week 3: Post-lab		
Duration (minutes)	20-30	10	55	10		60-90	5	0
Task	ISEQ →	Introduction & consent →	Warmup task	Main MDP task → Post-task survey ← Main CIS task	MDP Trust → Demographics ← CIS Trust	MDP Training → MDP peer/self assessment → CIS Training → CIS peer/self assessment	Task-survey →	GPA obtained
Context	Individual	Individual	Collaborative	Individual		Individual		Instructor

Figure 3:4 – Study Timeline

3:6.1 Pre-lab

Prior to the lab session the student cohort was informed of the possibility of taking part in the research study, and briefly what it would involve. They also completed the ISEQ (along with various other survey items), as part of their usual course. During this time, a computer lab at Maastricht University was selected, and Firefox (Mozilla, 2014) installed along with the customised Coagmento browser add-on (Shah, 2014a).

3:6.2 Lab session

Before each individual study session, PCs were logged on to a generic logon, with Firefox open, and on the 'login' page for the study. The browser cache was cleared and any extra windows or programs open closed. Each PC also had a paper copy of core instructions (Appendix 1), and the times for each task were written on a board at the front of the room for researcher and participant reference. Those times are given in Table 3:5.

Table 3:5 - Lab-session Timings

Time	Activity
0-5	Introduction to session from lab-assistants and primary researcher
5-10	Login, consent, basic familiarisation with Coagmento
10-20	Warmup task (3 minute warning given at end)
20-65	Main task (10 minute warning given at end)
65-75	Post-task survey

Each lab-session alternated between the MDP and CIS task, with logins associated to one or other task. In other words, participants were directed to either CIS or MDP task depending on the login they received at the beginning of the session. Each session was facilitated by the researcher (Simon Knight), and one lab assistant. The lab assistants were final year undergraduate students at Maastricht who are paid to assist in computer-lab sessions; because lab assistants rotate through sessions, occasionally two lab assistants would be present for part of any given session. At the start of each session, the lab-assistant outlined the context of the research study, and introduced the primary researcher to the class. The lab assistant reminded students they could participate in one of two tasks – the research task described in this thesis, or an optional alternative task – and gave any other course notices. The lab assistant also reminded students to

primarily work in English (which is the student's usual practice at Maastricht). The primary researcher then introduced the study in more detail, noting the three tasks – a warmup, the main task, and the at home assessment task – and surveys to be taken. It was highlighted that the tasks would likely be in an area the students knew little about, but that this choice was deliberate, to see how people use tools to find and evaluate information together, and that hopefully the information would be novel and interesting to the students. It was also noted that they might feel like there was more information than they could deal with in the time; again this was flagged as an intentional feature of the task, noting that the participants might need to make decisions about what resources to focus on. In the case of the MDP task students were told they should primarily use the documents provided, but that they could use the internet if necessary (for example for translation or to look up definitions – a practice they are encouraged in in their usual classes). For the CIS task, students were told they would be asked to search the internet for resources. In both tasks it was indicated that the research interest was in how people find and evaluate information together and using a tool to help support that. At this stage and throughout the session participant questions were answered with reference to the provided text instructions. Periodic software issues in the session were dealt with by refreshing the sidebar, or instructing the participant to click 'home' to reset the browser addon and reload the active task.

Following the introduction to the session students were given logins for the research-website, using separate PCs (seated roughly back to back), with a userID (used as the primary identifier throughout the research) and simple password; as part of the login process students gave their Maastricht student ID twice (matched to ensure no typographical errors). On login participants were instructed to read a briefing sheet and tick a box to confirm consent to participate, and then to make use of the instructions to continue on the tasks.

Once participants had passed the consent pages, the first task presented was a 'warmup' task – three short fact-retrieval questions as indicated below – proving a period in which partners could familiarise themselves with the tool and each other. For this warmup task, each question for each

pair was populated with a randomly selected country (per Appendix 4) from a pre-created list, to minimise the risk of copying the answers of those seated nearby.

> *Please type the answers to the following **three prompts** in your Task Pad (click in the bar at the top of the browser). **You may use the internet** to find the answers.*
>
> - *In 2010 what was the educational expenditure per primary student in [XXX] as a % of GDP?*
>
> - *In 2010 what was the total health expenditure as a % of GDP in [XXX]?*
>
> - *How much (in US dollars) does a big mac cost in [ZZZ]?*
>
> *If you find the warmup taking too long (over 10 minutes) but you feel you're now comfortable with using Coagmento, you should move on to Task 1.*

During this slot, the lab assistant and primary researcher addressed any queries, and assisted participants in identifying the various functions of Coagmento and the research-website. Specifically, participants were reminded to open their sidebar to view the 'chat' and queries of their partner, shown where the 'task pad' was located to write their answers, and shown the 'home' and 'active task' buttons so they could get back to the question prompt. The task pad was pre-populated with some guidance text for its use as in Appendix 5. Towards the end of this slot the 'submission' button was highlighted to participants, noting that both partners in the team would need to click 'submit' in order to move on to the main task. Participants were given a 3 minute warning, and encouraged to submit after 10 minutes on this task. The researcher and lab assistant ensured all participants started the main task with minimal difference, and participants spent approximately 45 minutes on this task (as described above in section 3:4.2.2), receiving a 10 minute warning before the end of that slot. The session ended with the short (under 10 minute) exit questionnaire described above (section 3:4.1.2). At the end of each session the procedure described above was followed to setup for the following group.

3:6.3   Post-lab

At the end of the lab-session week, all participants were emailed (by Dr Dirk Tempelaar, the course leader and Maastricht collaborator, as in Appendix 9) with a link to the second (at home) task, and asked to complete it within a week. This task involved logging in (using their Maastricht student ID), completing a training exercise or diagnostic assessment (described in 3:4.4.2 above),

followed by marking two peers' outputs in sequence, and then finally marking their own collaboratively authored output. This was followed by a short feedback survey (described in section 3:4.1.3), after which participants were thanked for taking part, and informed all tasks had been completed.

Following the second task, all participants were sent a debrief sheet giving further details of the study and inviting them to contact the researcher if they had any further questions. They were also sent a link to access the feedback from the assessment exercise at this time, and informed that they should contact the researcher or course leader if any feedback was inappropriate[20].

3:7 Chapter Summary

The preceding sections outline the materials and procedure used in this research. From this data I have previously (S. Knight, 2015) given a visual depiction of a simplified version of the dataset as in Figure 3:5. This figure depicts the survey data at the top, with the conceptual structure of the ISEQ (as a cognitive model) featuring strongly. In the figure, the ISEQ facets are represented thus: simplicity is represented by compartmentalization (of knowledge claims as isolated facts) versus connection (of knowledge claims as inter-connected); justification by rules of inquiry (science equipment; representing argument schemas and inquiry processes) versus observation (an eye; representing personal observation); certainty represents beliefs regarding how stable knowledge is (represented by the clock, indicating tentative versus stable knowledge); and source by belief in the 'given' versus active construction of knowledge (taking knowledge from the page, versus active probing of the source quality, and engagement with others to build knowledge). Below that we see various trace features, particularly page navigation (including querying) and the chat data. The outcomes then are the facets of the assessment rubric.

---

[20] No students contacted either the course leader or researcher for this reason. Unfortunately the system did not, at that time, track which students logged in to check their feedback, so we have no record of this behaviour.
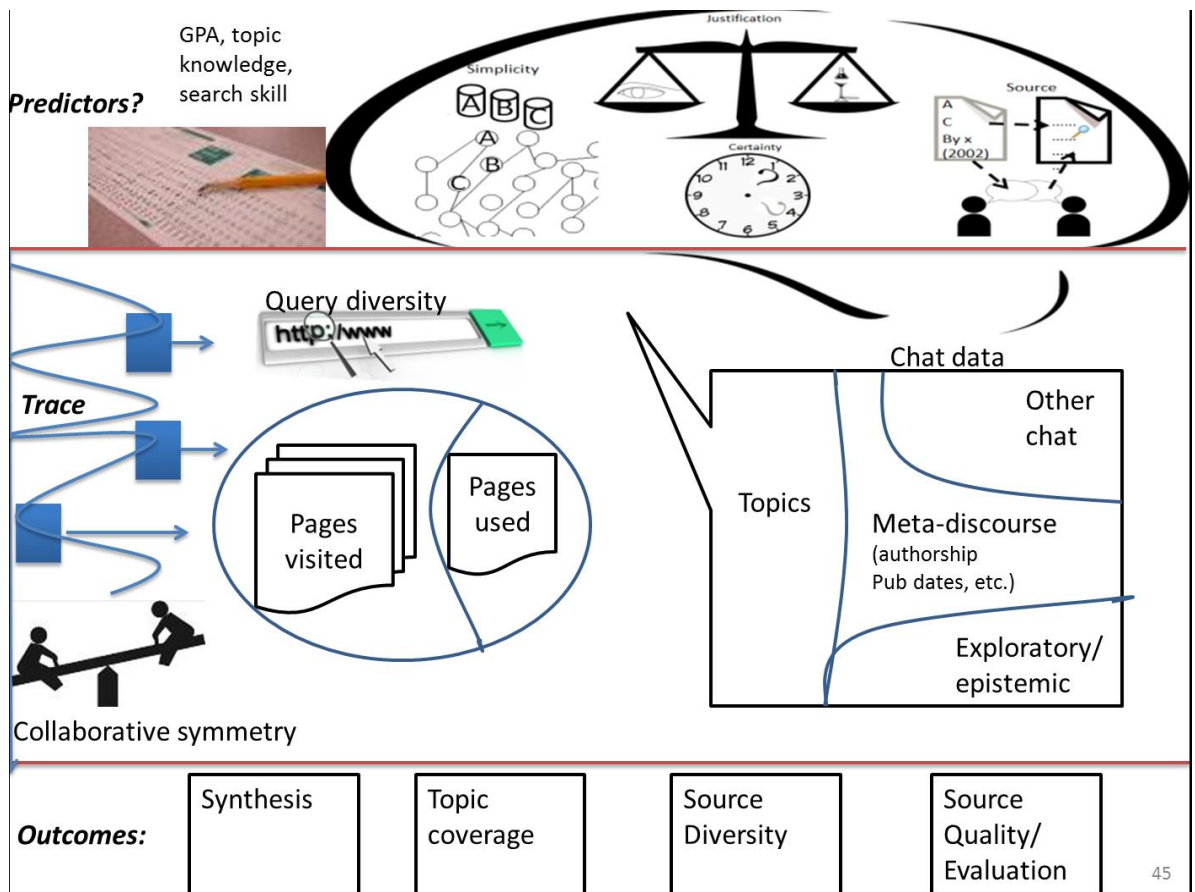
Figure 3:5 – Simplified Data Model

The results chapter will discuss this data in the context of the research questions. First contextual information will be given, using a subset of data from the survey instruments to describe the participant population in both tasks, and compare them to establish whether the participants in the two tasks can be considered part of the same population. In all cases participant numbers are given with reference to the particular analysis; due to varying missing data in both individuals and pairs providing a stable n across all analyses is not possible. Relatedly, analysis is conducted on both individually based and paired data; where paired data is analysed in some cases means are used (e.g. survey items) while in others aggregates are used (e.g. number of pages visited).

The first question regards outcome assessments. To address this question, first, the possibility of identifying 'high quality' assessors using the diagnostic/training exercise is discussed. Then data from the peer, self, and expert assessments of the textual outcomes is analysed for reliability. Descriptive statistics for the outcomes are provided and give contextual information on participant success in both tasks.

The second question regards the relationship between the ISEQ and outcomes. To address this question, first a factor analysis of the ISEQ is conducted, to establish its internal structure. Building on earlier research (see section 3:4.1.2.1) the ISEQ is then discussed in the context of the trustworthiness assessments, which provide a task-specific explicit insight into epistemic commitments. Finally, an analysis of the relationship between the ISEQ and outcome is given.

The third question regards the relationship between outcomes, and trace and survey data. To address this question, first a set of exemplifications of the epistemic nature of the tasks is given drawing on the chat data, to establish that the trace data has the kinds of epistemic properties anticipated in such a task. This analysis foregrounds both the potential of the development of a typology for analysis of chat data (through exemplifications of the anticipated types observed), and the nature of the rich data that is reduced through use of such a device. Descriptive statistics regarding other trace facets are then given, alongside discussion of key issues. A correlation analysis is conducted to identify relationships – and lack thereof – between key survey and trace variables. Finally, a regression analysis is conducted using the outcome variables as the dependent variable and trace and survey items as predictors.

Throughout the results section an understanding of the pedagogic context, and the theorised task design presented above, is important to understanding the analysis conducted. The provision of qualitative analysis in section 7:1.1.1is intended to further this contextualization of the research conducted through an exemplification of the epistemic nature of the tasks. The thesis thus far has described the motivation for developing learning analytics around epistemic cognition in information seeking, and a particular paradigm – collaborative information seeking – for doing so. The above methods describe an approach for this research. The following research questions reprint those given in the introduction, and literature review, drilling down further in the context of the specific methods described above.

1. Can epistemic markers of selection, evaluation, and integration of claims in a written output from a designed information seeking task be reliably identified?
    - Can the ways participants select sources (source diversity) and claims (topic coverage and claim clarity), evaluate (source quality and evaluation), and

integrate (synthesis) claims in a written output from MDP and CIS tasks be reliably identified by peers, self, and experts?

2. What relationships can be identified between facets of epistemic cognition on a psychometric instrument, and written outputs from a designed information seeking task (as in '1')

   - What relationships can be identified between scores on the ISEQ psychometric, and success on a written task (as in '1')

3. What is the relationship between trace behaviours on an information seeking task, epistemic properties of written outputs (as in '1'), and epistemic cognition measured on a psychometric instrument (as in '2')?

   - What relationships can be identified between trace features (described above) success on a written task (as in '1') and ISEQ scores (as in '2').

Chapter 4: INTRODUCTION TO RESULTS

4:1 Results overview and Analysis Plan

The following chapters present the primary results of analysis of the research data. In the

following sections I:

1. Present an overview of the participant survey data, which provides important context for and constraints on the analysis given

2. In Chapter 5, discuss RQ1, concluding that data from the peer-assessment task was not reliable, but that expert evaluation of the outputs on a modified rubric was reliable

3. In Chapter 6, discuss RQ2, outlining the ISEQ data used for this question, and its relationship to the outcome measures established by RQ1

4. In Chapter 7, discuss RQ3 by outlining relationships among the ISEQ, outcome, and trace data, both through the provision of statistical analysis, and through qualitative analysis of the chat data which exemplifies the epistemic nature of the tasks

In order to answer these questions, a set of analytic approaches were taken. Across this analysis

the specific context within which the data was gathered should be considered; that the tasks are

designed epistemic information seeking ones; that the participants are asked to interact with

resources in a tool mediated (social) environment to produce an authentic output text; and that

the participants are engaged in part of their own assessment practice through the use of a peer

and self-assessment tool.

To address the first question, the data from the diagnostic assessment and training was first

analysed, to investigate error rates in assessing text qualities against the reference values.

Building on this, inter-rater reliability analyses were conducted across the peer and self-

assessment data to establish whether the participants could reliably assess their own outputs,

using subset of participants including from 'all' to only those with relatively lower error rates on

the diagnostic. Following a deductive analysis of the written outputs and confirmation (and

amendment) of the rubric, a corresponding inter-rater reliability analysis was conducted between

expert raters.

To address the second question, first a factor analysis was conducted on the ISEQ data to

establish the psychometric properties of the instrument. Analysis of the trustworthiness ratings

was then conducted, to establish relationships between ratings of document types (in the MDP task) and descriptives in both tasks; these ratings were analysed for relationships to ISEQ scores. Finally, regression analyses were conducted to investigate relationships between ISEQ scores and outcomes (derived from analysis in RQ1). This analysis is used to draw comparison with earlier research conducted in different (generally, individual, self-report based) contexts, and to highlight the advantages and disadvantages of analysis of survey data, outcome measures (on the written output), and behavioural trace (including dialogue data).

The third question makes use of data from the first and second questions. Exemplifications of the epistemic nature of the task are given through analysis of the chat data. These exemplifications are drawn on to highlight the ways in which participants did (and did not) engaged in dialogue mediated interaction with each other and the information resources. This analysis also draws attention to the rich nature of the dialogue data, and the ways in which reduction to a typology offers a coarser grain analytic device (while nonetheless offering insight into the broad typology). Descriptive statistics, and exemplifications of the log-data (for example, frequent URLs) are given to illustrate the range of behaviours observed across groups. Correlation analyses were conducted between these behavioural traces, and the survey items (including the ISEQ). The chapter ends with a stepwise regression analysis to investigate relationships between the trace, survey data (including the ISEQ, derived from analysis in RQ2), and outcomes (derived from analysis in RQ1).

In each chapter, the discussion is broken into sections (as briefly described as the prelude to each chapter), with a separate 'results' and 'discussion' section for each question. The discussions are intended to foreground the main findings and implications of those findings for the particular analysis discussed in the chapter, along with the limitations of those results, and the future research direction indicated by the analysis. Further analysis across the questions is given in section 8:1.1 which discusses the research as a body of work more generally. The bulk of the analysis described here was conducted using the R statistical programming language (R Core Team, 2015), on RStudio (RStudio, 2013), hosted by the OU's cRunch project

(http://crunch.kmi.open.ac.uk). Where possible I have cited the individual package (collection of

functions) I have used, alongside the particular function used[21].

Qualitative analysis was facilitated by NVivo qualitative data analysis software (QSR International

Pty Ltd, 2012). Using this tool messages meeting chosen criteria can be selected such that all

messages from a group, task or containing a particular term may be viewed and analysed. In the

analysis presented here (section 7:1.1.1), a concordance style analysis is presented to exemplify

the ways in which terms in the typology (described in section 3:4.3.1.1.3) are used by participants;

as such, messages in which terms appear are identified and presented in full, to indicate the

context in which the terms appear across messages. In addition, a closer analysis of a subset of

groups is given alongside more detailed commentary describing the participants' interaction with

each other and their resources (including through the dialogue presented). In both analyses the

intent is both to illustrate the ways in which the terms targeted in the typology exemplify

epistemic dialogue, and the complexities of capturing the rich epistemic-commitments expressed

through such dialogue in the reduction of dialogue to a typology.

Where possible, effect sizes are given, typically in terms of the correlation coefficient r (or r^2 in

regressions) and Cramer's V for categorical data (Chi-square) which can be read in the same way

as r. Following Cohen's (1992) guidance on the magnitude of effect sizes, correlation coefficients

r of .10 are considered "small", of .30 "medium", and of .50 "large". Significance levels are given

to two decimal points alongside a notation, except in correlation matrices where notation alone

indicates the significance level. The following notation indicates the levels of significance

observed: <.001 = ***; <.01 = **; < .05 = *; <.1 = ^.

4:1.1   Background Demographic and Questionnaire Items

Following the study, a grade point average was computed for each student. Grades in the

Netherlands are given on a 1-10 scale, with a passing mark of 5.5. For participants in this study,

---

[21] In addition to the statistical packages referenced throughout, the RMySQL (James & DebRoy, 2012), plyr (Wickham, 2011), reshape2 (Wickham, 2007), zoo (Zeileis & Grothendieck, 2005), doBy (Højsgaard, Halekoh, Robison-Cox, Wright, & Leidi, 2012), and knitr (Xie, 2014, 2015) packages have all been invaluable in shaping and understanding my data, as have various community forums, particularly stackoverflow (see http://stackoverflow.com/users/2485335/sjgknight)

the average grade across 4 available marks (where available) was computed, giving M = 6.53 (SD = 1.35, n = 308) for CIS, and M = 6.43 (SD = 1.34, n = 262 i.e. 8 missing values) for the MDP task; unpaired t-test indicates there is no significant difference between the groups on this GPA measure $t(568) = 0.88$, $p = .38$, $r = .04$. There was also no significant difference in age between the CIS (n = 308, M = 19.01, SD = 1.32) and MDP (n = 270, M = 18.98, SD = 1.24) groups, $t(600) = 0.30$, $p = .70$, $r = .01$.

Table 4:1 indicates unpaired t-tests between the two tasks for partner familiarity, partner agreement, and prior knowledge, indicating significant differences on the final measure on a scale of 1-10, with a higher level of prior knowledge in the MDP task, although with a small to medium effect size (r .22). This indicates that there are no significant differences between the CIS and MDP populations with the exception of prior knowledge for which the MDP task had higher prior knowledge than the CIS group, perhaps because the specific topic of glyphosate has received more press coverage than red yeast rice (although of course the more general topic of statins has received much coverage).

Table 4:1 – Participant-partnership context

	CIS (n = 308)		MDP (n = 266[+])		
	M	SD	M	SD	t-test
Partner familiarity	3.69	2.86	3.61	2.90	$t(572) = 0.33$, p = .74, $r = .01$
Partner agreement	8.12	1.80	7.86	1.68	$t(572) = 1.78$, p = .08^, $r = .10$
Prior knowledge	1.95	1.87	2.87	2.20	$t(572) = 5.42$, p < .00***, $r = .22$

[+]4 missing values

Results of a t-test (given in Table 4:2) indicate a small ($r = .12$) significant difference in search experience between the CIS (M = 3.63, SD = 0.85) and MDP (M = 3.48, SD = .85) groups, with chi-square tests (and Fisher's Exact test for multiple frequencies less than five, denoted by an f in the row title) indicating no significant differences between CIS and MDP distributions on the other measures. Given the survey data was obtained following the task, it is possible that engaging in search in the CIS task prompted a higher self-reported level of search experience than in the MDP task in which such search was not a requirement.

Average daily search frequency

	1-3	4-6	7-10	10+	Occasionally	df	n	C2	p	V
CIS	42	93	66	89	18					
MDP	37	84	63	80	5	4	578	6.16	.19	.10

Search experience

	1	2	3	4	5	df	n	t	p	r
CIS	6	15	109	136	42					
MDP	4	26	103	111	26	576	578	2.10	.04*	.12

Most used search engine (f)

	Bing	Google	Yahoo	Other	df	n	p
CIS	1	269	01	30			
MDP	0	304			3	578	.20

Most used browser

	Chrome	Firefox	IE	Safari	Other	df	n	C2	p	V
CIS	143	48	16	96	5					
MDP	131	38	12	89	0	4	578	5.05	.28	.09

Most used operating system

	Linux	Mac	Windows	Other	df	n	C2	p	V
CIS	2	134	91	81					
MDP	1	122	78	69	3	578	.36	.95	.02

Following the at home task, students also completed a feedback element, as indicated in Table 4:3. Analysis of this data indicates generally positive levels of satisfaction regarding the collaboration (7.32 and 7.20), with a more neutral response to the task generally (6.05 and 5.68) and the browser addon (6.61 and 6.54) with a small (r = .10) significant difference in task satisfaction between the CIS and MDP tasks indicating a marginal preference in terms of task satisfaction for the CIS task than the MDP. Across all measures there large standard deviations can be observed (from 1.94 to 2.16) implying within group differences in responses to the questions. The general feedback (a text field) will not be reported in full here except insofar as it addresses discussion of the research questions in section 8:1. However it is worth noting here that some participants reported technical issues with the at home task, and having to repeat that task; this may be associated with the significantly lower task satisfaction rating. See the discussion for further consideration of this issue.

Table 4:3 – Feedback measures

	M	SD	n	df	t	p	r
Collaboration satisfaction (1-10)							
CIS	7.32	2.11	296				
MDP	7.20	2.16	247	541	0.63	.53	.03
Task satisfaction (1-10)							
	M	SD	n	df	t	p	r
CIS	6.06	1.94	296				
MDP	5.68	2.04	247	541	2.23	.03*	.10
Intuitiveness of browser add-on (1-10)							
	M	SD	n	df	t	p	r
CIS	6.61	2.02	296				
MDP	6.54	1.99	247	541	0.42	.67	.02

Correlational analysis, reported in Table 4:4, (using Pearson's method) on pair-mean scores (i.e. the average ratings given to, for example, task satisfaction) and individual scores indicated relationships between:

- task satisfaction and ratings of browser addon intuitiveness, suggesting that those who found Coagment more intuitive to use also found the task more satisfying;
- collaborative satisfaction and task satisfaction, suggesting that those who found the collaboration more satisfying also found the task more satisfying (or the converse) perhaps because their collaboration increased task satisfaction, or because those more satisfied with the task were more likely to engage in satisfying collaboration;
- finding the addon intuitive, and partner agreement, perhaps indicating the success of the awareness features of Coagmento in groups scoring higher on these items,
- for the MDP task only, a negative relationship between collaborative satisfaction and GPA, indicating that those with lower GPA found collaboration more satisfying (and the converse, that those with higher GPA were less satisfied by collaborating, or at least their experience of the collaboration). That this was true of MDP but not CIS indicates that the two tasks may be perceived differently, and that GPA may impact on this perception.
- Also for the MDP task only, a positive relationship between partner agreement and search experience, indicating that those who had higher search experience had higher levels of agreement, perhaps because they were superior at finding and identifying commensurate claims; it is interesting this was not true of the CIS task (although note the relationship is significant at the .10 level in the individual CIS data).
- And for the CIS task only, add-on intuitiveness appears to be weakly correlated with search experience, indicating that those who rate their search skills higher also found the addon more intuitive.

This data indicates that although there are small differences between the MDP and CIS participants, the populations appear to be generally homogenous along the dimensions described above. This implies that differences between CIS and MDP tasks in outcome, ISEQ, and trace, are

not grounded in differences in the above dimensions (although within-group differences may still

be related to these demographic factors).

Table 4:4 – Survey data correlation matrix

	age	search Experience	collaborative satisfaction	task satisfaction	addon intuitiveness	GPA	topic knowledge	partner agreement	partner familiarity
age		-.05 (-.09)	-.05 (.00)	.11 (.07)	.04 (.07)	.03 (.09)	.06 (-.01)	-.06 (-.06)	-.02 (-.01)
search Experience	-.13 (-.07)		.06 (.11^)	-.04 (-.03)	.16^ (.13*)	-.09 (-.01)	-.03 (-.01)	.12 (.10^)	.02 (.00)
collaborative satisfaction	-.12 (-.01)	-.04 (.06)		.24** (.26***)	.41*** (.45***)	.05 (-.05)	-.14^ (-.02)	.36*** (.28***)	-.03 (.01)
task satisfaction	-.13 (-.03)	-.01 (.10)	.41*** (.39***)		.43*** (.40***)	.12 (.06)	-.01 (.04)	-.01 (.01)	-.09 (-.02)
addon intuitiveness	-.12 (-.02)	.02 (.07)	.41*** (.41***)	.42*** (.39***)		-.05 (-.02)	-.02 (.02)	.02 (.00)	-.11 (-.06)
GPA	.00 (.03)	.12 (.03)	-.23* (-.17**)	-.16 (-.15*)	-.03 (-.05)		-.06 (-.09^)	.06 (.02)	-.01 (-.05)
topic knowledge	-.06 (-.04)	.14 (.13*)	.01 (.04)	.15 (.17**)	.01 (.04)	-.07 (-.14*)		-.11 (-.04)	.09 (.08)
partner agreement	-.03 (.00)	.24** (.15*)	.28** (.20**)	.11 (.08)	.19^ (.13*)	-.12 (-.05)	-.11 (-.04)		.04 (.08)
partner familiarity	.02 (.01)	.07 (.06)	.07 (.03)	.16 (.16*)	.10 (.10)	-.02 (.02)	.19* (.17**)	.16^ (.11^)	

*Unbracketed correlations refer to paired data, with individual data bracketed. Below the diagonal is based on MDP data (paired n = 102, individual n = 234), above on CIS data (paired n = 138, individual n = 294).

4:1.2   Comparison of Study Group to Excluded Data

As described in the participants section (section 3:2) a number of participants were excluded from

analysis. Data was obtained for a cohort of these participants (n = 298 plus 1 missing), providing a

comparison group for the n = 578 total cohort (although note that samples in the exclusion group

are pooled across task for comparison purposes). Analysis of survey data for both groups, shown

in Table 4:5 (t-tests) and Table 4:6 (Chi-squared and Fisher's exact tests) indicates no significant

differences in variance between the focus study group, and the excluded data group on most

measures; there is a small effect (V = .11) of group on search frequency (Table 4:6), with reduced

frequencies for the excluded group in the rightmost three categories. These categories are the

higher search frequencies plus 'occasionally'; given that 'search experience' is not significant (in

Table 4:5) and the small effect size indicated, this result is unlikely to be meaningful. This indicates

support for the claim that the group that is the primary focus of the analysis in this thesis is

homogenous with the participants excluded from this analysis, and that there is no selection bias

between the sessions that might impact on the generalizability of the analysis to the whole

cohort.

Table 4:5 – T-test comparison of focal group data and excluded data

	M	SD	n	df	t	p	r
GPA							
Study	6.48	1.35	569				
Excluded	6.32	1.28	247	500	2.00	.10	.06
Age							
Study	19.00	1.28	578				
Excluded	19.00	1.20	298	600	-0.20	.80	.00
Topic knowledge (1-10)*							
Study	2.38	2.08	574				
Excluded	2.79	2.87	28	30	-0.70	.50	.08
Search Experience (1-10)							
Study	3.56	0.85	578				
Excluded	3.52	0.80	298	600	0.70	.50	.02
Collaborative satisfaction (1-10)							
Study	7.27	2.13	543				
Excluded	6.54	2.71	28	30	1.00	.20	.15

Task satisfaction (1-10)						
	M	SD	n	df	t	p	r
Study	5.89	1.99	543				
Excluded	6.00	1.49	28	30	-0.40	.70	.03

Intuitiveness of browser add-on (1-10)						
	M	SD	n	df	t	p	r
Study	6.58	2.01	543				
Excluded	6.82	2.04	28	30	-0.60	.50	.06

Partner agreement (1-10)*						
	M	SD	n	df	t	p	r
Study	8.00	1.75	574				
Excluded	8.07	2.54	28	30	-0.10	.90	.02

Partner familiarity (1-10)						
	M	SD	n	df	t	p	r
Study	3.65	2.87	574				
Excluded	3.58	3.22	28	10	0.09	.90	.01

*⁺Bartlett's test of indicated significant differences between the variance of the study and excluded group's topic knowledge, $K^2(1) = 7.00$, p = .01; and partner agreement, $K^2(1) = 9.00$, p = .003; therefore Welch's correction was applied to these t-tests.

Table 4:6* - Chi Square and Fisher's Exact Test Comparison of focal group data and excluded data

Gender								
	Female	Male			df	n	C2	p	V
Study	225	353							
Excluded	127	171			1	876	1.11	.29	.04

Average daily search frequency								
	1-3	4-6	7-10	10+	Occasionally	df	n	C2	p	V
Study	79	177	130	169	23					
Excluded	48	107	58	77	3	4	871	10.00	.04*	.11

Most used search engine *(f)*						
	Bing	Google	Yahoo	Other	df	n	p
Study	1	573	1	3			
Excluded	1	295	1	1	3	876	1.00

Most used browser								
	Chrome	Firefox	IE	Safari	Other	df	n	C2	p	V
Study	274	86	28	5	185					
Excluded	48	13	6	1	31	4	677	.50	1.00	.03

Most used operating system *(f)*						
	Linux	Mac	Windows	Other	df	n	p
Study	3	256	318	1			
Excluded	0	132	165	1	3	876	.80

*Where frequency counts in many cells are less than five, Fisher's Exact Test is used (for which a test statistic and effect size are typically not reported), denoted by an (f) in the row title.

Chapter 5:   RESULTS AND DISCUSSION OF ANALYSIS OF RESEARCH QUESTION ONE

5:1 Research Question 1: Reliable Identification of Epistemic Properties in Written Outputs

5:1.1   Research Question 1: Results

The first research question concerns the written outputs created by each group in completing the tasks.

1. Can epistemic markers of selection, evaluation, and integration of claims in a written output from a designed information seeking task be reliably identified?
   - Can the ways participants select sources (source diversity) and claims (topic coverage and claim clarity), evaluate (source quality and evaluation), and integrate (synthesis) claims in a written output from MDP and CIS tasks be reliably identified by peers, self, and experts?

I first discuss the diagnostic assessment/training, describing the participants' performance on this task and its use in selecting 'high quality' assessors for the purposes of peer and self-assessment reliability analysis. I then give an analysis of the peer and self-assessment data, using inter-rater reliability analyses on a number of sub-sets of participant-raters, before describing analyses of expert inter-rater reliability. Finally the outcomes on the written output text are described using descriptive statistics and comparison (using Chi-squared analyses) between the CIS and MDP tasks. These sets of analyses focus on the possibility of reliable identification of markers of selection, evaluation, and integration of claims in the written outputs along the dimensions of the rubric described in section 3:4.4.3. The implications and limitations of these analyses are then discussed, along with some potential future research directions.

5:1.1.1 Diagnostic assessment

740 participants completed the diagnostic assessment/training (described in section 3:4.4.2); 390 in the CIS task and 350 in the MDP. All cases are reported here, as all participants engaged in the peer and self-assessment (thus outcomes on the diagnostic assessment relate to peer-scoring of the group of primary interest in this reporting). This task involved marking three texts on a rubric, such that for each rubric element one of the texts was in the 'high' range (7-9), one in the 'medium' range (4-6), and one in the 'low' range (1-3). Theses scores can thus be compared to a set of reference values for the texts, comparing researcher assigned values to those given by the participants.

Performance on the diagnostic assessment was measured in two ways. First, for each facet of the

rubric, the participant's ordinal ranking (1,2,3) was identified for each text, and an absolute

difference to the ordinal reference value computed[22]. Second, an absolute difference was

computed between the participant's positioning of the text on the reduced-scale (1,2,3 or low,

medium, high) and the reference values. Table 5:1 gives an example using the raw scores

indicating how the two methods provide small differences in results, indicated in the right most

two columns.

Table 5:1 – Example Diagnostic Performance Indicators

	Reference value	Order/ Position	Participant rating	Participant order	Participant position	Ordinal difference	Position difference
Text 1	1	1 (low)	4	1	med	0	1
Text 2	4	2 (med)	6	3	med	1	0
Text 3	8	3 (high)	5	2	med	1	1

Thus, the largest score possible on any facet is 6, or 4 if participants adhered to the instruction to

rank order the texts (because swapping the top and bottom rank gives two difference scores of 2,

and a 0), giving a total maximum score of 36, which would indicate maximal deviance from the

reference values. These differences were then summed across facets, and two total scores

created.

From these scores, 2 groups were identified for each method, those who scored '6' or under – i.e.,

they were 'off' by 1 point on 3 facets, or 2 points on 1 facet and 1 point on another – and those

who scored '12' and under. An overall score for each threshold was created by taking the lowest

of the two methods (ordinal or position difference). As Table 5:2 indicates, particularly for the

MDP task, this method reduces significantly the number of 'good' raters, particularly where the

stricter (error under 7) criterion is applied.

Table 5:2 – Error Frequencies in CIS and MDP groups

	n participants	
	CIS (n = 390)	MDP (n = 350)
Ordinal error =<6	25	1
Ordinal error =<12	156	84

---

[22] Note that due to a software problem some users submitted answers twice with minor differences between them; the average of their scores is taken in all analysis (hence an ordinal rating of 2.5, for example, is possible).

Position error =<6	106	12
Position error =<12	281	121
Lowest error =<12	281	130

The error frequency distributions for the ordinal (first 6), position (second 6), and lowest score of

the two (last 6) is shown for the CIS (pink, n = 390) and MDP (blue, n = 350) tasks in Figure 5:1,

created using the ggplot2 package (Wickham, 2009). These charts indicate that students had

higher error rates in identifying source quality evaluations, and claim clarity, in the MDP than the

CIS task.
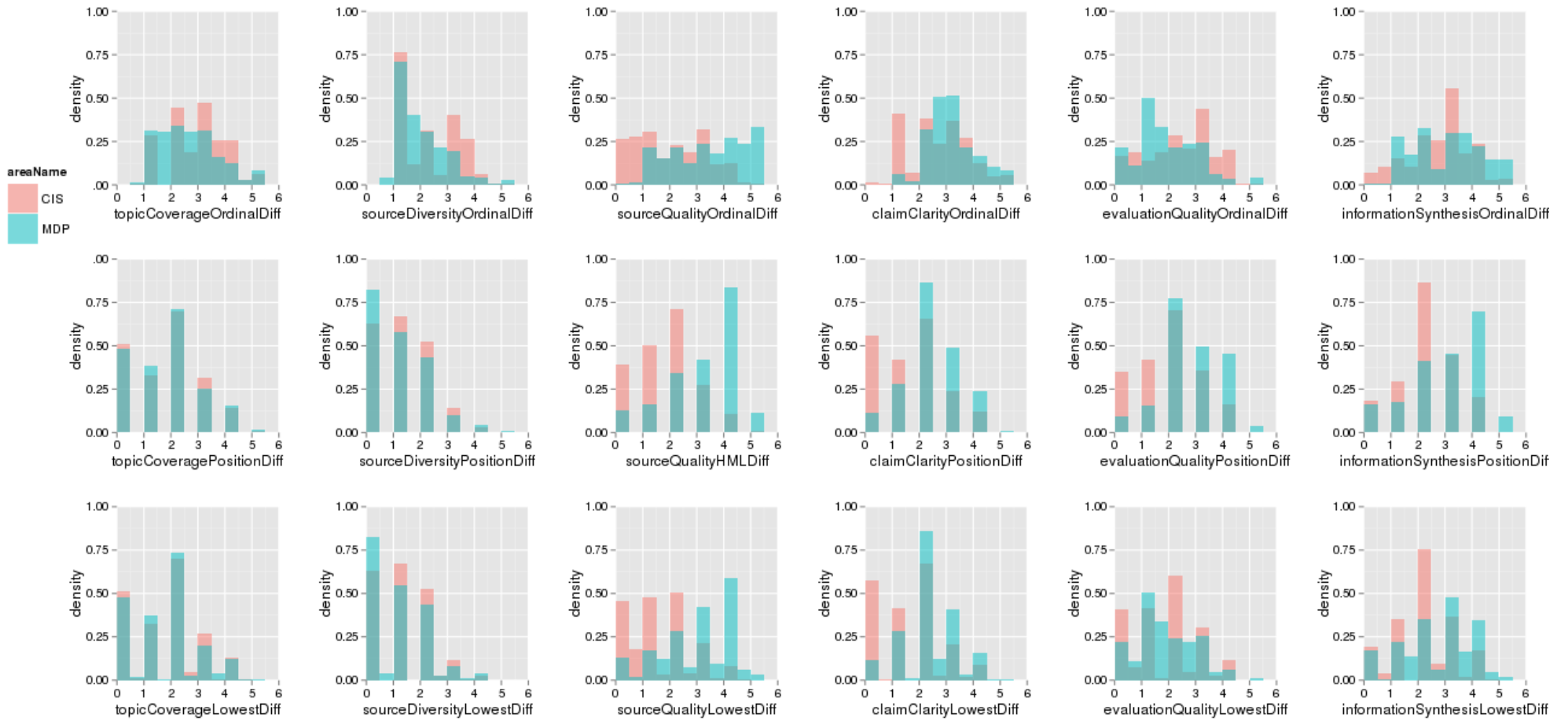
Figure 5:1 – Rates of Error on Diagnostic Assessment

5:1.1.2 Peer and Self-Assessment Reliability

A set of peer and self-assessments were collected for each text-output produced with each peer-rater assessing two random outputs from within their topic area, resulting in a mode of 3 peer, and 2 self-assessments[23]. Krippendorff's alpha (Hayes & Krippendorff, 2007) was computed using the irr package (Gamer, Lemon, Fellows, & Singh, 2012) based on the raw ratings (1-9; interval level data), and the reduced scores (1,2,3; ordinal level data)[24]. Krippendorff's alpha is a measure of agreement across raters, which can be used to assess agreement over any number of raters, and varying levels of measurement; it is thus highly suitable for this data in which outputs were assessed by different assessors, with unequal numbers of raters for each output. Krippendorff's alpha was computed for outputs, for all facets with the following rater subsets:

1. All raters, or subsetting to only raters with an error margin of: 1; 2; 3; 4; or 5, on the particular facet selected
2. All outputs rated, or subsetting to only those pads rated by at least: 1; 2; 3; or 4 raters
3. All raters, or subsetting to just peer-raters

And the following output subsets:

- All group sizes, or subsetting to group sizes of 1, 2 or 3
- All text lengths, or subsetting to only those texts with a length within: 1 standard deviation; 2 standard deviations; 3 standard deviations, of the mean text length
- Based on either CIS or MDP (treated separately)

These subsets were designed to maximise reliability, for example, on the assumptions that: more raters lead to more reliable outcomes; that raters with lower error rates in the diagnostic/training are likely to be better raters in the peer and self-assessment; that self-raters might have more bias in their ratings; that different group sizes might produce varying types of texts; and that inclusion of scores for very short or very long texts might reduce reliability.

---

[23] Note individuals and trios – not directly reported here – should have 1 and 3 self-assessments respectively. Note also that 4 peer assessments were intended for all outputs, due to software failure some outputs were produced individually, resulting in a larger n of outputs to distribute amongst peers and thus fewer peer-ratings per output.

[24] Some notes on developing this code are available http://sjgknight.com/finding-knowledge/2015/01/inter-rater-reliability-in-r/. Note Gwet's AC1/2 and Gwet's code for Krippendorff's alpha were also calculated (See, K. Gwet L., 2010; K. L. Gwet, 2011)

This results in approaching 3000 analysis iterations[25], which are not reported here. Outputs were visually inspected by the researcher, sorted by alpha; no consistent patterns in acceptably high alpha'swere observed, with all alpha's of over .6 observed in small subsets of data (returning fewer than 50 scored outputs). As a result, the peer and self-assessment rating data is not analysed further in this thesis. Analysis of the demographic data to confirm there was no selection bias in the raters was conducted as indicated in Table 5:3 and Table 5:4. This indicated that there was a small effect (r = .12) of GPA such that the cohort of students who did not complete the assessment had a slightly lower average GPA than those who did complete this component of the tasks. This may indicate that participants who did not complete this element encountered more difficulties in managing or/and completing the tasks, related to their general ability, than those who did complete these tasks. Analysis of gender distribution in the groups indicates a small effect (V = .11) such that a higher proportion of those who did not complete the assessment task were male.

Table 5:3 – T-test comparisons of those who did, and did not, complete the assessment task

	GPA						
	M	*SD*	*n*	*df*	T	p	*r*
Rated	6.48	1.31	692				
Missing	6.15	1.40	124	200	2.00	.01*	.12
	Age						
	M	*SD*	*n*	*df*	T	p	*r*
Rated	19.00	1.23	738				
Missing	19.10	1.39	138	200	-0.60	.50	.04
	Topic knowledge (1-10)*						
	M	*SD*	*n*	*df*	T	p	*r*
Rated	2.36	2.04	510				
Missing	2.61	2.52	92	100	-.90	.40	.05
	Search Experience (1-10)						
	M	*SD*	*n*	*df*	t	p	*r*
Rated	3.52	0.83	738				
Missing	3.64	0.85	138	200	-2.00	.10^	.07
	Collaborative satisfaction (1-10)						
	M	*SD*	*n*	*Df*	t	p	*r*
Rated	7.25	2.10	509				
Missing	7.10	2.64	62	70	0.40	.70	.03

---

[25] For example, looping through each combination of:2 topics, 4 datasets, 6 rubric-facets, 3 rater-quality evaluation methods, 4 group sizes, and 5 subsets of text-length data, gives 2880 possible combinations and data-iterations (although not all combinations are valid, e.g. due to missing data)

Task satisfaction (1-10)							
	M	*SD*	*n*	*df*	t	p	*r*
Rated	5.89	1.96	509				
Missing	5.92	2.03	62	80	-0.10	.90	.00

Intuitiveness of browser add-on (1-10)							
	M	*SD*	*n*	*df*	t	p	*r*
Rated	6.63	1.98	509				
Missing	6.29	2.24	62	70	1.00	.30	.08

Partner agreement (1-10)*							
	M	*SD*	*n*	*df*	t	p	*r*
Rated	7.99	1.79	510				
Missing	8.08	1.83	92	100	-0.40	.70	.02

Partner familiarity (1-10)							
	M	*SD*	*n*	*df*	t	p	*r*
Rated	3.56	2.87	510				
Missing	3.89	3.05	92	100	-1.00	.30	.06

*+Bartlett's test of indicated significant differences between the variance of the rater and missing group's: age, $K^2(1) = 4.00$, p = .05; collaborative satisfaction, $K^2(1) = 6.00$, p = .01; and topic knowledge, $K^2(1) = 8.00$, p = .005; therefore Welch's correction was applied to these t-tests.

Table 5:4 – Categorical comparison of those who did and did not complete the assessment task

	Gender									
	Female	*Male*				df	*n*	C2	*p*	*V*
Rated	314	424								
Missing	38	100				1	876	10	.001**	.11
	Average daily search frequency									
	1-3	*4-6*	*7-10*	*10+*	*Occasionally*	df	*n*	C2	*P*	*V*
Rated	99	249	162	201	27					
Missing	28	35	26	45	4	4	876	8.00	.08^	.10
	Most used search engine *(f)*									
	Bing	*Google*	*Yahoo*	*Other*		df	*n*		*p*	
Rated	1	732	2	3						
Missing	1	136	0	1		3	876		.40	
	Most used browser									
	Chrome	*Firefox*	*IE*	*Safari*	*Other*	df	*n*	C2	*p*	*V*
Rated	352	105	37	238	6					
Missing	65	21	10	39	3	4	876	4.00	.40	.07
	Most used operating system *(f)*									
	Linux	*Mac*	*Windows*	*Other*		df	*n*		*p*	
Rated	3	325	409	1						
Missing	0	63	74	1		3	876		.40	

*Where frequency counts in many cells are less than five, Fisher's Exact Test is used (for which a test statistic and effect size are typically not reported), denoted by an (f) in the row title.

5:1.1.3 Expert Assessment Inter-Rater Reliability

The 154 CIS outputs (n = 308) and 135 MDP outputs (n = 270) were rated by the primary

researcher (as described in section 3:4.4). The following stages of analysis were conducted:

1. Random subsets (30-60%) of the texts were read for distinguishing features, using the same kind of deductive approach described in the methods section
2. Random subsets (30-60%) of the texts were read for alignment with the original rubric
3. Informed by '1' and '2', the rubric was redesigned, such that content evaluation and source evaluation were scored on a single scale, and claim clarity was subsumed under topic coverage and synthesis; the outputs were then assessed using this rubric and a 1-3 scoring system[26]
4. A second independent researcher then graded a subset of 30 CIS outputs, using an iterative approach in which agreement was obtained on a subset of outputs (n = 5-10) before a further group were marked.
5. Inter-rater reliability was calculated on the raw values of the expert scores, using the IRR package (ibid) kappa2, with a (square) weighted disagreement, and Krippendorff's alpha (using the nominal function for those labelled T/F and 'Collapsed' and interval for those labelled 'scores')

The results of this analysis are given in Table 5:5 which indicates high reliability in identification of particular sub-topics (rows 1-3), and a marginally higher overall reliability for topic score. Reliability analysis also indicates that the synthesis score should be collapsed, to take a binary 'some synthesis present' versus 'no synthesis' indicator. In contrast, source quality is more reliably indicated on the full scale than using a binary. However, further analysis of the synthesis data, shown in Table 5:6, indicated that the distribution of scores is not even over the 3 values, and thus although the IRR for a collapsed score appears higher this is likely an artefact of the small n. As such, the 1-3 level scores are reported throughout this work.

Table 5:5 – Expert rater reliability indicators

	Krippendorff's alpha	Cohen's Kappa
Citrinin (T/F)	.773	.769 ($p < .0001$)
Concentration (T/F)	.800	.796 ($p < .0001$)
Statins (T/F)	.764	.762 ($p < .0001$)
Topic Score	.823	.820 ($p < .0001$)
Synthesis Score	.585	.583 ($p = .001$)
Source Diversity Score	.854	.852 ($p < .0001$)
Source Quality Score	.878	.876 ($p < .0001$)
Collapsed Synthesis (Some v none)	1	1 ($p < .0001$)
Collapsed Source Quality (some v none)	.614	.609 ($p = .0008$)
Collapsed Source Diversity (some v none)	.903	.902 ($p < .0001$)
Total Score	.867	.513 ($p = .0047$)

n raters = 2, n subjects = 30.

Table 5:6 – Expert rater analysis of synthesis scores

	Distributions	
	Primary researcher	Second assessor
Synth Score	2.53 (SD = 0.57)	2.40 (SD = .56)
Synth Score distribution	1 = 1; 12 = 2; 17 = 3	1 = 1; 16 = 2; 13 = 3
Collapsed Synthesis (Some v none)	False = 1; True = 29	False = 1; True = 29

[26] Note that 'topic' was scored based on the presence/absence of sub-topics, thus each subtopic receives a binary T/F and these are summed. This resulted in a small number of 0 Topic Scores where no sub-topics were adequately covered.

Subsequent analysis thus indicates means and standard deviations for this subset of double-marked outputs as indicated in Table 5:7. Indicating broad agreement across raters for each measure.

Table 5:7 – Descriptive statistics for expert rater scores

	Mean (standard deviation)	
	Primary researcher	Second assessor
Topic Score	2.07 (SD = 1.01)	2.13 (SD = 0.94)
Source Diversity Score	2.36 (SD = 0.81)	2.43 (SD = 0.86)
Source Quality Score	2.10 (SD = 0.57)	2.13 (SD – 0.63)
Synth Score	2.53 (SD = 0.57)	2.40 (SD = .56)
Total Score	9.07 (SD = 1.95)	9.10 (SD = 2.16)

5:1.1.4 Outcome descriptive statistics

Figure 5:2 indicates the distribution of scores for both tasks. Using Chi-square analysis it is perhaps unsurprising to see that topic and source diversity are significantly higher in the MDP task, given that the MDP groups were provided with a set of documents (perhaps implying they should use all of the materials) with a range of topics covered in them. However, given the presence of explicit conflicts in this document set, it might also be expected that the MDP groups would score higher on source quality – required to evaluate the 'best' claims from the competing documents – and synthesis – required to integrate the multiple claims from across documents. However this is not the case, with the distribution of scores in the CIS group tending to be higher on both synthesis and source quality; caution should be exercised in interpreting this difference given the possibility that the effect is an artefact of the task design (for example, the lack of synthesis might be due to the relatively reduced time to read and synthesise the documents), topic , or exposure to particular types of resources. As Table 5:8 indicates, the distribution of scores between CIS and MDP groups was significantly different, with CIS groups scoring significantly higher on source quality, and synthesis, and MDP groups scoring significantly higher on topic and source diversity. There was no overall difference in total scores between the two groups.

Table 5:8 – Comparison of rubric scores in the MDP and CIS tasks

	CIS (n = 154)		MDP (n = 135)		
	M	SD	M	SD	*Difference test (chi-squared & unpaired t-test)*
Topic Score	1.95	0.92	2.37	0.76	χ^2 (3, n = 289) = 16.74, *p* = .0008***, *V* = .24
Source Diversity Score	2.27	0.79	2.72	0.51	χ^2 (2, n = 289) = 35.19, *p* < .0001***, *V* = .35
Source Quality Score	2.04	0.79	1.74	0.78	χ^2 (2, n = 289) = 28.81, *p* < .0001***, *V* = .32
Synth Score	2.42	0.67	1.89	0.77	χ^2 (2, n = 289) = 10.41, *p* = .0055**, *V* = .19
Total Score	8.69	2.19	8.72	1.76	*t*(284.89)[+] = 0.130, *p* = .8968, *r* = .008

[+]Bartlett's test of indicated significant differences between the variance CIS and MDP total scores K^2(1) = 6.72, p = .010, therefore Welch's correction was applied to the t-test
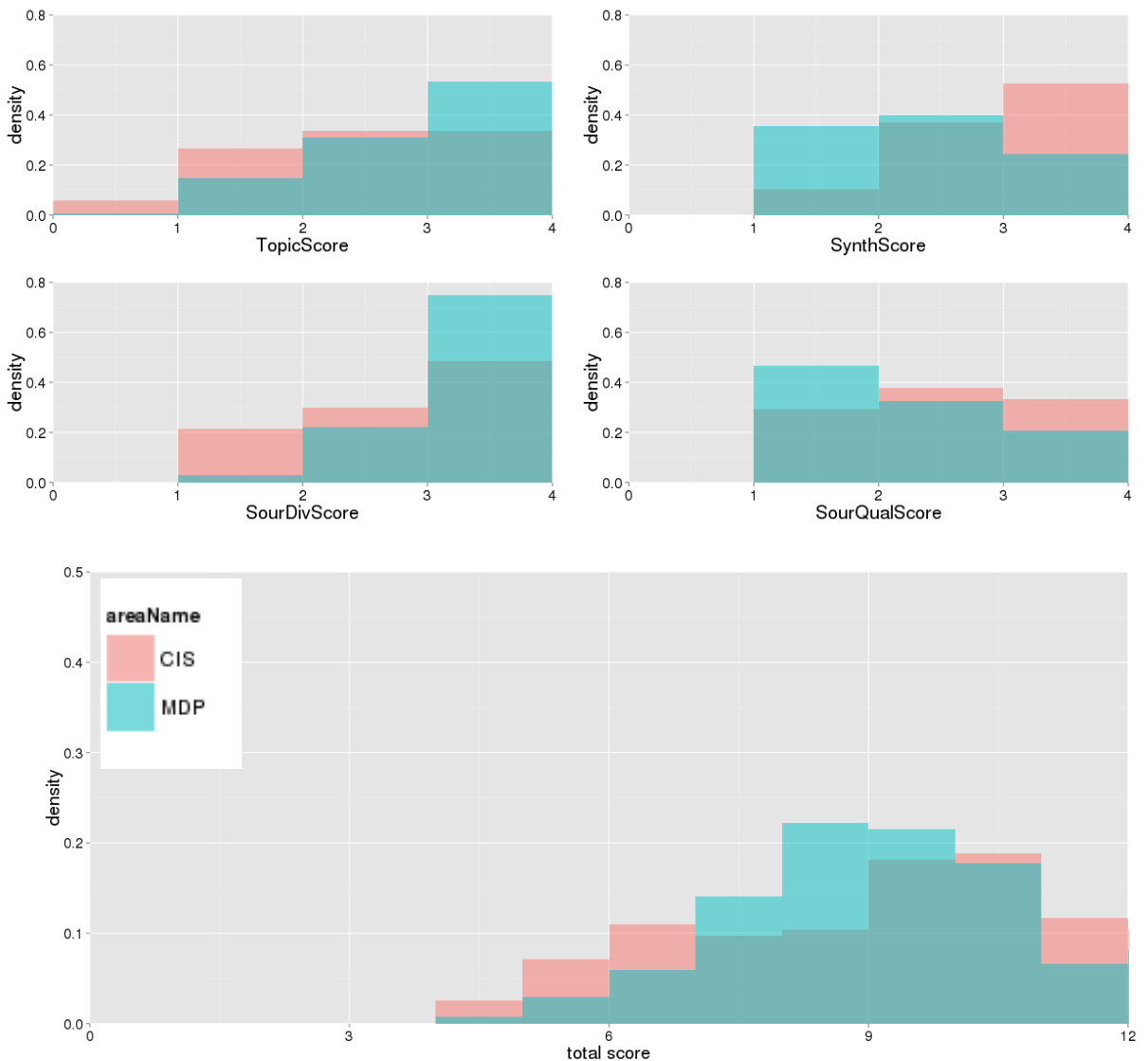
Figure 5:2 – Distribution of Scores in the CIS and MDP task

5:1.2   Research Question 1: Discussion

5:1.2.1 Findings and Implications

The first research question sought to establish the reliability with which epistemic facets of

written outputs could be identified. Results indicate that from an a priori model of six rubric

elements, a reduced model of four facets had reasonable inter-rater reliability for expert raters. However, results of inter-rater reliability assessment for peer and self-assessors indicated poor reliability. This was true of a variety of subsets of data, indicating that – in this study – participants could not readily identify the same writing features as expert raters. This suggests that the assessment task provides a level of challenge to many student participants, that more experienced raters (with experience of grading, and academic literacy skills) do not encounter, indicating the importance of work to develop effective peer and self assessment technologies in learning analytics approaches. While there may be an effect of lack of familiarity with the subject-material, or with the rubric design and assessment structure, or relate to a lack of motivation or time pressures, it is important to note that visual inspection indicates that many of the text-field responses were productively completed. As such, it is an important finding that many participants in this study find assessment of the outputs from their work challenging, similarly finding the training/diagnostic assessment difficult to accurately complete. Analysis of participants who did not complete the assessment task indicates a small effect for GPA, perhaps indicating that participants with lower GPA found the task more challenging to complete – an effect further research should probe in the group who did complete the assessment task.

In moving from the peer/self-assessment model, to expert ratings, two alterations were made to the rubric. First, based on a deductive reading of sample outputs two of the rubric facets were integrated into other elements; and second, the expert raters used a 1-3 scale, rather than the 1-9 scale given to participants. These changes represent an important component of such rubric use, ensuring the rubric best addresses the text features present in the outputs to which it is applied. The changes were intended to simplify the rubric, and make it more robust for further research, rather than to lose information. However, it should be noted that the alterations mean the participants and experts completed slightly different tasks, thus reducing comparability of results. It is also worth noting that while the rubric does not ostensibly assess the presentation of 'contrasting perspectives' or – of direct relevance to epistemic cognition – this facet might be considered related to the combination of source diversity, synthesis and evaluation. That is, that

inclusion, synthesis, and evaluation of a range of sources necessarily includes some presentation

and synthesis of contrasting ideas from a range of sources. Indeed, from the expert ratings,

participant total scores approached a normal distribution, with scores across the facets ranging in

their dispersions with CIS skewed to the higher scores in all but source quality/evaluation which

was relatively evenly distributed and MDP scores in both topic and source diversity scores but not

synthesis or source quality/evaluation which were relatively evenly distributed. This suggests that

the rubric elements had a good discriminatory power, that is, they distinguish between qualities

of the texts, and do not generally 'bundle' texts into a single score. As might be anticipated, the

MDP group scored higher on topic and source diversity with moderate effects of group on score

distribution (indicated in Table 5:8) – perhaps because this group perceived an expectation to

read each of the provided texts with a range of topics present in them. However, it is interesting

to note that the MDP group had lower synthesis and source quality scores, despite the presence

of conflicting claims (which might require resolution of inter-textual conflicts), and inter-textual

ties (which might require synthesis across texts). This result may be because participants

struggled to address the inter-textual ties across the 11 documents, while the CIS group could

make use of a smaller set of resources, including finding resources with pre-synthesised

information in them. The provision of documents containing inter-textual ties and conflicts alone

does not result in improved analysis of source qualities or synthesising behaviours. Across each

facet of the output-rubric, the participant's results were distributed across the range of scores,

and – with expert raters at least – reliably identifiable. This indicated that the selected facets were

a relatively content-neutral indicator of text qualities with discriminatory power, indicating

differences between text qualities. Indeed, following earlier multiple-document processing

research (S. R. Goldman, Lawless, et al., 2012) which found three clusters of students (satisficers,

who selected few sources; selectors who selected many sources but did not connect them; and

synthesisers who selected sources and integrated them), participants in this thesis research varied

across all dimensions, such that students who scored high or low on one facet did not necessarily

score commensurately on other facets. Further work is needed to explore patterns of scores across dimensions aligning with Goldman et al.'s clusters.

In addition, error rates in the diagnostic task do not indicate that any particular facet was more reliably assessed than another, as discussed further below. It is also interesting that in this case the diagnostic and peer assessments did not provide reliable tools for the assessment procedure. This is indicative of support for prior research (see sections 2:5.2 and 3:4.4) demonstrating the need for periods of clear training and feedback in peer and self-assessment tasks; although further research would be required to demonstrate that such support in fact increases reliability in this research design. Nonetheless, despite participants encountering some software-related technical issues in completion of the assessment stage (such that some participants could not complete this element), participants gave positive feedback about the process and the inclusion of peer-assessment in the exercise. This is particularly salient as this cohort of participants may not have encountered university-level peer and self-assessment yet (it being introduced in some second year modules). This indicates the potential of such a diagnostic and peer-self-assessment for learning analytics development to connect trace data to outcomes based on peer and self-assessment in a pedagogically grounded set of tasks. The evidence also supports the potential of such an approach for large scale deployment of peer and self-assessment; while technical issues are a limitation in this case, the potential of emerging pedagogic and computing tools for learning analytics is strong.

5:1.2.2 Limitations

Findings indicate reasonable reliability in identification of 4 key rubric-facets in the subset of written outputs dual-coded by expert raters. However, in the case of synthesis in particular, the observed levels of expert agreement were – while still acceptable – lower than desirable, perhaps indicating issues with the conceptualisation, or operationalization of this facet. That is, this result may point to difficulties with the conceptual definitions used in the rubric, or their application to the particular output texts; anecdotal feedback indicates that knowledge of the texts from which participants were drawing aided in assessment of synthesis for the researcher, while the

additional expert rater did not have this extra knowledge. There is potential here for further (perhaps computer-assisted) work to identify textual relationships between written outputs and the sources from which they draw (as in, for example, Hastings, Hughes, Magliano, Goldman, & Lawless, 2012).

In addition, these raters are experts relative to the participants being assessed (the thesis candidate and academic supervisor), but they are not subject experts in the science domains participants were asked to investigate. As such, the ratings of subject-experts may differ from those given. Indeed, it is also important to note that the diagnostic assessments were written and assigned reference values by a non-subject-expert. Although additional raters gave reference values, against which the researcher's marks were adjusted, these raters were also academic-experts rather than subject experts. As such, the samples may not have been representative of the written outputs produced by the participants (introducing a gap between training and assessment-task) and the reference values given may have differed from those a subject expert would have given (possibly introducing a bias in the selection of participant-raters). Finally, although as noted above much participant feedback on the assessment was positive, it is important to note the difficulties of running large scale assessment exercises on innovative technologies, in this case resulting in some data-loss and participant frustration that may have impacted on outcomes.

5:1.2.3 Future research
While the use of the diagnostic/training, followed by peer and self-assessment was not unproblematic in this research, the systems for deploying such an approach deserve further attention in the learning analytics literature. Expert validation of reference values on texts closely aligned with real participant outputs should be conducted. In addition, further work is needed to ensure participants understand the task requirements, and are capable of meeting them, and under what conditions (timing, number of rounds of training, feedback requirements, and so on).

Emerging language technologies raise further potential for work on the detection of features in output texts that are related to score-descriptors on the rubric facets. For example, analysis of the

written outputs for: rhetorical moves that are indicative of claims making, evaluation, and connecting (or synthesis) (see, for example, de Waard, Buitelaar, & Eigner, 2009; Groza, Handschuh, & Bordea, 2010; Simsek, Buckingham Shum, Sandor, De Liddo, & Ferguson, 2013); text cohesion (McNamara, Louwerse, McCarthy, & Graesser, 2010); and topic coverage (see, for example, Hastings et al., 2012).

For example, in the last of these studies (Hastings et al., 2012), students were asked to use three texts with relatively little semantic overlap to answer the inquiry question "In 1830 there were 100 people living in Chicago. By 1930, there were three million. Why did so many people move to Chicago?" They compared three methods to match source material to student writing outputs: Pattern matching approaches (i.e. looking for common text-strings); latent semantic analysis (LSA) to compare semantic-content at a sentence level across student outputs and assigned texts; and machine learning (using support vector machines) assigning student sentences to topic-classes assigned by human-raters. They found that LSA performed best in identifying explicit use of the assigned texts, while pattern-matching approaches were superior for detecting intra and inter-textual inferences (which could be characterised as synthesis). Such approaches are of interest to research such as that presented in this thesis given their potential to identify not only topics, but also the ways information from multiple sources are integrated (synthesis), and which sources information is drawn from (source diversity). The potential of language technologies, then, is to connect particular types or styles of language to epistemic characterisations described by the rubric; further work to connect computational outputs to human interpretable scores or feedback would then be required.

Chapter 6: RESULTS AND DISCUSSION OF RESEARCH QUESTION TWO

6:1 Research Question 2: Relationships Between Psychometric Profiles and Epistemic Outcomes

6:1.1 Research Question 2: Results

The second research question concerns epistemic cognition based on the ISEQ – a psychometric

instrument designed to measure internet specific epistemic cognition.

2. What relationships can be identified between facets of epistemic cognition on a psychometric instrument, and written outputs from a designed information seeking task (as in '1')

   - What relationships can be identified between scores on the ISEQ psychometric, and success on a written task (as in '1')

I first discuss the factor structure of the instrument, before giving descriptive statistics regarding

the target group's ISEQ scores. I then discuss the 'trustworthiness' assessments, which were

collected both as an explicit epistemic trace-marker to compare the ISEQ scores against (following

previous research), and as a possible additional indicator towards the outcome scores. Finally

regression analysis is conducted to assess the relationship between ISEQ scores and outcomes.

6:1.1.1 Internet Specific epistemological Questionnaire (ISEQ)

The 36 item ISEQ[27] was completed by 1003 students (all of whom had consented as participants in

the main study). The following procedure was taken for the analysis of the ISEQ results:

1. Reverse all items written to be scored negatively (as described in Bråten & Weinstein 2004)
2. Determine the number of factors to explore (following Courtney 2013) using:
   a. The R package:Psych's (Revelle, 2014) VSS (Very Simple Structure) function (Revelle, in draft; Revelle & Rocklin, 1979) which includes a variety of indicators of factor structure
   b. The R implementation of 'Exploratory factor analysis' (EFA) (Ruscio, 2012) described by Ruscio and Roche (2012)
   c. A scree plot (produced using the nScree function of nFactors, Raiche, 2010)
3. Explore a range of factor numbers (1-6 factors) for the items remaining, using the psych package's (Revelle, 2014) 'fa' (factor analysis) function, oblimin rotation, and manually analysing factor loadings for items strongly cross-loading (>.3) on more than one factor, or with only weak loads (<.3) on any factor.

I follow Bråten, Strømsø, and Samuelstuen (2005) ISEQ procedure, and Costello and Osborne's

(2005) general factor analysis guidance, in selecting oblimin rotation throughout (because we

---

[27] For copyright reasons only the ISEQ items within the discovered factor structure are reproduced here. My thanks to Ivar Bråten for kind permissions to reproduce the items listed here.

expect factors to be correlated), and a 'maximum likelihood' factor selection method. This process

was iterated four times until a satisfactory factor structure was discovered, with items removed

as follows:

1. First pass: exclude items 6,14,16,17,21,31,35.
2. Second pass: exclude items 6,14,16,17,21,31,35. And items 3,26,29
3. Third pass: exclude items 6,14,16,17,21,31,35. And items 3,26,29 and
   9,15,20,23,28,30,31,34,36

On no iteration was a third (or more) factor identified satisfactorily with such solutions resulting

in cross-loadings, low loadings, or single-item factors; a fourth iteration did not suggest an

improved factor structure. Nor was the factor structure improved by single factor models. Note

that the factor structure is almost identical to that of Bråten, Strømsø, and Samuelstuen (2005),

giving an identical 'Justification' factor, and 2 varying items: items 9 and 32, written to probe the

simplicity and source of knowledge respectively, were removed while items 13 and 7 – written to

probe the same respective constructs – were added.

The final two-factor solution gives 18 items with high loadings (>.4) and low overlap (<.3)[28]. The

two factors had eigenvalues of 6.12 and 2.65 respectively and explained 24.8% and 9.7 of the

sample variation respectively, or 34.4% in total. This compares to the 6.60 and 2.28 respective

eigenvalues and 47% sample variance explained in Bråten, Strømsø, and Samuelstuen (2005).

Item-to-factor loadings, eigenvalues and reliability estimates (Cronbach alphas) for both factors

are shown in Table 6:1; Cronbach's alpha may be compared favourably to Bråten, Strømsø, and

Samuelstuen's (2005) finding of .90 and .70 respectively. Note that for both factors, higher scores

should be interpreted as indicating more naïve perspectives.

---

[28] Note, Bråten, Strømsø, and Samuelstuen (2005) remove items with >.2 overlap.

Table 6:1 – ISEQ Factor Analysis

Variables/items	Factor loadings General Internet Epistemology α = .86	Justification for Knowing α = .72
1. The Internet contains accurate knowledge about the topics I study (certainty)	.74	
2. The Internet can provide me with most of the knowledge I need to succeed in my courses (source)	.67	
4. The truth about almost every issue raised in my classes is located on the Internet. (certainty)	.69	
5. On the Internet it is the richness of detail about what I am studying that is most prominent (simplicity)	.58	
7. For my studies, the Internet is a more important source of knowledge than my own thinking or reasoning (source)	.47	
8. On the Internet many different sources provide the correct answer to questions related to my course work (certainty)	.61	
11. The most important aspect of the Internet is that it contains so many specific facts about what I am studying in my classes (simplicity)	.54	
12. I am most confident that I have understood something for my classes when I have used the Internet as a source (source)	.52	
13. The Internet is above all an enormous reference work with specific facts about what I am studying (simplicity)	.51	
18. Most of what is true in my field of study is available on the Internet. (certainty)	.61	
19. The Internet is characterized by simple, concrete knowledge about issues related to my classes (simplicity)	.43	
24. The strength of the Internet is the vast amount of detailed information that is located there about what I am studying (simplicity)	.48	
25. The correct answer to questions related to my course work exists on the Internet (certainty)	.53	
32. Often, I doubt whether the Internet really is a good source for helping me understand the topics I am studying in my classes (source)*	.42	
Eigenvalue for General Internet Epistemology = 6.12		
10. To check whether the course-related knowledge I find on the Internet is reliable, I try to evaluate it in relation to other knowledge I have about the topic*		.59
22. I evaluate whether the course-related knowledge that I find on the Internet seems logical*		.59
27. I evaluate course-related knowledge claims that I encounter on the Internet by checking more knowledge sources about the same topic*		.60
33. To find out whether the course-related knowledge that I find on the Internet is trustworthy, I try to compare knowledge from multiple sources*		.70
Eigenvalue for Justification for Knowing = 2.65		

* item is reverse scored.

Note: The conceptual factor for which each item was written to probe is given in brackets in the 'general' factor; all justification items were written to probe that factor.

A confirmatory factor analysis was conducted (using the cfa function) on this model. Fit indices for this two-factor model indicate a reasonable fit between the model and the data, goodness of fit

index = .93, adjusted goodness of fit index = .90, comparative fit index = .89, root-mean-square error of approximation (RMSEA) = .063, with 90% confidence limits of .058 to .068. This compares to respective fit indexes of .86, .82, .95 and RMSEA of .067 (with confidence limits of 0.05 to 0.08) reported by Bråten, Strømsø, and Samuelstuen (2005).

6:1.1.1.1   ISEQ for the experimental groups

Descriptive values for the target group are indicated in Figure 6:1, which shows the scores on the general ISEQ factor for the CIS (far left) and MDP (one from left) groups; and for the justification ISEQ factor for the CIS (3rd) and MDP (far right) groups.). The boxplots used in this provide a visual depiction indicating the first and third quartile (forming the ends of the boxes), and the median (or second quartile, the horizontal line within the box). The mean is then additionally indicated by a diamond. The 'whiskers' extend to the furthest data point within 1.5 times of the inter-quartile range, with further data points (outliers) beyond that indicated by black circles.
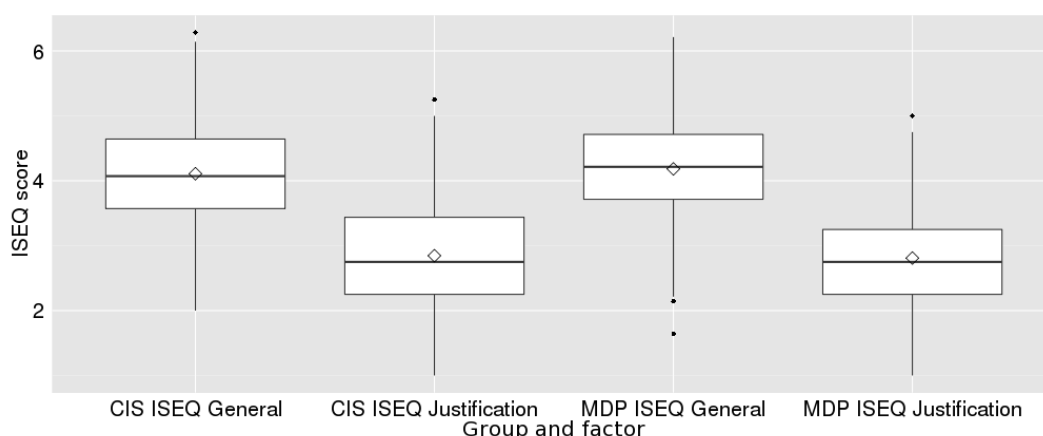
Figure 6:1 – Boxplot for CIS and MDP scores on ISEQ factors

An unpaired t-test indicates that there are no significant differences on the General factor between the CIS (M = 4.11, SD = 0.80) and MDP (M = 4.18, SD = 0.76) groups, $t(565) = 1.13$, $p = .26$, $r = .04$; nor did the CIS (M = 2.85, SD = 0.80) and MDP (M = 2.81, SD = 0.83) groups differ significantly on the Justification factor, $t(565) = 0.53$, $p = .60$, $r = .02$. Finally, comparison was made between the 567 results from the study group, and the remaining 436 (total n = 1003) results. Bartlett's test of homogeneity of variances indicated no significant differences between the General factor samples $K^2(1) = 2.86$, $p = .09$, or Justification factor samples $K^2(1) = 3.31$, $p = .07$. No significant differences were found on the General factor between the study-group (M =

4.12, SD = 0.75), and remaining sample (M=4.08, SD=0.72), $t(1001) = 1.21$, $p = .23$, $r = .03$; or between the study-group (M = 2.81, SD = 0.79), and remaining sample (M = 2.78, SD = 0.75) on the Justification factor, $t(1001) = 0.97$, $p = .33$, $r = .02$.

In addition, analysis of demographic variables was conducted to ensure there was no selection bias in the group of participants who completed the ISEQ (on which data the factor analysis was conducted). Given that most students on the module (n = 1003) did complete the ISEQ, those not completing the task may differ on some variables. As Table 5:3 indicates, t-test demonstrates a large effect for GPA ($r = .52$), such that the average GPA for the small group (n = 11) who did not complete the ISEQ was lower than those who did complete the ISEQ. This may suggest that some of the students with the lowest GPAs in the whole cohort are missing from the ISEQ and thus from its factor analysis. Table 6:3 shows the categorical demographic data, for which Fisher's Exact test is used as all frequencies are low; this analysis indicates that there were no significant differences in these technical areas between those who did and those who did not complete the ISEQ task.

Table 6:2 – T-test comparisons of those who did, and did not, complete the assessment task

	GPA						
	M	SD	n	df	t	p	r
ISEQ	6.46	1.31	805				
Missing	4.82	1.38	11	10	4.00	.003**	.52
	Age						
	M	SD	n	df	t	p	r
ISEQ	19.00	1.26	861				
Missing	18.60	1.18	15	10	1.00	.20	.16
	Topic knowledge (1-10)*						
	M	SD	n	df	t	p	r
ISEQ	2.38	2.11	590				
Missing	3.17	2.55	12	10	-1.00	.30	.17
	Search Experience (1-10)						
	M	SD	n	df	t	p	r
ISEQ	3.55	0.83	861				
Missing	3.20	0.94	15	10	1.00	.20	.19
	Collaborative satisfaction (1-10)						
	M	SD	n	df	t	p	r
ISEQ	7.24	2.16	560				
Missing	6.73	2.37	11	10	.70	.50	.05
	Task satisfaction (1-10)						
	M	SD	n	df	t	p	r
ISEQ	5.91	1.98	560				
Missing	5.09	1.30	11	10	2.00	.07^	.23

Intuitiveness of browser add-on (1-10)						
	M	SD	n	df	t	p	r
ISEQ	6.59	2.01	560				
Missing	6.82	1.94	11	10	-0.40	.70	.06

Partner agreement (1-10)*						
	M	SD	n	df	T	p	r
ISEQ	8.00	1.79	590				
Missing	8.17	1.85	12	10	-0.30	.80	.05

Partner familiarity (1-10)						
	M	SD	n	df	t	p	r
ISEQ	3.62	2.90	590				
Missing	3.25	2.56	12	10	0.50	.60	.07

Table 6:3 – Categorical comparison of those who did and did not complete the assessment task

Gender (f)						n	p
	Female	Male					
ISEQ	348	513					
Missing	4	11				876	.40

Average daily search frequency (f)							
	1-3	4-6	7-10	10+	Occasionally	n	p
ISEQ	125	279	185	241	31		
Missing	2	5	3	5	0	876	1.00

Most used search engine (f)						
	Bing	Google	Yahoo	Other		n	p
ISEQ	2	854	1	4			
Missing	0	14	1	0		876	.07^

Most used browser (f)						
	Chrome	Firefox	IE	Safari	Other	n	p
ISEQ	407	126	47	272	9		
Missing	10	0	0	5	0	876	.40

Most used operating system (f)						
	Linux	Mac	Windows	Other		n	p
ISEQ	3	382	474	2			
Missing	0	6	9	0		876	.80

Analysis of the ISEQ scores for the target group was conducted. Pearson's test of correlation on the paired ISEQ scores (i.e., taking the averages within each pair) indicates a small to moderate negative relationship between the two factors in the CIS group $r(150) = -.282$, $p = .000$; and the MDP group $r(124) = -.253$, $p = .004$. We also see a relationship between ISEQ Gen and partner familiarity $r(150) = -.181$, p = .026 in the CIS group, but not the MPD $r(122) = .085$, $p = .349$. Similarly, there is a relationship between the ISEQ general scores and search experience in the CIS group $r(150) = .197$, $p = .0145$, but not the MDP group $r(124) = .075$, $p = .407$; while a correlation between the justification factor and search experience holds for both CIS $r(150) = -.304$, p = .000,

and MDP $r(124) = -.230$, $p = .010$ indicating an association between more sophisticated views on the justification of knowledge (lower scores) and self-reported search experience.

A Pearson's test of correlation on the individual CIS data indicates a small positive correlation between the ISEQ general score and search experience $r(304) = .204$, $p = .000$; which had a small negative correlation with the ISEQ justification score $r(304) = -.182$, $p = .001$. Corresponding analysis of the MDP scores indicates a small negative correlation between the ISEQ justification score and search experience $r(259) = -.119$ $p = .056$ (significant at the .10 level), but not search experience and the ISEQ general score $r(259) = .087$ $p = .161$. This indicates a small relationship such that, in both tasks, as scores indicative of less sophisticated perspectives on the ISEQ justification factor increase (higher scores), search experience is decreased. The small positive relationship of ISEQ general scores to search experience is not aligned with expectation and may indicate that those with higher ISEQ general scores over-estimate their search experience. In addition there was a small correlation between the ISEQ justification factor and GPA $r(253) = -.159$ $p = .011$, which did not hold for the CIS data $r(304) = -.093$ $p = .104$.

Three approaches were taken to the ISEQ data. First, in section 7:1.1.2.3, relationships between individual ISEQ scores and trace data are reported. Then for paired analysis, two additional methods are used; first for each pair an average score is computed for both factors; giving the CIS task a mean of 4.11 (SD = 0.58) on the General factor and 2.84 (SD = 0.57) on the Justification factor, with the MDP task having a mean of 4.18 (SD = 0.54) on the General factor and 2.80 (SD = 0.57) on the Justification factor. These scores can be used in regression analysis. As a second approach to pair the data, the scores of each individual were categorised based on the distribution in the factor such that scores in the top quartile on either factor were labelled 'h', while those in the bottom quartile were labelled 'l' with the mid group labelled 'm' (particularly foregrounding the two extremes of variance). For each pair these labels were combined (unordered) such that for both factors 6 sub-groups can be identified: hh,hm,hl,mm,ml,ll. For information, the distribution is shown in Table 6:4 . Of key interest here is the potential to explore group-based ISEQ scores, and analyse meaningful sub-groups for significant differences in

behaviour and outcome, particularly for readily interpretable groupings (e.g. linear trends for hh, mm, ll). As Table 6:4 indicates, the spread is rather more complex, with few pairs falling into clear 'hh' 'mm' 'll' groups; given (as described below) the lack of linear trends in the averaged data, analysis using these categories is thus not reported in this thesis work.

Table 6:4 – Distribution of ISEQ group types in the CIS and MDP tasks

		hh	hm	hl	mm	ml	ll
CIS	General	8 (5.26)	30 (19.74)	21 (13.82)	37 (24.34)	41 (26.97)	15 (9.87)
	Justification	9 (5.92)	45 (29.61)	13 (8.55)	51 (33.55)	27 (17.76)	7 (4.61)
MDP	General	6 (4.76)	35 (27.78)	19 (15.08)	30 (23.81)	29 (23.02)	7 (5.56)
	Justification	6 (4.76)	33 (26.19)	16 (12.70)	35 (27.78)	27 (21.42)	9 (7.14)

*Note: Row percentages given in brackets

6:1.1.2 Trustworthiness Assessments

In both the CIS and MDP task trustworthiness assessments were collected for the documents encountered; either based on an average-highest-lowest 1-10 rating for resources visited (CIS), or an individual 1-10 ratings for each of the documents provided (MDP). Of the 308 CIS participants and 270 MDP participants 308 and 266 respectively completed the trustworthiness assessment survey (see section 3:4.1.2.1) – giving a 1-10 'trust' rating to each document (MDP) or to the most, least and average website trustworthiness observed (CIS).

Analysis of responses to the CIS trustworthiness survey indicated that some participants had very low standard deviations across their ratings, implying a universal or near universal rating strategy. Therefore any responses with a range of 0 (n = 10) or: giving a higher rating to the 'least trustworthy' or 'average' resource than 'most trustworthy' (n = 9, n = 11), or to the 'least trustworthy' than the 'average' (n = 20); was excluded (n = 29 cases, giving n = 39 excluded overall), leaving 269 valid cases remaining[29], summarised in Table 6:5.

Table 6:5 – Descriptive statistics for CIS trustworthiness ratings

	M	SD
Least trustworthy rating	4.67	1.75
Average trustworthiness rating	7.20	1.43
Most trustworthy rating	8.87	1.01

---

[29] Note, equal ratings for two out of three responses did not result in listwise exclusion.

A similar procedure was conducted for the MDP responses. This analysis identified 12 responses with a range of 0 (i.e. all documents were rated identically). Standard deviations were explored indicating a median standard deviation of 1.62 (M = 1.6), and a range from 0 to 3.91. All responses with a standard deviation of under 1 were eliminated (n = 57), in addition an (absolute) skew score was computed using the Psych package for all responses, with those with a skew over 1.5 (n = 9) removed; this removes responses in which single ratings artificially increase the range and standard deviation while the skew remains towards a rating that is otherwise universally applied (rating, for example: 10,10,10,10,2). This procedure left 206 valid responses from a total sample of 270, summarised in Table 6:6 .

Table 6:6 – Descriptive statistics for the MDP trustworthiness ratings

	M	SD
Document 1 rating	4.82	1.99
Document 2 rating	6.39	2.18
Document 3 rating	5.76	1.95
Document 4 rating	6.46	1.91
Document 5 rating	6.33	1.96
Document 6 rating	6.53	1.83
Document 7 rating	6.37	1.93
Document 8 rating	5.70	2.08
Document 9 rating	6.74	2.00
Document 10 rating	5.63	2.01
Document 11 rating	6.77	1.98

6:1.1.2.1   Relationships between ratings of trust

Analysis of the trust ratings was conducted to identify relationships between ratings on particular items.

A Spearman's rank correlation analysis was conducted between the three ratings provided by individuals in the CIS group. This indicated: a positive relationship between the average trustworthiness rating and the least trustworthy rating, $r(267) = .49$, $p < 0.0001$; a positive relationship between average trustworthiness rating and the most trustworthy rating, $r(267) = .57$, $p < 0.0001$; and a positive relationship between the least trustworthy rating and the most trustworthy rating, $r(267) = .22$, $p = .0003$. Correlations between trust ratings and scores on the General and Justification ISEQ factors were low, with some significant correlations as indicated in

Table 6:7. The optional question regarding the least or most trustworthy resource URL the participants found was completed by a large subset of the CIS participants (n = 215 for "most" and n = 235 for "least" respectively); a compiled list of URLs can be found in Appendix 17 with Appendix 18 indicating a frequency count for each domain name listed.

Table 6:7 – Correlation matrix for CIS trustworthiness ratings and ISEQ scores

	Spearman's Rho	
	ISEQ General	ISEQ Justification
Least trustworthy	.19 (p = .002**)	.11 (p = .083^)
Most trustworthy	-.14 (p = .026**)	-.08 (p = .216)
Average trustworthiness	.07 (p = .25)	-.08 (p = .186)

df = 267

Multiple regression analyses indicate small significant effects of ISEQ scores related to trustworthiness ratings. As indicated in the table, ISEQ scores have a small predictive effect for least trustworthiness scores such that higher scores (on both factors) are related to increases in trustworthiness assessments – that is, those with less sophisticated epistemic perspectives rate the least trustworthy pages they encounter more highly than those with more sophisticated (lower ISEQ score) epistemic perspectives. In contrast, in the model for ratings of the most trustworthy resource (which is marginal at p = .065), there is a small effect for the ISEQ general factor (significant at the p = .05 level) and a marginal small effect for the ISEQ justification factor (significant at the p = .10 level) such that higher scores on these factors are associated with lower ratings of the most trustworthy resource encountered. That is, participants with less sophisticated perspectives have lower ratings of the most trustworthy resources they encounter. The model for average trustworthiness is not significant at the .05 level. These findings indicate that scores on a self-report psychometric instrument (the ISEQ) can be predictive of the ways in which participants assess the trustworthiness of resources they encounter in a relatively uncontrolled, collaborative, naturalistic web-based task.

Table 6:8 - Multiple regression of Trustworthiness ratings by ISEQ scores

	t	p	B	β	F	df	p	adj. R^2
Least trustworthy								
Overall model					7.15	264	.001**	.044
ISEQ General	3.32	.001**	0.452	.204				
ISEQ Justification	2.48	.014*	0.334	.152				
Most trustworthy								
Overall model					2.76	264	.065^	.013
ISEQ General	-1.94	.053*	-0.150	-.121				
ISEQ Justification	-1.71	.089^	-0.133	-.107				
Average trustworthiness								
Overall model					1.30	264	.274	.002
ISEQ General	1.41	.16	0.160	.089				
ISEQ Justification	-0.46	.65	-0.051	-.029				

A Spearman's rank correlation matrix was also analysed for the MDP documents, with no

correlations over .3 identified, and most correlations below .2; no readily interpretable factor

structure could be identified across document ratings (for example, indicating covariance of

scores for particular types of documents or source). Correlations between trust ratings and scores

on the General and Justification ISEQ factors were low, with some significant correlations as

indicated in Table 6:9 . Given the lack of factor structure across document ratings, and that no

clear links between particular types of documents (for example, journal articles) and the observed

significant (weak) relationships, such correlations are not readily interpretable. As such, the

identification of patterns of 'trustworthiness ratings' is not possible, for example indicating that

individuals tended to rate particular sets of documents in similar ways, is not possible. As a result,

regression analyses on the relationship between document trust ratings and the ISEQ are not

reported here.

Table 6:9 – Correlations Between ISEQ Factors and MDP Document Trustworthiness Ratings

	Spearman's Rho	
	ISEQ General	ISEQ Justification
Document 1	.15 (p = .0289*)	.060 (p = .397)
Document 2	<.000 (p = .957)	-.032 (p = .648)
Document 3	-.05 (p = .471)	.031 (p = .659)
Document 4	-.028 (p = .700)	-.036 (p = .611)
Document 5	.064 (p = .366)	-.051 (p = .476)
Document 6	.158 (p = .025*)	-.145 (p = .040*)
Document 7	-.101 (p = .152)	.005 (p = .941)
Document 8	-.092 (p = .196)	-.121 (p = .087)
Document 9	-.141 (p = .046*)	.034 (p = .630)
Document 10	.066 (p = .350)	.140 (p = .048*)
Document 11	.119 (p = .092)	-.008 (p = .914)

df = 204

Given the small relationships between trustworthiness ratings and the ISEQ (as indicated in the

above correlation and regression analyses), their unclear structure (regarding the unclear factor

structure), and the uncertain nature of their external validity (particularly given the large number

of discarded data in the MDP case), this data is not reported on further in this work.

6:1.1.3 Relationship of ISEQ to Outcome

Finally, regression analyses were conducted to analyse relationships between the ISEQ and the

study cohort's outcomes as described in the discussion of RQ1. Analysis (reported below in Table

6:10 and Table 6:11), indicated that for each outcome variable (topic score, synthesis score,

source diversity score, source quality score, and total score respectively) the two ISEQ factors were not significant predictors (although for the MDP task, the model for source diversity approached significance, *p* = .068) indicating that ISEQ score are not predictive of outcome.

Table 6:10 – Multiple regression analysis for the relationship between outcomes and ISEQ factors in the CIS task

	t	*p*	B	β	*F*	*df*	*p*	adj. R^2
Topic Score								
Overall model					0.628	149	.535	-.005
ISEQ General	0.79	.429	0.107	.067				
ISEQ Justification	-0.54	.592	-0.073	-.046				
Synthesis Score								
Overall model					0.673	149	.512	-.004
ISEQ General	1.14	.257	0.112	.096				
ISEQ Justification	0.54	.588	0.054	.046				
Source Diversity Score								
Overall model					0.351	149	.705	-.009
ISEQ General	0.48	.632	0.056	.040				
ISEQ Justification	-0.52	.601	-0.622	-.044				
Source Quality Score								
Overall model					0.219	149	.804	-.011
ISEQ General	-0.28	.777	-0.033	-.024				
ISEQ Justification	-0.65	.514	-0.078	-.056				
Total Score								
Overall model					0.547	149	.580	-.006
ISEQ General	0.75	.450	0.242	.064				
ISEQ Justification	-0.49	.630	-0.159	-.042				

Table 6:11 - Multiple regression analysis for the relationship between outcomes and ISEQ factors in the MDP task

	t	*p*	B	β	*F*	*df*	*p*	adj. R^2
Topic Score								
Overall model					0.448	123	.640	-.009
ISEQ General	-0.11	.913	-0.014	-.010				
ISEQ Justification	-0.94	.350	-0.114	-.087				
Synthesis Score								
Overall model					0.852	123	.423	-.002
ISEQ General	-1.01	.317	-0.133	-.093				
ISEQ Justification	0.55	.583	0.069	.051				
Source Diversity Score								
Overall model					2.75	123	.068^	.027
ISEQ General	-0.90	.370	-0.078	-.082				
ISEQ Justification	-2.32	.022*	-0.191	-.212				
Source Quality Score								
Overall model					0.020	123	.98	-.016
ISEQ General	-0.14	.890	-0.019	-.013				
ISEQ Justification	0.11	.916	0.014	.010				
Total Score								
Overall model					.516	123	.598	-.008
ISEQ General	-0.82	.41	-0.244	-.076				
ISEQ Justification	-0.79	.43	-0.222	-.073				

As a result of analysis of the multiple regressions, stepwise regression analyses were conducted using the R stats package's (R Core Team, 2015) 'step' function making use of both forward and backward selection, using the ISEQ and other survey data (age, search experience, collaborative satisfaction, task satisfaction, addon intuitiveness, GPA, topic knowledge, partner agreement and partner familiarity) as predictors and the outcome scores as dependent variables. Using this method a set of models is constructed such that the function iteratively adds (forwards) and removes (backwards) independent variables at each 'step' based on a criterion indicating the variable's improvement of the model. In this thesis, the Akaike information Criterion (AIC) (Akaike, 1974) is used for variable and model selection. Thus, for each model an AIC is output in the stepwise processing, indicating the information lost under any model, selecting for better fitting models, but with a penalty for increasing numbers of variables in the model. Thus, variables included are ones which improve the model (although note, this does not equate to their having a significant independent effect within the model, thus variables may have p values over .05 yet still be included in the model).

Models with the smallest (signed) AIC are considered the 'best' model available (i.e., this should not be taken to mean they are a *good* model); the output models are presented in Appendix 15). In stepwise regression AIC is used to inform (but not decide) model selection; that is, selection of models should not just rely on statistical indicators (such as selecting the lowest AIC) but also on researcher judgement. As such, models (reported in Appendix 15, alongside multiple regression with all survey data included Appendix 14) were examined for a combination of a priori assumptions (regarding important features for any particular outcomes) and analysis of the variables included in each model, alongside the AIC for the model. Analysis of models with low AIC values indicated small differences in AIC across sets of models, and few models with variables of key interest – indeed, the 'best' models appear not well grounded theoretically (for example, in the MDP task indicating that the best model for source quality scores included only 'partner agreement as a predictor), or/and produce models which are not significant overall. Further research is required to investigate possible relationships here.

6:1.2 Research Question 2: Discussion

6:1.2.1 Findings and Implications

Factor analysis of the ISEQ indicated that the two factor structure previously reported – a justification factor, and a general factor – is supported by the data in this thesis research. Analysis indicated in both CIS and MDP groups that these factors were normally distributed, with lower scores in both groups for the justification factor than general, and no significant differences between the groups on either factor. This finding provides support from a large sample (n = 1003) for earlier research (Bråten et al., 2005) indicating that the conceptual model of factors regarding the source, justification, simplicity, and certainty of knowledge, is not supported by the empirical data. This replication with a larger sample and underlying two factor structure, indicates an empirical model which differs from the underlying conceptual model, suggesting need for new empirical and conceptual work in developing psychometrics for internet based epistemic cognition, including behavioural analyses such as that provided in this thesis.

Analysis of trustworthiness ratings indicated that most participants distinguished between the relative trustworthiness of the resources, with 206 of the 266 completers (total n = 270) in the MDP task kept following removal (as described above), and 269 of the 308 completers in the CIS participants kept. In the MDP data, the richness of the documents provided to the participants offers insight into the ways in which the participants deal with complex materials, but also introduces a complexity for analytic purposes such that analysis of covariance patterns in ratings of document trustworthiness is challenging. Similarly, the instruction to 'rate' the documents rather than 'rank' them – a decision made in order to reduce complexity for the participants and maintain 'interval' rather than 'ordinal' level data – also makes comparison of rankings more complex. That participants do not engage in implicit ranking behaviour, or/and rate documents such that similar documents are rated similarly (and thus covary), is an important finding for future epistemic cognition research making use of such rating or ranking exercises (see the general discussion section for some possible future research developments).In the CIS data the ISEQ general score had a small positive correlation to reports of 'least trustworthiness' and negative to reports of 'most trustworthiness' indicating that those with higher ISEQ scores

201

(indicating lower epistemic-sophistication) rated the least trustworthy resources higher, and the most trustworthy resources lower. This finding provides novel large-scale regression-based support for the previous research finding that those with less sophisticated views may not rate high quality resources as well as those with more sophisticated views, as Strømsø, Bråten, and Britt (Strømsø et al., 2011, p. 17) report "readers who believe that knowledge claims should be critically evaluated through logic and rules rated the science text as more trustworthy…. These effects hold true while controlling for readers' prior knowledge and text comprehensibility". This effect may be because students with less sophisticated evaluative capabilities cannot appropriately evaluate, and thus distrust inappropriately – an interpretation supported by Livingstone et al.'s (2005) survey research – as such we would expect to see that higher ISEQ scores (less sophistication) are associated with lower trust in the most trustworthy material encountered (as is found). While the observed relationship reported here is small, this finding in a large scale, collaborative, less-controlled environment supports the earlier self-report finding of Strømsø, Bråten, and Britt (2011) and Livingstone et al., (2005). The finding also indicates that students with higher ISEQ scores rated the poorest resource they encountered higher in trustworthiness, indicating a poorer ability to discriminate appropriately between the trustworthiness of high and poor quality resources. These relationships are identified on an individual level (i.e. using data from individual participants), although the tasks were collaborative in nature. That these relationships are sustained in individual contexts despite the activity and reading processes being collaborative is a novel finding, and one which future analysis should explore. This finding highlights an important strength of the ISEQ in providing predictive potential for the ways in which participants in fact engage in trustworthiness assessment of resources. This suggests that further work could explore the ways in which psychometric properties of beliefs in pairs might differ from those in individuals. Analysis of the relationship between the ISEQ data and outcomes on the written task indicated no clear relationships. In the MDP task regression analysis for source diversity scores, with the ISEQ factors as predictors, approaches significance at the .05 level (p = .07), with a significant effect of ISEQ justification scores such that lower ISEQ

scores are associated with higher source diversity scores – a theoretically grounded finding,

although one requiring further investigation given the low overall effect size (r^{2adj} = .03) and

significance level above the .05 level. Indeed, analysis of relationships between all survey data

(including ISEQ) and outcomes also indicated no clear relationship, with stepwise regression not

suggestive of further theoretically grounded models for exploration. To my knowledge this is the

first study to attempt to connect a psychometric instrument (specifically the ISEQ), to

theoretically related learning outcomes (in the form of the rubric scores on the written output).

This finding may indicate that the ISEQ does not exert direct effect on outcomes, although the

underlying structures may have indirect impact, perhaps via behaviour, which could be identified

in trace data. Previous research has connected psychometric instruments to other self-report

measures, with some limited research (as discussed in sections 2:4.5-2:5) analysing behavioural

indicators. The findings in analysis of this research question further highlight the need for analysis

of behavioural trace in understanding outcomes, and the need to research relationships between

psychometric properties, behaviour, and learning outcomes. Thus, subsequent sections

(discussing research question 3), will include the ISEQ as one predictor variable for further

analysis.

6:1.2.2 Limitations

This work provides a null-result for the association between self-report epistemic beliefs (in the

form of the ISEQ), and observed behaviours (in the form of scores on a written task). However,

some limitations should be noted in the observed data. First, as discussed above (section 5:1.2.2),

the findings here should be considered in the context of the limitations of the outcome data.

There may be a limitation insofar as associations are sought between the ISEQ and scores on a

defined (and thus, constrained) rubric; however the rubric underwent expert review and

refinement, and demonstrates reasonable reliability. Moreover, differences in scores can be seen

across participants indicating the rubric has discriminatory power. Despite this, the ISEQ displays

limited explanatory power for the outcome scores.

It is interesting that explicit solicitation of self-report data regarding trustworthiness assessments produced only small differences in such assessments in the CIS task. Results in the MDP task did not indicate patterns of trust across documents, indicating the need for research into the best ways to solicit such trustworthiness assessments from participants. Given a moderate number of assumed poor-quality results (indicated by low variance in trust ratings given) across both tasks, these results may indicate that contextual factors such as motivation and time pressures impacted on the participant results. In the MDP task in particular, many students did not read all of the documents prior to the trustworthiness survey, and thus may not have been able to respond appropriately to the survey. This suggests that future work could investigate both collaborative and individually based trustworthiness assessments; further extensions will be discussed below in section 8:1.1. The CIS data appears to be of better quality (with a lower exclusion rate). This might be expected given the task was relatively easier as it involved fewer questions which were based on individual participant's experience rather than the total set of documents. However, the positive results found in the CIS case should be treated with caution given the highlighted limitations, and their small effect sizes.

6:1.2.3 Future research

As noted above, there is a need for further research identifying the conceptual-theoretical and empirical implications of collaborative and social considerations in epistemic cognition research (for some preliminary work on this issue, see S. Knight & Littleton, in submissionb). Teasing apart individual and collaborative impacts on outcome (or trace indicators) is challenging in this research, for example, further work is needed to understand how pairs with similar and differing ISEQ profiles might interact. Moreover, understanding paired data, and the measurement or psychometric devices required to model paired, as opposed to or in addition to, individual cognitions is a challenge for measurement research. There is thus a need for continued measurement work on understanding how to assess epistemic cognition at the small group level. Furthermore, there is a need to develop conceptual work around understanding epistemic cognition as a social feature in information seeking contexts. Further research should thus apply

developing conceptual and measurement models to analyse behavioural implications of social epistemic cognition in information seeking tasks.

The null result presented here is important to epistemic cognition research, indicating that psychometric profiles captured through the ISEQ do not exert direct influence on outcomes in the specific large scale pedagogically designed collaborative information-based task described in this work. More controlled studies on both individual and collaborative tasks involving written outputs should aid in understanding the direct relationships between psychometric models of epistemic cognition and observed features in written outputs. Analysis of trace data in addition provides further data on the direct and indirect effects of epistemic cognition – at the behavioural, trace, level and in written outputs, as will be discussed in the next chapter.

Chapter 7:   RESULTS AND DISCUSSION OF RESEARCH QUESTION THREE

7:1 Research Question 3: Connecting Behavioural Trace, Outcomes, and Survey Data

7:1.1   Research Question 3: Results

The third research question concerns the development of an analytic approach grounded in the behavioural trace data, connecting the outcomes, background information regarding the participants, and their epistemic behaviour within the task:

3. What is the relationship between trace behaviours on an information seeking task, epistemic properties of written outputs (as in '1'), and epistemic cognition measured on a psychometric instrument (as in '2')?
    - What relationships can be identified between trace features (described above) success on a written task (as in '1') and ISEQ scores (as in '2').

I first provide an exemplification of the epistemic nature of the participant's interaction with the resources and each other, using an analysis of the terms in the typology of dialogue described in section 3:4.3.1.1.3, and a deeper analysis of extended excerpts from a subset of participant groups. This is followed by a discussion of the range of trace measures, and a discussion of the differences indicated therein using descriptive statistics. Section 7:1.1.3 then discusses the relationships (correlations) between trace indicators and survey data, indicating the ways in which particular behaviours, and individual differences (identified through survey data), are related. Section 7:1.1.4 builds on this analysis, to describe predictive relationships between the outcome data, and the trace and survey data, modelling the variance in outcome through exploration of the differences in behavioural indicators from the participants. The chapter discussion (section 7:1.2) then discusses these findings and their implications with regard to both the potential for analysis of behavioural data (both manually and computationally) and specifically with regard to computational approaches. The limitations of these analyses are discussed with implications motivating discussion of potential future research directions.

For both the CIS and MDP task a set of metrics was computed for the trace data from the main

task (i.e. preceding and subsequent data is excluded)[30], building off Table 3:3 in the methods, I

develop a set of trace indicators as in Table 7:1 below. In the following section I introduce these

trace indicators – which are used throughout this chapter – describing their incidence in both the

CIS and MDP task, using boxplots (as described p.191). In addition, where appropriate

exemplifications are given to illustrate the data.

Table 7:1 – Trace indicators

n	Variable name	Description	Data source	Grouped data	Individual data
1	Ug	number of unique pages visited	Pageviews	Y	Y
2	Urg	number of unique pages used, operationalised as those pages referred to in the chat, etherpad, or from which data was copied or snipped	Pageviews	Y	Y
3	f	An f measure, calculated as 2 * precision * recall /(precision + recall), where precision is the ratio of pages used to pages visited and recall is the ratio of the number of pages used by the individual group to the number of pages used by any group	Pageviews	Y	Y
4	-	Coverage – Ug/U (n of pages visited by group / n of total unique pages visited by any group); mirrors the distribution of '1', not separately reported here	Pageviews	Y	Y
5*	N query	number of queries	Pageviews	Y	Y
6*	N query terms	number of unique terms used in all queries	Pageviews	Y	Y
7*	Query vocabulary richness	6 / 5	Pageviews	Y	Y
8	Query depth	search engine results pages viewed past the 1st page	Pageviews	Y	Y
9[+]	ChatTaskTotal	number of chat messages sent	Chat	Y	Y
10[+]	Explor	number of chat messages containing an exploratory term	Chat	Y	Y
11[+]	Topic	number of chat messages containing a topic term	Chat	Y	Y
12[+]	SourQual	number of chat messages containing a sourcing term	Chat	Y	Y
16	Synthesis	number of synthesis messages, coordinating writing activity	Chat	Y	Y

---

[30] Note precision and recall, used in '3' were computed separately but are not used in this analysis. A likelihood of discovery metric was also computed (where pages visited by all groups are scored '1', and pages visited by few groups score approaching 0), but is not included in this analysis.

13	Nchar	length (in characters) of the text	Etherpad	Y	-
14	Contrib symmetry	contribution symmetry, a score indicating the percentage of the total text that can be attributed to a single author; closer to 0 indicating full symmetry (i.e. both authors contributed the same number of words), 1 indicating full asymmetry (i.e. one author contributed 0 words) [α]	Etherpad	-	Y
15	N touch points	number of touch points – an indicator of authorial transition in the etherpad, where 1 would indicate two chunks of authorial contributions, etc. (normalised and raw) [α]	Etherpad	-	Y

*Note: Not computed for the MDP task

[+] Note that chat data is considered in aggregate in project-based regressions only (i.e., no symmetry metric is computed). This is for two reasons: The chat data was relatively sparse, and as such frequencies are very low; chat messages are communicative in nature, aggregate frequencies account for both the sending and receiving of such communication. Individual counts are used in exploration of relationships between survey data and trace, for example individual ISEQ scores and individual chat frequencies.

[α] The symmetry score takes the difference in number of words contributed by each user individually (e.g. if 1 participant contributes 15 words, and the other 5 we calculate: 15-5 = 10), and computes a percentage difference over the total n of words (in that example, 66.66); due to incomplete data in etherpad's storage of author contributions, the total n of words contributed by all individual authors combined may be less than the total n of words present in the text. Code is available to extract the touchpoints and individual user word-level contributions from which these metrics are computed (See Mitsui, 2015 for some useable code).

7:1.1.1 Exemplification of the epistemic nature of the tasks

In order to confirm that the tasks were eliciting epistemic commitments of the kinds described in the literature review, the data was manually analysed, with a particular focus on exploration of the chat data for indications of the typology described in the methods chapter – of topic, source, exploratory, and (arising from that analysis) synthesis talk. The chat data is particularly foregrounded here as a core feature of the discursive environment in which the tasks were conducted, providing particular insight into the ways in which participants approached the problem presented, and made use of their mediating tools, including their shared dialogue.

7:1.1.1.1 Chat data analysis

Within the analysis sample, Table 7:2 indicates the (non-mutually-exclusive) number of chat messages sent with terms within each element of the typology (with a percentage indicator in brackets). This table shows that messages with one or more exploratory terms in them were most common in both tasks, with topic-terms and source-terms next most frequently occurring in CIS and MDP tasks respectively. Note that roughly two thirds of messages contained none of the terms targeted within the typology; some exemplifications of such messages are given below in

section 7:1.1.1.1.5. As will be indicated in that section, the typology was developed as an analytic device to capture salient distinctions between particular types of talk of interest to this research project namely: topic content; source qualities; exploratory dialogue; and, arising from the deductive process, synthesis chat pertaining to writing co-ordination.

Table 7:2 – Chat Typology Frequencies in CIS and MDP groups

	CIS (%)	MDP (%)
Explor	1276 (20.75)	1046 (18.63)
SourQual	320 (5.20)	528 (9.40)
Topic	620 (10.08)	243 (4.33)
Synthesis	528 (8.59)	392 (6.98)
No terms	4000 (65.06)	3867 (68.86)
Total messages	6148	5616

Messages including terms of each type were visually inspected, and exemplifications selected. These exemplifications are not intended to provide an exhaustive overview of the kinds of messages sent, nor to imply any kind of 'coding' or classification. Rather they are intended to give the reader insight into the kinds of epistemic chat engaged in, aligned with the typology in the methods. The messages are given without preceding or proceeding messages (one exception is noted), as discussed in the methods section (section 3:4.3.1.1.3). While this limits the salient context available for analysis, it provides insight into the level of data captured by key-term identification methods coding at the message level. Moreover, it facilitates the researcher to identify the ways in which terms are used across messages (although their wider structure is obscured) such that comparison between the use of terms in different individual messages may be made. These exemplifications thus illustrate the epistemic nature of the messages including the target-terms in the typology. A set of messages is selected to exemplify the kind of expressions the target-terms appear in; as such, the number of examples provided for each element of the typology, and task (MDP/CIS) varies.

7:1.1.1.1.1   *Exploratory chat*

Exploratory terms are intended to capture instances of partners explaining their ideas, engaging with the ideas of others, and attempting to build shared understanding. Instances of 'I think', 'because', 'so', 'maybe', are characteristic of the kinds of exploratory term of interest here. For

example, in the messages shown in Table 7:3 we see messages exhibiting coordinating (e.g. 1, 4, 9, 12, 18), explaining perspectives (2, 3, 5, 6, 15, 16) and explaining processes or procedures (e.g. 7, 11, 14). Note that only the second of these – explaining perspectives – might constitute the kind of co-construction of common knowledge implicated by exploratory dialogue. The other two types of example might be more accurately described as 'explanatory' in nature, and are perhaps more procedurally focussed. These exemplifications demonstrate the complexities of creating typologies of chat messages on a term or cue-phrase basis. However, they also provide a clear exemplification of the kinds of more exploratory, epistemically salient, dialogue anticipated in such tasks indicating the potential of such analysis in the context of CIS and MDP activities.

Table 7:3 – Samples of exploratory term messages

1	CIS	maybe one can do adventages and the other disadvanatges
2	CIS	i think the disadvantages are more important then the advantages
3	CIS	because i'm looking at an other website right now that talks about different sideeffects than i mentioned before
4	CIS	so how are we going to create a summary together
5	CIS	because of the active ingredient which they also use for drugs
6	CIS	I think Reuters is more reliable than that site
7	CIS	I think its better that we keep looking from different sources since we have more variety
8	CIS	yes i think we just just have to write a big conclusion on that rice
9	CIS	ok I will just make them in some sort of logical order
10	CIS	haha yeah we should but I think that those are the important main point about this red yeast rice
11	MDP	maybe we can try to find the best arguments in each article article by article so maybe we should start with reading government urged to act and on
12	MDP	if you read the first 5 links and I read the second 5
13	MDP	i'm finished and i think the best argument is
14	MDP	There are both claims that its harmful and that it is not so we should include both right
15	MDP	but I agree it's better than the 2nd to last article that I read
16	MDP	So I would say we leave out the first two because they support a different claim than the majority
17	MDP	k so what do you think our conclusion should be
18	MDP	I just put in everything which I think is important and rewrite it at the end ok

7:1.1.1.1.2   Source quality chat

Analysis of source-terms in chat messages (as in Table 7:4) indicated that sources were often referred to in reference to the specific claims within the source, but often using a generic term such as "a site" or "an article" rather than more specific metadata (e.g. 1, 18, 19), or referencing sources in passing (e.g. 9). Also evidenced are references to source and authorial qualities, for example particular pulication venues (e.g. 4,5,14,20,27), or authorial characteristics e.g. 'scientists', 24,25,26,28) or both (e.g. 11, and the combination of 21-22, which were sequential by

one pair). Finally, generic comments on source quality (e.g. 3,8,12) suggesting looking for 'good

sources' or citing a range of sources, are seen. Note that lines 21 and 22 were sequential, and by

different partners in a pair, indicating a small exemplar of source-based discussion. Again here we

see, along with some less-relevant messages, some very clear exemplifications of the kinds of chat

messages around source qualities and source meta-data of interest to this research. Throughout

these messages coordinated activity around sources and sourcing can be seen, with specific and

explicit evaluation of source qualities (generally or specifically) given in various places.

Table 7:4 – Samples of source quality term messages

1	CIS	this is what a site said
2	CIS	just have to copy everything from this site
3	CIS	remember to reference to reliable sources
4	CIS	is wikipedia enough reliable
5	CIS	I think Reuters is more reliable than that site
6	CIS	does the source where u got the animal study from say how high the doses were
7	CIS	all the sites that i found references the same site
8	CIS	I think its better that we keep looking from different sources since we have more variety
9	CIS	so u keep looknig reuters
10	CIS	I also find an article that is against the use of rice should i include it too or only articles that support the rice safety
11	CIS	lets back it up with scientific research rather than sum reuters news
12	CIS	lets collect the sources at the bottom and then were done i guess
13	CIS	its just a drugs company but they say they screen their yeast for impurities
14	CIS	kinda commercial site though it seems
15	CIS	but I'm missin something more scientific
16	MDP	shall we spilt the articles
17	MDP	after reading each article should we make a summary of it and then combine everything together
18	MDP	oh damn the next article is telling us that the first one isnt true hahaha this sucks
19	MDP	all 4 articles that i have read are pro claim
20	MDP	Yeah there's one of a scientific magazine which is not biased
21	MDP	I'm not sure whether the critisism from the science media center is scientific enough the way it is phrased just sound really colloquial
22	MDP	Yeah but it are all scientists
23	MDP	in the 3rd article they went to say that the average was 44 but the Uk scientists denounced the study as unrealiable Us biotech giant monsanto says that 40 years of independent assessments says it chemical doesn't pose a risk to human health
24	MDP	oh and some of the sourceses are crap one is written by a blogger we are only suppose to take those who seem accurate you know what i mean
25	MDP	but like all professors are claiming that the sample is not proper due to the fact that the real data was not published
26	MDP	what are Friends of the Earth and GM Freeze NGO s
27	MDP	I skipped one because it was a blogg
28	MDP	Article 10 is basically about a discover blogger named Keith Kloor who says that the research of one Carey Gillam is bullshit because she has done the reseach in a field which is not her expert field because she doesn't have any diplomas on that field or whatsoever

Analysis of messages including a topic-term indicated that topic-terms appeared in a range of

contexts indicated in Table 7:5. Some topic-terms were used in a rather general context (for

example, 4-7, 10, 13, 20), discussing the broad problem and theme of the task. At other times,

specific points are being made, drawing on general information rather than specific sources or

claims (for example, 1-3, 9, 11), while in other cases specific claims are drawn out and noted

or/and discussed (for example, 8, 12, 14-19, 21). These messages exemplifiy a more topic-

focussed content, with a more or less general target in each message. It is interesting that,

perhaps due to the length of the messages, these topic-focussed terms tend not to co-occur with

sourcing, exploratory, or synthesis terms. That is, for example, within individual messages there is

little evidence of explanation (exploratory terms) or weighing up specific or contested claims in

light of their role in the written output (synthesis terms) or their provinence (source quality

terms). As indicated below (section 7:1.1.1.2) this absence is perhaps due to the focus on

individual messages rather than longer extracts, however as the following section will illustrate,

the exemplifications provided here do in fact give a reasonable representation of the kinds of chat

that occurred.

Table 7:5 – Samples of topic term messages

1	CIS	it is dangerous as works simmilar to drugs affecting liver
2	CIS	so it may cause harm if somebody has liver problems etc
3	CIS	we can add sometrhing about contamination
4	CIS	do we like the red yeast rice
5	CIS	it is not the worst but definitely exposes one to health risks
6	CIS	did you ever hear of that rice D
7	CIS	i guess we can do something like first looking at the potential health risks
8	CIS	Animal studies have been conducted in China using high doses of red yeast rice products No damage to the kidneys liver or other organs were demonstrated in these studies
9	CIS	What do you already have I have that it can be dangerous if it contains monacolin K and the side effects of it
10	CIS	should we talk about the confusion drug or supplement
11	CIS	Yep it's some kind of toxic stuff that fucks up your kidney
12	MDP	wait this one That's not long o Glyphosates Suppression of Cytochrome P450 Enzymes and Amino Acid Biosynthesis by the Gut Microbiome Pathways to Modern Diseases
13	MDP	So do your articles state that the herbicide is damaging or harmless
14	MDP	The first two articles imply almost everyone in Europe has traces of glyphosate
15	MDP	well i kinda got it from disease severity but that might be wrong since its more about plant
16	MDP	The fifth article implies farmers are more susceptible
17	MDP	In the environmental benefits part it says 'environmentally benign''

18	MDP	it seems pretty unlikely it leads to ALS Alzheimer's MS cancer and autism
19	MDP	Glyphosate is less acutely toxic than common chemicals such as sodium chloride or aspirin with an LD50 for rats greater than 5 g kg1
20	MDP	it is used as weed killer
21	MDP	theres one document about the health risks to humans

7:1.1.1.1.4   Synthesis

Analysis of messages including synthesis-terms or text-coordinating language (shown in Table 7:6)

indicated references to the report structure – often noting introductions, or conclusions

specifically – (for example, 1,3,12,18,19), with attempts to coordinate writing (9,10,11) the taking

and sharing of notes (2,4,5,13,14,16,17,) or/and use of paraphrasing, copy-pasting, and citation

management (6,7,8,15). Messages within this facet of the typology exemplify the coordination of

activity, attempts to understand the task-context and requirements, and to address how well

those requirements were being met. As such, these messages are epistemically salient,

considering the kinds of information needed to present, how this information should be

presented and contextualised towards the task goals, and how to co-ordinate knowledge sharing

within the pairs – these considerations thus broadly relate to the simplicity/complexity of

knowledge, and the role of the self/other in sourcing knowledge.

Table 7:6 – Samples of synthesis term messages

1	CIS	how should we structure the report
2	CIS	I'm just going to write some random notes down then we can structure it once we have enough info
3	CIS	so we have the basic intro
4	CIS	we need to rewrite everything
5	CIS	lets not look for more information but rather rewrite vererything and the come to a conslusion
6	CIS	so we can copy paste the links to say we got the info from there
7	CIS	Do you think we need to 'summarise' it into our own words or what is the meaning of all this that we just know how to look things up
8	CIS	we are going to prison for plagarism
9	CIS	okay ill write it down
10	CIS	how much do we have to write
11	CIS	yeah do you want me to write something about it being illegal in USA
12	CIS	first introduction what it is then prositiv and negativ arguments and afterwards we judge it
13	MDP	after reading each article should we make a summary of it and then combine everything together
14	MDP	yeah its way more reliable because its a real experiment Dont read what they did try to find conclusions or at least thats what im doing
15	MDP	yeah but i dont think we should add any of our own words just phrase what is said by other
16	MDP	so should we start with just writing in the task pad what each article is about so we both know
17	MDP	do you want to copy some stuff into the taskapad
18	MDP	we could just make a short introduction stating what it is
19	MDP	What should the conclusion be

7:1.1.1.1.5   Messages without target-terms

In order to investigate the nature of the messages without any target keyterms in them, a sample of 10% of such messages (CIS n = 400, MDP n = 386) was visually inspected. Messages in this subset often had a number of (non-mutually-exclusive) features including that they:

1. Were very short (e.g. "yes", "good", "okay"), or omitted elements which would have been identified in the typology (e.g. "but this is short term")
2. Were not in English (frequently German or Dutch)[31]
3. Were off-topic (e.g. "im the coolest one int he [sic] room")
4. Included terms with a high incidence of false-positives (e.g. "okay you're right then just jump right into the arguments" is excluded because task instructions were frequently copied into the text-chat in their entirety, and included the term)
5. Included typographical errors (e.g. "i thinjk ")
6. Included alternative terms to those included in the typology (e.g. instead of toxic, "there is this poisonous thing in there sometimes")
7. Involved some task discussion (e.g. "this project is boring", "i have no idea what the hell we are doing here", "press submit")

Given the deductive process of term selection inevitably some less frequently used terms of interest have not been identified, or their inclusion would increase the numbers of false-positives. In some cases, this means that epistemically interesting messages are not identified (e.g. "let us collect a lot of facts and then in the end look for the strongest ones"). See the discussion chapter for further reflection on this aspect of the data.

7:1.1.1.1.6   *Section Summary*

As highlighted in Table 7:2, across the full set of message a substantial number (roughly 2/3) contained none of the terms contained within the typology types. The discussion above indicates some of the kinds of messages sent within this subset of messages, demonstrating frequent messages with less salience for the analytic focus in this work, alongside some less clear cases. Across the exemplifications shown in sections 7:1.1.1.1.1-7:1.1.1.1.4 there are clear indicators of epistemic dialogue, aligned with the typology described in the methods section (3:4.3.1.1.3), grounded in the theoretical discussion in section 2:4.4. These cases demonstrate the potential

---

[31] Initial analysis of chat data indicated very poor performance for automated language detection methods using the textcat package (Hornik et al., 2013) both through analysis of language frequencies (for example, 'Scots' had a rather high incidence, despite to my knowledge there being no Scots speaking students in the classes), and manual inspection of the classified languages. This is perhaps due to the short and informal nature of many messages. A superior analysis was subsequently conducted at a later stage (following other analysis reported here). This later analysis used aggregated chat such that each group's chat was collapsed into a single text string, and yielded better performance: 45 chat-sets were identified as non-English, with 32 manually confirmed (i.e. 71% correctly identified). However, no check is made on false negatives (i.e. non-English in the remaining text), and in fact many of the identified chats include whole messages in English, in addition to English terms, particularly target topic and source terms. Note that due to the performance of the classification, and late application of a superior method *all* chat data from the target CIS/MDP groups is included in this reporting, including non-English chat. See the discussion for relevant evaluation in this regard.

value of such term-based approaches in identifying the types of ways in which participants engage in dialogue around information seeking and processing tasks. However, across these cases we also see a depth of nuance that is not captured through simple summary counts of typology-aligned-term instances. Nonetheless such counts provide additional insight, allowing us to engage in both an analysis of raw chat messages, and quantitative methods (particularly counts) which may be taken to inform our understanding of the qualitative analysis (see, for example, Mercer, 2004). In addition, this close analysis of the messages indicates a number of what one might consider 'false positives' (in the exemplifications), and 'false negatives' (described briefly in section 7:1.1.1.1.5). As described in the methods, the intent of the typology development is not to capture all possible types of dialogue, or the breadth of nuance, but rather to develop an analytic device to distinguish between particular terms in the use of dialogue. Thus, additional types may emerge from the data (for example, as indicated above, off task talk) with less salience to the analytic approach. Finally, while the analysis provided here indicates clear epistemic dialogue, with differences between the ways in which terms are used apparent in the text, they offer coarser insight into the interactional nature of the dialogue – the ways in which the participants shared and built their knowledge together. To address this, the following section will engage in a closer analysis of a subset of groups, discussing their interaction with the tools (indicated through analysis of trace data), and each other (indicated through the chat data).

7:1.1.1.2   Manual analysis of trace data

In addition to concordance style analysis of chat data, a closer analysis was conducted on a small subset of projects' trace data both as a component of the deductive process, and, aligned with sociocultural approaches to discourse, as a means to draw on both fine and coarser grain analysis of the dialogue data in understanding the interactional processes. This analysis was conducted in order to examine, on a less fine grain level, the epistemic features of the task behaviour as foregrounded through the chat data. The analysed projects were selected as exemplifications of the ISEQ-outcome combinations possible, thus groups were selected to represent combinations of: High/low outcome, and high/low ISEQ (combining factors), for both CIS and MDP tasks, giving

a total of 8 possible combinations. The groups with the highest/lowest scores as appropriate were selected in each case. These groups are displayed in Table 7:7, which gives basic demographic information regarding the groups, and Table 7:8 which gives core trace data regarding the groups, intended to provide broader contextualisation to the nature of the research and these particular groups.

Table 7:7 – ISEQ and Outcome Scores for Chat-Analysis Target Groups

Task	ID	ISEQ group	Score group	Total score	Topic Score	Synth Score	Source Diversity Score	Source Quality Score	ISEQ Gen UID1	ISEQ Just UID1	ISEQ Gen UID2	ISEQ Just UID2
MDP	518	LL	L	6	1	2	2	1	3.29	4.75	3.43	1.75
MDP	414	HH	H	12	3	3	3	3	6.21	4.00	5.79	1.50
MDP	369	HH	L	4	1	1	1	1	5.64	2.50	5.50	2.50
MDP	325	LL	H	11	2	3	3	3	3.64	2.00	2.71	3.25
CIS	633	HH	L	4	1	1	1	1	4.21	3.75	4.50	4.00
CIS	586	LL	L	4	1	1	1	1	3.57	2.25	3.36	2.50
CIS	401	HH	H	11	2	3	3	3	5.57	2.50	4.93	3.25
CIS	245	LL	H	11	3	2	3	3	3.50	2.00	3.57	1.25

Table 7:8 – Trace Data Summaries for Chat-Analysis Target Groups

Task	ID	Ug	Urg	total chat	Exploratory chat	source chat	topic chat	synthesis chat	Ug diff	Urg diff	nchar	contrib symmetry	n touch points	n queries	N query diff~	Query term count~	Query Vocab richness~	Query depth
MDP	518	11	7	47	3	1	0	0	5	1	4415	0.01	3	------	------	------	------	0
MDP	414	10	6	59	14	10	12	3	2	3	2049	0.00	33	1	------	------	------	0
MDP	369	10	1	65	23	11	2	8	3	0	1223	0.32	3	2	------	------	------	0
MDP	325	11	8	73	4	3	1	6	5	2	6872	0.85	2	------	------	------	------	0
CIS	633	9	2	53	7	2	3	5	1	1	5538	0.27	6	27	23	31	1.15	0
CIS	586	17	5	41	6	0	5	2	12	1	5039	NA#	NA#	30	18	14	0.47	0
CIS	401	14	7	56	6	5	2	5	6	0	8676	0.14	9	29	19	12	0.41	0
CIS	245	18	9	12	2	4	4	1	4	1	4985	0.38	3	30	12	102	3.40	0

~ Queries are marked for both CIS and MDP tasks, other query metrics are not given for the MDP task here.

The NA here indicates authorial information for one participant was not available; this is due to data loss in the etherpad storage.

Clear differences between the groups are observed in analysis of their trace data, as indicated in Table 7:7 and Table 7:8, which illustrate the range of score/ISEQ combinations, and the commensurate range of trace indicators. The chat data, in particular, highlights marked differences between the groups. In the following sections this chat data is discussed in more detail, with each group described and exemplifications of the kind of chat they engaged in given. Readers should particularly note that the – as the tables illustrate – there are some broad similarities (e.g. overall levels of chat), and differences (e.g. levels of topic chat) across the groups. These differences and similarities in the trace are, as was the case in the preceding analysis, identifiable in the closer analysis of the data, which also reveals more nuanced differences in the ways in which the participants interact with the tools, each other, and the task. These examples highlight the epistemic nature of the tasks, and how this is seen in the chat data; they also foreground the complexity of analysis of such data as a form of behavioural indicator associated with learning outcomes.

7:1.1.1.2.1   18.1.2.1 MDP groups

Group: 325

In the MDP task, group 325 was a high scoring group (total score = 11, missing 1 point on 'topic'), with a low overall ISEQ score (indicating more sophisticated epistemic cognition). The group wrote a relatively long text (number of characters = 6872) and appear to have been highly asymmetrical in their writing (symmetry score = .85, indicating 85% of the words attributable to either author, were written by one of the participants, perhaps due to one participant pasting larger chunks of text into the pad). The group have only two touch points, indicating that the author transitions only twice (suggesting each author focussed on their own "block" of text). This group made use of 8 articles, referring to them in the written output, chat, or copying directly from the pages, with one partner reading 5 more articles than the other.

The group began the task segment expressing uncertainty regarding the task:

userID	projectID	localTime	Content
610	325	17:39:50	what to we actually have to do
611	325	17:40:04	i have no clue haha

610	325	17:40:13	thats unfortunate
610	325	17:40:15	Meneithe
610	325	17:40:17	R
611	325	17:40:20	read the articles

After about 8 minutes exploring the task pad, and each looking at documents and the
instructions, one suggests:

userID	projectID	localTime	Content
611	325	17:47:20	do you want to do the first 5 and i do the last 5

They then engage in a stretch of reading time. Towards the end of the task the pair then discuss
how best to represent and summarise the information found, including how to organise 'pro' and
'con arguments and whether to synthesise information or primarily use quotation:

userID	projectID	localTime	Content
610	325	18:01:06	do you want to sort yout stuff to my list of pros and cons
611	325	18:02:42	did you summarize them
611	325	18:02:49	or copy passages
610	325	18:03:00	copy passages
610	325	18:03:13	no way gonna summarize that takes way to long
610	325	18:03:19	or are we supposed to do that
611	325	18:04:16	i ahve no clue

Then they discuss how to weigh up the articles, and collate them into groups to give their opinion
on. In this end stage they also note an inter-textual tie between two articles, saying: "well one of
the pros references to one of the con articles". We thus see some attempts to synthesise
information (identifying ties between articles, and grouping claims), with a concern to evaluate or
give opinions on the claims given.

userID	projectID	localTime	content
610	325	18:05:33	we have to write why we think there the best too
610	325	18:05:37	damn
611	325	18:05:53	the arguments or what
610	325	18:05:58	yeah i think so
610	325	18:06:08	but we cant do that for all of them
611	325	18:06:17	maybe as a conclusion
610	325	18:07:02	check the task pad
610	325	18:07:07	i put the description
610	325	18:07:28	we have to decide which ones are the best arguments
611	325	18:07:36	yeah but still for every single argument
611	325	18:07:38	thats too much work
611	325	18:07:42	oh alright
611	325	18:08:33	well one of the pros references to one of the con articles
610	325	18:08:50	we have to grouop them

| 611 | 325 | 18:10:01 | ehm for which aspects |
| 610 | 325 | 18:11:34 | lets stat putting the arguments into groups and wirte our opinions |

Group: 369

Also in the MDP task, group 369 are on the opposing ends of the ISEQ and total scores to group

325 scoring 4 overall, with a high overall ISEQ score (indicating relatively less sophisticated

epistemic cognition). The group wrote roughly 1/6 the number of characters that group 325 did,

with a smaller asymmetry (.32) score, and again a small number of touch points (n = 3) suggesting

little authorial transitioning in the text. This group made use of only 1 article, referring to it in the

written output, chat, or copying directly from the pages.

The pair start off reading the articles, both reading the first two articles independently. They then

discuss dividing the articles between them, briefly noting that the second article (the scientific

report for Friends of the Earth) is better than the first (the Friends of the Earth press release; see

Table 3:2 for the list of documents):

userID	projectID	localTime	Content
692	369	10:11:31	what do you think of these articles so far
693	369	10:12:08	do we need to read all of them
692	369	10:12:54	I think so but we can divide the articles between each other
693	369	10:13:15	Nice
692	369	10:13:41	which articles have you read so far I've the first 2
692	369	10:13:51	Read
693	369	10:13:54	mee too
693	369	10:14:32	so we can divide other articles
692	369	10:14:34	second one is better than the first one would you agree
692	369	10:14:56	Yh
693	369	10:14:59	Yes
693	369	10:15:21	argree1
693	369	10:15:31	Agree
692	369	10:16:19	ok so how shall we divide it I'll read the bottom 4 and you read the rest
693	369	10:17:28	Okbut i don't know what will we do after finish the reading
693	369	10:17:40	give a summary
692	369	10:17:49	of each one
693	369	10:18:04	maybe the best one
692	369	10:18:40	I think because we only have until 1045

Each then flags a 'best' article from the articles they have read:

userID	projectID	localTime	content
692	369	10:25:17	ok I've read the bottom 4 and the best one out of them all is the 2nd to last article
693	369	10:25:39	I read the fifith article it looks nice

They then discuss how to deal with the length given the short amount of time, deciding to:

userID	projectID	localTime	content
692	369	10:38:35	I think we should go for article 9 and copy and paste the main claim from the article

We thus see a focus on a single article here, rather than on trying to find (synthesise and evaluate) claims from multiple different sources. This suggests that the single source is taken as the authoritative resource from which to draw material, yet there is little discussion of the qualities of that source (in terms of metadata or content), and the ways in which the source relates or does not relate to other sources (and their metadata and content).

Group: 414

Group 414 in the MDP task scored high overall (total score = 12), with a relatively high ISEQ score (indicating lower sophistication in epistemic cognition). The group wrote a text of 2049 characters – roughly 1/3 the length of group 325's, and double that of group 369's. The pair's contributions to the text appear to be exactly equal (contribution symmetry = 0), with a very large number of touch points (n = 33) indicating authorial intermingling, suggesting that the pair edited each other's text, and were integrating their contributions. The pair made use (referred to in the chat, text output, or directly copied from) 6 articles, suggesting a selectivity in their article use; both read over half the articles, with 10 opened, and a difference in viewing of 2 articles.

Participant 776 starts the task by suggesting their partner should just read the abstract ('because it is a complete summarz [sic] of the essay'), going on to explain what an abstract is.

userID	projectID	localTime	content
776	414	12:41:12	only read the abstract if there is one because it is a complete summarz of the essay
777	414	12:41:55	huh what
777	414	12:42:03	what is an abstract D
777	414	12:42:18	Im reading all these texts
776	414	12:42:38	its a short text at the beginning of an essay stating whatit is about and what the result is

They then negotiate both reading everything (indeed, one reads 8 articles the other 10), planning for time at the "15minutes left if not more" in order to develop "our final argument".

userID	projectID	localTime	content
777	414	12:48:54	read everything first
777	414	12:48:58	I have 5 articles to go
777	414	12:49:01	for scanning
776	414	12:49:14	okay and then
777	414	12:49:26	At the end we'll have like 15minutes left if not more
777	414	12:49:30	for our final argument

Part way through this period, one of them (participant 777) uses their prior knowledge ("I used to be very interested in this topic") to discuss the risks of glyphosate:

userID	projectID	localTime	content
777	414	12:50:18	One thing that doesnt radiate very clearly from the articles is that if you use roundup
777	414	12:50:36	You need genetically modified GM crops which you also buy from Monsanto
777	414	12:50:50	It causes a dependency on their products its really clever
777	414	12:51:09	I used to be very interested in this topic
777	414	12:51:36	it basicly kills everything that is not part of the GM product you are raising
776	414	12:51:43	it is but i think we only have to write a small empfehlung* to the minister if to use it or not
777	414	12:51:49	And its a health hazard

*Note: empfehlung can be translated from German as 'recommendation' in English.

Their partner (participant 776) points out that there are contradictions between the texts "some say its health hazard some say it isn't", indicating attempts to identify inter-textual ties and synthesise the information provided. Interestingly, the participant with prior knowledge suggests that an article (number 7, critiqued by 8 – see Table 3:2) supportive of the participant's prior knowledge is "really good":

userID	projectID	localTime	content
777	414	12:52:09	right assessing our literature
777	414	12:52:12	We can quote
777	414	12:52:24	I didnt get at all how to quote though those instructions werent really clear
777	414	12:52:42	lets continue reading mate
776	414	12:52:48	the articles are contradictorysome say its health hazard some say it isn't
776	414	12:52:55	yeah lets do that
777	414	12:53:34	http//edusearchcoagmentoorg/instruments/mdp/1020Samsel20Seneffhtm
777	414	12:53:56	this ones really good saying it is particularly dangerous to western people because of their diet
777	414	12:54:07	enhancing the negative effects on their health
777	414	12:54:33	10 samsel seneff at the end of the link

Then they negotiate a conclusion, led by participant 777's prior content knowledge and the task requirements (with little apparent focus on source qualities).

userID	projectID	localTime	content
777	414	13:03:12	I wouldnt agree We can say that this is still debatable but that our advice would be to stay away from the crops Doses found in humans being small just means that they are not very exposed to the glyphosate But they are still exposed
777	414	13:04:44	Id agree on the final con being effect on humans not yet studied well enough and that there is a risk that there is an alleged link between Roundup and many civilizational diseases which are more present in the west obviously like Alzheimers and Diabetes
776	414	13:04:49	yes but this exposure does not have any short term health dangers for humans
776	414	13:05:51	i mean i also am against this bs but as professionals we should nevermind state some pros too so that the minister has an overview and can decide on his own
777	414	13:06:39	Ok how about we admit to nonexistant shortterm effects
776	414	13:07:04	thats good
777	414	13:07:31	But would not advise the ministry to opt for the use of GMO and Roundup as the longerm effects on society considering the risk of civilisational diseases are not yet certain

Group: 518

Group 518 scored relatively low-mid (total score = 6) with a low ISEQ score overall (indicating more sophisticated epistemic cognition). The group wrote the second longest of the MDP example group's texts (number of characters = 4415) with a very symmetrical contribution (contribution symmetry = .01) and low number of touch points (n = 3) indicating few authorial transitions. The pair made use (referred to in the chat, text output, or directly copied from) 7 articles, suggesting a selectivity in their article use; one of the partners read 5 more articles than the other, suggesting an asymmetry in article viewing.

This group talked in German for the whole task, therefore excerpts of their chat are not presented here. Translation (using google translate) indicates that the pair divided the articles between them, and discussed looking for pros/cons in the articles, or the best supported claims (drawing on the task instructions). It is interesting to note that despite the usual practice of working in English, and specific request to do so in this session that some groups did chat in other languages while completing English language written outputs.

Group: 401

In the CIS task, group 401 scored high (total score = 11) and had relatively high ISEQ scores (indicating lower sophistication in epistemic cognition). The group wrote the longest text (number of characters = 8676) with a small asymmetry (contribution symmetry = .14) and relatively high number of touch points (n = 9) indicating moderate amounts of authorial intermingling or/and editing of each other's contributions. The pair viewed 14 pages, making use (referred to in the chat, text output, or directly copying from) 7 pages. The group issued 29 queries, with a relatively low number of 12 unique terms in them.

The group begin by negotiating how to write their text ("we can both edit", "make our summary", etc.) agreeing to write in the taskpad area so that they have a shared awareness (see section 2:2.5) of their work, and can then delete unneeded text.

userID	projectID	localTime	Content
753	401	11:21:46	that taskpad is our shared document where we both can edit our documents
753	401	11:23:22	there we make ourn summary
753	401	11:23:24	and add information
754	401	11:24:14	yes and if i have found something do i just copy past it in the task pad or do i send you a link
753	401	11:24:35	int he document
753	401	11:24:41	we both see what we are doing
754	401	11:25:05	okay and then we delete what we dont need

The pair flag that they should 'judge the claims with the strongest support', going on to suggest they should summarise what they find first

userID	projectID	localTime	Content
754	401	11:31:24	we have to judge the claims with the strongest support
754	401	11:31:43	but we dont have to judge about red yeast rice
754	401	11:31:59	so we have to find claims
753	401	11:32:05	lets just summarize everything we find first
754	401	11:32:53	okay i search for claims now
754	401	11:33:37	then we summaries everything

The pair switch to chatting in Dutch, extracts from which are not given here (google translate indicates this discussion involves some negotiation over finding good sources and whether they

need to write why they choose sources or not, deciding that no conclusion is needed but they should say how much the claims have been supported).

Group: 586

Group 586 in the CIS task, scored low (total score = 4), with lower ISEQ scores (indicating higher epistemic sophistication). The group wrote a text of 5039 characters, with no authorial information data recorded. The pair viewed 17 pages, referring to 5 in the chat or task pad or through copying from those pages. This group issued 30 queries with 14 unique terms in them, as with group 401, indicating a relatively low query vocabulary richness.

The group starts by suggesting that they should create a short summary of the nature of red yeast rice:

userID	projectID	localTime	Content
1100	586	10:11:50	maybe we should get a quick summary together of what red yeast rice is
1099	586	10:12:22	alright

Going on to suggest they might summarise in the task pad (or 'main board'), using quotations to do so:

userID	projectID	localTime	Content
1100	586	10:18:25	so it also has positive effects
1099	586	10:18:52	lets just summ them up on the main board
1100	586	10:19:32	ok
1100	586	10:19:55	u know how to quote stuff
1100	586	10:20:05	so then we can just copz paste it
1099	586	10:21:46	lets do just like this

They send some messages about what red yeast rice is, concluding with the message:

userID	projectID	localTime	Content
1100	586	10:39:21	its kiond of a medicine with some side effects

Group: 633

Group 633 in the CIS task, scored low (total score = 4), with higher ISEQ scores (indicating lower epistemic sophistication). Similar to group 586 and 245 they wrote a text of 5538 characters, with a moderate asymmetry in contribution (contribution symmetry = .27) and a moderate amount of

authorial transitions/intermingling (number of touch points = 6). The pair viewed a small number

of pages (n = 9), only making use (through the chat, task pad, or copying from) 2 pages, which

were shared early on in the chat data. In contrast while they issued a similar number of queries to

other groups (n = 27), there were more unique terms in these (n = 31) indicating a wider

vocabulary use across the queries issued (query vocabulary richness = 1.15).

The group started by both finding and sharing a useful article. They briefly discus which article to

draw from but make a selection without any discussion or explicit evaluation of the source quality

or its content:

userID	projectID	localTime	Content
1187	633	12:47:22	i found a good article from the university of maryland medical center
1187	633	12:48:15	http//ummedu/health/medical/altmed/supplement/redyeastrice
1187	633	12:49:03	read and we can start to write
1188	633	12:49:10	http//wwwivlproductscom/HealthLibrary/HealthConcerns/HeartHealth/ RedYeastRiceforCholesterolTheProsandCons/
1188	633	12:49:18	i found this
1187	633	12:50:07	ok i will read
1188	633	12:50:30	can we just copy and paste D
1188	633	12:50:47	Hahaha
1187	633	12:50:48	from you site or my
1188	633	12:50:55	as you want
1188	633	12:51:00	yours is better i think
1187	633	12:51:18	okay let's make a good copy oopen the task pad
1188	633	12:51:36	we can choose the most intersting topics and just past them

Finally, they divide the content of the site they are drawing on, negotiating that they will each

summarise some of the points made ("I do the first three points and you the other two"):

Then they divide the points they want to make, so one does 3 the other 2

userID	projectID	localTime	content
1187	633	13:06:37	because it's to long on te website
1188	633	13:07:06	ok let's do a summary
1187	633	13:07:25	maybe just put the name
1188	633	13:07:31	i do the first three points and you the other two

Group: 245

Group 245 in the CIS task, scored high (total score = 11, missing one mark on synthesis), with

lower ISEQ scores (indicating higher epistemic sophistication). The group wrote the shortest of

the texts at 4985 characters, although this was a similar length to two of the other groups (586,

633). The pair had a moderate asymmetry in contribution (contribution symmetry = .38) which was the largest of the CIS groups; they had a small number of touch points (n = 6) indicating relatively few authorial transitions or intermingling, suggesting that each author focussed their writing on a single block of text. The pair viewed the most pages (n = 18), and used the most (through the chat, task pad, or copying from) at n = 9. While they issued a similar number of queries to other groups (n = 30), they had a large number of unique terms in these queries (n = 102). One possible explanation for this is that they queried using large quotations; this is confirmed by analysis of the log data, with for example one query as follows:

> *Red yeast rice is often sold as an herbal supplement. There are no regulated manufacturing standards in place for many herbal compounds and some marketed supplements have been found to be contaminated with toxic metals or other drugs. Herbal/health supp - Google Scholar*
>
> UserID 460 Query at 14:28:44

This group was notable for its very low incidence of chat (12 messages) in the main task. The chat was mostly used to ask "did you find anything interesting", with some link sharing, and a few clarificatory remarks (e.g. "nice could you send also the websites addresses for reference", "I think we are done").

7:1.1.1.2.3   Section Summary

The preceding analysis indicates that across a small subset of groups, clear differences in their interaction with the tool, and each other, can be identified. These differences in part relate to the ways they choose to work (for example, a relative lack of chat in some groups), or their task-orientation (how they make sense of the task, off-topic talk, etc.), but in many cases the dialogue evidenced in the chat is epistemic in nature. The analysis thus demonstrates some of the insight the trace can offer, including the chat (amongst others), although as noted in the preceding section, term-based-typology analysis is by its very nature entails a shift from fine to coarse grain analysis

Across the wider variety of trace indicators differences are observed in the participant's interaction with the tools and each other, in the balance of activity (page views, text contributed,

etc.) and in the group-summary activity. Discussion of these differences above suggests that these indicators provide insight into the epistemic nature of the tasks. Moreover, the analysis provided demonstrates that the task design does indeed probe epistemic constructs, engaging participants in tool mediated epistemic activity oriented towards the task.

As in the discussion above (section 7:1.1.1.1) it is important again to highlight that while this analysis provides evidence of the range of activity participants engage in, and behaviours indicative of epistemic-commitments that may be tracked by behavioural trace analysis, the analysis also indicates the nuance of much of the data. This indicates the complexity of making connections between trace (including dialogue) and outcome; for example, some groups engaged in very little chat, with seemingly little impact on their outcome scores or other measures. Thus, while this close analysis indicates a range of responses with a set of data provided for each group, it is challenging to identify patterns across this subset. The next section, then, provides a more quantitative description of the range of trace observed, giving descriptive statistics for the range of behavioural trace indicators, alongside further exemplifications of the kinds of behaviour engaged in (such as websites visited).

7:1.1.2 Trace descriptives

7:1.1.2.1   CIS Trace
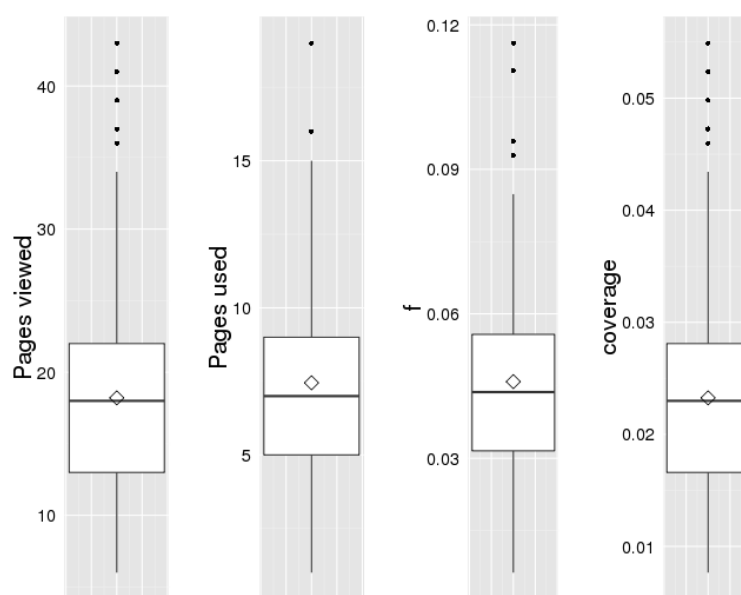
Figure 7:1 – Boxplots of Page Related Trace Indicators

As Figure 7:1 indicates most groups visited under 20 pages (Ug), using around 8 (Urg) of those pages; as the discussion above highlighted, in some cases a single page was relied on for all information. Groups visited just over 2% of all pages visited (coverage) and f scores (a measure of accuracy, in this case strongly related to the n of pages used by an individual group, compared to the set of all pages used by any group) were rather low (all < .12). Groups with higher 'f' values will have made use of more of the pages they visited, and used more of the total subset of pages used by any group. Groups with a higher coverage will have visited more of the pages visited by any group. Note that each page (rather than domain, e.g. 'medicinenet') is considered; inevitably this somewhat increases the number of unique visited pages compared to unique used pages where the latter tend to refer to higher-level domain names (for example, not referring to a subpage indicated '/page1.html', see Table 7:9). Table 7:9 indicates some of the most commonly viewed, and used pages. We see a sharp drop-off in the percentage of groups using each page, such that within the top 20 most commonly visited under 20% of groups visited some pages. We also see that the most commonly viewed page was not used by all groups who viewed it perhaps indicating it was less directly useful than other pages such as the umm.edu page (viewed by ~60% of groups) which was used by most groups who viewed it, and was the second most commonly used page. We also see that many commonly visited pages share a domain (for example, a number of 'webmd.com' pages are present), or are sub-pages (for example, the pages: medicinenet.com/red_yeast_rice_and_cholesterol/page[1-5].htm), which indicates that in some cases pages may be visited while a higher-level (or, alternative) page from the same domain is used.

Table 7:9 - Illustrative examples of pages used and visited[32]

rank	url	project count	% visited	% used	used rank	Website type
1	http://www.medicinenet.com/red_yeast_rice_and_cholesterol/article.htm	138	89.61	62.99	1	Lay health advice
2	http://www.webmd.com/cholesterol-management/red-yeast-rice	99	64.29	43.51	4	

---

[32] This data can be compared to the reported most and least trustworthy URLs reported by a subset of participants, shown in Appendix 17 (which gives the raw lists) and Appendix 18 (which compiles a frequency table of each domain, indicating a small number of sites were listed by different participants as the 'most' and 'least' trustworthy website).

3	http://umm.edu/health/medical/altmed/supplement/red-yeast-rice	90	58.44	56.49	2	
4	http://www.webmd.com/vitamins-supplements/ingredientmono-925-red%20yeast%20rice%20(red%20yeast.aspx?activeingredientid=925&activeingredientname=red%20yeast%20rice%20(red%20yeast	85	55.19	33.12	6	
5	http://articles.mercola.com/sites/articles/archive/2009/09/10/why-you-should-avoid-red-rice-yeast.aspx	79	51.30	29.22	8	Alternative medicine
6	http://en.wikipedia.org/wiki/Red_yeast_rice	79	51.30	22.08	12	Encylopedia
7	http://www.medicinenet.com/red_yeast_rice_and_cholesterol/page4.htm	64	41.56	43.51*	3	Lay health advice
8	http://nccam.nih.gov/health/redyeastrice	63	40.91	36.36	5	Alternative medicine
9	http://www.mayoclinic.org/drugs-supplements/red-yeast-rice/safety/hrb-20059910	59	38.31	24.03	10	
10	http://www.emedicinehealth.com/drug-red_yeast_rice/article_em.htm	57	37.01	18.83	13	Lay health advice
11	http://www.medicinenet.com/red_yeast_rice_and_cholesterol/page4.htm#how_safe_are_red_yeast_rice_products	57	37.01	25.97	9	
12	http://www.reuters.com/article/2008/07/09/us-contamination-common-idUSCOL97022820080709	54	35.06	31.17	7	News
13	http://www.drugs.com/mtm/red-yeast-rice.html	46	29.87	16.88	14	Lay health advice
14	http://altmedicine.about.com/od/herbsupplementguide/a/redyeastside.htm	41	26.62	16.23	15	Alternative medicine
15	http://www.medicinenet.com/red_yeast_rice_and_cholesterol/page2.htm	39	25.32	3.90	36	Lay health advice
16	http://www.nlm.nih.gov/medlineplus/druginfo/natural/925.html	35	22.73	23.38*	11	Government advice
17	http://www.medicinenet.com/red_yeast_rice_and_cholesterol/article.htm#what_is_red_yeast_rice	33	21.43	11.69	17	
18	http://www.webmd.boots.com/cholesterol-management/guide/red-yeast-rice	31	20.13	11.69	18	Lay health advice
19	http://www.medicinenet.com/red_yeast_rice_and_cholesterol/page5.htm	30	19.48	6.49	25	
20	http://www.nutraingredients-usa.com/Suppliers2/Red-rice-yeast-supplements-raise-contamination-issues	29	18.83	11.04	19	Nutrition news

* Note, the higher level of 'use' than 'viewing' may be as a result of strings matching in the use case without having been viewed (for example, by manually typing 'page4' on the end of a url that has not, in fact been visited). This discrepancy may also be a result of errors in the log data. The 'website type' column provides the researcher's assessment of the kind of information and authorship of each given resource.
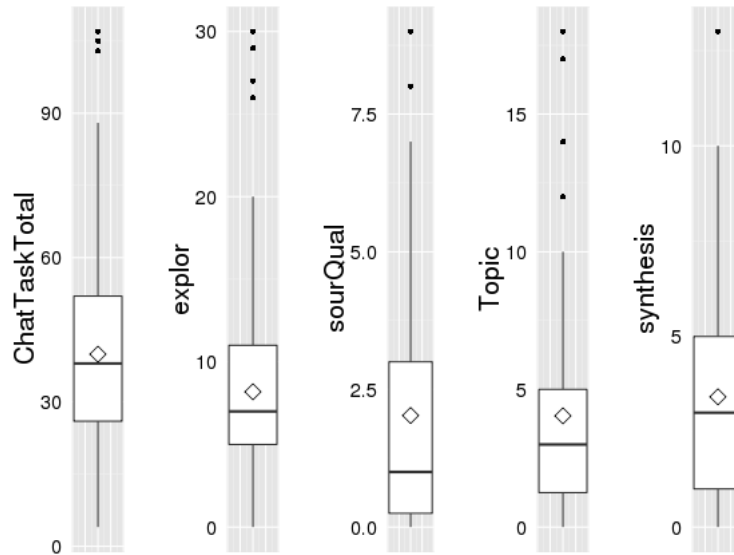
Figure 7:2 – Boxplot of Chat Related Trace Indicators

Analysis shown in Figure 7:2 shows that groups tended to send slightly over 50 messages within

the main task-slot, with 20% involving exploratory types of terms, very few (under 2) involving

sourcing meta-discourse, and under 5 topic-based or synthesis-writing terms. The section 7:1.1.1

gives some of the key exemplifications from the chat data, with section 7:1.1.1.1 indicating the

ways in which particular typology-terms were used in messages, and 7:1.1.1.2 the broader nature

of the dialogue, including the variability in the nature of the chat even where the overall quantity

of chat was similar.
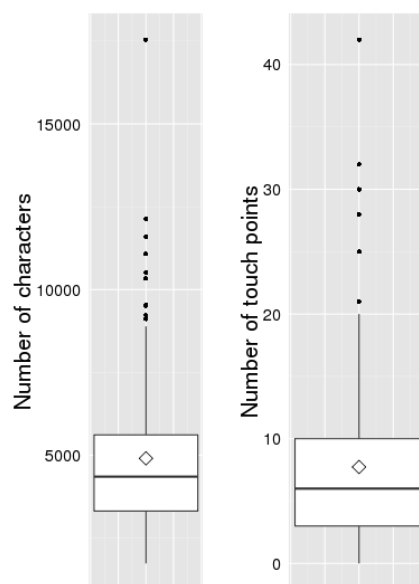
Figure 7:3 – Boxplots of Etherpad Related Trace Indicators

Figure 7:3 shows that the CIS groups wrote an average (M) of 4907 characters in their texts, with a

median of 4356 and SD of 2356 – indicating a wide range of text lengths written. The CIS pairs had

a median of 6 touch points, with a maximum of 42 (n = 1) and a minimum of 0 (n = 6, indicating

data loss in authorial attributions for those groups).
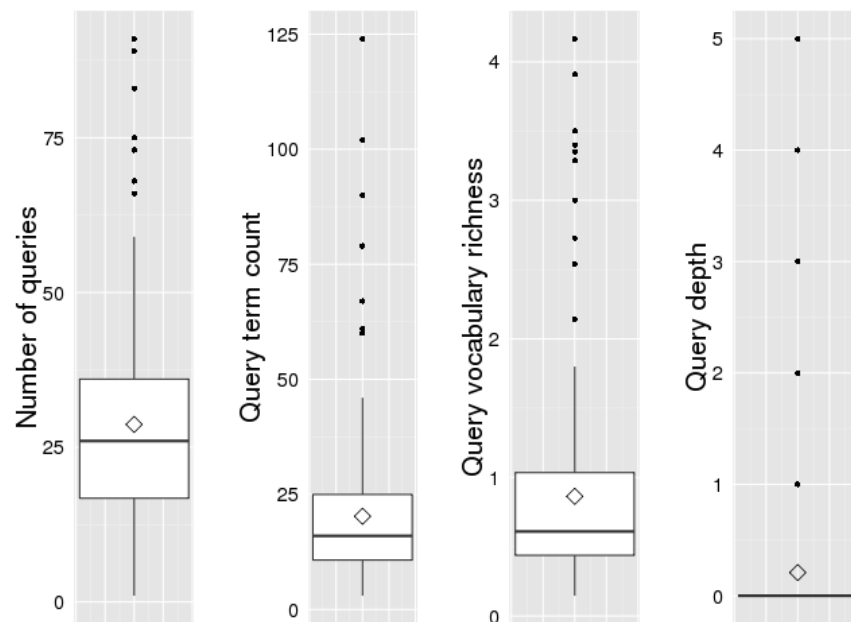
Figure 7:4 – Boxplots of Query Related Trace Indicators

Figure 7:4 shows that groups issued an average of just over 25 queries, using an total number of

terms in those queries numbering in the high teens. As the third boxplot indicates, for many

groups the n of unique terms was equal to the n of queries (which is what '1' indicates), while for

some many more terms than queries were used (>1) and others more queries than terms (<1).

We also see that most groups did not go past the first page of query results (query depth). Indeed,

analysis of a frequency table indicates that 139 groups visited only the first page of query results,

with 9 groups going beyond the first page once, 1 group both twice and thrice, and 2 groups both

four and five times. Note depth here indicates any page beyond the first – so going to page 2 of

results, twice or going to both page 2 of results then page 3 would both be marked as 'query

depth 2'. Table 7:10 indicates the most common unique queries (i.e. not taking into account a

group repeatedly using the same query) across all groups. This suggests that most groups

explored pages from a small number of queries, although as indicated by the boxplot, these query pages were repeatedly visited (through reissuing the same query, use of the back button, or re-activating the query tab in the browser window).

Table 7:10 – Most common queries in the CIS task

Query	Project count	%
red yeast rice	135	87.66
red yeast rice risks	29	18.83
red yeast rice contamination	26	16.88
red yeast rice safety	15	9.74
risks of red yeast rice	13	8.44
safety of red yeast rice	12	7.79
red yeast rice safety concerns	11	7.14
Translate	11	7.14
red yeast rice benefits	10	6.49
red yeast rice france	10	6.49
statin drugs	9	5.84
Citrinin	8	5.19
google scholar	8	5.19
red yeast rice side effects	8	5.19805
red yeast rice#q=red yeast rice risks	8	5.19
safety red yeast rice	8	5.19805
what is red yeast rice	7	4.54
Lovastatin	6	3.89
red yeast rice pros and cons	6	3.89
benefits of red yeast rice	5	3.24

As Figure 7:5 indicates, some asymmetry was common in page viewing and use, etherpad contributions, and querying – as was also demonstrated in closer discussion of the subset of participants discussed in section 7:1.1.1.2. The boxplots represent a median asymmetry of 4 pages viewed (with a maximum of 18) and, 1 used (with a maximum of 8). Contribution symmetry indicates a median difference of 142 words contributed, with a maximum of 1482 (likely due to copy-pasting chunks of text). Generally groups were fairly symmetrical in the activity (with few asymmetries indicating one partner conducting the majority of the work).
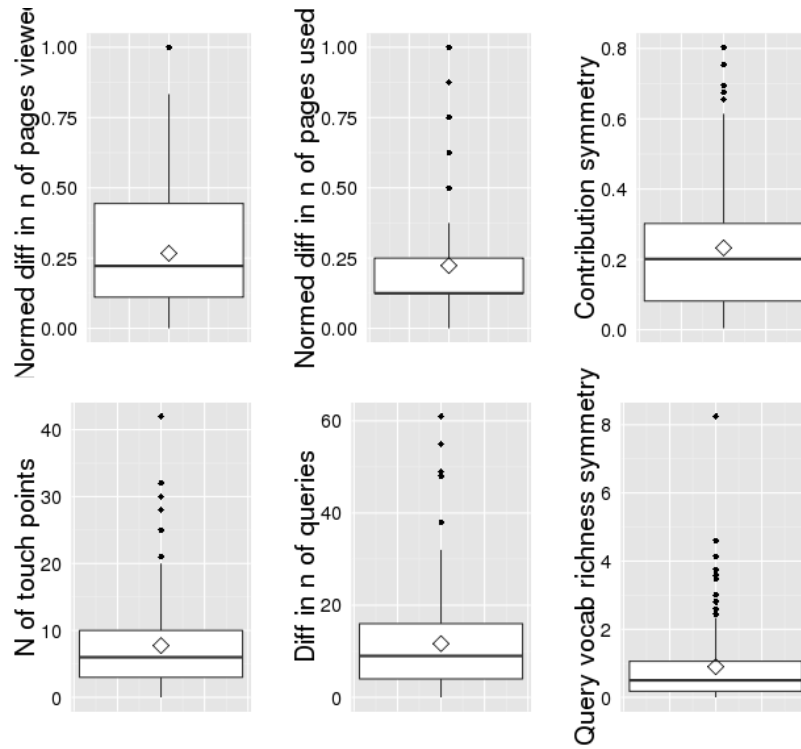
Figure 7:5 – Boxplots of Symmetry Indicators

7:1.1.2.2   MDP Trace

As Figure 7:6 indicates, most (n = 102) groups viewed all 11 pages provided to them (note that for

analysis of page views (Ug), additional pages viewed were excluded), with only 2 groups viewing

fewer than 8 pages (5 and 7 respectively), and 31 groups viewing 8-10 pages (inclusive). Groups

used a larger range of other pages (where all urls are analysed, this results, for example, from

sharing links to translations or other background materials).
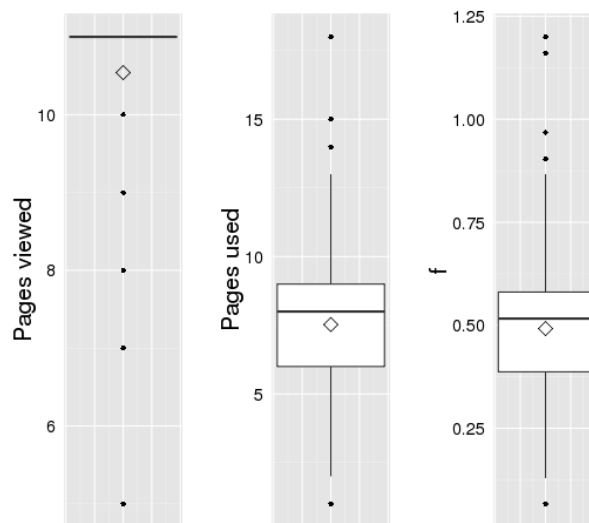
Figure 7:6 – Boxplot of Page Related Trace Indicators
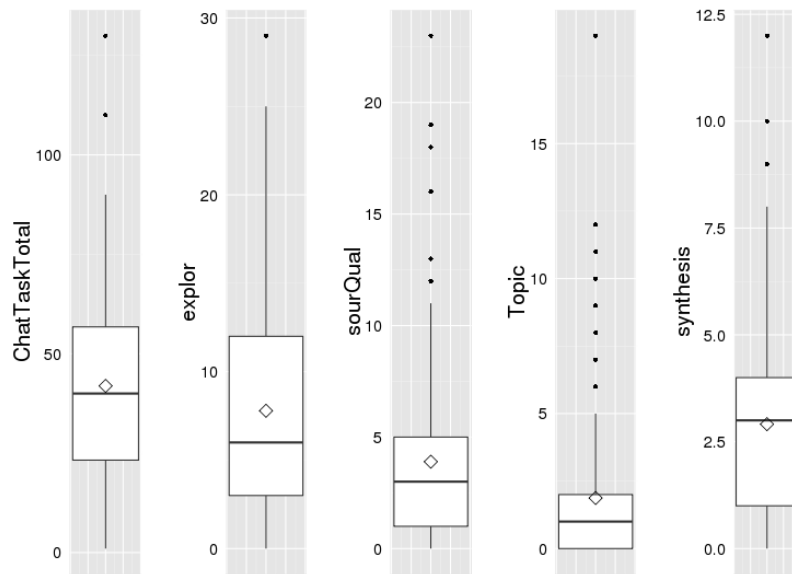
Figure 7:7 – Boxplots of Chat Related Trace Indicators

As in Figure 7:7, similar amounts of chat are seen in the MDP task as in the CIS (Figure 7:2), although slightly less topic chat. Again, extended exemplifications of this kind of chat can be found in the earlier section.

In Figure 7:8 we see a somewhat longer text measured by number of characters in the MDP than CIS task (Figure 7:3) (MDP M = 5965, SD = 3065; compared to CIS M = 4907, SD = 2356). There are fewer touch points in the MDP task (median of 3, compared to 6 in the CIS task).
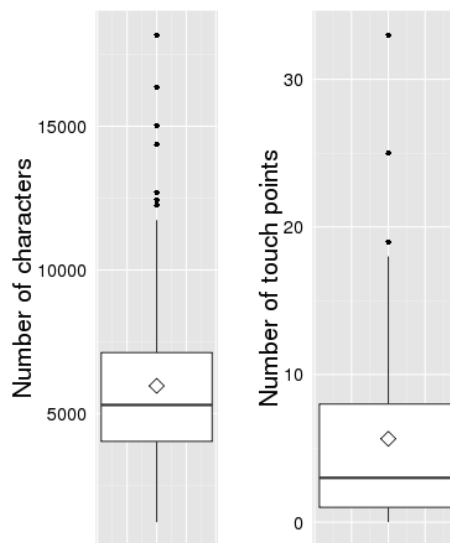
Figure 7:8 – Boxplots of Etherpad Related Trace Indicators

Although querying behaviour (indicated in Figure 7:9) was not of direct interest to the MDP task (and is not used in the regression analysis) it is interesting to note that despite not needing to query, 66 groups did issue some queries (while 69 issued none). Visual inspection of the queries indicates most queries were for key terms (e.g. 'glyphosate', 'roundup') or/and translation (e.g. searching for 'google translate'), further highlighting the salience of search to student learning, and the importance of analysis of CIS and collaborative MDP in learning contexts.
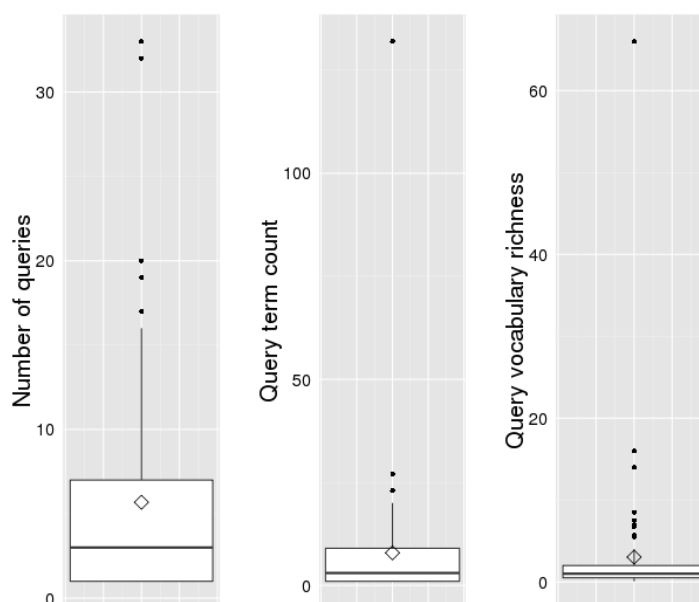
Figure 7:9 – Boxplots of Query Related Trace Indicators

As Figure 7:10 indicates, small asymmetry was common in page viewing and use and etherpad contributions. The boxplots represent a median asymmetry of 2 pages viewed (with a maximum of 9) and, 1 used (with a maximum of 6). Contribution symmetry indicates a median difference of 167 words contributed, with a maximum of 56216 (due to copy-pasting chunks of text). Generally partner activity was symmetrical (with few asymmetries indicating one partner conducting the majority of the work).
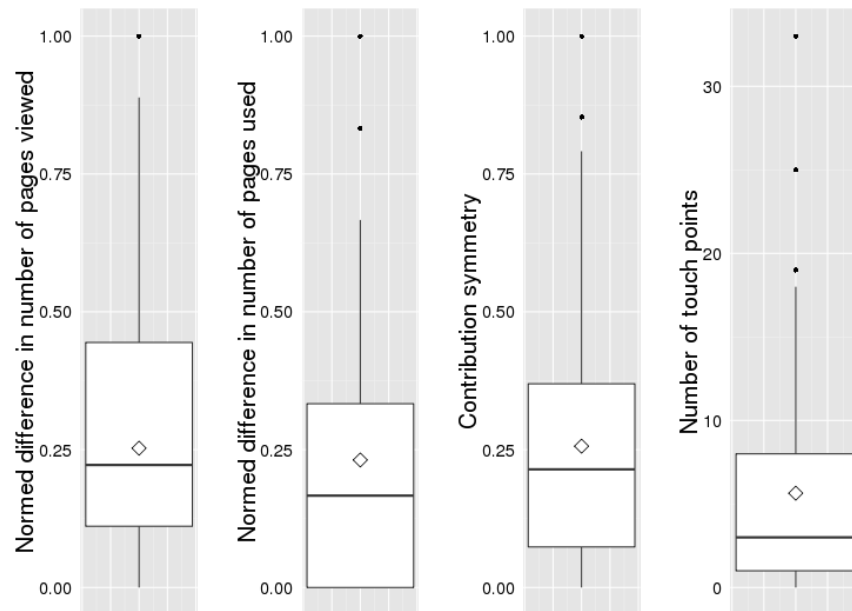
Figure 7:10 – Boxplots of Symmetry Trace Indicators

7:1.1.2.3   Section Summary

These analyses of the trace data provide indicate summaries of the ways in which participants

interact with each other and their information environment on some key indices. Through the

provision of means, medians and ranges they indicate the general levels of activity, which may be

considered in comparison to prior research (which has typically not used as rich a log-data set),

and with regard to understanding the kinds of behaviours most participants engaged in. In

addition, the range of behaviours taken – including less common behaviours, indicted by outliers

– can be observed and investigated. The range of observed behaviours is indicative of the

potential analytic role of such indicators, however, as in the case of the chat data closer analysis

of the specific resources visited and used, queries made, and so on also informs our

understanding of the participants' behaviours and their nuance. In addition, the ways in which

various behaviours interact or co-vary is of interest – for example, suggesting relationships among

trace, an issue to which I now turn.

7:1.1.3 Trace Correlations

Pearson's correlation matrices as indicated in Table 7:11 were examined for significant

relationships using casewise exclusion. Correlations on the paired data were computed using

averages for demographic and survey data (e.g. average age of the partners) and aggreagtes for

the trace (e.g. total number of pages visited). For individual data correlations used individual summary data, excluding group-based variables (such as symmetry scores). For brevity the table excludes age, and some symmetry measures (which are not normally distributed). As in the table above (Table 4:4), the table should be read as CIS in above the diagonal, and MDP below, with paired data unbracketed and individual data bracketed. Note that not all significant reuslts will be discussed below, for example where correlations are between measures of the same type (e.g. the chat kinds), where only paired-data indicates a correlation (but not individual), or where correlations are theoretically spurious (for example, the significant CIS correlation between ISEQGen and partner familiarity).

7:1.1.3.1   CIS Trace Correlations

In the CIS data, there are a set of significant correlations between the number of pages viewed and used, and querying – including the number of queries, their vocabulary richness, and their depth. This indicates that querying and page behaviours are associated, suggesting for example that there was relatively little incidence of groups browsing through large numbers of pages from relatively few queries. There was also a moderate correlation between number of pages used and: number of characters (reflected in the individual data, but not reported in the table, by a relationship between number of pages used and number of words attributed to that author $r(288)$ = .19 p < .001). Chat data is in various ways related to other trace, indicating that – as exemplified in section 7:1.1.1 – chat is associated with other trace activities such as querying, reading, and writing. Specifically we see that the number of touch points is negatively associated with source chat $r(152)$ = -.173 p = .032 and positively with synthesis chat $r(152)$ = .218 p = .007, indicating that this chat may indeed have been used to engage in coordinated writing efforts as indicated by the chat, and authorial mingling. We also see a small correlation between source chat and query vocabulary richness $r(150)$ = .20 p = .01, also reflected in the individual data $r(261)$ = .13 p = .03, and between topic chat and query term count $r(150)$ = .18 p = .03, perhaps reflecting an association between talking about sources and using a wider range of query terms (an association which could be bi-directional).

In relation to the ISEQ data, it is important to note very few statistically significant relationships between either partner or individual ISEQ scores and trace data. The small correlations between the ISEQ factors and self report search experience are of interest, and suggest that for the justification factor higher search experience was associated with lower ISEQ score (and vice-versa), and the converse for the general factor particularly given the absence of other correlations (which might imply, for example, a participant level response bias towards high/low scoring across self-report items). This suggests a relationship between justification sophistication and search capability, and that those with less sophisticated perspectives on the general factor perhaps overrate their own search capability. Given higher scores in both indicate less sophisticated epistemic perspectives this impact on self-report of search experience may deserve further investigation.

In addition, other small relationships between self-report data (e.g. search expertise) and trace, are demonstrated in the data, indicating self-report may offer insight into behavioural markers in trace, specifically:

- A small correlation between pair-search experience and number of pages visited, $r(152) = .16$ $p = .047$, also reflected in the individual data $r(299) = .13$ $p = .021$; and a small correlation between pair-search experience and query depth $r(152) = .19$ $p = .020$, also reflected in the individual data $r(300) = .13$, $p = .03$. Implying a small relationship between self report search experience and some search-behaviours, perhaps indicating that self-reports of search experience provide a valid insight into more advanced search use (such as browsing to deeper search engine results pages).
- Small correlations between partner familiarity, partner agreement, topic knowledge, and total chat: $r(152) = .20$, $p = .01$, and in the individual data $r(300) = .28$, $p = .00$; $r(152) = .20$ and $r(300) = .20$ $p = .00$, in the individual; and $r(152) = .21$, $p = .01$ and $r(300) = .12$, $p = .05$ in the individual respectively. Perhaps indicating that chat was used to communicate ideas and contribute to levels of agreement.
- Small correlations between pair-collaborative satisfaction with total chat $r(140) = .274$ $p = .001$. Implying that pairs who were satisfied with their collaboration had more chat; note that paired data (exploratory, topic, and synthesis chat) are significant here, this may imply that the important relationship is not between individual collaborative-satisfaction and messages sent individually but pair-collaborative-satisfaction and the pooled set of messages sent/received.
- A small correlation between pair-topic knowledge and contribution assymmetry $r(146) = .23$, $p = .01$, indicating that pairs with higher (average) topic knowledge have more asymmetry in contribution (likely because the increase in topic knowledge comes from one partner not both, and that partner contributes more).
- A small negative correlation between pair-task-satisfaction and the number of touch points $r(140) = -.168$ $p = .05$; visual inspection indicated this was due to outliers, removal of the 95th percentile removed this effect ($r(133) = -.02$, $p = .90$). Given the small number

of outliers, further investigation is required to establish whether high numbers of touch points (indicating more authorial intermingling) are associated with lower task satisfaction, perhaps implying a quadratic relationship.

Given the correlation sizes, multiple correlations, and mixed paired/individual nature of the analysis, some caution should be taken in drawing conclusions. However, evident in this data is a link between chat and partner-common-knowledge (through partner familiarity, partner agreement, collaborative satisfaction, and topic knowledge), highlighting the importance of chat data as a data source. We also see that self-report search experience is related to behavioural markers of information seeking behaviours (page viewing and query depth) indicating a validation of the self-report measure.

7:1.1.3.2 MDP Trace Correlations

In the MDP data, it is interesting to note – parallelling the CIS data – the relationship between the number of pages used and the number of characters in the text output (although this was not reflected in the individuals-word-contributions for the MDP group $r(300) = -.06$, $p = .04$). Individual exploratory and source chat was positively associated with visiting pages, and negatively with using pages perhaps indicating the use of chat in – respectively – sharing knowledge and evaluation, and filtering out poor articles. While topic chat was negatively associated with the number of characters in the ouput and both visiting and using pages on a pair, but not individual, level. The number of touch points was positively associated with topic, exploratory, and total chat, (but –in contrast to the CIS data – not synthesis) again perhaps indicating coordinating behaviour.

As in the CIS case, the justification factor of the ISEQ was negatively correlated with search experience (although unlike the CIS data, it was not positively associated with the general factor). ISEQ justification also appears to have a small negative relationship to GPA. There was also one small negative correlation between the justification factor and the number of pages visited, perhaps indicating that those with more sophisticated perspectives visited more of the pages provided to them. Again, it is important to note that the ISEQ is not clearly related to any other trace data in this task.In relation to other self-report data, we see:

- Small correlations between pair-collaborative satisfaction and: exploratory chat $r(109) = .225$ $p = .018$; topic chat $r(109) = .252$ $p = 008$; synthesis chat $r(109) = .189$ $p = .047$; and total chat $r(109) = .189$ $p = 046$; with individual data reflecting a correlation between collaborative satisfaction and exploratory chat $r(227) = .130$ $p = .050$; and total chat $r(227) = .132$ $p = .046$. As in the CIS case indicating a relationship between collaborative satisfaction and chat levels.
- Small correlations between pair-addon intuitiveness and source chat $r(109) = .253$ $p = .007$, and exploratory chat $r(109) = .199$ $p = .036$; with individual data indicating a small correlation between addon intuitiveness and use of source chat $r(227) = .161$ $p = .015$. This indicates that those who found the addon more intuitive were more likely to engage in chat using the addon (or, conversely, that those who engage in more chat behaviours were more likely to find the addon intuitive)
- Individual data indicated a small correlation between GPA and use of source chat $r(240) = .193$ $p = .003$, indicating that those with higher GPA made more specific references to source qualities.

7:1.1.3.3 Section summary

In both the CIS and MDP task, all observed correlations are small, and thus some caution should be taken in drawing conclusions. Collaborative satisfaction appears to be related to chat levels in both tasks– while, in the MDP but not CIS data, there is no significant correlation to partner familiarity. For the MDP task there was a small correlation between the individual ISEQ data and number of pages viewed, indicating that participants with less sophisticated perspectives on the justification of knowledge opened fewer pages while no such correlation was observed in the CIS data. Despite the relationships between the ISEQ and internet behaviours, we see very few relationships, e.g. in the amount of chat, or pageview/pageuse behaviours, or queries (and query vocabulary richness). This is notable given that we might expect those with high general ISEQ scores to uncritically source, synthesise relatively little, and make less use of their partner (resulting in more asymmetry, less chat, and fewer touch points); and those with low justification ISEQ scores to corroborate more, make more use of their partners (resulting in more symmetry, and more chat), and critique sources (resulting in more source chat) more identifying connections across them (and thus synthesising).

With regards to trace correlations, the correlation in both tasks of page-use to output-length is of interest, suggesting participants draw on and elaborate more information when drawing from a range of sources (although recall that the url characters make up part of the additional character length in the text outputs). While in the MDP task chat data seems to be associated with page viewing and use, and with the number of touch points, only the latter relationship holds for the

CIS data. Moreover, in this latter case touch points in the CIS task were associated negatively with source chat and positively with synthesis, while in the MDP case touch points were associated positively with topic, exploratory and total chat (but not synthesis). The CIS data provides the additional data-source of querying behaviour, which indicates relationships among number of queries, query depth, and query vocabulary, and both chat and number of pages viewed.

The next step, then, is to explore relationships among trace, survey data, and outcome variables – to which the next section turns.

	search Experience	collaborative Satisfying	task Satisfying	addon Intuitive	GPA	ISEQ Gen	ISEQ Just	topic Knowledge	partner Agreement	partner Familiarity	Pages viewed	Pages used	f	contrib symmetry	N touch points	nchar	explor	sourQual	Topic	synthesis	ChatTask Total	n queries	Query term count	Query richness	query depth
search Experience						.20* (.20***)	-.30*** (-.18**)	-.03 (-.01)	.12 (.10^)	.02 (.00)	.16* (.13*)	.04 (.03)	.03 (.06)	-.02	-.01	-.04	.06 (.06)	.12 (.07)	.12 (.05)	.08 (.06)	.04 (.02)	.11 (.13*)	-.06 (.03)	-.06 (.02)	.19* (.13*)
collaborative Satisfying						.02 (.02)	-.03 (.04)	-.14^ (-.02)	.36*** (.28***)	-.03 (.01)	-.03 (-.04)	-.14^ (-.12^)	-.14^ (-.12*)	-.10	.03	.04	.18* (.07)	.04 (.01)	.21* (.09)	.19* (.11^)	.27*** (.12^)	-.08 (-.06)	-.08 (-.06)	.02 (.05)	.02 (-.01)
task Satisfying		*See* Table 4:4, p. 170				.01 (-.05)	.03 (.02)	-.01 (.04)	-.01 (.01)	-.09 (-.02)	.02 (.08)	.01 (.12^)	.01 (.12^)	-.11	-.17*	.07	-.06 (-.04)	.09 (-.03)	-.03 (-.03)	-.02 (-.02)	-.02 (-.07)	.00 (.01)	.04 (.03)	.11 (.05)	-.13 (-.02)
addon Intuitive						-.01 (.02)	-.02 (.03)	-.02 (.02)	.02 (.00)	-.11 (-.06)	.16^ (.10)	.05 (.06)	.04 (.06)	-.02	-.12	.12	-.10 (-.12*)	.02 (-.03)	.05 (-.07)	-.08 (-.02)	-.08 (-.11^)	.07 (.00)	.09 (.01)	.08 (.03)	.11 (.03)
GPA						.02 (.00)	-.15^ (-.09)	-.06 (-.09^)	.06 (.02)	-.01 (-.05)	.02 (.05)	.05 (.01)	.05 (.01)	-.05	-.04	.11	-.08 (-.02)	.10 (.1-)	.01 (.06)	-.20* (-.04)	-.03 (.00)	-.01 (.03)	.11 (.06)	.04 (-.06)	-.06 (.02)
ISEQ Gen	.07 (.09)	.09 (-.03)	.03 (.00)	.15 (.08)	-.13 (-.08)		-.28*** (-.22***)	-.08 (-.01)	.11 (.06)	-.18* (-.12*)	.08 (.01)	.03 (.00)	.02 (.00)	-.03	.07	.11	-.08 (-.10)	.13 (.00)	0 (-.06)	-.04 (-.09)	.00 (-.06)	.07 (.06)	-.08 (.03)	-.14^ (-.01)	-.02 (.04)
ISEQ Just	-.23** (-.12^)	.06 (.11^)	.15 (-.01)	-.02 (.01)	-.12 (-.16*)	-.25** (-.21***)		.04 (.07)	-.02 (-.01)	-.01 (-.01)	-.03 (.02)	.00 (.02)	.00 (.01)	.02	-.08	.02	-.15^ (-.06)	-.15^ (-.02)	-.17* (-.05)	-.09 (-.02)	-.13^ (-.04)	.05 (.01)	-.01 (-.01)	.06 (.08)	.04 (.09)
topic Knowledge	.14 (.13*)	.01 (.04)	.15 (.17**)	.01 (.04)	-.07 (-.14*)	.08 (.02)	.04 (.04)		-.11 (-.04)	.09 (.08)	.04 (.02)	-.02 (.00)	-.02 (.01)	.23**	-.12	.08	.18* (.11^)	.00 (.06)	.08 (.06)	.10 (.05)	.21* (.12^)	.04 (.03)	-.02 (.03)	-.05 (-.02)	-.04 (.00)
partner Agreement	.24** (.15*)	.28** (.20**)	.11 (.08)	.19^ (.13*)	-.12 (-.05)	.00 (.01)	-.07 (-.11^)	-.04 (-.04)		.04 (.08)	-.04 (.00)	-.07 (-.09)	-.07 (-.10^)	-.23**	.07	-.06	.24** (.21***)	.05 (.10)	.11 (.12*)	.08 (.06)	.20* (.20***)	-.01 (.01)	-.02 (-.04)	.06 (-.03)	.06 (-.02)
partner Familiarity	.07 (.06)	.07 (.03)	.16 (.16*)	.10 (.10)	-.02 (.02)	.08 (.10^)	-.03 (.01)	.19* (.17**)	.16^ (.11^)		-.01 (-.02)	.09 (.06)	.1 (.06)	.07	.03	.03	.07 (.13*)	.09 (.15*)	.18* (.15*)	.10 (.14*)	.26** (.28***)	.01 (-.02)	.05 (.00)	.05 (.03)	-.09 (-.09)
Pages viewed	.07 (-.01)	.09 (.02)	.05 (.02)	.09 (.05)	-.07 (-.06)	-.12 (.08)	-.05 (-.18**)	-.02 (-.02)	.03 (.03)	.00 (-.06)		.51*** (.37***)	.47*** (.34***)	-.07	-.11	.11	-.08 (-.05)	.01 (.03)	.12 (.04)	-.15^ (-.08)	-.09 (-.06)	.65*** (.57***)	.20* (.17**)	-.29*** (-.21***)	.48*** (.39***)
Pages used	.10 (.06)	.06 (.04)	-.08 (-.08)	-.13 (-.12^)	.01 (-.08)	-.08 (-.04)	-.12 (.03)	.15^ (.10^)	-.09 (-.01)	-.01 (-.01)	.22* (.10)		1.00*** (1.00***)	-.14^	-.09	.29***	-.04 (.01)	-.09 (.09)	.05 (.08)	.04 (.05)	-.03 (-.02)	.34*** (.21***)	.16* (.04)	-.13^ (-.10^)	.24** (.16**)
f	.09 (.06)	.06 (.04)	-.08 (-.08)	-.14 (-.12^)	.02 (-.08)	-.08 (-.05)	-.12 (.04)	.15^ (.10^)	-.1 (-.01)	-.01 (-.01)	.15^ (.08)	1.00*** (1.00***)		-.14^	-.08	.29***	-.04 (.01)	-.10 (.09)	.04 (.07)	.05 (.04)	-.02 (-.02)	.32*** (.19**)	.15^ (.03)	-.13 (-.10^)	.21** (.15*)
contrib symmetry	-.03	.10	.04	-.02	.14	-.08	-.01	.03	.01	.04	.00	.06	.06		.05	.16^	-.05	-.03	0	-.09	.04	-.13	-.02	.07	-.07
N touch points	-.03	.14	.15	.01	-.10	.12	-.04	.04	-.02	-.04	.07	-.05	-.05	-.03		.08	.10	-.17*	-.03	.22**	.06	0	.02	-.08	-.08
nchar	-.10 (-.09)	.02 (.22*)	-.06 (.01)	-.12 (.20*)	.08 (.11)	-.07 (.05)	-.12 (-.08)	-.11 (-.08)	.03 (-.04)	-.12 (-.08)	.03 (.01)	.34***	.35***	.15^	.01		-.07	.05	-.09	.06	-.01	.1	.05	-.11	.06
explor	-.09 (-.03)	.22* (.13*)	.01 (-.01)	.20* (.08)	.12 (.07)	.11 (.05)	-.08 (-.08)	-.08 (-.02)	-.04 (-.02)	-.08 (-.08)	.01 (.19**)	-.16^ (-.12^)	-.16^ (-.12^)	.11	.20*	-.10		.18*	.46***	.57***	.70***	.01 (.01)	.01 (.01)	.01 (.01)	-.07 (-.07)
sourQual	-.04 (-.01)	.13 (.06)	-.05 (-.05)	.25** (.16*)	.12 (.19**)	.02 (-.01)	-.10 (-.07)	-.02 (.01)	.00 (-.04)	.05 (.01)	-.06 (.14*)	-.16^ (-.20**)	-.16^ (-.20**)	.07	.03	-.13	.68*** (.37***)		.37***	.06	.24**	-.07 (-.05)	.08 (.05)	.20* (.13*)	.04 (.01)
Topic	.09 (.05)	.25** (.09)	.13 (.03)	.14 (.05)	-.05 (.07)	.12 (.07)	-.05 (-.04)	.08 (.10)	.17^ (.07)	.10 (.07)	-.19* (.08)	-.19* (-.08)	-.18* (-.09)	-.04	.22**	-.18*	.44*** (.46***)	.56*** (.42***)		.28***	.53***	.11 (.06)	.18* (.08)	.09 (.03)	-.06 (-.06)
synthesis	-.06 (-.04)	.19* (.07)	.04 (-.05)	.14 (.00)	-.05 (.09)	.13 (.08)	.02 (.01)	-.11 (-.05)	-.05 (-.01)	-.08 (-.07)	.06 (.06)	-.11 (-.11^)	-.12 (-.11^)	.01	.14	-.04	.54*** (.56***)	.30*** (.30***)	.28*** (.28***)		.51***	-.07 (-.12^)	0 (-.03)	.06 (.09)	-.05 (-.07)
ChatTask Total	-.06 (-.03)	.19* (.13*)	.09 (.04)	.12 (.05)	-.08 (.02)	-.01 (.00)	.17^ (.06)	-.10 (-.01)	.11 (.08)	.04 (.02)	-.03 (.14*)	-.06 (-.04)	-.06 (-.04)	.10	.23**	.03	.53*** (.74***)	.47*** (.42***)	.49*** (.52***)	.52*** (.58***)		.05 (0)	-.01 (.02)	-.07 (-.06)	
n queries																							.34*** (.36***)	-.36*** (-.36***)	.43*** (.43***)
Query term count																								.55*** (.43***)	.03 (.05)
Query richness																									-.12 (-.11^)
query depth																									

Table 7:11 – Correlation Matrix for Trace and Survey Data

* Note age and symmetry measures (which we would not expect to have direct linear correlations) are removed from this table. Sample sizes vary as reported in discussion of the measures. Unbracketed correlations refer to paired data, with individual data bracketed. Below the diagonal is based on MDP data, above on CIS data.

7:1.1.4 Regression Models: From Trace and ISEQ to Outcome

In order to analyse relationships between outcome, trace indicators, and survey-item variables (with a particular interest in IESQ scores) multiple linear regressions were conducted.

Each of the metrics reflects potentially epistemically salient features: Page viewing and use; information querying and behaviours; chat within the typology presented; writing metrics and division of labour (or symmetry) scores. As such neither the variables present in, their order, nor their interactions, are driven by an a priori model.

Analysis of skewness and kurtosis (measures of distribution symmetry and distribution shape) was conducted to ensure variables were normally distributed. The natural logarithm of variables with a non-normal distribution is used to compensate by normalizing the distribution (or reducing the 'long tail') of results as follows:

- CIS: nchar, query vocabulary richness, topic knowledge
- MDP: nchar, topic knowledge

In order to (a) ensure only the best theoretically grounded variables are included and (b) reduce covariance within the independent variables, a number of variables were removed as follows:

- Asymmetry is captured in contribution asymmetry on the basis that asymmetries elsewhere should be borne out in the written text, or/and relate to markers in the chat data.
- Number of queries and query terms are combined in the measure of query vocabulary richness
- Page views and use are combined in the 'f' measure
- Some of the survey data we would expect to relate directly to outcomes (e.g. search skill, and other survey items) while others might, if any effect, have indirect effect which we would anticipate observing via other variables (age and GPA); these latter variables were removed for this stepwise analysis.

No remaining variables covaried. Note that in some cases (for example, topic knowledge) analysis could be conducted to exclude pairs, rather than to include the variable as a predictor; this may be productive for future research, but at this exploratory stage the interest is in the ways that the data available are predictive of outcome in interaction with behavioural trace. A set of stepwise models were constructed (as described further on p.200) with the R stats package's (R Core Team, 2015) 'step' function making use of both forward and backward selection. In this analysis, at each

'step' the algorithm adds or removes variables, favouring smaller residual error with a penalty for the addition of extra variables (in order to avoid overfitting). Thus, models may not be independently significant nor the variables within them, but the models with the smallest residual error are typically examined. These may be taken as exploratory models indicative of potential for future research; in some cases variable effect sizes are small or p values high, these remain in the model according to the criteria described above, but should (as in any regression) be treated with caution.

As in the earlier stepwise regressions, the Akaike information Criterion (AIC) (Akaike, 1974) has been used to inform (but not decide) model selection. Thus, models were examined with consideration to a priori assumptions (regarding theoretically important independent variables for prediction of any particular outcome), and analysis of the variables included in each model, alongside the AIC for those models. Recall (p.200) that AIC provides an indicator of information lost under any model, selecting for better fitting models, but with a penalty for increasing numbers of variables in the model. Models with the smallest (signed) AIC are considered the 'best' model available (i.e., this should not be taken to mean they are a *good* model). In this reporting I follow the rule of thumb that models within 0-2 AIC of the minimum AIC model are have substantial empirical support, with those 4-7 having considerably less support, and greater than 10 essentially none (Burnham & Anderson, 2002, pp. 70–71); I thus manually identify models within 5 of the minimum AIC with a preference for those within 3 of the minima; these models are presented in Appendix 16). In each table below, the models are presented in order of decreasing AIC (where smaller AIC indicates 'better' models). The model effects are given by adjusted R^2 with β – the standardised regression coefficient – given alongside B, the unstandardized coefficient. The standardised coefficient indicates the difference (in standard deviations) made on the dependent variable per standard deviation increase in the predictor variable; they can be read such that that the largest β is the variable with the biggest impact on the outcome. The unstandardised coefficient (B) expresses the same relationship, expressed in

the original unit of measurement, such that the indicated quantity (B) shows the amount the

dependent variable increases for each unit increase of the independent variable.

7:1.1.4.1 CIS Outcome Regressions
7:1.1.4.1.1 Topic score

Table 7:12 – Multiple Linear Regressions for CIS Topic Scores from Stepwise Process

	t	p	B	β	F	df	P	adj. R^2
Topic Score								
Model One (AIC = -69.1)					10.70	141	.000***	.284
Contrib symmetry	-1.12	.266	1.162	-.081				
Nchar	7.40	.000***	0.003	.523				
ChatTaskTotal	0.63	.530	0.003	.054				
Synthesis chat	-1.60	.111	0.047	-.132				
Topic knowledge	-1.81	.073^	-0.219	-.133				
Partner familiarity	1.85	.066^	0.047	.135				

Following the guidance described above, one model was investigated for CIS topic scores,

indicating a significant (p < .0001) model accounting for a moderate amount of variance (r^{2adj} =

.28). The number of characters in the text output is the best predictor of topic score (and the only

effect significant at the .05 level), with possible smaller effects from other variables, including a

negative effect of synthesis chat (p = .11), contribution symmetry (p = .17), and topic knowledge

(p = .07), and a positive effect of partner familiarity (p = .07). This suggests that output length is

the key variable: participants who wrote more, included more topics in what they wrote.

Variables which are not significant should be investigated further, but may indicate that

participants with higher familiarity included more topics (perhaps because they could co-ordinate

and share their knowledge better). Other variables indicate that higher topic knowledge tended

to reduce the number of topics covered (perhaps because they believed them to be less relevant,

or because their focus was on the topics of their prior knowledge), and that those who co-

ordinated their writing more (synthesis chat) also covered fewer topics; although contrary to this

finding is that more asymmetry in contribution was related to poorer outcomes (although note p

= .27 for contribution symmetry).

Table 7:13 – Multiple Linear Regressions for CIS Synthesis Scores from Stepwise Process

	t	*p*	B	β	*F*	*df*	*p*	adj. R^2
Synthesis Score								
Model One (AIC = -106)					2.27	143	.032*	.056
F	-1.52	.132	-4.429	-.123				
Exploratory chat	-2.50	.014*	-0.032	-.271				
Topic chat	1.44	.153	0.026	.129				
Synthesis chat	1.38	.170	0.035	.135				
Query vocabulary richness	-1.33	.184	-0.106	-.108				
TopicKnowledge	2.76	.007**	0.275	.225				
Partner agreement	1.12	.265	0.044	.092				
Model Two (AIC = -107)					2.49	146	.025*	.056
F	-1.31	.193	-3.762	-.104				
Exploratory chat	-2.43	.016*	-0.031	-.261				
Topic chat	1.51	.132	0.026	.135				
Synthesis chat	1.25	.214	0.031	.120				
TopicKnowledge	2.94	.004**	0.289	.238				
Partner agreement	1.13	.262	0.044	.092				

Two models were investigated for the CIS synthesis scores, both of which are significant at the .05 level, explaining a small amount of the score variance (r^{2adj} = .06). A significant positive effect of topic knowledge (*p* = .01) and a negative effect of exploratory chat (*p* = .01) are seen, with a larger beta for topic knowledge (indicating a larger effect size). That topic knowledge had an effect may suggest that those with higher topic knowledge were better able to spot inter-textual ties and condense these into a synthesised output. The negative effect of exploratory chat is surprising, as the use of talk to build common knowledge would be expected to have an association with synthesis. As indicated in section 7:1.1.1.1.1, this effect may relate to the various ways in which terms related to the typology used may be used – some of which are more procedural and may relate to task-confusion, while others are more closely associated with the kinds of co-construction of interest in this work. In addition some non-significant relationships may be identified for investigation in further work, namely a positive association to topic chat (*p* = .13 & *p* = .15), synthesis chat (*p* = .21 & *p* = .17) and partner agreement (*p* = .26 & *p* = 27) perhaps indicating that the sharing of specific claims (topic chat) and their co-ordination (synthesis chat) to build agreement is related to synthesising in the written outputs. Negative relationships to f (*p* = .19 & *p* = .13) and query vocabulary richness (in one model, *p* = .18) may indicate a relationship between poorer ability to synthesise found information and behaviours which might indicate less

effective use of information (use of larger number of terms in queries, unfocussed use of a large number of documents retrieved).

7:1.1.4.1.3 Source diversity score

Table 7:14 - – Multiple Linear Regressions for CIS Source Diversity Scores from Stepwise Process

	t	p	B	β	F	df	p	adj. R^2
Source Diversity Score								
Model One (AIC = -101)					11.2	144	.000***	.290
ISEQ General	-0.46	.647	-0.046	-.033				
ISEQ Justification	-1.10	.275	-0.111	-.080				
F	3.67	.000***	11.221	.264				
Nchar	5.35	.000***	0.726	.389				
Exploratory Chat	-2.20	.030*	-0.030	-.215				
ChatTaskTotal	2.47	.015*	0.010	.240				
Model Two (AIC = -101)					13.5	145	.000***	.294
ISEQ Justification	-1.01	.315	-0.097	-.070				
F	3.69	.000***	11.237	.265				
Nchar	5.35	.000***	0.719	.385				
Exploratory Chat	-2.16	.032*	-0.029	-.209				
ChatTaskTotal	2.45	.015	0.010	.237				
Model Three (AIC = -102)					17.3	148	.000***	.300
F	3.74	.000***	11.291	.265				
Nchar	5.44	.000***	0.723	.388				
Exploratory Chat	-2.13	.035*	-.028	-.200				
ChatTaskTotal	2.56	.012	0.010	.240				

Three models were investigated for the CIS source diversity scores, all of which are significant at the .05 level , explaining a moderate amount of variance in the scores (r^{2adj} = .29-.30). Across the models, f, the number of characters, and the total chat levels were positively predictive of source diversity scores with exploratory chat negatively associated (all significant at the .05 level). As noted above, with regard to the exploratory chat this may indicate the complexity of term-based typologies, such that the terms capture many different kinds of chat. Other variables indicate that visiting and making use of pages, while writing longer texts and engaging in more chat with collaborators is related to higher scores. In the upper models (which have slightly higher AIC, implying slightly less support), the ISEQ factors are included with negative effects for both (p > .05). While the p values for these effects are not significant, these variables are theoretically grounded and should be further investigated, their inclusion may imply that participants with less sophisticated perspectives on justification of knowing (for example, establishing source authority

and corroborating), and the general ISEQ factor, draw on and explicitly cite fewer sources in their written outputs.

7:1.1.4.1.4 Source quality score

Table 7:15 – Multiple Linear Regressions for CIS Source Quality Scores from Stepwise Process

	t	p	B	β	F	df	p	adj. R^2
Source Quality Score								
Model One (AIC = -87)					5.77	130	.000***	.236
ISEQ General	-1.32	.189	-0.143	-.102				
Nchar	5.32	.000***	0.783	.413				
Exploratory Chat	-2.56	.012*	-0.032	-.220				
Source Quality Chat	0.71	.479	0.022	.058				
Topic Chat	1.26	.211	0.027	.114				
Search Experience	2.06	.041*	0.215	.159				
Collaborative Satisfaction	-1.23	.223	-0.050	-.102				
Topic Knowledge	1.93	.056*	.214	.149				
PartnerAgreement	-1.11	.270	-.054	-.091				
Model Two (AIC = -88.3)					6.32	131	.000***	.235
ISEQ General	-1.43	.155	-.155	-.111				
Nchar	5.46	.000***	.800	.422				
Exploratory Chat	-2.79	.006**	-0.034	-.237				
Source Quality Chat	0.72	.472	0.022	.059				
Topic Chat	1.35	.181	0.029	.122				
Search Experience	1.98	.050*	0.207	.153				
Collaborative Satisfaction	-1.72	.089^	-0.066	-.134				
Topic Knowledge	1.99	.049*	0.220	.154				
Model Three (AIC = -88.8)					7.18		.000***	.237
ISEQ General	-1.36	.176	-0.146	-.104				
Nchar	5.48	.000***	0.802	.423				
Exploratory Chat	-2.77	.006**	-0.034	-.234				
Topic Chat	1.67	.098^	0.033	.143				
Search Experience	2.05	.042*	0.213	.157				
Collaborative Satisfaction	-1.76	.080^	-0.067	-.137				
Topic Knowledge	1.98	.050*	0.219	.153				

Three models were investigated for the CIS source quality scores, all of which are significant at the .001 level, explaining a moderate amount variance in the scores (r^{2adj} = .24). Across the models there is again a negative effect of exploratory chat (see discussion above), and a positive effect of text length, search experience and topic knowledge (all significant at the .05 level). Again, this indicates that the largest effect is for text length, such that those who wrote more, engaged in more evaluative writing. We also see that topic knowledge has a positive effect such that those with higher topic knowledge score higher, indicating they bring their knowledge to bear on the sources and information found; it is interesting that this effect is seen on source quality, but not topic coverage scores. Search experience also has a positive effect indicating that those with

higher self-report search experience also engage in more evaluation, perhaps because they find

more diverse results, or because their search experience is related to their ability to effectively

critique and weigh up sources. Small but not significant (at the .05 level) positive effects are also

seen for topic chat and source quality chat, indicating that groups who talked more about the

topic content, and the sources from which it was drawn, engaged in more source quality

evaluation and scored higher on the written outputs. This association is important given it is

strongly related to the theorised model presented in section 8. We also see small negative effects

(not significant at the .05 level) for the ISEQ general factor, partner agreement and collaborative

satisfaction. The ISEQ association is important given the expected association between ISEQ

scores and outcome, and indicates that those with less sophisticated epistemic perspectives

scored lower in the written outputs. The small negative effects of partner agreement and

collaborative satisfaction may be indicative of acquiescence in source evaluation from some

groups with higher levels of agreement and satisfaction (such that no disagreements regarding

source quality were discussed); future work should further investigate this small effect.

7:1.1.4.1.5   Total score

Table 7:16– Multiple Linear Regressions for CIS Total Scores from Stepwise Process

	t	p	B	β	F	df	p	adj. R^2
Total Score								
Model One (AIC = 165)					10.8	144	.000***	.282
ISEQ General	0.03	.974	0.009	.002				
F	1.65	.102	13.973	.120				
Nchar	6.23	.000***	2.366	.461				
Exploratory Chat	-2.79	.006**	-0.107	-.277				
ChatTaskTotal	2.04	.043	0.023	.209				
Topic Chat	0.25	.801	0.013	.021				
Model Two (AIC = 163)					12.8	147	.000***	.279
F	1.54	.126	13.103	.112				
Nchar	6.35	.000***	2.40	.467				
Exploratory Chat	-2.47	.015*	-0.093	-.238				
ChatTaskTotal	1.71	.089^	0.019	.173				
Topic Chat	0.12	.907	0.006	.010				
Model Three (AIC = 163)					16.0	148	.000***	.284
F	1.57	.119	13.221	.113				
Nchar	6.45	.000***	2.393	.465				
Exploratory Chat	-2.48	.014*	-0.092	-.237				
ChatTaskTotal	1.86	.064^	0.020	.177				

Three models were investigated for the CIS total scores, both of which were significant at the .05

level, explaining a moderate amount variance in the scores (r^{2adj} = .28). Across models there is

significant positive effect of number of characters and negative effect of exploratory talk. In addition, total chat is significant or approaches significance ($p < .10$) across models indicating that participants who engaged in more chat scored higher. We also see non-significant effects of f, the ISEQ general factor and topic chat; as noted above, these effects should be further investigated, and in particular their contribution to individual facets (of which the total score is composed) should be identified through further exploratory work. However, particularly in the case of topic chat and the ISEQ general factor, given the very high p values (>.8) and small effects these variables are unlikely to yield direct predictive value for outcomes.

7:1.1.4.2   MDP Outcome Regressions
7:1.1.4.2.1   Topic score

Table 7:17 – Multiple Linear Regressions for MDP Topic Scores from Stepwise Process

	t	p	B	β	F	df	p	adj. R^2
Topic Score								
Model One (AIC = -95.6)					11.8	120	.000***	.258
ISEQ Justification	0.15	.880	0.016	.012				
F	3.50	.001***	1.147	.294				
Nchar	3.66	.000***	0.462	.307				
ChatTaskTotal	-1.98	.049*	-0.005	-.156				

One model was investigated for the MDP topic scores, significant at the .001 level, and explaining a moderate amount of the variance in scores (r^{2adj} = .26). Within this model there was a significant positive effect of the number of characters (indicating those who wrote longer texts included more topics) and f (indicating that groups who used more pages included more topics). In addition, there was a small negative effect of total chat, indicating that chat had a small negative effect on scores. Finally, there was a non-significant (p = .88) positive effect of the ISEQ justification factor such that less sophisticated perspectives were associated with higher scores; given the very high p value (p = .88) it is likely that this variable is not predictive of topic score.

7:1.1.4.2.2   Synth score

Table 7:18 – Multiple Linear Regressions for MDP Synthesis Scores from Stepwise Process

	t	p	B	β	F	df	p	adj. R^2
Synthesis Score								
Model One (AIC = -50.7)					2.33	113	.037*	.063
ISEQ General	-1.22	.22	-0.154	-.110				
Contribution symmetry	1.27	.21	0.422	.115				
Nchar	-1.54	.13	-0.226	-.143				
Source Quality Chat	1.98	.05*	0.035	.181				
Topic Knowledge	1.17	.25	0.136	.104				
Partner Agreement	1.45	.15	0.078	.129				
Model Two (AIC = -51.5)					1.96	114	.089^	.039
ISEQ General	-1.19	.235	-0.152	-.109				
Contribution symmetry	1.54	.126	0.511	.140				
Nchar	-1.96	.052*	-0.285	-.180				
Topic Knowledge	1.15	.253	0.135	.104				
Partner Agreement	1.35	.181	0.074	.122				
Model Three (AIC = -52.2)					2.12	115	.083^	.036

	t	p	B	β
ISEQ General	-1.20	.234	-0.153	-.109
Contribution symmetry	1.57	.119	0.523	.143
Nchar	-2.06	.042*	-0.298	-.189
Partner Agreement	1.22	.223	0.067	.110

Three models were investigated for the MDP synthesis scores, one of which was significant at the .05 level (with the remainder p = .09 and p = .08 respectively), and explaining a small amount of the variance in scores (r^{2adj} = .04-.06). There are no consistent significant effects across the models. Number of characters appears to be have a significant negative effect (at the .05 level in two models), indicating – as in the CIS case – that those who wrote longer texts synthesised the information less. The first model (which is significant overall) indicates that source quality chat also has a small positive effect (p = .05) indicating that those who talked more about the sources scored higher on the synthesis facet. In addition we see small positive (but non-significant) effects of partner agreement (indicating greater partner agreement was associated with better synthesis outcomes – perhaps due to more co-ordinated writing), contribution symmetry (perhaps indicating that groups in which one partner took the lead and did more writing synthesised better), and topic knowledge (perhaps indicating topic knowledge was used to identify and consolidate topics in a synthesised output). A small negative (non-significant) effect was observed for the ISEQ general factor, indicating that those with less sophisticated perspectives on knowledge synthesised less.

7:1.1.4.2.3 Source diversity score

Table 7:19 – Multiple Linear Regressions for MDP Source Diversity Scores from Stepwise Process

	t	p	B	β	F	df	p	adj. R^2
Source Diversity Score								
Model One (AIC = -160)					8.56	98	.000***	.304
ISEQ Justification	-1.02	.308	-.083	-.086				
F	1.69	.095^	0.419	.152				
Nchar	4.88	.000***	0.465	.436				
Exploratory Chat	-1.86	.066*	-0.014	-.160				
Collaborative Satisfaction	-1.31	.193	-0.044	-.122				
Addon Intuitiveness	1.07	.289	0.037	.099				
Model Two (AIC = -161)					11.1	105	.000***	.314
F	1.85	.067^	0.442	.159				
Nchar	5.47	.000***	.493	.470				
Exploratory Chat	-1.81	.073^	-.013	-.150				
Collaborative Satisfaction	-1.10	.275	-.030	-.097				
Addon Intuitiveness	0.90	.368	.030	.080				

Two models were investigated for the MDP source diversity scores, both of which were significant at the .001 level, explaining a moderate amount of the variance in scores (r^{2adj} = .30-.31). In both models there is a positive significant (at the .05 level) effect of number of characters, indicating (as in the CIS case) that those who wrote longer texts drew on and explicitly cited more sources. There are also non-significant positive effects of f (in the MDP case likely indicating that those who used more pages scored higher) and add-on intuitiveness (indicating that those who found

the addon intuitive scored higher). There were also negative effects for exploratory chat (which is discussed above), collaborative satisfaction (requiring further investigation) and (in one model) the ISEQ justification factor; the last of these is non-significant but of theoretical interest, and follow-up research should explore this relationship further.

7:1.1.4.2.4 Source quality score

Table 7:20 – Multiple Linear Regressions for MDP Source Quality Scores from Stepwise Process

	t	*p*	B	B	*F*	*df*	*p*	adj. R^2
Source Quality Score								
Model One (AIC = -47)					3.28	122	.014*	.068
Contribution Symmetry	1.19	.236	0.370	.103				
Source Quality Chat	2.32	.022*	0.041	.209				
Synthesis Chat	-0.78	.435	-0.022	-.070				
Partner Agreement	2.46	.015*	0.131	.212				
Model Two (AIC = -47.3)					4.18	123	.007**	.070
Contribution Symmetry	1.20	.233	0.373	.103				
Source Quality Chat	2.19	.030*	0.037	.189				
Partner Agreement	2.49	.014*	0.133	.214				
Model Three (AIC = -47.6)					3.76	124	.026*	.057
Contribution Symmetry	1.36	.178	0.426	.118				
Partner Agreement	2.37	.019*	0.128	.207				

Three models were investigated for the MDP source diversity scores, all of which were significant at the .05 level, explaining a small amount of the variance in scores (r^{2adj} = .04-.07). Across all and two models respectively there was a significant positive effect of partner agreement and source quality chat, indicating that those who talked more about the sources given to them scored higher, and perhaps that those who tried to build consensus (building partner agreement, requiring reconciling differences across sources) scored higher. A small non-significant effect was also seen of contribution symmetry, indicating that higher asymmetry in contribution was related to better outcomes, and a small non-significant negative effect of synthesis chat indicating that more chat about writing-coordination was associated with poorer scores. Further research should investigate these non-significant effects, which are not readily interpretable.

7:1.1.4.2.5 Total score

Table 7:21 – Multiple Linear Regressions for MDP Total Scores from Stepwise Process

	t	*p*	B	β	*F*	*df*	*p*	adj. R^2
Total Score								
Model One (AIC = 111)					2,55	129	.023*	.069
F	1.25	.212	1.190	.121				
Contribution Symmetry	1.44	.154	1.024	.125				
Nchar	2.00	.047*	0.700	.195				
Topic Knowledge	1.02	.309	0.269	.091				
Partner Agreement	1.20	.232	0.149	.106				
Partner Familiarity	-0.40	.693	-0.025	-.035				
Model Two (AIC = 111)					3.05	121	.013*	.075
F	1.25	.21	1.187	.121				
Contribution Symmetry	1.42	.16	1.009	.123				

Nchar	2.08	.04*	0.718	.199
Topic Knowledge	0.98	.33	0.256	.089
Partner Agreement	1.16	.25	0.143	.101

Two models were investigated for the MDP total scores, both of which were significant at the .05 level, explaining a small amount of the variance in scores (r^{2adj} = .07). Across models the only significant (at the .05 level) effect is for the number of characters, indicating (as in the CIS case) that those who wrote longer texts scored higher. We also see positive effects of f (in the MDP task most likely indicating groups who used more pages did better), contribution symmetry (indicating an unexpected effect of asymmetry related to better scores), topic knowledge (indicating those with higher topic knowledge scored higher) and partner agreement (indicating those who had consensus with their partner, perhaps because they reconciled differences across texts, scored higher). We also see a small, not readily interpretable, negative effect of partner familiarity.

7:1.2   Research Question 3: Discussion

7:1.2.1 Findings and Implications

Qualitative analysis exemplifies the epistemic nature of the two tasks. Throughout the chat data in both MDP and CIS tasks, clear and subtle epistemic concerns are expressed. The use of a typology distinguishing: exploratory; topic; source; and synthesis chat appears to capture, non-mutually-exclusive, thematic differences in the chat data. Trace data indicates differences between the groups in the ways they navigate to, browse, select information from, and write about, documents, with some within-group asymmetry being common in activity levels. It is thus clear from this research that behavioural markers of epistemic commitments may be observed in activity log data obtained from designed CIS and MDP tasks. In answering both the first and third research questions, clear differences are observed between the participant pairs, and these differences have limited relationships to the ISEQ data analysed in discussion of research question two – again pointing to the importance of behavioural trace in analysis of epistemic commitments.

In line with earlier research (Hargittai et al., 2010; Kobayashi, 2014) participants did not make extensive reference to source features in their chat, nor did they extensively evaluate source features in the written outputs (as indicated by the distribution of scores on the source quality/evaluation rubric facet). Contrary to expectations, based on prior research on productive

educational dialogue and claims regarding the need for meta-discourse in productive literacy (see section 2:4.4), terms related to exploratory dialogue seem to be negatively associated with various outcome measures. This may be because those terms were in fact capturing a broader set of activities including task-negotiation and explanation, rather than co-construction of knowledge, as indicated in the exemplifications in section 7:1.1.1.1.1. Further research is required to investigate this issue, and potential means to go beyond term-based methods for distinguishing various kinds of productive learning dialogue from each other, and less productive dialogue.

Earlier research (Bråten et al., 2005) indicating a relationship between ISEQ scores and self-report internet-learning behaviours is supported by the small negative correlation between the ISEQ justification factor and self-report search experience for both the CIS and MDP task indicating an association between more sophisticated views on the justification of knowledge (lower scores) and self-reported search experience. Furthermore, a reported association (Strømsø & Bråten, 2010) between lower (more sophisticated) ISEQ scores and help-seeking behaviour may be supported by the increase in topic-chat in CIS but not MDP tasks. Finally, small associations in the MDP task between the ISEQ justification factor and viewing pages may provide further evidence for: Strømsø and Bråten's (2010) finding that students with beliefs in fact checking and reasoning engage in more self-regulatory strategies; Kammerer et al.'s (2013) finding that students without such beliefs tend towards one-sided representations; and Salmerón and Kammerer's (2012) finding that students with more sophisticated beliefs select a more diverse array of results in search tasks. However, it is surprising that there were few other relationships between ISEQ scores and (on an individual or collective level) trace data, or outcomes. Indeed, in the associations indicated above, the relationships held only for one task rather than both. Moreover, other relationships were not observed in any data. For example, no relationships between the ISEQ and source evaluation (in the chat or written outputs) was identified – contra Kammerer et al., (2013), nor were there readily identifiable associations to comprehension of conflicting sources such as synthesis and source evaluation chat or outcomes, contra Barzilai and Eshet-Alkalai (2015). This indicates the need for further exploration of these relationships and

conceptual work to better understand relationships between information seeking and epistemic commitments in collaborative contexts, and 'messy' information seeking environments such as those used in this research.

Analysis of correlations indicated a number of small relationships between facets of trace data, and between trace data and survey data. Of particular interest are that in the CIS group touch points were positively associated with synthesis, while in the MDP task touch points were associated positively with topic, exploratory and total chat (perhaps indicating coordinating chat not captured by 'synthesis'). We also see that in the CIS task querying behaviours (number of queries, vocabulary and depth) were associated with both number of pages viewed, and the amount of chat, suggesting that groups who chatted more worked to explore more queries and pages, or/and that more exploration was associated with more chat, perhaps supportive of Lazonder's (2005) claim that collaboration aids in overcoming a 'vocabulary problem' in information seeking activities.

Stepwise regression models should be used as an exploratory indicator for further research. Across the regression models, the number of characters in the written output is the only relatively consistent significant predictor of outcome score, indicating that those who wrote longer texts scored higher (a not uncommon finding in analysis of this crude predictor for essay grades, see for examples, Chodorow & Burstein, 2004; Mellor, 2011). Further work should investigate whether this is best modelled as a quadratic relationship, such that very long, or very short texts score lower. Text length is, in this case, likely a proxy for the level of detail, number of topics or themes, and number of sources included. It is worth noting that the number of characters was not significant in the CIS synthesis score models, and had a negative effect in the MDP models indicating a small improvement in synthesis score for shorter texts (perhaps indicating consolidation of information from across the large number of pre-assigned documents).

The amount of variance explained for each facet was low to moderate. In both CIS and MDP tasks topic coverage (~26-28% variance explained) and source diversity (~29-31% variance explained)

were well explained. However, synthesis was relatively poorly explained in both (~4-6% variance explained), with source quality and the total score better explained in the CIS task (~24 and 28% variance explained respectively), than the MDP task (~8% variance explained in both). These figures indicate the strong potential of analysis of trace data arising from information seeking tasks in explaining variance in learning outcomes, such that outcomes may be modelled based on behavioural trace data. Across tasks, there were some consistencies in predictor variables for outcome regression models. Specifically:

- Topic score models included the amount of chat, and number of characters, indicating small positive effects for both predictors such that writing longer texts and chatting more to one's partner is related to better outcomes
- Synthesis score models included partner agreement (small positive effect), and topic knowledge (small negative effect), such that higher levels of agreement with lower (prior) topic knowledge are associated with better outcomes
- Source diversity models included the ISEQ justification factor (small negative effect), f, number of characters, and exploratory talk (small negative effect), such that more sophisticated ISEQ justification perspectives, the viewing and use of a range of resources, writing longer texts and fewer exploratory terms are associated with better outcomes
- Source quality models included source quality chat (small positive effect), such that those who used more source terms had better outcomes
- Total score models included f, number of characters, and total chat, such that the viewing and use of a range of resources, writing longer texts, and engaging in more chat were associated with better outcomes

Note that these variables are not the only features, nor the ones with the largest effect sizes, in the models, but they are consistent across task. Moreover, in some cases the model does not indicate that they are significant indicators, although model selection retains them; given the small effect sizes, and lack of significance these results should be treated with caution, however their consistent presence indicates potential for further investigation. Such investigation could involve follow-up research focussing on these particular indicators, perhaps including experimental manipulation to distinguish different types of behaviours, and the triangulation of summary-metrics and a closer qualitative analysis of the processes and practices engaged in.

This section has focussed on trace data obtained in a set of single-session, single-task classes during a single-week in Maastricht. As such, the trace is highly specific to the particular context – a context which has been described in detail in the Literature Review and Methods chapters of

this thesis. However, it is important to recognise the specificity of the findings, and the particular 'messiness' of tasks designed as authentic, collaborative, problem solving performance exercises. It is also important to note that while reasonable levels of variance are explained, substantial variance remains unexplained, particularly with regard to synthesis models (CIS $r^{2 \, adj}$ = .06; MDP r^{2} ^{adj} = .04-.06) and MDP source quality model ($r^{2 \, adj}$ = .04-.07), and total score ($r^{2 \, adj}$ = .07-.08). Given the more controlled nature of the MDP task, with documents across which sub-topics, source qualities, and inter-textual ties (and disagreements) are known, this difference deserves further investigation, and is especially amenable to experimental manipulation. Results across both tasks evidence the complexity and challenge of design of robust behavioural indicators – such as those used in this research, which were drawn from established literature – that may be associated with learning outcomes. This points to a challenge for learning analytics research generally; further work is needed to identify appropriate tasks to elicit behaviours to be identified through key indicators which may then be associated with conceptually grounded outcome variables which are of an appropriate granularity.

It is, however, interesting to note an effect (albeit small) of source chat on source quality scores. This variable is uniquely present in this facet's models, and has clear theoretical grounding. Also notable is a small negative effect (which is not significant at the .05 level) of the ISEQ general factor on source quality scores in the CIS task, a theoretically grounded effect. Similarly, we see a small positive effect (not significant at the .05 level) of synthesis chat on synthesis scores in the CIS task but not the MDP task. Search experience has a small positive effect on source quality in the CIS task, indicating that those with higher self-reported search skill evaluated sources better. In the MDP task partner agreement has a small positive effect on source quality scores, perhaps because these partners reconciled differences between the source-texts through evaluation processes, which led to greater partner agreement and evaluation in the outputs. There is a small negative effect of topic knowledge in the topic score model for CIS; this implies that, unexpectedly, those who had higher topic knowledge performed worse, perhaps because they did not identify the sub-topics (about which they have lacked topic knowledge); however, topic

knowledge has a small positive effect on synthesis scores in the CIS task, suggesting a relationship between synthesis and topic knowledge.

7:1.2.2 Limitations

Indeed, it is also important to note that, while some individual differences including the ISEQ and topic knowledge were controlled for in this research, other differences were not. This is important given, for example, the role cultural background (such as nationality) may play in collaborative learning contexts. Similarly, there may be further sources of process data which are not adequately accounted for in the models presented. The etherpad data is one obvious source of data for further analysis. Deeper content analysis may facilitate development of further indicators, for example, further exploration of the type of website being visited and the attention given to them would enrich analysis in the CIS case. Other data sources should also be explored, or 'designed in' to the task or the browser addon.

It is important to highlight the potential risks of stepwise analysis for model selection. As described in the results section, various authors have raised concerns of over or under-fitting using such model selection, issues which are not necessarily compensated for by the use of researcher discretion and application of theory in selecting models.

Of course this is not to say that one would expect clear linear relationships between a variable's observed frequency and outcomes, and indeed this linear approach – as taken in this work – should be questioned. Mere presence of markers is likely a shallow indicator, and moreover associations between outcomes and *more* activity of type 'x' may not mean that doing more 'x' is a good thing. For example, some behaviour indicators (such as chat quantities) may be indicators of general engagement rather than any more nuanced epistemic indicator – drawing out these nuances is a concern for all learning analytics work, to which this thesis contributes an exploratory analysis.

Indeed, this work also highlights the complexities of making use of the rich but 'messy' behavioural trace data from large-scale innovative software tools (Coagmento in this research).

For example, in this work (as noted in footnote 31) analysis of the chat data indicated a small incidence of non-English messages. Observations within the study sessions also indicated that on occasion the etherpad collaborative text editor was being used to communicate between partners, data which is not analysed. In other instances, as has been noted, there is data loss related to software issues. Across these points, it is important to note that the use of innovative software facilitative of collaborative learning behaviours is an important part of the development of models of learning processes, but raises challenges for analytics in terms of data loss and cleaning that go beyond providing a stable tool to learners (which for most participants in this study, was effectively achieved).

7:1.2.3 Future research

There are a plethora of other possible data sources to draw on, or trace indicators to develop from the data available in tasks such as those used in this research. The present work has made use of a core set of indicators from prior work (as discussed in section 2:5.3 and 3:4.3.1.1) alongside some theoretically motivated analysis of chat data, in a novel theoretically grounded set of tasks. Developing work should include analysis of temporal components of information seeking, include duration and sequence analysis. For example, further work could analyse pageview durations as a measure of engagement (controlling for text length and difficulty, see, Smucker & Clarke, 2012).

In addition, the etherpad data provides a rich source of data which is presently providing only limited indicators (through contribution symmetry and touch points); existing work provides a grounding for development of indicators in this research. Relatedly, the version of Coagmento used in this research tracked copy actions, but not paste actions (which the most recent version additionally does track); further research could explore both and in particular operationalise 'use' such that pasting of the information be required in addition to copying. The combination of these approaches might particularly aid in further identification of synthesis and other behaviours as described in relation to Goldman et al., (2012) and Hastings et al., (2012) in chapter 6 above such

that sections of text in a written output could be associated with text that had been copied and pasted.

Although indicators were developed to assess symmetry in participant activity, these are taken as individual measures rather than in aggregate or overview. This allows analysis of individual measures, and it is unclear how measures could be scaled for aggregation (is adding an extra word the same as viewing an extra page?), however, some insight into division of labour is lost through this analysis. Because the current measure lacks direction (the absolute value is used), the directionality of partner effort is lost such that the measure is opaque regarding whether an asymmetry across multiple measures is because a single partner did more on all measures, or the pairs split activities each focussing on particular tasks. Future research could develop methods to understand not whether individual tasks were divided evenly, but whether individual tasks were divided among the participants such that, for example, one participant might undertake more "searching and viewing" while the other might undertake more "writing and analysis", while in other pairs we might see both participants engaging in the same behaviour.

Chapter 8:   GENERAL CONCLUSION AND DISCUSSION

8:1 General Discussion and Conclusion

As outlined in the introductory sections to this thesis, learning analytics research has had an increasing focus on connecting analytic models to the learning sciences. In that section (building on published work, S. Knight, Buckingham Shum, et al., 2014) I argued that learning analytics should be framed in terms of epistemology, pedagogy and assessment, indicating that analytic devices (as assessments) implied particular modes of thinking about knowledge (epistemology), and were fundamentally associated with pedagogic strategies. This thesis work thus sought to develop part of a theoretically grounded exemplification of a learning analytic approach, considering a pedagogically motivated construct (epistemic cognition) related to literacy particularly in rich multimedia environments such as those encountered in mature internet use. I presented a particular epistemological position for my own work, and have outlined a novel – socialised – account of epistemic cognition in the context of information seeking tasks that builds on this epistemological stance, and draws out its learning implications (for social, dialogue mediated, learning).

The epistemic cognition field has, typically, tended towards smaller studies (single-class size, typically around 30 students) using self-report methods (psychometrics and interviews), and almost exclusively individually based tasks. However, as I argued in the literature review, collaboration in information seeking tasks does occur even unprompted in information seeking, and has pedagogic potential in addition to offering access to a form of trace data – language in use – for analysis of epistemic *commitments.* Alongside analysis of participant's tool mediated interaction (through which trace was gathered) in their information seeking, this stance on information seeking provided the grounding for my approach to epistemic commitments as the target of learning analytics. In developing this approach, a rich set of learning, information, and computer sciences, literatures and their theoretical positions were drawn together to develop a theoretically grounded, interdisciplinary approach.

8:1.1   General Discussion

This thesis thus set out to investigate patterns of information seeking as epistemic processes to inform the development of learning analytics in these complex literacy-based practices, using a pedagogically grounded shared model of CIS and MDP to do this.

 The empirical work has demonstrated that:

1. Textual epistemic features can be identified reliably (by experts), although work is needed to validate and develop systems for students to peer and self-assess effectively
2. The ISEQ psychometric instrument has some small significant relationships to trace markers, self-report search experience, and (in CIS) general ratings of resource trustworthiness, but is largely not related to outcome in a CIS/MDP context
3. Trace data can give insight into epistemic commitments in a collaborative online information seeking task, and that these findings – in a novel authentic task, making use of trace analysis – can be aligned with prior research findings[33]

These findings demonstrate that epistemic properties of texts, associated with literacy-based concerns, can be reliably identified in written texts – student's topic coverage, evaluation, source diversity, and information synthesis. Peer and self-assessment results point to challenges in this area, yet there is clear potential for the development of tasks to bring students into their own assessment, at scale, using models such as those developed in this work. The work indicates that self-report measures – which are relatively simple to deploy and assess – have some clear associations to some trace-measures, and student assessments of resource 'trustworthiness', but are not clearly related to outcome, or a wide range of trace markers. This highlights the benefit of analysis of behaviour, and cognition 'in action', through interaction with tools and collaborative dialogue engaged in. Analysis of the trace indicates its potential over and above self-report measures; the trace (and in particular the dialogue) shows evident epistemic markers, and quantitative analysis indicates clear differences in groups' behaviours, with predictive relationships identified suggesting potential for models of outcome predicted by behavioural markers.

---

[33] For some key alignments with prior work in tabulated form see Appendix 19

The research, in its conceptual and empirical explication, has addressed an analytic model that maps outcomes, trace indicators and behaviours, and theorised constructs. As described in the introductory sections, this approach asks, "What are students supposed to do when they study multiple documents? And what kind of mental representation of such materials do they form?" (Rouet, 2006, p. 65), or moreover in my context, what kinds of dialogue do they engage in, how do they interact with each other and their tools when they seek information, and how do these behaviours relate to outcomes. This thesis research has provided an exemplification of potential productivity in the 'middle space' at the confluence of learning science and analytic techniques, with a particular focus on theoretically grounded tasks, well theorised psychological constructs, and analytic approaches drawn from established literatures. Commonalities and alignments were drawn from a rich set of literatures, and this thesis work serves as an exemplification of the potential of learning analytic work at the confluence of learning, information, and computer sciences.

The research has been grounded in a broad pragmatic-sociocultural approach (see section 1:1.4), analysing a learning context in which trace markers in participant's tool mediated information-seeking (including their dialogue) were identified to develop a model associating epistemic commitments to outcome. This analysis was conducted based on an authentic epistemic task, involving the use of a peer and self-assessment, to bring students into their own assessment. Aligned with the pragmatic-sociocultural stance advanced in the introductory sections, this research, building on prior epistemic cognition and information seeking work, is original in: (1) being behaviourally oriented, (2) to authentic information problems in (3) social contexts, (4) connected to an outcome measure including peer-assessment.

The work has aimed to foreground contextual features – not only in terms of understanding individual trace-data-points, but in understanding the wider task and student context. As indicated in the discussion of each research question, even with theoretically grounded and contextualised tasks, interpreting trace indicators is a challenge. Samples of data – particularly

chat – are given as exemplifications, however due to the scale of the research, and the nature of the log data, close contextual-analysis of this kind of trace data is challenging. Although qualitative analysis of the chat data indicates clear epistemic markers (see section 7:1.1.1), in some cases in both the chat and other trace, these are not explicit enough for automated or semi-automated analysis. In addition, although a theoretically grounded account of the participant's context and the tasks in which they are engaged is given, the use of (generally) summary statistics and 'code and count' strategies loses some of the rich context present in temporal analyses of data (Suthers, 2006). Nonetheless, there are clear observable relationships between some elements of trace, survey, and outcome data. Moreover, the tasks have clear epistemic features, and as demonstrated through close analysis of small numbers of individual group's trace data, there are observed differences in approaches to addressing the task-problem; these differences can be characterised as epistemic commitments, suggesting a site for research into epistemic strategies across contexts (of temporality, domain, collaborative-setting and so on). In addition, the results indicate differences between the two tasks (CIS and MDP), in behavioural indicators related to task outcome, indicating that while both tasks present a similar question, and share outcomes on the rubric, they prompt varying behaviours in students; a further site for potential research is thus in exploring differences between CIS and MDP contexts. On a fine grained level, stark differences within and between pairs can be identified; the aggregation of this data for behavioural traces occludes some of the complexity, despite which predictive models can be identified; nevertheless, these models are context dependent to the task desk, and close analysis of the trace data highlights the need for highly contextualised models of epistemic commitments, in the context of such specified tasks. The development of these tasks highlights the potential of information science research into information seeking (and CIS) to draw on the learning sciences in their analysis of collaboration, and literacy practices (such as claim selection, evaluation, and synthesis) aligned with psychological constructs (such as epistemic cognition).

The complexities of the design alongside the use of between rather than within subjects design, raises challenges for distinguishing key differences between the tasks. However, future work

should continue the exploratory investigations given here to understand alignments between MDP and CIS models of information seeking and literacy. Rich, naturalistic tasks were used in this research – in the CIS task, involving information seeking across unknown documents, and in the MDP task, providing a range of documents of varying qualities across which there was clear imperative to prioritise reading, a known set of sub-topics, and a set of inter-textual ties (and disagreements); that behavioural indicators can be observed related to outcome in such rich contexts is an important finding, more controlled tasks may seek to increase this explained variance (but at the loss of naturalistic context). Future research should further analyse these trace differences, and explore the foregrounding of between and within-group activity, for example through student dashboards to feedback their activity indicators for self-reflection both subsequent to, and integrated into the tool-design and process of the tasks (see Bateman, Teevan, & White, 2012 in the context of search dashboards; and Duval, 2011 for a discussion of visualisation in learning analytic tools). The research has also pointed to the potential of diagnostic assessment and peer and self-assessment; although in this research reliability was low for the student participants, with reliable expert ratings. Both visualisations for reflection, and peer and self-assessment have potential to bring students into their assessment practices, aligning well with the sociocultural nature of the work.

As noted in the discussion of research question 3, further conceptual work, and measurement work is required to develop models of epistemically salient indicators in designed information seeking tasks. A collaborative understanding of epistemic commitments introduces conceptual and empirical challenges, especially the interaction between individual and collaborative features is to be accounted for; this work points to the need for further research in the area. The use of a collaborative context in this research was grounded in: a theoretical understanding of the importance of dialogue for learning; the methodological advantages of analysis of collaboration as a site for learning; and the empirical evidence indicating the potential of collaboration as a site for, epistemically salient, information-seeking-oriented dialogue – a potential which is currently underexplored. This theoretical and empirical grounding is supported by the empirical evidence in

this thesis, indicating that participants do indeed engage in dialogue around information seeking, that this dialogue is epistemically salient, and that some forms of dialogue have associations with learning outcomes.

Pointing to the potential of tasks and innovative technologies such as those used in this thesis (and the interdisciplinary work from which they emerge) is the participant written feedback, which was frequently positive in nature (as indicated quantitatively in Table 4:3). Visual inspection of the written feedback of the participants focussed on in this thesis indicated a number of considerations that were varyingly stated positively or negatively by different participants, and a set of issues encountered particularly with the use of the browser addon. For example, in line with quantitative measures of task-satisfaction (Table 4:3), some participants expressed that they enjoyed the collaborative element of the task with some particularly noting the use of the etherpad element as beneficial, while others expressed a preference for working alone on such tasks. Similarly, others indicated that a topic of more direct interest (or relevance) to them and their studies would be preferred, while a smaller number explicitly noted enjoying the different topic and appreciating that it was a topic few people in the classes would know about. In both tasks some participants also flagged the time constraints imposed; however these are introduced as a design decision serving both a pragmatic ends (to keep the task within a single session) and an experimental one (to ensure prioritisation of activity was necessary).

Pointing to the challenges involved in development of analytic approaches with innovative tools were a number of comments regarding the chat element of the Coagmento tool, which participants noted had a 'lag' in sending and receiving messages, thus impacting on their working together. This may impact on the nature of the trace data available – although 'between groups' analysis still provides insight into how different groups interacted with each other and the tool. Indeed even in cases in which participants were generally positive about the use of a tool such as Coagmento, some flagged the chat as a particular problem, highlighting the difficulties of 'scaling up' with emerging technologies. One other user experience issue was reported with regard to the

'post-lab' assessment element software, however, despite this issue, a number of participants gave explicit positive comments with regard to the assessment element appreciative of the fact that they would be receiving feedback on their work in the lab-session; this again points to the potential of such assessment tasks in learning analytics research. In both these cases the technological issues appear to be related to the scale of concurrent use, rather than specific study-design concerns, indicating the need for multi-disciplinary teams in developing large-scale analytic tools to be deployed effectively.

In the case of task design, some participants indicated that the task instructions could have been more explicit. This was sometimes specifically noted in the context of the assessment tasks, which were the stage at which the assessment criteria were made explicit through the use of the rubric. As noted in the discussion of the assessment development, previous research indicates that use of a rubric at both the task-completion and task-assessment time is optimal, and indeed some participants indicated that they did in fact evaluate sources; one participant said:

> *Also we didn't know the criteria so we didn't take a look at sources, source evaluation and making it into a coherent piece of information. We didn't do this so we had really little point on it which is a pitty because it would have been easy points for us since we did use the spurces and in our head evaluated them.* (general feedback, UID 493).

However, in this thesis research a core interest was to explore between-group differences across the elements assessed by the rubric. Certainly exploring nuance in task design, instructions and assessment criteria is important. However, grounded in the extensive literature reviewed in earlier sections, the research reported in this thesis gave the instruction to provide the "best supported claims" and "explain why"; the research interest, then, is precisely in understanding how participants interpret such instructions, and what information they provide to support their claims (their epistemic commitments around claim selection, provision of sourcing information, and connection of inter and intra-textual ties). Following on from this initial work, further research should investigate the potential of varying instruction types on within and between-group differences in task behaviour.

Cutting across the participant feedback were three implementation issues related to pragmatic concerns. The first two of these are common in research contexts, and in part relate to the need to minimise disruption and keep experimental time to a maximum of a single session, while the last points to the challenges of large-scale adoption of innovative technology-based study design:

1. Participants had a short period to familiarise themselves with Coagmento, primarily through the instructions at the beginning, and the warmup task, with a self-report control item regarding how intuitive they found the browser addon
2. Participants had a short period to familiarise themselves with their partner, many of whom were not well known previously. Again the use of the warmup task provided a short period of familiarisation session, with self-report items controlling for their prior familiarity with the partner, and level of agreement with them
3. There was a level of data loss over the whole week, with a number of software frustrations (outlined above) experienced even within the subset of participants reported in this thesis. The general feedback and satisfaction survey items give insight into these issues. Although the major software concerns which resulted in data loss from whole sessions were addressed over the course of the lab-week, the chat appears to have been a consistent (although more minor) problem, and the assessment task had separate technical issues.

In the first two cases, alternative designs (e.g. multi-session tasks) and self-selected groups may alter the feedback, and lead to differing task behaviours. A longer project involving use of Coagmento with a sustained partner would provide an interesting site for future research. Moreover, given quantitative feedback regarding the browser addon was largely positive (Table 4:3), a longer period of familiarisation may offer limited gains. The third point is concerning, although as indicated in Table 4:5 and Table 4:6 it is reasonable to assume that the data loss has not resulted in a biased data-set here, and that the large sample size remaining is robust. This third point indicates the challenges of use of innovative technologies which are both research-projects in their own right, and being developed for use in research. It is important that learning analytics researchers make use of the emerging capabilities and affordances of such tools, while at the same time recognising that innovative approaches might lead to unexpected results.

It is also important to note the impact that use of tools, such as Coagmento or other CSCL technologies, have on learning behaviours. For example, Gagnière, et al., (2012) 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 and prompts for evaluation might have in supporting epistemic dialogue in CIS contexts, although in Gagnière et al.'s research the task design itself was more fact-finding oriented rather than 'exploratory' in nature. Moreover, the interest of the research presented in this thesis was in participant's relatively unprompted epistemic commitments. As such, specific prompting or structuring environments are of interest for future research, rather than a focus in this work. Despite this, the tools provided do mediate participants interaction in important – if less structured – ways; although CIS and CSCL research indicates the positive potential of awareness and dialogue tools (as discussed in section 2:2.5), some research indicates potential pitfalls with such systems. In addition, the ways groups are constructed is likely to impact on the nature and success of the collaboration, including by impacting on the type of dialogue 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). Thus, in CSCL environments, it should be noted that:

> *In spite of [the] 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). 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)

Interdisciplinary research in information seeking and CSCL is well placed to explore these issues, drawing on the work (and task design) presented in this thesis. Design and development of

structuring tools should thus explore the implications of prior work in this area for CIS and epistemic commitments. It is also important that we consider the types 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. Understanding the nuance in task design, tool use, and outcome measures as features of the research context is thus central to interpreting observed behaviours in the kinds of information seeking contexts of interest to this research.

Building on the findings of the research presented in this thesis, there are a number of targets for future work which address the overall concerns of the research (while the individual 'future research' sections under each research question are more targeted). These might be characterised by four general themes, each mediating potential interactions in different ways:

1. CSCL characterisation
2. Collaborative dynamics
3. Task design
4. Analytic approaches

As implied above, the nature of the CSCL environment – comprising elements of the task design, website structure, and Coagmento structures – may mediate both collaborative interaction, and other task-relevant behaviours. Future research should investigate the impact of CSCL design choices, for example, through the foregrounding of elements of the trace processes to participants, potentially facilitating sensemaking around their partner's work while they are working. The use of 'search trails' (visual maps of query-pageview behaviours) are being investigated (Capra, personal communication (2015), see also Shah, Capra, & Hansen, 2015) as a means of collaborative, or distributed sensemaking (Fisher, Counts, & Kittur, 2012) through which participants share not only the pages they have found, but also their sensemaking and organisation of the information across the pages.

In this latter work, Fisher et al., (2012) investigated how participants used other participant's problem-solving to aid them in addressing a problem; as such, these participants were 'distributed' and not engaged in explicit collaboration. This hints at another area for future research – investigation of the collaborative dynamics. In the research reported in this thesis the

intent was to allow participants to negotiate their own working, indeed engagement in such

processes was of interest both in terms of exploratory chat (indicative of making explicit one's

ideas, and listening to other's ideas), and coordination of writing (synthesis chat). The explicit

assigning of roles, manipulation of group-composition and size, or addition of instructions to

collaborate in more specific ways, may all impact not only on the collaborative dynamics, but also

on the task-outcomes.

Relatedly, manipulations of the task design would perhaps, alongside impacting on task-

completion behaviours, impact collaborative behaviour. Such manipulations might include the use

of more/less explicit task instructions particularly regarding success criteria for the outcome, or

expectations regarding desirable behaviours, perhaps including CSCL support or prompts for

'evaluative annotation' of resources. Of course, the varying of task-topic, or/and target-output is

also important for understanding the epistemic-context of differing designs, for example, asking

participants to write a policy recommendation on economic policy, rather than an evidence

briefing on a herbicide/food supplement. There is thus a range of possible output-targets with

implications across the literatures drawn on in information seeking, collaboration, and epistemic

cognition tasks. However, as noted earlier, this thesis work provides one of very few tasks

connecting information seeking, epistemic cognition, and task outcomes, with the range of

possible target-outputs and their impact on information seeking, collaboration, and epistemic

cognition underexplored.

Across these possible modifications to the research design, further development is required on

the analytic approaches. As highlighted under the individual research questions, approaching

individual and paired data in both the survey and trace is a challenge. There may be modifications

to the CSCL environment or task design which would assist in analysis – one of which will be

discussed further below. In addition, in this research the focus has been, building on prior work,

largely on summary data. However these kinds of 'code and count' approaches are problematic

(Reimann, 2009; Suthers, 2006) insofar as they gloss temporal considerations such as sequence,

accretion and knowledge building over time, durations of significant events, and 'interaction

trajectories' (Furberg & Ludvigsen, 2008) – the trajectory of interactions over a time period. This

issue of temporality is important and although oft overlooked in learning research (Littleton,

1999; Mercer, 2008), developing learning analytics for temporal analysis (see, for example, S.

Knight, Wise, Chen, & Cheng, 2015) provide ripe space for further interdisciplinary work here.

Furthermore, as noted with regard to research question 2, further work is required to understand

how different groups tackle information seeking and literacy problems. In the context of RQ1 the

concern was with whether clusters of students might be identified through analysis of writing

patterns, following Goldman et al.'s (2012) work. The addition of trace data during the writing

process in this research provides scope to extend Goldman et al.'s (2012) research, to explore not

only the sources inserted into a text, but which sources were opened, how they were talked

about, and other patterns in finding, processing and writing about resources. As such, further

work attempting to cluster the behaviour of participants may yield productive results.

In addition, as noted in the limitations above, analysis indicates a number of areas for alterations

in the current design. While the peer and self-assessment in this research yielded poor reliability,

the reliability of the expert assessment indicates that with further training, instructions, and

perhaps exemplars (such as those used in the diagnostic/training session) peer and self-

assessment may be effective (as discussed in sections 2:5.2, 3:4.4 and 5:1.2). A parallel

development to this component of the research would be to go beyond the extensive piloting

used, and more directly validate the assessment task, rubric, and the sample texts used in the

diagnostic process, which should be written both to reflect differences across the rubric, and to

align with real participant outputs. Validation of the rubric by subject-experts would also provide

scope to include content accuracy measures, and subject-specific source-evaluation accuracy

indicators in the rubric. This would facilitate a closer attention to the inclusion of 'contrasting

perspectives', where the current research uses implicit measures of perspective taking through

scoring based on source diversity, synthesis and evaluation. There is scope, following these

alterations, to also explore the ways in which epistemic factors relate to success on the assessment tasks (for example, indicating relationships between poor performance at identifying 'synthesis' in the diagnostic, and particular behaviours and outcomes in the main task).

As noted in the discussions above, this exploratory work may point to areas for more controlled work to improve models to understand the role of epistemic commitments in complex literacy tasks (such as the CIS and MDP task developed herein). Thus, one means through which to simplify analysis is through the use of more experimental design. The use of MDP tasks, where the range of topics and source qualities are known and can be controlled, provides considerable potential for such designs. However, as alluded to in the discussion of research question 2, in this research a challenge for interpretation was encountered in analysis of the MDP trustworthiness ratings and understanding interactions with a rich complex set of documents. In addition, while the work reported in this thesis provided a relatively naturalistic study in which various contextual features were known, this has resulted in uncertainty with regards to understanding of task requirements. Of course, between-group differences in their treatment of the task is precisely the concern of interest, however as noted above further research investigating the use of more explicit instructions offers potential for additional insight. For example, an experimental study could be designed which asked participants to, in stages:

a. Read a set of documents provided by the researcher for an assigned task
b. select their 'top' documents, giving insight into the number selected by each participant, and their coverage of topics and source-types
c. rank the selected documents (rather than rating them), giving a clear insight into their perceived relative value and evaluation of the sources.
d. give an explanation for their rankings, giving further insight into source evaluation
e. summarise the documents, giving insight into their synthesis and ability to give multiple perspectives (identification of inter and intra textual ties)

The use of a 'staged' design such as this facilitates a clearer analysis by restricting the points at which particular behaviours are solicited, and thus analysed at. Thus, the skills of corroboration, sourcing and integration might be separately 'staged' for different analytic periods. Indeed such a design provides opportunity for insight into both collaborative and individual insights, with potential to draw more on explicit outputs (solicited explanations and rankings), than more

nuanced trace markers, at the cost of a less naturalistic task. To such a structure, scaffolding or scripting techniques to prompt feedback or particular types of activity might also be added, to further investigate changes in epistemic commitments. In focussing on tool mediated interaction (with information resources, and collaborators), this – and less constrained contexts such as that in this thesis work – offer a novel approach to epistemic cognition, focussing on social, behavioural and context-sensitive models of cognition 'in action'.

In broader terms, there is potential in approaching analytic problems through analysis of learning contexts, and the development of behaviour indices (even based on existing metrics) aligned with pedagogically motivated task design. Understanding the mediating roles of tools in this process – as an interactional tool, data tracker (through log-files), and element of the task design – is important, as is developing understanding of the social environment in which learning occurs.

8:1.2   Concluding Remarks: The Epistemic Role of Learning Analytics

This work has argued that an important component of literacy lies in understanding how students deal with information, how they conceptualise information problems and seek to address these problems using various tools – including search engines. This understanding provides space for a research paradigm which analyses the in-action decisions students make around their search behaviour, in order to investigate their epistemic commitments under various tool-mediated (social) contexts. This thesis has thus investigated patterns of information seeking as epistemic processes. Empirical results have provided evidence that log data do distinguish different approaches to the information problems used in this research.

Distinct from earlier models of epistemic cognition and beliefs, this work has foregrounded a particular stance on epistemic commitments, as behavioural tool-mediated, social ways of treating knowledge. This stance was aligned with the epistemological position described in the introductory sections in the context of learning analytics, focussing on 'cognition in action', through behavioural trace and collaborative dialogue. 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. Analysis of the data collected in the

information tasks in this research, including the collaborative dialogue, indicates the need for

theorised work on a social account of epistemic cognition.

Recall the 2009-2015 PISA definition (noted on p.29) that: "Reading literacy is understanding,

using, reflecting on and engaging with written texts, in order to achieve one's goals, to develop

one's knowledge and potential, and to participate in society." (OECD, 2013, p. 9). This thesis has

engaged in exploratory work to meet calls for performance based assessment and evidence

centred design for literacy assessments (S. R. Goldman, Lawless, et al., 2012; Lawless et al., 2012).

The research has made use of authentic performance assessments, using "an assortment of

documents, such as tables, figures, graphs, newspaper reports and photographs" (Benjamin et al.,

2009, p. 3) and making use of the internet to deliver "richer versions of the tasks" (ibid). It has

done this, following evidence centered design (Mislevy et al., 2012), by (1) identifying a high-level

construct, namely epistemic cognition, (2) identifying behaviours indicative of particular means of

approaching epistemic tasks, through an exposition of epistemic commitments, (3) developed

theoretically grounded tasks (MDP and CIS) and task contexts (collaboration, and a selected-

technology) to elicit these salient indicators. In this novel large scale context, support has been

provided for prior research into epistemic cognition, indicating the value of the kind of task-design

approach, and model of epistemic commitments, taken in this thesis.

The work has thus sought to address the learning analytics frameworks and models (see Table 2),

and the learning analytics lifecycle (Clow, 2012), such that: selection of data is targeted at

particular aims or goals; this student-level data regarding the learners and their activities is then

collected; and aggregated or processed into metrics or indicators; which might then inform

interventions – a later stage which is not addressed in this work. The work has indicated that

literacy skills for 'mature internet use', of integration, sourcing, and corroboration, can be

identified in written outputs, with some relationships to trace indicators including chat. However,

further work is required to develop deeper understanding of the role of particular task and collaborative contexts in such epistemic processes, the ways in which trace-indicators inform our understanding of epistemic-commitments, and the ways in which to best support and intervene in student epistemic commitments.

Of course one consideration in the development and theoretical grounding of novel assessment models is their epistemological positioning. It is important here to note that of course search engines themselves are a socio-technic, epistemic tool (S. Knight, 2014). 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. Indeed, in the case of CSCL environments it is precisely the power of technological tools to support students in their reasoning, learning, and collective action which motivates 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.

Just as search engines are socio-technical systems, so too are learning analytics. As I argued in Knight, Buckingham Shum and Littleton (2014), 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 is not a new claim – new tools and means of interaction have been embedded in education throughout history from books to various writing

systems. Learning Analytics research should consider not only the analytic device (for example, numeric outputs) or the context of feedback and intervention into which they are deployed, but also the wider task design, curricula context, and model of assessment, epistemology and pedagogy (as exemplified by the Danish example of permitting internet access in exams). This thesis work has provided an exemplification of a kind of theorised task design to engage the 'middle space', for development of learning analytics within a particular (collaborative information seeking) environment, for an educationally significant learning science construct.

REFERENCES

Acosta, A. (2013, November 20). *The Role of Epistemic Cognition in Complex Collaborative Inquiry*

Curricula (Thesis). Retrieved from https://tspace.library.utoronto.ca/handle/1807/42610

Acquavella, J. F., Alexander, B. H., Mandel, J. S., Gustin, C., Baker, B., Chapman, P., & Bleeke, M.

(2004). Glyphosate biomonitoring for farmers and their families: results from the Farm

Family Exposure Study. *Environmental Health Perspectives*, *112*(3), 321–326.

http://doi.org/PMCID: PMC1241861

Adler, J. (2014). Epistemological Problems of Testimony. In E. N. Zalta (Ed.), *The Stanford*

Encyclopedia of Philosophy (Spring 2014). Retrieved from

http://plato.stanford.edu/archives/spr2014/entriesestimony-episprob/

Admiraal, W., Huisman, B., & Van de Ven, M. (2014). Self- and Peer Assessment in Massive Open

Online Courses. *International Journal of Higher Education*, *3*(3).

http://doi.org/10.5430/ijhe.v3n3p119

Ahola, S. (2009). Measurement Issues in Studying Personal Epistemology. *Psychology & Society*,

2(2), 184–191.

Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE*

Transactions on, *19*(6), 716–723. http://doi.org/10.1109/TAC.1974.1100705

Alexander, P. A. (2006). What Would Dewey Say? Channeling Dewey on the Issue of Specificity of

Epistemic Beliefs: A Response to Muis, Bendixen, and Haerle (2006). *Educational*

Psychology Review, *18*(1), 55–65. http://doi.org/10.1007/s10648-006-9002-7

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.

http://doi.org/10.1145/1520340.1520547

Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, *32*(2), 159–181. http://doi.org/10.1080/02602930600801928

Anmarkrud, Ø., Bråten, I., & Strømsø, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences*, *30*, 64–76. http://doi.org/10.1016/j.lindif.2013.01.007

ANSES. (2014). *Opinion of the French Agency for Food, Environmental and Occupational Health and Safety: on the risks associated with the presence of "red yeast rice" in food supplements* (No. Request No 2012-SA-0228). ANSES. Retrieved from https://www.anses.fr/fr/documents/NUT2012sa0228EN.pdf

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?

Axtell, G. (2000). *Knowledge, Belief, and Character: Readings in Contemporary Virtue Epistemology*. Rowman & Littlefield.

Baehr, J. (2011). *The Inquiring Mind*. Oxford University Press. Retrieved from http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199604074.001.0001/acprof-9780199604074

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

Balfour, S. P. (2013). Assessing writing in MOOCs: Automated essay scoring and calibrated peer review. *Research & Practice in Assessment*, *8*(1), 40–48.

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

Baron, J., Lynch, G., Maziarz, N., Bilkstein, P., Merceron, A., & Siemens, G. (2015). *LAK '15: Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. New York, NY, USA: ACM.

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., & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction*, *36*, 86–103. http://doi.org/10.1016/j.learninstruc.2014.12.003

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. http://doi.org/10.1080/07370008.2011.636495

Barzilai, S., & Zohar, A. (2014). Reconsidering Personal Epistemology as Metacognition: A Multifaceted Approach to the Analysis of Epistemic Thinking. *Educational Psychologist*, *49*(1), 13–35.

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 Information Review*, *13*(5), 407–424. http://doi.org/10.1108/eb024320

Beetham, H., McGill, L., & Littlejohn, A. H. (2009). Thriving in the 21st century: Learning literacies for the digital age (LLiDA project). JISC. Retrieved from http://elearning.jiscinvolve.org/wp/2009/06/11/thriving-in-the-21st-century-learning-literacies-for-the-digital-age/

Benjamin, R., Chun, M., Hardison, C., Hong, E., Jackson, C., Kugelmass, H., … Shavelson, R. (2009). Collegiate Learning Assessment. Retrieved from http://www.plu.edu/assessment/widgets/documents-forms/items/Returning%20to%20Learning%20Document.pdf

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. http://doi.org/10.1007/s11092-008-9068-5

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. http://doi.org/10.1177/0267323102017001607

Boseley, S. (2014, August 2). British Medical Journal acted correctly in controversy over statins. *The Guardian*. Retrieved from http://www.theguardian.com/society/2014/aug/02/british-medical-journal-correct-articles-mistaken-fear-statins

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. http://doi.org/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.

http://doi.org/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. http://doi.org/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.

http://doi.org/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., Anmarkrud, Ø., Brandmo, C., & Strømsø, H. I. (2014). Developing and testing a model of

direct and indirect relationships between individual differences, processing, and multiple-

text comprehension. *Learning and Instruction*, *30*, 9–24.

http://doi.org/10.1016/j.learninstruc.2013.11.002

Bråten, I., Braasch, J. L. G., Strømsø, H. I., & Ferguson, L. E. (2014). Establishing Trustworthiness

when Students Read Multiple Documents Containing Conflicting Scientific Evidence.

Reading Psychology, *0*(0), 1–35. http://doi.org/10.1080/02702711.2013.864362

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. http://doi.org/10.1080/00461520.2011.538647

Bråten, I., & Ferguson, L. E. (2014). Investigating cognitive capacity, personality, and epistemic beliefs in relation to science achievement. *Learning and Individual Differences*, *36*, 124–130. http://doi.org/10.1016/j.lindif.2014.10.003

Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (2013). Justification beliefs and multiple-documents comprehension. *European Journal of Psychology of Education*, *28*(3), 879–902. http://doi.org/10.1007/s10212-012-0145-2

Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (2014). 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*, *84*(1), 58–85. http://doi.org/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. http://doi.org/10.1007/s11409-007-9004-y

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

Bråten, I., & Strømsø, H. I. (2006b). Epistemological beliefs, interest, and gender as predictors of Internet-based learning activities. *Computers in Human Behavior*, *22*(6), 1027–1042. http://doi.org/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. http://doi.org/10.1080/01638530902959646

Bråten, I., & Strømsø, H. I. (2010). When law students read multiple documents about global warming: Examining the role of topic-specific beliefs about the nature of knowledge and knowing. *Instructional Science*, *38*(6), 635–657. http://doi.org/10.1007/s11251-008-9091-4

Bråten, I., Strømsø, H. I., & Salmerón, L. (2011). Trust and mistrust when students read multiple

information sources about climate change. *Learning and Instruction*, *21*(2), 180–192.

http://doi.org/10.1016/j.learninstruc.2010.02.002

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. http://doi.org/10.2190/E763-X0LN-6NMF-CB86

Bråten, I., & Weinstein, C. E. (2004). Internet-specific Epistemological Questionnaire (ISEQ).

Austin, TX: Department of Educational Psychology, University of Texas at Austin.

British Educational Research Association. (2011). Ethical Guidelines for Educational Research.

British Psychological Society. Retrieved from https://www.bera.ac.uk/researchers-

resources/resources-for-researchers

British Psychological Society. (2013). Conducting Research on the Internet: Ethics Guidelines for

Internet-mediated Research. British Psychological Society. Retrieved from

http://www.bps.org.uk/publications/policy-and-guidelines/research-guidelines-policy-

documents/research-guidelines-poli

British Psychological Society. (2014). Code of Human Research Ethics. British Psychological

Society. Retrieved from

http://www.bps.org.uk/system/files/Public%20files/code_of_human_research_ethics_de

c_2014_inf180_web.pdf

Britt, M. A., & Gabrys, G. L. (2001). Teaching advanced literacy skills for the World Wide Web. In C.

Wolfe (Ed.), *Learning and teaching on the World Wide Web* (p. Chapter 4). Orlando,

Florida: Academic Press. Retrieved from http://psycnet.apa.org/psycinfo/2001-06609-003

Broder, A. (2002). A taxonomy of web search. *SIGIR Forum*, *36*(2), 3–10.

http://doi.org/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. http://doi.org/10.1007/s11409-009-9053-5

Bromme, R., Scharrer, L., Stadtler, M., Hömberg, J., & Torspecken, R. (2014). Is it believable when it's scientific? How scientific discourse style influences laypeople's resolution of conflicts. *Journal of Research in Science Teaching*, n/a–n/a. http://doi.org/10.1002/tea.21172

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., & Deakin Crick, R. (2012). Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. New York: ACM Press. Retrieved from http://oro.open.ac.uk/32823/

Buckingham Shum, S., & Ferguson, R. (2012). Social Learning Analytics. *Educational Technology & Society*, *15*(3), 3–26.

Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer Science & Business Media.

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., Chen, A. T., McArthur, E., & Davis, N. (2013). Searcher Actions and Strategies in Asynchronous Collaborative Search. Presented at the ASIST. Retrieved from https://asis.org/asist2013/proceedings/submissions/papers/91paper.pdf

Capra, R., Marchionini, G., Velasco-Martin, J., & Muller, K. (2010). Tools-at-hand and learning in multi-session, collaborative search (p. 951). ACM Press. http://doi.org/10.1145/1753326.1753468

Carlson, P. A., & Berry, F. C. (2003). Calibrated peer review and assessing learning outcomes (Vol. 2). Presented at the Frontiers in Education, Boulder, CO: IEEE. http://doi.org/10.1109/FIE.2003.1264740

Case, P. (2013, June 13). Scientists dismiss "unreliable" glyphosate study. Retrieved from

http://www.fwi.co.uk/arable/scientists-dismiss-unreliable-glyphosate-study.htm

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. http://doi.org/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. http://doi.org/10.1080/00461520.2011.587722

Chodorow, M., & Burstein, J. (2004). BEYOND ESSAY LENGTH: EVALUATING E-RATER®'S

PERFORMANCE ON TOEFL® ESSAYS. *ETS Research Report Series*, *2004*(1), i–38.

Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of scaffolded peer

assessment of writing from instructor and student perspectives. *Journal of Educational*

Psychology, *98*(4), 891. http://doi.org/10.1037/0022-0663.98.4.891

Cho, Y. H., & Cho, K. (2011). Peer reviewers learn from giving comments. *Instructional Science*,

39(5), 629–643. http://doi.org/10.1007/s11251-010-9146-1

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.

http://doi.org/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. http://doi.org/10.1037/10096-006

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

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155.

http://doi.org/10.1037/0033-2909.112.1.155

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. http://doi.org/10.1002/asi.21541

Cole, C. (2012). *Information need*. American Society for Information Science and Technology by Information Today, Incorporated.

Cook, D. A., & Beckman, T. J. (2009). Does scale length matter? A comparison of nine-versus five-point rating scales for the mini-CEX. *Advances in Health Sciences Education*, *14*(5), 655–664. http://doi.org/10.1007/s10459-008-9147-x

Costello, A. B., & Osborne, J., W. (2005). Getting the most from your analysis. *Practical Assessment Research & Evaluation*, *10*(7). Retrieved from http://pareonline.net/pdf/v10n7.pdf

Courtney, M. G. R. (2013). Determining the number of factors to retain in EFA: Using the SPSS R-Menu v2. 0 to make more judicious estimations. *Practical Assessment, Research & Evaluation*, *18*(8), 1–14.

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

Crook, C., & Lewthwaite, S. (2010). Technologies for formal and informal learning. In K. Littleton, C. Wood, & J. K. Staarman (Eds.), *International Handbook of Psychology in Education. Emerald* (pp. 435–461). Emerald Group Publishing.

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.
http://doi.org/10.1287/mnsc.32.5.554

Darbyshire, P., & McDonald, H. (2004). Choosing response scale labels and length: Guidance for
researchers and clients. *Australasian Journal of Market Research*, *12*(2), 17–26.

Darling-Hammond, L., & Adamson, F. (2010). Beyond basic skills: The role of performance
assessment in achieving 21st century standards of learning. In *Stanford Center for
Opportunity Policy in Education (SCOPE), Stanford University, School of Education.
Retrieved from http://edpolicy. stanford. edu*. Retrieved from
https://edpolicy.stanford.edu/sites/default/files/beyond-basic-skills-role-performance-
assessment-achieving-21st-century-standards-learning-report_0.pdf

Davies, P. (2000). Computerized Peer Assessment. *Innovations in Education & Training
International*, *37*(4), 346–355. http://doi.org/10.1080/135580000750052955

Davis, A. (1998). 3: Understanding and Holism. *Journal of Philosophy of Education*, *32*(1), 41–55.
http://doi.org/10.1111/1467-9752.00076

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. http://doi.org/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. http://doi.org/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.

Dawson, S. (2010). "Seeing" the learning community: An exploration of the development of a resource for monitoring online student networking. *British Journal of Educational Technology*, *41*(5), 736–752. http://doi.org/10.1111/j.1467-8535.2009.00970.x

Deakin Crick, R., Broadfoot, P., & Claxton, G. (2004). Developing an effective lifelong learning inventory: The ELLI project. *Assessment in Education: Principles, Policy & Practice*, *11*(3), 247–272. http://doi.org/10.1080/0969594042000304582

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. http://doi.org/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

de Waard, A., Buitelaar, P., & Eigner, T. (2009). Identifying the epistemic value of discourse segments in biology texts. In *Proceedings of the Eighth International Conference on Computational Semantics* (pp. 351–354). Stroudsburg, PA, USA: Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=1693756.1693802

Dewey, J. (1938a). *Experience and education*. Kappa Delta Pi. Retrieved from http://ruby.fgcu.edu/courses/ndemers/colloquium/ExperiencEducationDewey.pdf

Dewey, J. (1938b). The theory of inquiry. *New York: Holt, Rinehart & Wiston*. Retrieved from http://www.secondroad.net/foundations/practical/01DeweyLogicOfInquiry.pdf

Dillenbourg, P., Fox, A., Kirchner, C., Mitchell, J., & Wirsing, M. (2014). Massive open online courses: current state and perspectives. *Dagstuhl Manifestos*, *4*(1), 1–27.

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. http://doi.org/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. http://doi.org/10.1016/j.erap.2012.03.004

Dong, Y., Liang, J.-C., Yu, Y.-Y., Wu, J.-C., & Tsai, C.-C. (2014). The relationships between Chinese

higher education students' epistemic beliefs and their judgmental standards of searching

for literature online: undergraduate versus graduate comparisons. *Interactive Learning*

Environments, *0*(0), 1–17. http://doi.org/10.1080/10494820.2014.948460

Downes, S. (2007, February 5). What Connectivism Is. Retrieved from

http://www.downes.ca/post/38653

Duell, O. K., & Schommer-Aikins, M. (2001). Measures of people's beliefs about knowledge and

learning. *Educational Psychology Review*, *13*(4), 419–449.

http://doi.org/10.1023/A:1011969931594

Duval, E. (2011). Attention please! Learning analytics for visualization and recommendation. In

Proceedings of LAK11: 1st International Conference on Learning Analytics and Knowledge

(pp. 9–17). Retrieved from

http://dl.acm.org/ft_gateway.cfm?id=2090118&ftid=1075994&dwn=1

Dweck, C. (2006). *Mindset: The new psychology of success*. Random House.

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.

http://doi.org/10.1080/07370008.1993.9649021

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. http://doi.org/10.1002/asi.10133

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115. http://doi.org/10.1111/j.1365-2648.2007.04569.x

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215. http://doi.org/10.1037/0033-295X.87.3.215

Evans, B. M., & Chi, E. H. (2010). An elaborated model of social search. *Information Processing & Management*, *46*(6), 656–678. http://doi.org/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. http://doi.org/10.1016/j.ipm.2009.12.001

FDA. (2007, August 9). FDA Warns Consumers to Avoid Red Yeast Rice Products Promoted on Internet as Treatments for High Cholesterol [WebContent]. Retrieved from http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/2007/ucm108962.htm

Ferguson, L. E. (2014). Epistemic Beliefs and Their Relation to Multiple-Text Comprehension: A Norwegian Program of Research. *Scandinavian Journal of Educational Research*, *0*(0), 1–22. http://doi.org/10.1080/00313831.2014.971863

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. http://doi.org/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. http://doi.org/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., & 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., & Sharples, M. (2014). Innovative Pedagogy at Massive Scale: Teaching and Learning in MOOCs. In C. Rensing, S. de Freitas, T. Ley, & P. J. Muñoz-Merino (Eds.), *Open Learning and Teaching in Educational Communities* (pp. 98–111). Springer International Publishing. Retrieved from http://link.springer.com.libezproxy.open.ac.uk/chapter/10.1007/978-3-319-11200-8_8

Fisher, K., Counts, S., & Kittur, A. (2012). Distributed sensemaking: improving sensemaking by leveraging the efforts of previous users. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (pp. 247–256). Retrieved from http://dl.acm.org/citation.cfm?id=2207711

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. http://doi.org/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. http://doi.org/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. http://doi.org/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. http://doi.org/10.1080/09500690701543617

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. http://doi.org/10.1016/j.erap.2011.12.005

Gallagher, J. (2014, July 18). Statins: Millions more to get drugs in controversial plans. *BBC News*. Retrieved from http://www.bbc.co.uk/news/health-28352290

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). irr: Various Coefficients of interrater Reliability and Agreement (Version 0.84). Retrieved from http://CRAN.R-project.org/package=irr

Gardner, J. (2011). *Assessment and learning* (2nd ed.). SAGE.

Garrison, J. (1995). Deweyan Pragmatism and the Epistemology of Contemporary Social Constructivism. *American Educational Research Journal*, *32*(4), 716–740. http://doi.org/10.3102/00028312032004716

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.

Gielen, S., Dochy, F., & Onghena, P. (2011). An inventory of peer assessment diversity. *Assessment & Evaluation in Higher Education*, *36*(2), 137–155.

Gillam, C. (2013, April 25). Heavy use of herbicide Roundup linked to health dangers-U.S. study. *Reuters*. Retrieved from http://www.reuters.com/article/2013/04/25/roundup-health-study-idUSL2N0DC22F20130425

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. http://doi.org/10.1080/02702710902733600

GM Freeze, & Friends of the Earth. (2013, June 13). Government Urged to Act After Weedkiller Traces Found in Britons. Retrieved from http://www.gmfreeze.org/news-releases/225/

Goldberg, S. C. (2010). *Relying on others: An essay in epistemology*. Oxford University Press.

Goldman, A. (2010). Social Epistemology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2010). Retrieved from http://plato.stanford.edu/archives/sum2010/entries/epistemology-social/

Goldman, A., & Whitcomb, D. (2011). *Social epistemology: essential readings*. Oxford University Press.

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. http://doi.org/10.1002/RRQ.027

Goldman, S. R., Lawless, K., Pellegrino, J. W., & Gomez, K. (2012). A Technology for Assessing Multiple Source Comprehension: An Essential Skill of the 21st Century. In M. Mayrath (Ed.), *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research*. Information Age Publishing (IAP).

Goldman, S. R., & Scardamalia, M. (2013). Managing, Understanding, Applying, and Creating Knowledge in the Information Age: Next-Generation Challenges and Opportunities. *Cognition and Instruction*, *31*(2), 255–269. http://doi.org/10.1080/10824669.2013.773217

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

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*. http://doi.org/10.1016/j.ipm.2012.12.008

Gordon, R., Y. (2010). Marked Variability of Monacolin Levels in Commercial Red Yeast Rice Products: Buyer Beware! *Archives of Internal Medicine*, *170*(19), 1722–1727. http://doi.org/10.1001/archinternmed.2010.382

Greco, J., & Turri, J. (2013). Virtue Epistemology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2013). Retrieved from http://plato.stanford.edu/archives/win2013/entries/epistemology-virtue/

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. http://doi.org/10.1080/00461520802178458

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. http://doi.org/10.1080/00461520.2010.515932

Greene, J. A., & Yu, S. B. (2014). Modeling and measuring epistemic cognition: A qualitative re-investigation. *Contemporary Educational Psychology*, *39*(1), 12–28. http://doi.org/10.1016/j.cedpsych.2013.10.002

Greeno, J., Collins, A., & Resnick, L. B. (1992). Cognition and learning. In B. Berliner & R. Calfee

(Eds.), *Handbook of Educational Psychology*. New York. Retrieved from

http://www.wou.edu/~girodm/611/GCR.pdf

Gress, C. L. Z., Fior, M., Hadwin, A. F., & Winne, P. H. (2010). Measurement and assessment in

computer-supported collaborative learning. *Computers in Human Behavior*, *26*(5), 806–

814. http://doi.org/10.1016/j.chb.2007.05.012

Grice, H. P. (1975). Logic and conversation. In H. Geirsson & M. Losonsky (Eds.), *Readings in*

Language and Mind (pp. 41–58). Oxford, UK: Blackwell.

Groza, T., Handschuh, S., & Bordea, G. (2010). Towards automatic extraction of epistemic items

from scientific publications. In *Proceedings of the 2010 ACM Symposium on Applied*

Computing (pp. 1341–1348). New York, NY, USA: ACM.

http://doi.org/10.1145/1774088.1774377

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

Gwet, K., L. (2010). *agree.coeff3.raw.r: functions in this script file compute various agreement*

coefficients and their standard errors when dealing with multiple raters. Advanced

Analytics. Retrieved from http://www.agreestat.com/r_functions.html

Gwet, K. L. (2011). *On the Krippendorff's alpha coefficient*. Retrieved from

http://www.agreestat.com/research_papers/onkrippendorffalpha_old.pdf

Hacker, K. L., & Mason, S. M. (2003). Ethical gaps in studies of the digital divide. *Ethics and*

Information Technology, *5*(2), 99–115. http://doi.org/10.1023/A:1024968602974

Haddock, A., Millar, A., & Pritchard, D. (2010). *Social Epistemology*. Oxford University Press.

Retrieved from

http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199577477.001.0001/

acprof-9780199577477

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. http://doi.org/10.1111/j.1467-9817.2012.01536.x

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. http://doi.org/10.1207/S15327809JLS1201_3

Hancock, G. R., & Klockars, A. J. (1991). The effect of scale manipulations on validity: Targetting frequency rating scales for anticipated performance levels. *Applied Ergonomics*, *22*(3), 147–154.

Hanrahan, S. J., & Isaacs, G. (2001). Assessing Self- and Peer-assessment: The students' views. *Higher Education Research & Development*, *20*(1), 53–70. http://doi.org/10.1080/07294360123776

Hansen, P., & Järvelin, K. (2005). Collaborative Information Retrieval in an information-intensive domain. *Information Processing & Management*, *41*(5), 1101–1119. http://doi.org/10.1016/j.ipm.2004.04.016

Hansen, P., Shah, C., & Klas, C.-P. (Eds.). (Forthcoming). *Collaborative Information Seeking: Best practices, New Domains, New Thoughts*. Springer. Retrieved from http://collab.infoseeking.org/cisbook2015/

Harding, A. (2008, July 9). Contamination common in red yeast rice products. *Reuters*. Retrieved from http://www.reuters.com/article/2008/07/09/us-contamination-common-idUSCOL97022820080709

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. http://doi.org/10.1111/j.1475-682X.2009.00317.x

Hargittai, E., Fullerton, L., Menchen-Trevino, E., & Thomas, K. Y. (2010). Trust Online: Young Adults' Evaluation of Web Content. *International Journal of Communication*, *4*(0), 27.

Harlen, W. (2007). *Assessment of learning*. London, UK: SAGE.

Harris, L. R., & Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation. *Teaching and Teacher Education*, *36*, 101–111. http://doi.org/10.1016/j.tate.2013.07.008

Harris, P. L. (2012). *Trusting What You're Told*. Harvard University Press.

Hastings, P., Hughes, S., Magliano, J. P., Goldman, S. R., & Lawless, K. (2012). Assessing the use of multiple sources in student essays. *Behavior Research Methods*, *44*(3), 622–633. http://doi.org/10.3758/s13428-012-0214-0

Hatch, J. A. (2002). Chapter 4: Analyzing Qualitative Data - Typographical Analysis. In *Doing Qualitative Research in Education Settings* (pp. 152–161). SUNY Press.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89.

Haythornthwaite, C., & de Laat, M. (2010). Social Networks and Learning Networks: Using social network perspectives to understand social learning. In *7th International Conference on Networked Learning*. Retrieved from http://celstec.org.uk/system/files/file/conference_proceedings/NLC2010_Proceedings/abstracts/PDFs/Haythornwaite.pdf

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.

Heilman, J. M., Kemmann, E., Bonert, M., Chatterjee, A., Ragar, B., Beards, G. M., … others. (2011). Wikipedia: a key tool for global public health promotion. *Journal of Medical Internet Research*, *13*(1). Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/pmc3221335/

Hendry, G. D., Armstrong, S., & Bromberger, N. (2012). Implementing standards-based assessment effectively: Incorporating discussion of exemplars into classroom teaching. *Assessment & Evaluation in Higher Education*, *37*(2), 149–161.

Hertzum, M. (2008). Collaborative information seeking: The combined activity of information seeking and collaborative grounding. *Information Processing & Management*, *44*(2), 957–962. http://doi.org/10.1016/j.ipm.2007.03.007

Hirsch, B., Hitt, G. W., Powell, L., Khalaf, K., & Balawi, S. (2013). Collaborative learning in action. In *2013 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)* (pp. 283–288). http://doi.org/10.1109/TALE.2013.6654447

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. http://doi.org/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. http://doi.org/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

Højsgaard, S., Halekoh, U., Robison-Cox, J., Wright, K., & Leidi, A. A. (2012). doBy: doBy-Groupwise summary statistics, general linear contrasts, population means (least-squares-means), and other utilities (Version 4.5-13). Retrieved from http://cran.open-source-solution.org/web/packages/doBy/

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/

Horgan, álne, & Sweeney, J. (2012). University Students' Online Habits and Their Use of the Internet for Health Information: *CIN: Computers, Informatics, Nursing*, *30*(8), 402–408. http://doi.org/10.1097/NXN.0b013e3182510703

Hornik, K., Mair, P., Johannes, R., Geiger, W., Buchta, C., & Feinerer, I. (2013). The textcat Package for n-Gram Based Text Categorization in R. *Journal of Statistical Software*, *52*(6), 1–17.

Hsieh, H.-F., & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, *15*(9), 1277–1288. http://doi.org/10.1177/1049732305276687

Hsu, C.-Y., Tsai, M.-J., Hou, H.-T., & Tsai, C.-C. (2013). Epistemic Beliefs, Online Search Strategies, and Behavioral Patterns While Exploring Socioscientific Issues. *Journal of Science Education and Technology*, 1–10. http://doi.org/10.1007/s10956-013-9477-1

Hutchison, P., & Hammer, D. (2010). Attending to student epistemological framing in a science classroom. *Science Education*, *94*(3), 506–524. http://doi.org/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. http://doi.org/10.1016/j.compedu.2007.09.011

Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics*, *30*(4), 437–455. http://doi.org/10.1016/S0378-2166(98)00009-5

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. http://doi.org/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. http://doi.org/10.1016/j.ipm.2008.05.007

James, D. A., & DebRoy, S. (2012). RMySQL: R interface to the MySQL database (Version 0.9-3). Retrieved from http://cran.r-project.org/web/packages/RMySQL/index.html

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

Jenkins, G. D., & Taber, T. D. (1977). A Monte Carlo study of factors affecting three indices of composite scale reliability. *Journal of Applied Psychology*, *62*(4), 392.

Johal, G. S., & Huber, D. M. (2009). Glyphosate effects on diseases of plants. *European Journal of Agronomy*, *31*(3), 144–152.

Johnson, J. D., & Meischke, H. (2006). A Comprehensive Model of Cancer-Related Information

Seeking Applied to Magazines. *Human Communication Research*, *19*(3), 343–367.

http://doi.org/10.1111/j.1468-2958.1993.tb00305.x

Jucks, R., & Paus, E. (2013). Different Words for the Same Concept: Learning Collaboratively From

Multiple Documents. *Cognition and Instruction*, *31*(2), 227–254.

http://doi.org/10.1080/07370008.2013.769993

Kammerer, Y., Amann, D. G., & Gerjets, P. (2015). When adults without university education

search the Internet for health information: The roles of Internet-specific epistemic beliefs

and a source evaluation intervention. *Computers in Human Behavior*, *48*, 297–309.

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.

http://doi.org/10.1016/j.chb.2012.10.012

Kammerer, Y., & Gerjets, P. (2011). Effects of search interface and Internet-specific epistemic

beliefs on source evaluations during Web search for medical information: an eye-tracking

study. *Behaviour & Information Technology*, *31*(1), 83–97.

http://doi.org/10.1080/0144929X.2011.599040

Karimi, M. N., & Atai, M. R. (2014). ESAP Students' Comprehension of Multiple Technical Reading

Texts: Insights from Personal Epistemological Beliefs. *Reading Psychology*, *35*(8), 736–761.

http://doi.org/10.1080/02702711.2013.802753

Kärkkäinen, E. (2006). Stance taking in conversation: From subjectivity to intersubjectivity. *Text &

Talk*, *26*(6), 699–731. http://doi.org/10.1515/TEXT.2006.029

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://workshops.fxpal.com/cscw2010cis/submissions/tmp10.pdf

Katz, S. (2000). Competency, epistemology and pedagogy: curriculum''s holy trinity. *Curriculum

Journal*, *11*(2), 133–144. http://doi.org/10.1080/09585170050045164

Kean, J. (2012). Show AND tell: using peer assessment and exemplars to help students understand quality. *Practitioner Research in Higher Education*, *6*(2), 83–94.

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. http://doi.org/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. http://doi.org/10.1037/0022-0663.90.1.134

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. http://doi.org/10.1207/s15326985ep3901_2

Kitchener, R. F. (2002). Folk epistemology: An introduction. *New Ideas in Psychology*, *20*(2), 89–105. http://doi.org/10.1016/S0732-118X(02)00003-X

Klamma, R. (2013). Community Learning Analytics – Challenges and Opportunities. In J.-F. Wang & R. Lau (Eds.), *Advances in Web-Based Learning – ICWL 2013* (pp. 284–293). Springer Berlin Heidelberg. Retrieved from http://link.springer.com.libezproxy.open.ac.uk/chapter/10.1007/978-3-642-41175-5_29

Knight, S. (2012a). Methodological justification: A framework for the analysis of students' collaborative use of search engines for information retrieval in the classroom. University of Cambridge.

Knight, S. (2012b, July). *Finding Knowledge - The Role of Dialogue in Collaborative Information Retrieval in the Classroom* (Master's). University of Cambridge, Cambridge.

Knight, S. (2013a). Appendix C7.1: Resources for Searching with the Internet. 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). 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. (2014). Finding Knowledge: What Is It To "Know" When We Search? In *Society of the Query Reader: Reflections on Web Search*. Amsterdam, The Netherlands: Institute of Network Cultures. Retrieved from http://networkcultures.org/wpmu/portal/publication/society-of-the-query-reader-reflections-on-web-search/ individual chapter from http://oro.open.ac.uk/39307/

Knight, S. (2015, March). *Collaborative Information Seeking Tasks as Complex Performance Assessments*. Presented at the Library and Information Science Speaker Series, School of Communication and Information, Rutgers.

Knight, S., Arastoopour, G., Williamson Shaffer, D., Buckingham Shum, S., & Littleton, K. (2014). Epistemic Networks for Epistemic Commitments. In *International Conference of the Learning Sciences*. Boulder, CO: International Society of the Learning Sciences. Retrieved from http://oro.open.ac.uk/39254/

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., 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. http://doi.org/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., Buckingham Shum, S., & Littleton, K. (2014). Epistemology, assessment, pedagogy: where learning meets analytics in the middle space. *Journal of Learning Analytics*, *1*(2). Retrieved from http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3538

Knight, S., & Littleton, K. (in submissiona). Dialogue as data.

Knight, S., & Littleton, K. (in submissionb). Epistemic Cognition Socialised.

Knight, S., & Littleton, K. (in press). Learning through Collaborative Information Seeking. In P. Hansen, C. Shah, & C.-P. Klas (Eds.), *Collaborative Information Seeking: Best practices, New Domains, New Thoughts*. Springer.

Knight, S., & Littleton, K. (2015a). Developing a multiple-document-processing performance assessment for epistemic literacy. Presented at the The 5th International Learning Analytics & Knowledge Conference (LAK15): Scaling Up: Big Data to Big Impact, Poughkeepsie, NY, USA (Forthcoming). Retrieved from http://lak15.solaresearch.org/

Knight, S., & Littleton, K. (2015b). Thinking, Interthinking, and Technological Tools. In R. Wegerif, L. Li, & J. C. Kaufman (Eds.), *The Routledge International Handbook of Research on Teaching Thinking* (p. section 7(al)). Routledge. Retrieved from http://www.routledge.com/books/details/9780415747493/

Knight, S., & Mercer, N. (Forthcoming). The role of collaborative, epistemic discourse in classroom information seeking tasks. *Technology, Pedagogy and Education*.

Knight, S., & Mercer, N. (2015). The role of exploratory talk in classroom search engine tasks. *Technology, Pedagogy and Education*, *24*(3), 303–319. http://doi.org/10.1 08 0/14 75 939 X .2014.931884

Knight, S., Wise, A. F., Chen, B., & Cheng, B. H. (2015). It's About Time: 4th International Workshop on Temporal Analyses of Learning Data. Presented at the The 5th International Learning Analytics & Knowledge Conference (LAK15): Scaling Up: Big Data to Big Impact, Poughkeepsie, NY, USA (Forthcoming). Retrieved from http://lak15time.github.io/

Kobayashi, K. (2014). Students' consideration of source information during the reading of multiple texts and its effect on intertextual conflict resolution. *Instructional Science*, *42*(2), 183–205.

Kotzee, B. (2013). *Education and the Growth of Knowledge: Perspectives from Social and Virtue Epistemology* (1 edition). Malden, Massachusetts: Wiley-Blackwell.

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. http://doi.org/10.1016/S0747-5632(02)00057-2

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. http://doi.org/10.1002/(SICI)1097-4571(199106)42:5<361::AID-ASI6>3.0.CO;2-#

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, *15*(3), 309–328. http://doi.org/10.1016/S0885-2014(00)00030-7

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. http://doi.org/10.1016/j.compedu.2008.11.010

Ku, K. Y. L., Lai, E. C. M., & Hau, K. T. (2014). Epistemological Beliefs and the Effect of Authority on Argument-counterargument Integration: An Experiment. *Thinking Skills and Creativity*, *in press*. http://doi.org/10.1016/j.tsc.2014.03.004

Kulkarni, C., Wei, K. P., Le, H., Chia, D., Papadopoulos, K., Cheng, J., … Klemmer, S. R. (2015). Peer and Self Assessment in Massive Online Classes. In H. Plattner, C. Meinel, & L. Leifer (Eds.),

Design Thinking Research (pp. 131–168). Springer International Publishing. Retrieved from

http://link.springer.com.libezproxy.open.ac.uk/chapter/10.1007/978-3-319-06823-7_9

Lackey, J. (2008). *Learning from Words: Testimony as a Source of Knowledge*. Oxford University

Press.

Lackey, J., & Sosa, E. (2006). *The Epistemology of Testimony*. Oxford University Press.

LAK11. (2011). 1st International Conference on Learning Analytics and Knowledge 2011 |

Connecting the technical, pedagogical, and social dimensions of learning analytics.

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. http://doi.org/10.1016/S0306-4573(01)00034-6

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

University Press.

Lawless, K. A., Goldman, S. R., Gomez, K., Manning, F., & Braasch, J. (2012). Assessing multiple

source comprehension through evidence-centered design. In J. Sabatini P. ., T. O'Reilly, &

E. Albro (Eds.), *Reaching an understanding: Innovations in how we view reading

assessment* (pp. 3–17). Rowman & Littlefield.

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. http://doi.org/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. http://doi.org/10.1006/ijhc.2001.0512

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. http://doi.org/10.1089/cpb.2007.9969

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. http://doi.org/10.1080/09500690701613733

Linell, P. (1998). *Approaching dialogue: Talk, interaction and contexts in dialogical perspectives* (Vol. 3). John Benjamins Publishing.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, Performance-Based Assessment: Expectations and Validation Criteria. *Educational Researcher*, *20*(8), 15–21. http://doi.org/10.3102/0013189X020008015

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. http://doi.org/10.1016/S0959-4752(02)00042-7

Littlejohn, A., Beetham, H., & McGill, L. (2012). Learning at the digital frontier: a review of digital literacies in theory and practice. *Journal of Computer Assisted Learning*, *28*(6), 547–556. http://doi.org/10.1111/j.1365-2729.2011.00474.x

Littleton, K. (1999). Productivity through interaction. In K. Littleton & P. Light (Eds.), *Learning with computers: Analysing productive interaction* (p. 179). Psychology Press.

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. http://doi.org/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.

Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *4*(2), 73–79. http://doi.org/10.1027/1614-2241.4.2.73

Lu, Z., Kou, W., Du, B., Wu, Y., Zhao, S., Brusco, O. A., … Capuzzi, D. M. (2008). Effect of Xuezhikang, an Extract From Red Yeast Chinese Rice, on Coronary Events in a Chinese Population With Previous Myocardial Infarction. *The American Journal of Cardiology*, *101*(12), 1689–1693. http://doi.org/10.1016/j.amjcard.2008.02.056

Macrae, F. (2008, June 10). Quack medicine: Peking duck is "better for your heart than statins." *Mail Online*. Retrieved from http://www.dailymail.co.uk/health/article-1025277/Quack-medicine-Peking-duck-better-heart-statins.html

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*.

Maloney, J., Resnick, M., Rusk, N., Silverman, B., & Eastmond, E. (2010). The Scratch Programming Language and Environment. *Trans. Comput. Educ.*, *10*(4), 16:1–16:15. http://doi.org/10.1145/1868358.1868363

Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, *49*(4), 41. http://doi.org/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. http://doi.org/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. http://doi.org/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. http://doi.org/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. http://doi.org/10.1007/s11251-008-9089-y

Mayring, P. (2000). Qualitative Content Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *1*(2). Retrieved from http://www.qualitative-research.net/index.php/fqs/article/view/1089

McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix: Capturing Linguistic Features of Cohesion. *Discourse Processes*, *47*(4), 292–330. http://doi.org/10.1080/01638530902959943

Medizinisches Labor Bremen. (2013). *Determination of Glyphosate residues in human urine samples from 18 European countries* (No. Glyphosate MLHB-2013-06-06). Bremen,

Germany: Medizinisches Labor Bremen. Retrieved from

http://www.foeeurope.org/sites/default/files/glyphosate_studyresults_june12.pdf

Mellor, A. (2011). Essay length, lexical diversity and automatic essay scoring. *Memoirs of the Osaka Institute of Technology*, *55*(2), 1–14.

Mercer, N. (2004). Sociocultural discourse analysis: analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, *1*(2), 137–168. http://doi.org/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. http://doi.org/10.1080/10508400701793182

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

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. http://doi.org/10.1007/s11217-007-9071-1

Michaels, S., O'Connor, M. C., Hall, M. W., & Resnick, L. (2002). Accountable talk: classroom conversation that works. *Pittsburg: University of Pittsburgh*. http://doi.org/10.1007/s11217-007-9071-1

Mislevy, R., Behrens, J. T., Dicerbo, K. E., & Levy, R. (2012). Design and discovery in educational assessment: evidence-centred design, psychometrics, and educational data mining. *Journal of Educational Data Mining*, *4*(1), 11–48.

Mitsui, M. (2015). Etherpad contributions and touchpoints code (Version commit 052f940fab). Retrieved from https://github.com/mmitsui/Miscellaneous-Bits/tree/master/etherpad

Mokhtari, H. (2014). A Quantitative Survey on the Influence of Students' Epistemic Beliefs on Their General Information Seeking Behavior. *The Journal of Academic Librarianship, in press*. http://doi.org/10.1016/j.acalib.2014.04.001

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. http://doi.org/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

Moses, K. (2013, June 17). Do we need conclusive scientific proof to become concerned about an issue? | Kara Moses [Environment blog]. Retrieved from http://www.theguardian.com/environment/blog/2013/jun/17/wait-conclusive-science-concerned-issue

Mostert, M., & Snowball, J. D. (2013). Where angels fear to tread: Online peer-assessment in a large first-year class. *Assessment & Evaluation in Higher Education*, *38*(6), 674–686. http://doi.org/10.1080/02602938.2012.683770

Mozilla. (2014). *Firefox*. Retrieved from https://www.mozilla.org/en-US/firefox/new/

Muis, K. R. (2007). The Role of Epistemic Beliefs in Self-Regulated Learning. *Educational Psychologist*, *42*(3), 173–190. http://doi.org/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. http://doi.org/10.1007/s10648-006-9003-6

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.

Newman, K., Knight, S., Elbeshausen, S., & Hansen, P. (in press). Situating CIS – The importance of

Context in Collaborative Information Seeking. In P. Hansen, C. Shah, & C.-P. Klas (Eds.),

Collaborative Information Seeking: Best practices, New Domains, New Thoughts. Springer.

NHS Choices. (2008, June 9). Chinese red yeast rice - Health news - NHS Choices. Retrieved from

http://www.nhs.uk/news/2008/06June/Pages/Redyeastriceandheartdisease.aspx

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.

http://doi.org/10.1145/572020.572033

Nishisato, S., & Torii, Y. (1971). Effects of categorizing continuous normal variables on product-

moment correlation. *Japanese Psychological Research*. Retrieved from

http://psycnet.apa.org/psycinfo/1973-03523-001

Nulty, D. D. (2011). Peer and self-assessment in the first year of university. *Assessment &

Evaluation in Higher Education*, *36*(5), 493–507.

http://doi.org/10.1080/02602930903540983

O'Duke, S., & Powles, S. B. (2008). Glyphosate: a once-in-a-century herbicide. *Pest Management

Science*, *64*(4), 319–325. http://doi.org/10.1002/ps.1518

OECD. (2013). PISA 2015: Draft reading literacy framework. OECD Publishing. Retrieved from

http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Reading%20Framewo

rk%20.pdf

Office of Public Sector Information. Data Protection Act 1998 (1998). Retrieved from

http://www.legislation.gov.uk/ukpga/1998/29/contents

Olson, D. R., & Bruner, J. S. (1996). Folk psychology and folk pedagogy. *The Handbook of

Education and Human Development*, 9–27.

Orsmond, P., Merry, S., & Reiling, K. (1996). The importance of marking criteria in the use of peer assessment. *Assessment & Evaluation in Higher Education*, *21*(3), 239–250. http://doi.org/10.1080/0260293960210304

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

O'Toole, R. (2013, May 22). Pedagogical strategies and technologies for peer assessment in Massively Open Online Courses (MOOCs). University of Warwick, Coventry, UK: University of Warwick. Retrieved from http://wrap.warwick.ac.uk/54602/

Palermos, S. O., & Pritchard, D. (2013). Extended knowledge and social epistemology. *Social Epistemology Review and Reply Collective*, *2*(8), 105–120.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129–144. http://doi.org/10.1016/j.edurev.2013.01.002

Panadero, E., Romero, M., & Strijbos, J.-W. (2013). The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*, *39*(4), 195–203. http://doi.org/10.1016/j.stueduc.2013.10.005

Papamitsiou, Z., & Economides, A. A. (2014). Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Journal of Educational Technology & Society*, *17*(4), 49–64.

Pardo, A., & Teasley, S. (2014). Learning Analytics Research, Theory and Practice: Widening the Discipline. *Journal of Learning Analytics*, *1*(3), 4–6.

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. http://doi.org/10.1145/1518701.1518974

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

Paus, E., & Jucks, R. (2012). Common Ground? How the Encoding of Specialist Vocabulary Affects Peer-to-Peer Online Discourse. *Discourse Processes*, *49*(7), 565–598. http://doi.org/10.1080/0163853X.2012.711671

Pellegrino, J. W. (2013). Measuring What Matters in a Digital Age: Technology and the Design of Assessments for Multisource Comprehension. In D. G. Sampson, P. Isaias, D. Ifenthaler, & J. M. Spector (Eds.), *Ubiquitous and Mobile Learning in the Digital Age* (pp. 259–286). New York, NY, USA: Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4614-3329-3_17

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. http://doi.org/10.1017/S0140525X04000056

Piech, C., Huang, J., Chen, Z., Do, C., Ng, A., & Koller, D. (2013). Tuned Models of Peer Assessment in MOOCs. *arXiv:1307.2579 [cs, Stat]*. Retrieved from http://arxiv.org/abs/1307.2579

Pieschl, S., Stallmann, F., & Bromme, R. (2014). High School Students' Adaptation of Task Definitions, Goals and Plans to Task Complexity – The Impact of Epistemic Beliefs. *Psihologijske Teme*, *23*(1), 31–52.

Pistilli, M., Willis, J., Kock, D., Arnold, K., Teasley, S., & Pardo, A. (2014). *LAK '14: Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*. New York, NY, USA: ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2567574

Porsch, T., & Bromme, R. (2010). Effects of epistemological sensitization on source choices. *Instructional Science*, *39*(6), 805–819. http://doi.org/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

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, *104*(1), 1–15. http://doi.org/10.1016/S0001-6918(99)00050-5

Prins, F. J., Sluijsmans, D. M. A., Kirschner, P. A., & Strijbos, J. (2005). Formative peer assessment in a CSCL environment: a case study. *Assessment & Evaluation in Higher Education*, *30*(4), 417–444. http://doi.org/10.1080/02602930500099219

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

QSR International Pty Ltd. (2012). NVivo qualitative data analysis software (Version 10) [Windows]. Retrieved from https://www.qsrinternational.com/products_nvivo.aspx

Raeburn, P. (2013, April 26). Discover blogger Keith Kloor stumbles into nest of questionable studies and reporting on GMOs and multiple ailments. Retrieved from https://ksj.mit.edu/tracker/2013/04/discover-blogger-keith-kloor-stumbles-ne/

Raiche, G. (2010). *nFactors: an R package for parallel analysis and non graphical solutions to the Cattell scree test*. Retrieved from http://CRAN.R-project.org/package=nFactors

R Core Team. (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435–448. http://doi.org/10.1080/02602930902862859

Redecker, C., & Johannessen, Ø. (2013). Changing Assessment — Towards a New Assessment Paradigm Using ICT. *European Journal of Education*, *48*(1), 79–96. http://doi.org/10.1111/ejed.12018

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. http://doi.org/10.1007/s11412-009-9070-z

Resnick, L. B. (2001). Making America smarter: The real goal of school reform. In C. Arthur, L. (Ed.), *Developing minds: A resource book for teaching thinking* (3rd ed., pp. 3–6). Alexandria, VA: Association for Supervision and Curriculum Development.

Revelle, W. (in draft). *An introduction to psychometric theory with applications in R (in preparation)*. Springer. Retrieved from Draft chapters available at http://personality-project.org/r/book/

Revelle, W. (2014). psych: Procedures for Personality and Psychological Research (Version 1.4.8). Evanston, Illinois, USA: Northwestern University. Retrieved from http://CRAN.R-project.org/package=psych

Revelle, W., & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research*, *14*(4), 403–414. http://doi.org/10.1207/s15327906mbr1404_2

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

Ridker, P. M., & Cook, N. R. (2013). Statins: new American guidelines for prevention of cardiovascular disease. *The Lancet*, *382*(9907), 1762–1765. http://doi.org/10.1016/S0140-6736(13)62388-0

Rienties, B., Willis, A., Alcott, P., & Medland, E. (2013). Student experiences of self-reflection and peer assessment in providing authentic project-based learning to large class sizes. In *Facilitating learning in the 21st century: leading through technology, diversity and authenticity* (pp. 117–136). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-94-007-6137-7_7

Roberts, R. C., & Wood, W. J. (2007). *Intellectual Virtues*. Oxford University Press. Retrieved from http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199283675.001.0001/acprof-9780199283675

Robinson, R. (2001). Calibrated Peer Review™: an application to increase student reading & writing skills. *The American Biology Teacher*, *63*(7), 474–480.

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

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. http://doi.org/10.1207/s15327809jls1502_4

Rouet, J.-F. (2006). *The Skills of Document Use: From Text Comprehension to Web-Based Learning* (First Edition edition). Mahwah, NJ: Routledge.

Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19–52). Information Age Publishing (IAP). Retrieved from http://www.niu.edu/britt/recent_papers/pdfs/RouetBritt_chapter_for_McCrudden.pdf

RStudio. (2013). RStudio: Integrated development environment for R (Version 0.97.551) [Server]. Boston, MA, USA: RStudio. Retrieved from http://rstudio.com

Ruscio, J. (2012, November 28). EFA with Comparison Data. Retrieved April 7, 2015, from

http://www.tcnj.edu/~ruscio/taxometrics.html#EFAwithCD

Ruscio, J., & Roche, B. (2012). Determining the number of factors to retain in an exploratory factor

analysis using comparison data of known factorial structure. *Psychological Assessment*,

24(2), 282. http://doi.org/10.1037/a0025697

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.

Salmerón, L., Kammerer, Y., & García-Carrión, P. (2013). Searching the Web for conflicting topics:

Page and user factors. *Computers in Human Behavior*, *29*(6), 2161–2171.

http://doi.org/10.1016/j.chb.2013.04.034

Salomon, G. (1996). *Distributed Cognitions: Psychological and Educational Considerations* (1st Pbk.

Ed). Cambridge University Press.

Samsel, A., & Seneff, S. (2013). Glyphosate's Suppression of Cytochrome P450 Enzymes and

Amino Acid Biosynthesis by the Gut Microbiome: Pathways to Modern Diseases. *Entropy*,

15(4), 1416–1463. http://doi.org/10.3390/e15041416

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. http://doi.org/10.1016/S0747-5632(01)00048-6

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. http://doi.org/10.1080/07370000902797379

Schmidt, C. W. (2012). Trending now: using social media to predict and track disease outbreaks. *Environ Health Perspect*, *120*(1), a30–a33. http://doi.org/10.1289/ehp.120-a30

Schmidt, H. G., Molen, H. T. V. D., Winkel, W. W. R. T., & Wijnen, W. H. F. W. (2009). Constructivist, Problem-Based Learning Does Work: A Meta-Analysis of Curricular Comparisons Involving a Single Medical School. *Educational Psychologist*, *44*(4), 227–249. http://doi.org/10.1080/00461520903213592

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, *82*(3), 498–504. http://doi.org/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. (2001). Current Themes and Future Directions in Epistemological Research: A Commentary. *Educational Psychology Review*, *13*(4), 451–464. http://doi.org/10.1023/A:1011922015665

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., Bendixen, L. D., & Dunkle, M. E. (2002). Development and validation of the Epistemic

    Belief Inventory (EBI). Retrieved from http://psycnet.apa.org/psycinfo/2001-18187-012

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. http://doi.org/10.1080/713838244

Science Media Centre. (2013, June 13). expert reaction to press release from FotE and GM Freeze

    about glyphosate in urine. Retrieved from http://www.sciencemediacentre.org/expert-

    reaction-to-press-release-from-fote-and-gm-freeze-about-glyphosate-in-urine/

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. (2009). Lessons and challenges for collaborative information seeking (CIS) systems

    developers. In *Collaborative Information Behaviour Workshop at GROUP*. Sanibel Island,

    FL, USA. Retrieved from

    http://comminfo.rutgers.edu/~chirags/papers/Shah_CIB_Workshop.pdf

Shah, C. (2010). Coagmento-a collaborative information seeking, synthesis and sense-making

    framework. *Integrated Demo at CSCW*, 6–11.

Shah, C. (2012a). Coagmento: A Case Study in Designing. In E. Currás & N. L. Romero (Eds.),

    Systems Science and Collaborative Information Systems: Theories, Practices and New

    Research: Theories, Practices and New Research (p. 242). Hershey, PA: IGI Global.

Retrieved from http://www.igi-global.com/chapter/coagmento-case-study-designing-user/61295

Shah, C. (2012b). *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. http://doi.org/10.1002/asi.22819

Shah, C. (2014a). *Coagmento*. Retrieved from http://coagmento.org/

Shah, C. (2014b). Evaluating collaborative information seeking – synthesis, suggestions, and structure. *Journal of Information Science*, 0165551514530651. http://doi.org/10.1177/0165551514530651

Shah, C., Capra, R., & Hansen, P. (2015). Report on the Workshop on Social and Collaborative Information Seeking (SCIS). DIMACS, Rutgers. Retrieved from http://dimacs.rutgers.edu/Workshops/SCIS/SCIS2015Report.pdf

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. http://doi.org/10.1002/meet.14504701211

Shah, C., & González-Ibáñez, R. (2011). Evaluating the synergic effect of collaboration in information seeking. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval* (pp. 913–922). New York, NY, USA: ACM. http://doi.org/10.1145/2009916.2010038

Shah, C., Hendahewa, C., & González-Ibáñez, R. (2015). Rain or Shine? Forecasting Search Process Performance in Exploratory Search Tasks. *Journal of the Association for Information Science and Technology*. http://doi.org/10.1002/asi.23484

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

Shah, C., Marchionini, G., & Kelly, D. (2009). Learning design principles for a collaborative

information seeking system. In *CHI'09 Extended Abstracts on Human Factors in Computing

Systems* (pp. 3419–3424). ACM. Retrieved from

http://dl.acm.org/citation.cfm?id=1520496

Sharples, M., Adams, A., Ferguson, R., Gaved, M., McAndrew, P., Rienties, B., … Whitelock, D.

(2014). *Innovating Pedagogy 2014*. Retrieved from http://www.editlib.org/p/149392/

Sharples, M., McAndrew, P., Weller, M., Ferguson, R., FitzGerald, E., Hirst, T., & Gaved, M. (2013).

Innovating Pedagogy 2013 (Innovation Report No. 2). Milton Keynes: The Open

University. Retrieved from

http://www.open.ac.uk/iet/main/sites/www.open.ac.uk.iet.main/files/files/ecms/web-

content/Innovating_Pedagogy_report_2013.pdf

Sharples, M., McAndrew, P., Weller, M., Ferguson, R., FitzGerald, E., Hirst, T., … Whitelock, D.

(2012). *Innovating Pedagogy 2012* (Innovation Report No. 1). Milton Keynes: The Open

University. Retrieved from

http://www.open.ac.uk/personalpages/mike.sharples/Reports/Innovating_Pedagogy_rep

ort_July_2012.pdf

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.

Simsek, D., Buckingham Shum, S., Sandor, A., De Liddo, A., & Ferguson, R. (2013). XIP Dashboard:

visual analytics from automated rhetorical parsing of scientific metadiscourse. Presented

at the 1st International Workshop on Discourse-Centric Learning Analytics, Leuven,

Belgium. Retrieved from http://lakconference2013.wordpress.com/

Slade, S., & Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. *American

Behavioral Scientist*, *57*(10), 1510–1529. http://doi.org/10.1177/0002764213479366

Smucker, M. D., & Clarke, C. L. (2012). Time-based calibration of effectiveness measures. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval* (pp. 95–104). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2348300

Snow, C. (2002). *Reading for Understanding. Towards an R&D Program in Reading Comprehension*. DTIC Document. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA402712

Southavilay, V., Yacef, K., Reimann, P., & Calvo, R. A. (2013). Analysis of Collaborative Writing Processes Using Revision Maps and Probabilistic Topic Models. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 38–47). New York, NY, USA: ACM. http://doi.org/10.1145/2460296.2460307

Spence, P. R., & Reddy, M. (2012). Beyond practices: A field study of the contextual factors impacting collaborative information seeking. *Proceedings of the American Society for Information Science and Technology*, *49*(1), 1–10. http://doi.org/10.1002/meet.14504901131

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. http://doi.org/10.1007/s11412-007-9015-3

Stecher, B. (2010). Performance assessment in an era of standards-based educational accountability. *Standford Center for Opportunity Policy in Education*. Retrieved from http://edpolicy.stanford.edu/sites/default/files/publications/performance-assessment-era-standards-based-educational-accountability_0.pdf

Strijbos, J.-W., & Sluijsmans, D. (2010). Unravelling peer assessment: Methodological, functional, and conceptual developments. *Learning and Instruction*, *20*(4), 265–269. http://doi.org/10.1016/j.learninstruc.2009.08.002

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. http://doi.org/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. http://doi.org/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. http://doi.org/10.1080/01443410.2010.538039

Suen, H. K. (2014). Peer assessment for massive open online courses (MOOCs). *The International Review of Research in Open and Distributed Learning*, *15*(3). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1680

Sundin, O., & Francke, H. (2009). In search of credibility: pupils' information practices in learning environments. *Information Research*, *14*(4). Retrieved from http://www.informationr.net/ir/14-4/paper418.html

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. http://doi.org/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. (2006). Technology affordances for intersubjective meaning making: A research

agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning*, *1*,

315–337. http://doi.org/10.1007/s11412-006-9660-y

Suthers, D. D., Lund, K., Rosé, C. P., Teplovs, C., & Law, N. (Eds.). (2013). *Productive Multivocality*

in the Analysis of Group Interactions (Vol. 15). Springer. Retrieved from

http://www.springer.com/education+%26+language/learning+%26+instruction/book/978

-1-4614-8959-7

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. http://doi.org/10.1145/2460296.2460298

Tempelaar, D. T., Niculescu, A., Rienties, B., Gijselaers, W. H., & Giesbers, B. (2012). How

achievement emotions impact students' decisions for online learning, and what precedes

those emotions. *The Internet and Higher Education*, *15*(3), 161–169.

http://doi.org/10.1016/j.iheduc.2011.10.003

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

The Open University. (2014, September). Policy on Ethical Use of Student Data for Learning

Analytics. The Open University. Retrieved from

http://www.open.ac.uk/students/charter/essential-documents/ethical-use-student-data-

learning-analytics-policy

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.

http://doi.org/http://dx.doi.org/10.1016/j.learninstruc.2006.10.006

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of*

Educational Research, *68*(3), 249–276. http://doi.org/10.3102/00346543068003249

Topping, K. J. (2010). Methodological quandaries in studying process and outcomes in peer assessment. *Learning and Instruction*, *20*(4), 339–343. http://doi.org/10.1016/j.learninstruc.2009.08.003

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. http://doi.org/10.1111/j.0007-1013.2004.00411.x

Tsai, C.-C. (2001). A Review and Discussion of Epistemological Commitments, Metacognition, and Critical Thinking with Suggestions on Their Enhancement in Internet-Assisted Chemistry Classrooms. *Journal of Chemical Education*, *78*(7), 970. http://doi.org/10.1021/ed078p970

Tsai, C.-C. (2004). Information commitments in Web-based learning environments. *Innovations in Education and Teaching International*, *41*(1), 105–112. http://doi.org/10.1080/1470329032000172748a

Tsai, C.-C., & Liang, J.-C. (2009). The development of science activities via on-line peer assessment: The role of scientific epistemological views. *Instructional Science*, *37*(3), 293–310. http://doi.org/10.1007/s11251-007-9047-0

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. http://doi.org/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. http://doi.org/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. http://doi.org/10.1016/S0306-

4573(97)00040-X

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%20septemb

er%202010.ashx

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. http://doi.org/10.1016/S1479-3687(02)80010-4

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

van Zundert, M., Sluijsmans, D., & van Merriënboer, J. (2010). Effective peer assessment

processes: Research findings and future directions. *Learning and Instruction*, *20*(4), 270–

279. http://doi.org/10.1016/j.learninstruc.2009.08.004

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

Vygotsky, L. S. (1987). *The Collected works of L.S. Vygotsky*. (R. W. Rieber & A. S. Carton, Eds.) (Vol.

1). New York: Plenum Press.

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. http://doi.org/10.1016/j.chb.2007.01.030

Watson, W., & Ishiyama, J. (2012). Using Calibrated Peer Review to Facilitate Writing Assignments

in Large Political Science Classes. In *APSA 2012 Teaching & Learning Conference Paper*.

University of North Carolina (UNC) at Chapel Hill. Retrieved from

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1997612

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. http://doi.org/10.1007/s11412-006-6840-8

Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response

styles: The number of response categories and response category labels. *International*

Journal of Research in Marketing, *27*(3), 236–247.

http://doi.org/10.1016/j.ijresmar.2010.02.004

Wells, G. (1999). *Dialogic inquiry: towards a sociocultural practice and theory of education*.

Cambridge University Press.

Wells, G. (2002). The role of dialogue in activity theory. *Mind, Culture, and Activity*, *9*(1), 43–66.

http://doi.org/10.1207/S15327884MCA0901_04

Wells, G., & Claxton, G. (2002). *Learning for life in the 21st century: sociocultural perspectives on*

the future of education. Wiley-Blackwell.

Wenger, E. (1998). *Communities of practice: learning, meaning, and identity*. Cambridge University Press.

Weng, L.-J. (2004). Impact of the Number of Response Categories and Anchor Labels on Coefficient Alpha and Test-Retest Reliability. *Educational and Psychological Measurement*, *64*(6), 956–972. http://doi.org/10.1177/0013164404268674

Wertsch, J. (1993). *Voices of the Mind: a Sociocultural Approach to Mediated Action* (Reprint). Harvard University Press.

Whitelock, D. (2010). Activating assessment for learning: are we on the way with web 2.0? In *Web* (Vol. 2, pp. 319–342).

White, R. W. (2013). Beliefs and Biases in Web Search. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*. Retrieved from http://research.microsoft.com/en-us/um/people/ryenw/papers/WhiteSIGIR2013.pdf

Whitmire, E. (2003). Epistemological beliefs and the information-seeking behavior of undergraduates. *Library & Information Science Research*, *25*(2), 127–142. http://doi.org/10.1016/S0740-8188(03)00003-3

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, *21*(12), 1–20.

Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. New York: Springer. Retrieved from http://had.co.nz/ggplot2/book

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, *40*(1), 1–29.

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. http://doi.org/10.1145/2391224.2391228

Williams, G. M., Kroes, R., & Munro, I. C. (2000). Safety Evaluation and Risk Assessment of the

    Herbicide Roundup and Its Active Ingredient, Glyphosate, for Humans. *Regulatory*

    Toxicology and Pharmacology, *31*(2), 117–165. http://doi.org/10.1006/rtph.1999.1371

Williams, K. (1998). Assessment and the challenge of scepticism. In D. Carr (Ed.), *Education,*

    knowledge, and truth: beyond the postmodern impasse. London, UK: Routledge.

Williamson, B. (2015). Governing software: networks, databases and algorithmic power in the

    digital governance of public education. *Learning, Media and Technology*, *40*(1), 83–105.

    http://doi.org/10.1080/17439884.2014.924527

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%2

    0II.pdf

Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational*

    Psychologist, *45*(4), 267–276. http://doi.org/10.1080/00461520.2010.517150

Wood, P., & Kardash, C. (2002). Critical elements in the design and analysis of studies of

    epistemology. *Personal Epistemology: The Psychology of Beliefs about Knowledge and*

    Knowing, 231–260.

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.

    http://doi.org/10.1016/j.learninstruc.2010.02.003

Wu, C.-H., Hwang, G.-J., & Kuo, F.-R. (2014). Collab-Analyzer: An environment for conducting web-

    based collaborative learning activities and analyzing students' information-searching

    behaviors. *Australasian Journal of Educational Technology*, *30*(3). Retrieved from

    http://ascilite.org.au/ajet/submission/index.php/AJET/article/view/998

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. http://doi.org/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. *Educational Technology & Society*, *10*(2), 120–132.

Xiao, Y., & Lucking, R. (2008). The impact of two types of peer assessment on students' performance and satisfaction within a Wiki environment. *The Internet and Higher Education*, *11*(3–4), 186–193. http://doi.org/10.1016/j.iheduc.2008.06.005

Xie, Y. (2014). knitr: a comprehensive tool for reproducible research in R. In V. Stodden, F. Leisch, & R. Peng D. . (Eds.), *Implementing Reproducible Research*. Chapman and Hall/CRC.

Xie, Y. (2015). knitr: A General-Purpose Package for Dynamic Report Generation in R (Version 1.9). Retrieved from http://cran.r-project.org/package=knitr

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

Yue, Z., Han, S., & He, D. (2013). An Investigation of the Query Behavior in Task-based Collaborative Exploratory Web Search. *Proceedings of the American Society for Information Science and Technology*, *50*(1). http://doi.org/10.1002/meet.14505001091

Yue, Z., Jiang, J., Han, S., & He, D. (2012). Where do the query terms come from?: an analysis of query reformulation in collaborative web search. In *Proceedings of the 21st ACM international conference on Information and knowledge management* (pp. 2595–2598). Retrieved from http://dl.acm.org/citation.cfm?id=2398700

Zeileis, A., & Grothendieck, G. (2005). zoo: S3 infrastructure for regular and irregular time series. *arXiv Preprint math/0505527*. Retrieved from http://arxiv.org/abs/math/0505527

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. http://doi.org/10.1002/tea.20368

Zillien, N., & Hargittai, E. (2009). Digital Distinction: Status-Specific Types of Internet Usage*. *Social Science Quarterly*, *90*(2), 274–291.

APPENDICES

Appendix 1 – Task instructions – (provided to students printed and on screen)

Today's task: Getting setup

Make sure you're paired up and have a login slip

Find seats in your pairs back to back

Open Firefox & login using the details we've given you

In Firefox the sidebar has a chat window – use that to work with your partner *(we're interested in how this chat tool is used!)*

Introduction: What tasks you will be doing

5 minutes – pair up, login

10 minutes to read instructions & do warmup task

45 minutes for the main task

5 minutes at end for a survey

 You'll research an unfamiliar topic together

There's more information available than you can read in the time! You'll need to prioritise what you read, or how deeply you read it

You'll use an 'etherpad' to create a summary report

To start, just follow the instructions on screen – there's a short intro task (10 mins) then the main tasks (45 mins)

Introduction: How you should complete the tasks

Use the information you've been given to create a summary with your partner

Use the etherpad to write it

You should come to consensus when you're done, and both click 'submit'

It should take ~45 minutes; leave 5 minutes at the end of the session

There's a short post-survey, we'll email you with a further task

Introduction: What happens after today

Thanks for taking part!

We'll email you this week with another task:

There's a short survey – complete at home if not today

We'd like you to do some peer assessment, with a short bit of training, and then look at colleague's work from other sessions (before October 17th)

We'll send you feedback, and tell you a bit more about what we're interested in after October 20th

Instructions

One of the transferable skills universities help you learn, is how to work together to find information, and convey that information to others. The tasks here are about trying to help a government minster understand the best supported evidence around a scientific issue – you'll need to work with your partner to find the best supported claims, and then write a summary document for the minister. Don't worry if you don't understand everything you find, focus on the best supported of the claims.

Toolbar instructions

On the right of the screen you'll see a sidebar. You should use the chat tool to communicate with your partner, you can also see what searches you've both made using the 'history' tab (if you can't see the sidebar go to 'view>sidebar' and make sure 'coagmento' is ticked).

At the top of the window, you'll see a row of icons.

- Connect (logs in to both toolbar and study website)
- Snip
- Active Task (links to the 'active task' page on the study website)
- Task Pad (links to an etherpad for each active task)
- Submit Task

Working with your partner

For most of the tasks you'll need to work with a partner. You should:

Share pages to discuss as you visit them

Share specific information or arguments from pages by using the chat box, the snippet tool, or by cutting and pasting extracts from the pages into the shared document ("Editor")

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 Editor to collate this information. There's more than you can read – you'll need to prioritise what you read & how deeply.

Once you've agreed as a pair that you've found enough information, work together to use the Editor to collate the best supported information, you can use URLs to link to particular pages where you refer to them.

Completing the tasks

There are three warmup questions, followed by a main task and then a short survey. You'll be guided through the process as you submit each question.

 If the task requires your partner, you'll be held on an instructions page until they're online too.

Once you've reached consensus with your partner, you should click 'submit', your partner will then have to confirm the submission before you can both move on to the next task.

 Next week, you'll get a chance to give and receive some feedback on your tasks – we'll send you more details on that at the end of the session

CLICK 'ACTIVE TASK' to get started

Appendix 2 – Full Ethics Briefing

Participant Information Sheet

Study title:

Designing a tool to support seeking and evaluating information together

What is the purpose of the study?

We are inviting you to take part in a study evaluating an experimental web-browser to help teams find and evaluate information on the web.

Why have I been approached?

For the purposes of the study I need to recruit a number of pairs of adult participants.

Do I have to take part?

No. Participation is entirely voluntary. If you change you mind about taking part in the study you can withdraw at any point during the sessions and at any time in the two weeks following that session. If you decide to withdraw all your data will be destroyed and will not be used in the study. There are no consequences to deciding that you no longer wish to participate in the study.

What happens during the study?

You will be able to complete this study from any quiet location where you won't be disturbed; you will be working in pairs for some of the tasks. You should make sure you are working in a different location to this partner.

 The study will involve one main task and a warmup:

The first of task will be a practice to get used to the software, we will ask you to find the answer to a question by searching for information on the internet with your partner, this should take no longer than 10 minutes

Then you will work with your partner to find some information, and write a summary document from it. This task should take no longer than 45 minutes.

You'll be using Firefox for these tasks, with an add-on installed (Coagmento) which you will use to collaborate with your partner; you should only use the chat tool on Coagmento to communicate with your partner. While the browser add-on is logged in, it will log the following data:

- Your chat messages;
- the webpages you visit;
- the searches you make in search engines such as Google and Bing;
- snippets you make using the tool

In addition, the tool is linked to a webpage with a shared document space (called an etherpad) which logs your edits. To assist your collaboration, your partner will also be able to access all of this data (except the webpage visits). In collaboration with Dr Dirk Tempelaar we will collate anonymised demographic and psychometric data for analysis about how students learn together.

What are the possible disadvantages and risks of taking part?

We do not anticipate any risks associated with participation in this study.

What are the possible benefits of taking part?

You will gain an insight into how a psychology research project is conducted and what it is like to be a participant in such a study. The tasks are relevant to all students as they are about the kind of transferable skills around collaboration, and finding information, that all graduates should have.

Will my taking part in this study be kept confidential?

Yes, no personally identifying information will be shared.

What will happen to the results of the research study?

You will be given a personal login for the browser add-on. Identifying information will be kept passworded, and will not be associated with the anonymous data collected.

This research forms part of the primary investigator's (Simon Knight) Doctoral research at the Open University supervised by Professors Simon Buckingham Shum and Karen Littleton, Dr Bart Rienties and Mr Fridolin Wild. The research is being undertaken in collaboration with Dr Chirag Shah at Rutgers University, who is developing the Coagmento tool, and Dr Dirk Tempelaar at Maastricht University.

The standard terms of use of Coagmento, including details of the use of data, can be found at: http://www.coagmento.org/terms.php. Data will be shared between the OU team and Dr Shah for research, educational, and dissemination purposes only. All data will be available to Dr Dirk Tempelaar in his capacity as course leader.

Deanonymised data will not be shared other than within the OU supervisory and external-examiner team except where we are legally bound to do so. Pseudonyms will be used in reporting, and any identifying information (such as email addresses) mentioned in chat, search, or page-view logs will be redacted.

The data will be kept in full for the duration of the primary investigator's PhD research or until January 2016 (whichever is later), after which it will be reviewed with a view to full anonymisation (i.e. we will delete any personal data we still hold on you, but keep the anonymised data). The data will be reviewed after that point with a view to full deletion when appropriate. Data stored will be kept in a password protected file in accordance with the Data Protection Act.

Who is organising and funding the research?

The research is organised by Mr Simon Knight, who is a research student at the Open University's Knowledge Media Institute. It is not externally funded.

Who has reviewed the study?

The Open University Ethics Committee has reviewed and approved this study.

Contact for Further Information

Mr Simon Knight

Knowledge Media Institute, Walton Hall, Milton Keynes, MK7 6AA

Email: simon.knight@open.ac.uk

Consent

This study and these materials have been designed in accordance with the British Psychological Society code of ethics (2010), and (subordinate to that code) the guidelines for internet mediated research (2013) in addition to the Open University's ethics and data protection guidance and procedures and the associated regulatory frameworks.

By agreeing to take part in this study, you acknowledge that you have received guidance on what will be requested of you, and give permission for the researchers to use data you provide as described.

 Note that:

- You may withdraw from the study at any time with no adverse consequences
- You may request that your data is destroyed at any time (until such time as we destroy the identifying information associating the records with your information; after January 2016)
- Should you have any questions during the study, you may contact the principal investigator (Simon Knight) or his supervisor Professor Simon Buckingham Shum
- Following the study you will be given a full debrief
- We do not anticipate any risks associated with participation in this study.

Appendix 3 – Participant Debrief

Dear [student]

Thanks for taking part!

You can login to http://edusearch.coagmento.org/instruments/get_my_feedback.php using your student ID as the login to view the feedback you were given for you written text. Remember, this feedback is just one set of perspectives, and we have not vetted the comments. If you feel anything in these is inappropriate, you can email to report the comment.

A key part of the tasks you have worked on over the last two weeks has been literacy and information seeking skills. These are key transferable skills, they are important to all disciplines, and will play a role in whatever professional pathway you choose. You might find the following links useful to explore.

- Information skills materials:
  http://onlinelibrary.maastrichtuniversity.nl/service/information-skills/

- Advice on finding information:
  http://onlinelibrary.maastrichtuniversity.nl/support/finding-information/

- Materials & activities on accessing, finding & reviewing information:
  http://www.open.ac.uk/safari/

Below you will find some information about why we set these particular tasks, and what we are researching.

[NOTE a link to the assessment feedback was given here]

Debrief sheet

Thank you for taking part in this study.

Purpose of the study:

Now that you have completed the task we can give you more information about the study aims. We aim to investigate the processes people go through when using the experimental web browser when evaluating claims on the internet, and seeing how different bits of information are connected. We call these processes 'epistemic commitments', and we are interested in how we can explore them by using a collaborative information seeking system. For example, we are interested in whether we can see processes that indicate people corroborating information across multiple sources, or using 'authoritative' websites when they look for information. We want to explore these processes so that we can understand different ways of working, and how we can support student's research skills using such new tools.

In this study, we asked groups of undergraduate students to complete two types of tasks. In addition, we asked all the groups to write a summary of the best supported claims they had found, and these will be used to compare facets of those written outputs for associations with particular process data traces. We also collected some participant data, and survey instrument data; these will also be explored for associations with task process and written output data. Finally, two types of task were given, one with assigned documents and one which involved searching on the web. In the former, the researchers use their prior knowledge about the documents to give a deeper insight into behaviour in that task, which will then be compared to the open web search trace. As noted at the outset of the study, your personal information will remain secure.

As this is an experimental study it was important to control aspects of the information you were given to work with. Therefore a number of the articles some students were asked to read were modified to: Reduce their length; remove extraneous information; vary document content such that key claims from each document could be identified. As such the specific claims made in each document, and the metadata associated with them, may not reflect their original status. We made these changes to ensure that each document was manageable, and so that we could investigate how each particular document was being used. A full document list with changes marked is available upon request. We deliberately chose topics you might not be familiar with to explore how groups find and evaluate novel information.

<u>Findings</u>

We anticipate that by analysing the processes people go through when they look for information we can develop a scalable tool to support people's collaborative information seeking. We would be happy to share final reports or a summary of findings with you when the research is complete, feel free to contact us.

Thank you again for your participation.

Simon Knight

Appendix 4 – Warmup Task

Please type the answers to the following **three prompts** in your Task Pad (click in the bar at the top of the browser). **You may use the internet** to find the answers**.**

- In 2010 what was the educational expenditure per primary student in [XXX] as a % of GDP?

- In 2010 what was the total health expenditure as a % of GDP in [XXX]?

- How much (in US dollars) does a big mac cost in [ZZZ]?

Warmup – variables,

- XXX= France, Finland, Ethiopia, Estonia, The Gambia, Guinea, Guyana, Guatemala, El Salvador, Ecuador (list of countries not too far from top of alphabet, with data for 2010)

- Yyy = France, Norway, Switzerland, Sweden, Finland, Italy, Ireland, Canada, Greece, Turkey, Japan, Egypt, Russia, (sublist of countries on big-mac index)

If you find the warmup taking too long (over 10 minutes) but you feel you're now comfortable with using Coagmento, you should move on to Task 1.

Appendix 5 – Etherpad Default Text

Welcome to your Task Pad! Use this pad to write your answer:

1) You and your partners need to reach consensus when you're finished and all click 'Submit'.

2) Avoid chatting in this space. Use the chat for that!

3) Remember, we're looking for a report of the best supported claims.

4) Use the evidence you have (you'll need to decide how much to read) to create a summary for the minister.

5) Use 'bold' to make headings

6) you can insert URLs for references, or like (Templaar, 2014)

Good luck!

Appendix 6 – CIS Post-Survey

– Please rate your knowledge of the issues and topics raised in this task prior to today (10=I had expertise in this topic prior to today, 1=I had never heard of these topics prior to today)

– Please rate the level of agreement with your partner regarding the best supported claims in the information you found (10=we strongly agreed about all the best supported claims, 1=we strongly disagreed about all the best supported claims)

– Please rate on a level of 1-10 how well you know your partner? (1 would indicate you hadn't really met them prior to today, 10 would indicate you know them very well)

– Please rate how trustworthy the information you found in this task was on average (10 = The sources were high quality, and the information was very credible, 1=The sources were low quality and the information lacked credibility)

– Please give an estimate rating for the *least* trustworthy page you found where 1 is 'not at all trustworthy' and 10 is 'very trustworthy' (if you can remember it, please post the URL here)

– Please give an estimate rating for the *most* trustworthy page you found where 1 is 'not at all trustworthy' and 10 is 'very trustworthy' (if you can remember it, please post the URL here)

– Please give any other feedback or commentary on examples of types or sources of documents you found, and how trustworthy you found them in the space provided.

Appendix 7 – MDP Post-Survey

– Please rate your knowledge of the issues and topics raised in this task prior to today (10=I had expertise in this topic prior to today, 1=I had never heard of these topics prior to today)

– Please rate the level of agreement with your partner regarding the best supported claims in the information you found (10=we strongly agreed about all the best supported claims, 1=we strongly disagreed about all the best supported claims)

– Please rate on a level of 1-10 how well you know your partner? (1 would indicate you hadn't really met them prior to today, 10 would indicate you know them very well)

– Please rate how trustworthy you thought Document X was (10 = This source was high quality, and the information was very credible, 1=This source was low quality and the information lacked credibility) Repeated for each document, with a link to the document

Appendix 8 – Task Sourcing

For the MDP task, a set of documents were selected. These sources were read, and sections which were unique to that document or provided extraneous detail (particularly detail which would require further research or prior knowledge to understand) were removed.

The topics chosen must be:

- Confined to a topic researchable in a short period of time
- Accessible to a lay person with no disciplinary knowledge of the topic
- Genuinely contested
- Not widely known about
- Topics with a variety of sources (i.e. not only news, or scholarly articles) written about them

To source documents for the MDP task a number of strategies were taken as follows:

1) First, the "most controversial" topics on English Wikipedia (as computed by the number of edit reversions to each article; available http://wwm.phy.bme.hu/) were explored, selecting broadly scientific articles for close analysis.
    a) This gave the following set as good candidate articles: Global warming; chiropractice; homeopathy; Prem Rewat; evolution; creation science
    b) However, many of these cases are, although controversial, not *scientifically contested.* Moreover, the articles cover rather large swathes of information, while for a research project a smaller chunk of information would be preferable.
2) Secondly, articles under the 'Category:Scientific controversies' on Wikipedia were explored. In this case issues were similar to the first; the articles were very general, and in many cases discussed resolved (mostly historic, and many implausible) and often obvious controversies (i.e. scandals) as opposed to disputes (i.e. areas of contested claims).
3) Thirdly, the child categories of Category:Scientific controversies were explored including 'Psychiatry controversies'; 'environmental controversies'; 'medical controversies'. Some possibilities
    a) From Environmental controversies:
        i) https://en.wikipedia.org/wiki/Atrazine_controversy#Controversy altrazine (this would be a good option, but is rather too close to Glyphosate issue below, one or other should be chosen).
        ii) https://en.wikipedia.org/wiki/Criticism_of_the_IPCC_Fourth_Assessment_Report A good sub-topic article on climate change debates
        iii) https://en.wikipedia.org/wiki/Fracking_controversy#Environmental_and_human_health_concerns Fracking would be a very current case, but might be too well known about.
    b) Medical controversies
        i) https://en.wikipedia.org/wiki/Red_yeast_rice#Clinical_evidence see case two below
        ii) https://en.wikipedia.org/wiki/Aspartame_controversy This Wikipedia article is reasonably comprehensive on a part of this controversy (a viral campaign), and the main Aspartame article gives a more comprehensive overview. However, interestingly google SERP gives many very critical top hits. This gives opportunity for interesting critical discussion around Wikipedia value versus other sources (and using search v. link following in Wikipedia to source authoritative articles)
        iii) https://en.wikipedia.org/wiki/Bisphenol_A This would be a good case, but the Wikipedia article is exceptionally comprehensive
    c) No suitable candidates for exploration from other categories
4) A fourth strategy was to explore the Guardian newspaper's "controversies in science" science subsection at http://www.theguardian.com/science/controversiesinscience. The first 375 of

573 article headlines (25 pages of headlines, going back to December 2008) were scanned for suitable content with article views made for suitable candidates.

a) The role of null findings in psychology (and some specific examples, including the Bystander effect) and their relationship to biased reporting of psychology findings is discussed in this article
http://www.theguardian.com/science/blog/2013/feb/27/psychologists-bmc-psychology

b) The role of tentative findings in making policy decisions is discussed in this article (which provided sourcing for further articles as below)
http://www.theguardian.com/environment/blog/2013/jun/17/wait-conclusive-science-concerned-issue

c) GM-food debate http://www.theguardian.com/environment/2010/jun/02/gm-food-public-dialogue

d) Acupuncture for migraines and the placebo effect
http://www.theguardian.com/science/2009/jan/20/acupuncture-headaches-migraine-placebo

e) Chiropracy and the Simon Singh libel case
http://www.theguardian.com/science/2010/apr/15/simon-singh-libel-case-dropped1 see also e.g. links around (or/and controversial edits on) the Wikipedia article:
https://en.wikipedia.org/wiki/Talk:Chiropractic#Research_status

Document Format Process for MDP

- Download/save html pages and files (and inspect the first set of links, using one of these tools to retrieve those: http://gnu.huihoo.org/wget-1.5.3/html_chapter/wget_3.html or http://www.httrack.com/)
- Edit documents appropriately to remove extraneous information, ensure consistency of content, and to vary key claims across documents
- Create a website with documents on for browsing, removing surface features for a fairly standardised appearance across. For copyright reasons this should be password protected.
- Develop document model indicating the key claims and metadata information in each document. In the first instance compare documents for claims so that key variables can be modified to differ across sources, after that step tf;idf can be used to provide a model of document variance and relationship between claims made in discourse and similarity to those in documents.
  - Use

Document Download Process for CIS

- Download each page visited http://gnu.huihoo.org/wget-1.5.3/html_chapter/wget_3.html or http://www.httrack.com/
- Use metadata extraction tool to extract metadata
- Use summary/keyword extraction to extract keyterms (use these in dialogue processing)
- Use tf;idf across set to create index; this offers potential scope to associate claims with documents (or at least, to give a confidence for any claim made)

Case 1: glyphosate (MDP task)

The original Guardian article http://www.theguardian.com/environment/blog/2013/jun/17/wait-conclusive-science-concerned-issue was used to source further articles (although it is unlikely to be a key article for participants in its own right). This article led directly to article (1) and (2) below. Further research around "glyphosate" AND "urine", and following links through stories.

1) Original 'study' press release http://www.gmfreeze.org/news-releases/225/ and the study itself:
http://webcache.googleusercontent.com/search?q=cache:ZUfJbfZ5hKUJ:www.foeeurope.org/sites/default/files/glyphosate_studyresults_june12.pdf+&cd=1&hl=en&ct=clnk&gl=uk&client

pdf: http://www.foeeurope.org/sites/default/files/glyphosate_studyresults_june12.pdf)

2) http://www.fwi.co.uk/articles/13/06/2013/139515/scientists-dismiss-39unreliable39-glyphosate-study.htm which is also in http://www.sciencemediacentre.org/expert-reaction-to-press-release-from-fote-and-gm-freeze-about-glyphosate-in-urine/ (see the 'about' for why these are interesting pairings)

3) http://blogs.wsj.com/brussels/2013/06/13/study-youre-in-trouble-roundup/ news coverage

4) http://ksj.mit.edu/tracker/2013/04/discover-blogger-keith-kloor-stumbles-ne heavy criticism of story which attributes various diseases to the substance...interestingly I found this article by searching for the quote "It has been very, very extensively studied." which I got from the above WSJ article and had similar concerns about to this author (attribution).

5) http://www.nytimes.com/2013/09/20/business/misgivings-about-how-a-weed-killer-affects-the-soil.html?pagewanted=all&_r=2& nice summary of use as fertilizer but probably not for use

6) Review of literature on mineral/soil damage of fertilizer finding generally little evidence for it http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3479986/

7) http://blogs.prevention.com/inspired-bites/2013/04/26/a-new-study-highlights-the-risks-of-genetically-modified-foods-and-the-chemicals-used-on-them/ critical blog, makes use of "MIT study" (see ksj.mit.edu post above)

8) http://mobile.reuters.com/article/idUSL2N0DC22F20130425?irpc=932 The Reuters article referred to above and the study itself http://www.mdpi.com/1099-4300/15/4/1416 note the same (MIT) author cites Andrew Wakefield in an MMR study...favourably http://blogs.prevention.com/inspired-bites/2013/04/26/a-new-study-highlights-the-risks-of-genetically-modified-foods-and-the-chemicals-used-on-them/ as an aside there's also an interesting wiki discussion on this https://en.wikipedia.org/wiki/Wikipedia_talk:Identifying_reliable_sources_%28medicine%29/Archive_7#Review_of_Monsanto.27s_Roundup_herbicide

9) http://www.globalresearch.ca/monsanto-roundup-glyphosate-weedkiller-in-our-food-and-water/5339244 long article broadly supportive of FoE stance on Roundup

10) http://stopogm.net/sites/stopogm.net/files/webfm/plataforma/KremerPreface2009.pdf journal article on agricultural impact

11) http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1241861/ risk

12) http://naldc.nal.usda.gov/download/17918/ a review document of roundup risk hosted on US govt agricultural library, published in "Pest Management Science"

Case 2: Red Yeast Rice (CIS task)

In this case:

The minister has heard that French officials have raised some concerns about the safety of 'Red Yeast Rice' and would like a briefing on its potential risk.

To seed the search, a few of the articles cited in http://www.nhs.uk/news/2008/06June/Pages/Redyeastriceandheartdisease.aspx could be given (Daily Express, etc.) These are now rather low in the SERP so unlikely to be returned otherwise. There's also a rather brilliant Daily Mail headline (Quack medicine: Peking duck is 'better for your heart than statins'. **Daily Mail**, June 10 2008).

This is an interesting case because:

1) The Wikipedia article is not particularly high quality

2) Google searches for 'Red Yeast Rice' bring up rather different content to 'Monascus purpureus' (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).

Case 3: Aspartame (Potential CIS task)

https://en.wikipedia.org/wiki/Aspartame_controversy This Wikipedia article is reasonably comprehensive on a part of this controversy (a viral campaign), and the main Aspartame article gives a more comprehensive overview. However, interestingly google SERP gives many very critical top hits. This gives opportunity for interesting critical discussion around Wikipedia value versus other sources (and using search v. link following in Wikipedia to source authoritative articles)

Appendix 9 – Peer Assessment Introduction

Dear students

A key skill for students, and professionals, is being able to read and evaluate written texts. If you completed the 'alternative student project' task in labs this week, the 2nd part of that project is to engage in some evaluation of written texts using a 'rubric' a scoring sheet to give scores to different qualities of a text.

Please go to http://edusearch.coagmento.org/index.php to complete this task, no later than Monday 13th October. You must complete the 'lab' and this 'at home' tasks to pass the 'alternative student project'.

Please use your Student ID to login

Appendix 10 – Marking Guidance for Student Assessors

To help you understand how you will be marking, we've written some example answers to help you work with the marking-rubric. This part of the task involves you reading those texts and marking them. For each of the questions, read all three texts and use the form to submit your ratings.

We anticipate this part of the task taking 30-45 minutes.

You'll be assessing 3 texts, remember the prompt for these was "***A review is coming up for the license of Glyphosate, the official would like to know the best supported claims around its risks***" and a set of documents were given to source from.

You'll be assessing 3 texts, remember the prompt for these was "***The official has heard that French officials have raised some concerns about the safety of 'Red Yeast Rice' and potential contamination, and would like a briefing on its potential risk.***" and the responders were asked to search the internet to find documents to source from.

For each element of the rubric below, one of the texts will be in the low (1-3) range, one in the medium (4-6) range, and one in the high (7-9) range:

1. Topic coverage – The text covers a range of different topics and relates them to the question (the risks of the substance)
2. Range of sources coverage – The text uses a range of sources
3. Quality of sources – The text evaluates the quality of sources cited
4. Clarity of the claims – The text states claims clearly. It gives quotations, numeric indications, and figures where appropriate
5. Quality of evaluation – The text evaluates evidence
6. Synthesis of information – The text synthesizes information from across sources

For each text you will also indicate three improvements that could be made.

Topic coverage – The text covers a range of different topics and relates them to the question (the risks of the substance)

A score of 1 indicates that the text is very focused on a small number of concepts or topics, a score of 9 indicates that the text includes a range of core concepts or topics. To score 9 at least 3 general themes, with key concepts described would need to be present.

Some example themes from the glyphosate question include: the relevance of glyphosate in urine; the risks of glyphosate for human health; the agricultural risks of glyphosate use.

Some example themes from the red yeast rice question include: the risks of red yeast rice as a statin product; the risks of poor-labelling of the quantity of active ingredient; the risks of red yeast rice contamination (especially with citrinin).

1. – The text has a very narrow focus, on one main topic with little breadth to it
2. –
3. –
4. –
5. – The text discusses one topic in depth, or gives superficial coverage of at least three topics
6. –
7. –
8. –
9. – A range (three or more) of topic areas are discussed in depth and related to the question

Range of sources coverage – The text uses a range of sources

1. No sources are cited
2. –
3. –
4. –
5. –The text cites only a small number of sources (roughly 4)
6. –
7. –
8. –
9. A range of core sources are cited

Quality of sources – The text evaluates the quality of the sources cited

A score of 1 indicates that the text only cites poor quality sources, a score of 9 indicates that the text evaluates sources specifically (it makes reference to the authors, publication venues, date of publication, etc.) and uses high quality sources.

1. – There is no evaluation of source quality at all
2. –
3. –
4. –
5. – The text only evaluates some sources or/and gives only partial evaluation of sources
6. –
7. –
8. –
9. – The text makes specific reference to evaluation of sources (it makes reference to the authors, publication venues, date of publication) and uses high quality sources

Clarity of the claims – The text states claims clearly. It gives quotations, numeric indications, and figures where appropriate

A score of 1 indicates that the text makes only very general and unclear claims. A score of 9 indicates that the text uses precise language backing up claims with appropriate use of quotations, etc. and clearly stating the core concepts being used.

1. – The text is very unclear and general, no keyterms or concepts are referred to, and the points made are very general. No headings, tables, figures, numeric data, or quotations are given.
2. –
3. –
4. –
5. – The text is too general in places, or/and does not always use keyterms and concepts. There might be limited use of headings, tables, figures, numeric data, or quotations.
6. –
7. –
8. –
9. – The text uses keyterms and concepts clearly, there is a coherent structure organising the specific concepts around general themes. Headings, tables, figures, numeric data, or quotations are used effectively. (note, not all need to be used, but at least some should be given).

Quality of evaluation – The text evaluates evidence

A score of 1 indicates that no evaluation is made, claims might be stated but they are not judged. A score of 9 indicates that all the claims made are evaluated, and credibility assessments are stated.

1. – Claims made in the text are not evaluated at all for their credibility
2. –
3. –
4. –
5. – Claims made in the text are partially, or inconsistently evaluated
6. –
7. –
8. –
9. – Claims made in the text are well evaluated, and credibility assessments are clearly stated

Synthesis of information – The text synthesizes information from across sources

A score of 1 indicates that individual claims - from single, or multiple documents - are stated without being related. A score of 9 indicates that there is a narrative thread through the text, claims are integrated from across multiple sources, and relationships between claims are made explicit.

1. – The text makes no attempt to synthesise or relate claims. It might read like a list, with claims stated but the relationships single or multiple documents left unstated.
2. –
3. –
4. –
5. – The text gives partial, or inconsistent synthesis.
6. –
7. –
8. –

9. – The text builds a narrative thread, claims are integrated from across multiple sources, and the relationships between them are made explicit.

Appendix 11 – CIS Exemplar Texts

8:1.3   Text 1

8:1.4   Text 2

8:1.5   Text 3

8:1.6   Introduction to red yeast rice

Red yeast rice is a red fermented rice cultivated with the mould Monascus purpureus. https://en.wikipedia.org/wiki/Red_yeast_rice. It is used to lower cholesterol levels. A study published in the American Journal of Clinical Nutrition (1999) in a double blind placebo controlled 12 week trial found significant decrease in cholesterol in those taking the red yeast rice supplements http://ajcn.nutrition.org/content/69/2/231.full . Red yeast rice might be a good alternative for people who cannot take statins according to a 2009 article reporting a study which found that those who could not take other statins still experienced positive effects, but without many of the side effects, when taking red yeast rice instead http://news.health.com/2009/06/15/red-yeast-rice-supplements-cholesterol/ Animal studies with high doses of red yeast rice show no damage, with only minor side effects in human trials. Furthermore, given that red yeast rice is used so widely and has been for thousands of years in Asian countries, scientists consider red yeast rice to be safe in the long term. http://www.medicinenet.com/red_yeast_rice_and_cholesterol/article.htm

8:1.7   Statin risks

However, because the supplements have the same active ingredient as statins and thus the same risks of side effects, so the American Heart Association warns against using it until further studies are conducted. http://www.webmd.com/vitamins-supplements/ingredientmono-925-Red%20Yeast%20Rice%20%28RED%20YEAST%29.aspx?activeIngredientId=925&activeIngredientName=Red%20Yeast%20Rice%20%28RED%20YEAST%29
 In 2013, the French agency ANSES held a consultation in response to 25 reports of problems (muscle & liver damage) from taking red yeast rice, and suggested that: people taking statins, intolerant to them, pregnant or breastfeeding, children or adolescents, over 70, consuming large amounts of grapefruit, etc. should not take red yeast rice https://www.anses.fr/en/content/nutrivigilance-anses-launches-consultation-today-its-opinion-food-supplements-containing-red

Red yeast rice is an Asian ingredient and supplement, it is a bright reddish purple fermented rice, which acquires its colour from being cultivated with the mold Monascus purpureus. The substance is used to lower cholesterol levels, with evidence from a double blind placebo controlled 12 week trial (published in the American Journal of Clinical Nutrition (1999)), indicates its potential to significantly decrease cholesterol in those taking the supplement http://ajcn.nutrition.org/content/69/2/231.full. This is because red yeast rice contains the same active ingredient as many statins (lovastatin), and according to a meta-analysis of 91 RCTs of over 12 weeks with almost 70,000 participants statins reduced LDL-cholesterol by 24-49% (See http://www.biomedcentral.com/1471-2296/4/18).Both of these studies provide strong support for the use of substances such as red yeast rice to control cholesterol, especially given that both articles are peer reviewed and based on strong methods – randomised controlled blind trials, and a meta-analysis which reviews multiple findings. However, the first article is now rather old (1999), while the latter is not specifically on red yeast rice, but rather statins in general.Moreover, red yeast rice has a variety of side effects including muscle and liver damage. Because red yeast rice can contain the same active ingredient as statins, the American Food and Drug Administration regulates red yeast rice, so in the US many of them do not contain much of the active ingredient, with a large variation across them (from the FDA: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/2007/ucm108962.htm)One article (http://www.medicinenet.com/red_yeast_rice_and_cholesterol/page4.htm#how_safe_are_red_yeast_rice_products) does suggest that animal studies with high doses of red yeast rice indicate its safety, and that because red yeast rice is used so widely in Asia and has been for thousands of years, scientists consider red yeast rice to be safe in the long term. However, although they say the article is reviewed by a medical Doctor (in January 2014), no

Red yeast rice is an Asian ingredient and supplement, it is a bright reddish purple fermented rice, which acquires its colour from being cultivated with the mold Monascus purpureus; in Japanese it is referred to as koji – a grain or bean with mould culture.It has been used for centuries in Asian countries to lower cholesterol levels; because it has been used for so long it is seen as fairly safe by most people, and it is used in both cooking (including to colour Peking duck) and natural medicine by many people. A study published in the American Journal of Clinical Nutrition (1999) in a double blind placebo controlled 12 week trial found significant decrease in cholesterol in those taking the red yeast rice supplements.Red yeast rice might be a good alternative for people who cannot take statins according to a 2009 article reporting a study which found that those who could not take other statins still experienced positive effects, but without many of the side effects, when taking red yeast rice insteadThe benefits of red yeast rice may partly be because it contains several compounds rather than pharmaceutical products which focus on a single active ingredient to have its effect. This may be especially true is a good brand is used (the author of this article suggests Solaray) then issues about quality control of red yeast rice can be addressed.
In a daily mail article (Quack medicine: Peking duck is 'better for your heart than statins' from 2008) they note

An article suggests that because red yeast rice is like a statin, it is also risky. It also suggests that cholesterol is not 'bad', if it does need lowering then other methods (exercise and diet) are better, but they claim cholesterol is used by the body to help your body heal and repair, and that very low cholesterol levels are also bad. http://articles.mercola.com/sites/articles/archive/2009/09/10/why-you-should-avoid-red-rice-yeast.aspx

8:1.8 Active ingredient quantity control risks

However, suppliers may also be adding additional statin ingredients rather than just those naturally occurring https://en.wikipedia.org/wiki/Red_yeast_rice. In June 2014, red yeast rice was recalled over chemical contamination, with a big lot containing undeclared lovastatin (the active ingredient in statins). http://www.digitaljournal.com/life/health/red-yeast-rice-recalled-over-chemical-contamination/article/387813.
According to Reuters (2008), a ConsumerLab report shows active ingredient levels varies dramatically (with some showing much lower levels than would have clinical effect), and others containing contaminants (4 out of 10). Although all of the products were labeled as containing 600 milligrams of red yeast rice, their lovastatin content varied from 0.1 milligrams, found in Walgreen's Finest Natural Red Yeast Rice, to 10.6 milligrams. The four least potent formulations, including the Walgreen's product — contained citrinin. Other contaminated products included Natural Balance Red Yeast Rice Concentrated Extract, Solaray(r) Red Yeast Rice, and VegLife(r) 100% Vegan Red Yeast Rice. People who want to get the cholesterol lowering benefits of red yeast rice should consume enough to get 5 to 15 milligrams of lovastatin, based on their initial cholesterol level, according to Cooperman. http://www.reuters.com/article/2008/07/09/us-contamination-common-idUSCOL97022820080709

8:1.9 Active ingredient and toxicity risks

According to science daily (2010), reporting on JAMA and Archives Journals reports, the active ingredients in red yeast rice varies substantially, with 1/3 of the 12 tested also containing a toxic compound (citrinin) http://www.sciencedaily.com/releases/2010/10/101025161036.htm
An examiner article (2012) According to the Berkeley Wellness Alerts article, "Working with the testing company ConsumerLab.com, researchers from the University of Pennsylvania analyzed 12 red yeast rice products from major manufacturers. The results, in the *Archives of*

citations are given to back up the claims. Furthermore, given the potential for variation in red yeast rice, even if it is used widely in Asia it does not mean it is safe to consume without medical consultation.Indeed, authoritative organisations such as the American Heart Association warn against using it until further studies are conducted.In fact, in 2013 the French Agency for Food, Environmental and Occupational Health & Safety (ANSES) held a consultation in response to 25 reports of problems (muscle & liver damage) from taking red yeast rice, and suggested that: people taking statins, intolerant to them, pregnant or breastfeeding, children or adolescents, over 70, consuming large amounts of grapefruit, etc. should not take red yeast rice https://www.anses.fr/en/content/nutrivigilance-anses-launches-consultation-today-its-opinion-food-supplements-containing-red This indicates concern about its risk, especially to particular groups of people, and reports of health impact from taking red yeast rice. That an official organisation thinks further research (and potentially regulation) is required should be a concern.There are some claims in the popular press about the benefits of red yeast rice – for example, the Daily Mail article: Quack medicine: Peking duck is 'better for your heart than statins' from 2008 – claimed that the risk of death from heart disease was cut by a third, and cancer by two thirds over a 5 year placebo-controlled study according to the article in the American Journal of Cardiology. According to the University of Pennsylvania researchers, the effects could not be explained by the statin content of the red yeast rice alone. http://www.dailymail.co.uk/health/article-1025277/Quack-medicine-Peking-duck-better-heart-statins.html However, these results will need further testing, and no side effects were reported n in the article. In particular, although it might be the case that red yeast rice has other active ingredients (in addition to lovastatin), the isolation of these ingredients and their impact on health requires further research, as indicated in the original article http://www.sciencedirect.com/science/article/pii/S0002914908003536 – a good reminder that tabloid newspapers are likely to often report rather simplified single-sided perspectives on health stories.

Therefore, despite claims that red yeast rice may have positive impact, it is not without concerns of risks, and further research is needed before concluding what regulation should be implemented. The

that red yeast rice consumption in a many-year placebo-controlled study published in the American Journal of Cardiology was related to a reduced risk of death from heart disease, and of cancer. The authors of the research (at University of Pennsylvania) claim that these effects cannot be explained by the statin content of the red yeast rice alone.
However, suppliers may also be adding additional statin ingredients rather than just those naturally occurring. In June 2014, red yeast rice was recalled over chemical contamination, with a big lot containing undeclared statin active ingredient. Given the risks of statins, this is especially concerning. Indeed, a report shows active ingredient levels varies dramatically (with some showing much lower levels than would have clinical effect), and others containing contaminants. Also, even though all of the bottles were the same size, their active ingredient varied from very smalll amounts to a few milligrams. Moreover, the four least potent formulations, contained a toxin. In order to get benefits from red yeast rice people should consume enough to get a few milligrams of the active ingredient a day. According to science daily (2010), reporting on JAMA and Archives Journals reports, the active ingredients in red yeast rice varies substantially too. In addition, some of the dozen tested varieties of red yeast rice also contained a toxic compound.
One article (2010) notes that the idea 'natural' products are safer is not true, but that because of lobbying the supplements industry can sell whatever they like as long as they don't make health claims on the bottle.
However, quality control is an issue with varying amounts of the active ingredient in it. The

Internal Medicine in late 2010, indicated "dramatic variation in active ingredients" among the capsules. "The amount of lovastatin in a daily recommended dose ranged from 0.2 to 14 milligrams. Not only are those low doses, compared to the 20 or 40 milligrams in each prescription lovastatin pill, but the extreme variability in lovastatin levels means the manufacturers' recommendations are quite arbitrary. In addition, one third of the products were contaminated with citrinin, a potential kidney toxin."

http://www.examiner.com/article/are-patients-finding-red-yeast-rice-without-toxins

strong peer-reviewed studies noted in the first paragraph give some evidence for the potential positive effect of red yeast rice, but further research is required, and people should be made aware of its medical nature before consuming it.

growth of the fungus also leads to contamination, for example with a toxin.

————————————————

Appendix 12 – MDP Exemplar Texts

8:1.10   Text 1

8:1.11   Text 2

8:1.12   Text 3

8:1.13   Agriculture factors

Glyphosate is very widely used for weed control. However, there are some concerns of "substantial evidence accumulated on multiple adverse effects on crop health and productivity and soil-plant-microorganism interactions mediated by this herbicide" (Journal of Agronomy preface). Because of this, there is a suggestion that switching to alternative weed control systems might reduce risks (Journal of Agronomy preface).

Although it is generally considered effective and "toxicologically and environmentally safe" (mini-review), "the use of this virtually ideal herbicide is now being threatened by the evolution of glyphosate-resistant weeds" (mini-review). The rise in crops genetically modified to be resistant to glyphosate raises the risk of glyphosate resistant weeds evolving; diversity in weed management is recommended (mini review).

8:1.14   Presence in urine

There have been recent concerns about the presence of glyphosate markers in human urine (GM Freeze press release), finding "urine tests of 182 volunteers in 18 countries across Europe, found that on average 44 per cent of samples contained traces of glyphosate. Of the 10 samples taken in the UK, seven had weedkiller traces." (GM freeze press release). They raise the concern that the samples were from city dwellers, and we don't really know how the glyphosate got into their bodies or what effect it has. They were reporting a study (glyphosate

Recent concerns have been raised regarding the presence of glyphosate markers in human urine. These concerns seem to be based on a report produced by a German lab which measured glyphosate trace in samples from across Europe, finding large variations in its presence which they say might be associated with diet. Seven of the samples taken in the UK had weedkiller traces, and there is some suggestion that there might be potential damage to wildlife and people. The report itself claims to have used a validated method for detection of glyphosate markers (although it does note that baseline levels vary and these ones are taken only for Bremen). They report sample concentrations for all samples individually, noting 10-90% of samples from each EU state detected glyphosate, with 12 (6.6%) over the reference value/baseline.

However, the sources raising concerns around the safety of glyphosate, and its presence clearly have an agenda to push, and are interested in reducing use of herbicides in farming techniques. It also isn't clear how independent the scientific reviewer (glyphosate report) is – and the report was commissioned by those groups with an agenda to push (GMFreeze, etc.). In addition, the results appear to have been taken out of context. For example, only 6.6% of samples were over the reference value, but the results are reported to suggest a much larger issue. In addition, the report of every sample, but not of total means and standard deviations is very unusual in scientific reports.

These concerns are backed up by evidence from a variety of scientists who do not appear to have any association with the producers of glyphosate, or with

8:1.16   Glyphosate in urine

Recent concerns have been raised by Friends of the Earth regarding the presence of glyphosate markers in human urine (GM Freeze press release). They claim that a high proportion of urine samples taken in the EU contained glyphosate markers and that we should be concerned about damage to wildlife and human health. This is based off a report (glyphosate report) which used a validated method to detect glyphosate in urine, and found a high presence across the samples. Scientists in the Science Media Centre raise various concerns about the report for not publishing a proper methodology, and not publishing in a peer review journal. They suggest that it's hard to interpret the results and that the levels of glyphosate in urine are probably not a major concern. There seems to be less bias in the scientific group (Science Media Centre) given they are from a range of academic institutions, than the Friends of the Earth piece which is only from one perspective.

8:1.17   Health risks of glyphosate

In another study (Entropy article), though, the claim is made that glyphosate is associated with lots of medical conditions including cancer, heart disease, Alzheimer's,

report) produced by a German lab which measured glyphosate trace in samples from across Europe, finding large variations in its presence which they say might be associated with diet (glyphosate report).

These results were rejected by scientists (reported in Farmer's Weekly) who say the research is not very good "it is not scientific, and cannot be taken seriously by anyone" (Farmer's Weekly). They say "It is not surprising to find glyphosate in urine…Glyphosate is not metabolised by the human body but excreted into the urine and faeces" (Farmer's Weekly).

8:1.15 Association with disease

Another study says glyphosate's "negative impact on the body is insidious and manifests slowly over time as inflammation damages cellular systems…consequences are most of the diseases and conditions associated with a Western diet, which include gastrointestinal disorders, obesity, diabetes, heart disease, depression, autism, infertility, cancer and Alzheimer's disease" (Entropy article). According to Reuters the author (Seneff and Samsel) are both scientists, employed by MIT and (previously) by the government respectively (Reuters). Another piece by Robin O'Brien notes that genetically engineered crops mean enhanced "profitability of the chemical companies by enabling the increased sale of their chemical treatments and weed killers" (O'Brien), suggesting negative effects might only be emerging now. Another blog talking about the paper (Kloor) suggests the authors of the study aren't qualified to make the claims they've made in the paper.

groups campaigning against herbicide use. These scientists criticise the claims raised around risks of glyphosate presence in urine for not properly publishing methodology, not publishing in a peer review journal, and for the results which indicate glyphosate levels are "vastly lower than the level at which it would be cause for concern" or be of risk to health. While certainly some people using this critique (e.g. farmers, the producers of glyphosate) might have a vested interest in continuing its use and therefore be somewhat biased sources, the scientists perspective appears to be from a range of academic institutions and with no tie to glyphosate.

In addition, there appears to be other evidence from peer review sources indicating that even farmers – who would have frequent exposure to glyphosate – do not have dangerous amounts of glyphosate in their urine, and a significant proportion (40%) did not have detectable levels of glyphosate in their urine, with those using precautions such as rubber gloves having lower levels. The various sources come from different perspectives, and some of them appear to be biased as outlined above. There is thus different evidence from sources of varying quality indicating a low risk with regard to the presence of glyphosate in human urine.

etc. According to Reuters this study (Entropy article) comes from two scientists – one at MIT, and one ex-government environmental contractor and member of the Union of Concerned Scientists.. Reuters say that in lots of countries consumer groups and scientists are warning that heavy use of glyphosate causes problems for plants, people and animals, and that there is a deadline of 2015 to decide if glyphosate should be regulated. But Monsanto, who sell a herbicide with glyphosate in it, says its safety is very good and that's why it's so popular.

8:1.17.1 lyphosate use

According to Reuters, farmers like glyphosate and the genetically modified crops because it means that they can spray the herbicide directly onto the crops to kill weeds, but without harming the crops. Use of glyphosate has doubled in the last 6 years. However, another piece by Robin O'Brien notes that genetically engineered crops mean enhanced "profitability of the chemical companies by enabling the increased sale of their chemical treatments and weed killers" (O'Brien). O'Brien suggests that negative effects might only be emerging now after a few generations of spraying crops with glyphosate, so we urgently need to monitor its effects and control its use. However, the mini-review suggests that glyphosate is very safe.

There are therefore various risks people should be aware of, there are potential risks around its health impact although its presence in urine is probably not a major concern. The evidence so far (particularly contained in the 'mini-review') indicates that it is safe, although it should be monitored over time to ensure new risks do not emerge.

8:1.18

8:1.19

Appendix 13 – Marking Scheme Developed for Expert Assessors

The coding scheme consists of a set of binary classes. Note that Synthesis, Sdiv, and Squal might be additive – that is, the T/F is given to the highest present class, although lower classes might also be present they're marked as F. E.g. if FullSynth is displayed, mark T for that but not for Lists or SomeSynth. The scheme is:

1. **Topic** –
   1.1. **CIS/MDP** Citrinin – is citrinin contamination mentioned as a risk of red yeast rice or not/ Urine – is the presence of glyphosate (or associated markers) in urine mentioned
   1.2. **CIS/MDP** Concentration – is the fact that the red-yeast-rice active ingredient often varies in concentration or strength (or other synonym) mentioned/ Disease – is the risk (or not) of glyphosate for human health mentioned
   1.3. **CIS/MDP** Statins (or lovastatin) – is the fact that RYR has the same active ingredient as statins mentioned and any of the associated risks/ Agricultural – is the risk of glyphosate on crops, ecosystems, and agricultural generally mentioned
2. **Synthesis** –
   2.1. Lists – claims are listed, perhaps listing each claim found in each source used sequentially, with little or no attempt at thematic organisation, consolidation, synthesis or expansion to general claims
   2.2. SomeSynth – Some claims are listed but there might be some attempt at synthesis or organisation, e.g. connecting claims from multiple sources, giving concluding remarks summarising the listed the information, etc. Listing claims and then providing a basic single sentence summary does not warrant a '2'.
   2.3. FullSynth – The text gives flowing prose, relating claims from across multiple sources, and perhaps (although not necessarily) giving a conclusion or summary
3. **Sdiv** (source diversity) - note repetition of the same source counts only once –
   3.1. 0-1 sources are directly or indirectly referred to in the text
   3.2. 2-3 sources are directly or indirectly referred to in the text
   3.3. 4+ - 4 or more sources are directly or indirectly referred to in the text
4. **SQual** (source quality)
   4.1. EvalNone – no evaluation of sources or information is given; claims are stated uncritically with no juxtaposition, reference to source properties (e.g. scientific authority, methods, publication date, etc.), or information-credibility (e.g. credibility of claims made, balance of claims across sources, etc.)
   4.2. EvalSome – patchy evaluation is given, perhaps referring to the qualities of 1 or 2 sources, or critiquing claims based on source properties. Evaluation might be implicit in places, e.g. presenting two pieces of contrasting evidence alongside each other with no explicit comparison or evaluation of their relative properties.
   4.3. EvalFull – Evaluation is balanced and consistent leading to a conclusion regarding information credibility

Appendix 14 – Multiple Regressions for Survey Variables Predictive of Outcome

All survey items

CIS

	t	p	β	F	df	p	adj. R^2
Total Score							
Overall model				.779	128	.660	-.0178
Age	-1.31	.19	-0.26				
Search experience	.0.37	.71	0.13				
Collaborative satisfaction	-0.77	.44	-0.11				
Task satisfaction	0.47	.64	0.07				
Addon intuitiveness	1.48	.14	0.26				
GPA	0.69	.49	0.14				
ISEQ General	1.03	.31	0.36				
ISEQ Justification	0.43	.67	0.16				
Topic knowledge	0.72	.47	0.10				
Partner agreement	-0.83	.41	-0.13				
Partner familiarity	1.17	.25	0.09				
Topic Score							
Overall model				1.387	128	.1865	.02973
Age	1.074	.28	-0.09				
Search experience	-1.14	.25	-0.14				
Collaborative satisfaction	-0.43	.67	-0.02				
Task satisfaction	1.15	.25	0.07				
Addon intuitiveness	1.42	.16	0.10				
GPA	0.90	.37	0.07				
ISEQ General	1.40	.17	0.20				
ISEQ Justification	0.41	.68	0.06				
Topic knowledge	-1.42	.16	-0.08				
Partner agreement	-0.71	.48	-0.04				
Partner familiarity	2.34	.0208*	0.07				

Synthesis Score

Overall model				0.8037	128	.6361	-0.01578
Age	-0.77	.44	-0.05				
Search experience	1.23	.22	-.13				
Collaborative satisfaction	-0.66	.51	-0.03				
Task satisfaction	0.31	.76	0.01				
Addon intuitiveness	0.29	.77	0.02				
GPA	0.30	.77	0.02				
ISEQ General	0.81	.42	0.09				
ISEQ Justification	0.64	.53	0.07				
Topic knowledge	2.11	.0369*	0.09				
Partner agreement	0.92	.36	0.04				
Partner familiarity	-0.23	.82	-0.01				

Source Diversity Score

Overall model				0.8238	128	.9811	-0.05716
Age	-0.95	.34	-0.07				
Search experience	-0.36	.72	-0.06				
Collaborative satisfaction	-0.06	.95	-0.00				
Task satisfaction	-0.12	.91	-0.01				
Addon intuitiveness	1.25	.21	0.08				
GPA	-0.02	.98	-0.00				
ISEQ General	0.67	.50	0.09				
ISEQ Justification	-0.14	.89	-0.02				
Topic knowledge	0.30	.76	0.02				
Partner agreement	-0.43	.67	-0.02				
Partner familiarity	0.58	.56	0.02				

Source Quality Score

Overall model				1.165	128	.3177	0.0129
Age	-0.74	.46	-0.05				
Search experience	2.48	.14	0.19				
Collaborative satisfaction	-1.05	.30	-0.05				
Task satisfaction	-0.11	.92	-0.01				

	T	p	β
Addon intuitiveness	1.00	.32	0.06
GPA	0.68	.50	0.05
ISEQ General	-0.07	.95	-0.02
ISEQ Justification	0.35	.73	0.05
Topic knowledge	1.47	.14	0.07
Partner agreement	-1.85	.0667*	-0.20
Partner familiarity	0.24	.24	0.01

MDP

	T	p	β	F	df	P	adj. R^2
Total Score							
Overall model				.9134	90	.5311	-.0095
Age	-0.106	.916	-0.022				
Search experience	0.070	.944	0.022				
Collaborative satisfaction	-0.277	.783	-0.039				
Task satisfaction	-0.114	.910	-0.017				
Addon intuitiveness	0.688	.493	0.104				
GPA	-0.215	.830	-0.044				
ISEQ General	-1.688	.095	-0.586				
ISEQ Justification	-1.321	.190	-0.460				
Topic knowledge	1.696	.093	0.206				
Partner agreement	1.447	.151	0.218				
Partner familiarity	-1.498	.138	-0.105				
Topic Score							
Overall model				1.1168	90	.3577	.0126
Age	-1.313	.192	-0.110				
Search experience	-1.299	.197	-0.164				
Collaborative satisfaction	-0.340	.735	-0.019				
Task satisfaction	0.091	.928	0.005				
Addon intuitiveness	-0.118	.906	-0.007				
GPA	0.581	.563	0.0478				
ISEQ General	-0.749	.455	-0.104				

ISEQ Justification	-2.803	.006*	-0.391				
Topic knowledge	1.467	.146	0.0713				
Partner agreement	0.062	.951	0.004				
Partner familiarity	-0.931	.354	-0.026				
Synthesis Score							
Overall model				0.944	90	.5028	-.0061
Age	1.477	.143	0.139				
Search experience	1.303	.196	0.183				
Collaborative satisfaction	0.085	.932	0.005				
Task satisfaction	-0.547	.586	-0.037				
Addon intuitiveness	0.905	.368	0.061				
GPA	-0.705	.483	-0.064				
ISEQ General	-1.395	.166	-0.216				
ISEQ Justification	1.039	.301	0.162				
Topic knowledge	0.818	.415	0.044				
Partner agreement	1.099	.275	0.074				
Partner familiarity	-0.524	.601	-0.016				
Source Diversity Score							
Overall model				1.3982	90	.1875	.0416
Age	-2.214	.029	-0.141				
Search experience	-1.318	.191	-0.126				
Collaborative satisfaction	-0.825	.412	-0.035				
Task satisfaction	0.895	.373	0.041				
Addon intuitiveness	-0.929	.355	-0.043				
GPA	1.097	.276	0.068				
ISEQ General	-1.067	.289	-0.112				
ISEQ Justification	-2.564	.012	-0.270				
Topic knowledge	1.451	.150	0.053				
Partner agreement	0.586	.559	0.027				
Partner familiarity	-0.906	.367	-0.019				
Source Quality Score							
Overall model				1.0058	90	.4478	.0006

Age	0.957	.341	0.090
Search experience	0.907	.367	0.128
Collaborative satisfaction	0.158	.875	0.010
Task satisfaction	-0.394	.695	-0.026
Addon intuitiveness	1.363	.176	0.093
GPA	-1.036	.303	-0.095
ISEQ General	-0.984	.327	-0.153
ISEQ Justification	0.251	.802	0.039
Topic knowledge	0.676	.501	0.037
Partner agreement	1.680	.096	0.114
Partner familiarity	-1.374	.173	-0.043

Appendix 15 – Raw Stepwise Regression for Survey Data Models Predictive of Outcome

CIS

Stepwise multiple linear regression
```
> predictors <-
c("ISEQGen","ISEQJust","age","searchExperience","collaborativeSatisfying","t
askSatisfying","addonIntuitive","GPA","topicKnowledge","partnerAgreement","p
artnerFamiliarity")
> no.na.data <- na.omit(x[c(predictors, group_vars[4:8])])
> no.na.data[c("age","topicKnowledge")] <-
c(log(no.na.data$age),log(no.na.data$topicKnowledge))
> forms <-
lapply(c(paste(group_vars[4:8],"~age+searchExperience+collaborativeSatisfyin
g+taskSatisfying+addonIntuitive+GPA+ISEQGen+ISEQJust+topicKnowledge+partnerA
greement+partnerFamiliarity",sep="")),FUN=formula)
> model <- lapply(forms,function(dep){lm(formula=dep,
+            no.na.data)})
> lapply(model, function (dv) step(dv,direction="both"))
Start:  AIC=-18.9
TopicScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerAgreement + partnerFamiliarity

                             Df Sum of Sq RSS    AIC
- collaborativeSatisfying  1      0.10 103 -20.7
- ISEQJust                 1      0.12 103 -20.7
- partnerAgreement         1      0.40 104 -20.3
- GPA                      1      0.65 104 -20.0
- searchExperience         1      0.82 104 -19.8
- age                      1      0.90 104 -19.6
- taskSatisfying           1      1.10 104 -19.4
- topicKnowledge           1      1.27 104 -19.1
<none>                                  103 -18.9
- ISEQGen                  1      1.53 105 -18.8
- addonIntuitive           1      1.54 105 -18.8
- partnerFamiliarity       1      4.29 107 -15.2

Step:  AIC=-20.7
TopicScore ~ age + searchExperience + taskSatisfying + addonIntuitive +
    GPA + ISEQGen + ISEQJust + topicKnowledge + partnerAgreement +
    partnerFamiliarity

                             Df Sum of Sq RSS    AIC
- ISEQJust                 1      0.13 103 -22.6
- GPA                      1      0.63 104 -21.9
- partnerAgreement         1      0.66 104 -21.8
- searchExperience         1      0.80 104 -21.6
- age                      1      0.86 104 -21.6
- taskSatisfying           1      1.06 104 -21.3
- topicKnowledge           1      1.22 104 -21.1
- addonIntuitive           1      1.46 105 -20.8
<none>                                  103 -20.7
- ISEQGen                  1      1.53 105 -20.7
+ collaborativeSatisfying  1      0.10 103 -18.9
- partnerFamiliarity       1      4.29 108 -17.0
```

```
Step:  AIC=-22.6
TopicScore ~ age + searchExperience + taskSatisfying + addonIntuitive +
    GPA + ISEQGen + topicKnowledge + partnerAgreement + partnerFamiliarity

                          Df Sum of Sq RSS    AIC
- GPA                     1      0.56  104  -23.8
- partnerAgreement        1      0.63  104  -23.7
- age                     1      0.86  104  -23.4
- taskSatisfying          1      1.08  104  -23.1
- searchExperience        1      1.13  104  -23.0
- topicKnowledge          1      1.24  104  -22.9
- ISEQGen                 1      1.41  105  -22.6
- addonIntuitive          1      1.46  105  -22.6
<none>                                 103  -22.6
+ ISEQJust                1      0.13  103  -20.7
+ collaborativeSatisfying 1      0.11  103  -20.7
- partnerFamiliarity      1      4.20  108  -19.0

Step:  AIC=-23.8
TopicScore ~ age + searchExperience + taskSatisfying + addonIntuitive +
    ISEQGen + topicKnowledge + partnerAgreement + partnerFamiliarity

                          Df Sum of Sq RSS    AIC
- partnerAgreement        1      0.57  104  -25.0
- age                     1      0.80  105  -24.7
- searchExperience        1      1.21  105  -24.2
- addonIntuitive          1      1.28  105  -24.1
- taskSatisfying          1      1.35  105  -24.0
- topicKnowledge          1      1.36  105  -24.0
- ISEQGen                 1      1.44  105  -23.9
<none>                                 104  -23.8
+ GPA                     1      0.56  103  -22.6
+ collaborativeSatisfying 1      0.08  104  -21.9
+ ISEQJust                1      0.06  104  -21.9
- partnerFamiliarity      1      4.29  108  -20.1

Step:  AIC=-25
TopicScore ~ age + searchExperience + taskSatisfying + addonIntuitive +
    ISEQGen + topicKnowledge + partnerFamiliarity

                          Df Sum of Sq RSS    AIC
- age                     1      0.74  105  -26.1
- addonIntuitive          1      1.22  106  -25.4
- topicKnowledge          1      1.25  106  -25.4
- ISEQGen                 1      1.31  106  -25.3
- taskSatisfying          1      1.34  106  -25.2
- searchExperience        1      1.36  106  -25.2
<none>                                 104  -25.0
+ partnerAgreement        1      0.57  104  -23.8
+ GPA                     1      0.50  104  -23.7
+ collaborativeSatisfying 1      0.31  104  -23.4
+ ISEQJust                1      0.04  104  -23.1
- partnerFamiliarity      1      4.10  108  -21.6

Step:  AIC=-26.1
TopicScore ~ searchExperience + taskSatisfying + addonIntuitive +
    ISEQGen + topicKnowledge + partnerFamiliarity

                          Df Sum of Sq RSS    AIC
```

367

```
- ISEQGen                       1          1.16 106 -26.5
- taskSatisfying                1          1.18 106 -26.5
- searchExperience              1          1.22 106 -26.4
- addonIntuitive                1          1.23 106 -26.4
- topicKnowledge                1          1.32 106 -26.3
<none>                                          105 -26.1
+ age                           1          0.74 104 -25.0
+ partnerAgreement              1          0.50 105 -24.7
+ GPA                           1          0.45 105 -24.6
+ collaborativeSatisfying       1          0.23 105 -24.4
+ ISEQJust                      1          0.04 105 -24.1
- partnerFamiliarity            1          4.14 109 -22.6

Step:  AIC=-26.5
TopicScore ~ searchExperience + taskSatisfying + addonIntuitive +
    topicKnowledge + partnerFamiliarity

                              Df Sum of Sq RSS   AIC
- searchExperience             1         0.80 107 -27.5
- addonIntuitive               1         1.06 107 -27.1
- taskSatisfying               1         1.26 108 -26.9
- topicKnowledge               1         1.44 108 -26.6
<none>                                        106 -26.5
+ ISEQGen                      1         1.16 105 -26.1
+ age                          1         0.59 106 -25.3
+ GPA                          1         0.49 106 -25.2
+ partnerAgreement             1         0.39 106 -25.0
+ collaborativeSatisfying      1         0.21 106 -24.8
+ ISEQJust                     1         0.00 106 -24.5
- partnerFamiliarity           1         3.52 110 -23.9

Step:  AIC=-27.5
TopicScore ~ taskSatisfying + addonIntuitive + topicKnowledge +
    partnerFamiliarity

                              Df Sum of Sq RSS   AIC
- addonIntuitive               1         0.76 108 -28.5
- topicKnowledge               1         1.34 108 -27.7
- taskSatisfying               1         1.51 109 -27.5
<none>                                        107 -27.5
+ searchExperience             1         0.80 106 -26.5
+ ISEQGen                      1         0.73 106 -26.4
+ GPA                          1         0.55 107 -26.2
+ partnerAgreement             1         0.51 107 -26.1
+ age                          1         0.51 107 -26.1
+ collaborativeSatisfying      1         0.22 107 -25.8
+ ISEQJust                     1         0.08 107 -25.6
- partnerFamiliarity           1         3.48 111 -25.0

Step:  AIC=-28.5
TopicScore ~ taskSatisfying + topicKnowledge + partnerFamiliarity

                              Df Sum of Sq RSS   AIC
- topicKnowledge               1         1.38 109 -28.7
<none>                                        108 -28.5
+ addonIntuitive               1         0.76 107 -27.5
+ ISEQGen                      1         0.69 107 -27.4
+ age                          1         0.53 107 -27.2
+ searchExperience             1         0.50 107 -27.1
```

```
+ partnerAgreement          1        0.45 107 -27.1
+ GPA                       1        0.39 108 -27.0
+ ISEQJust                  1        0.06 108 -26.6
- taskSatisfying            1        3.07 111 -26.6
+ collaborativeSatisfying   1        0.02 108 -26.5
- partnerFamiliarity        1        3.28 111 -26.3


Step:  AIC=-28.7
TopicScore ~ taskSatisfying + partnerFamiliarity


                          Df Sum of Sq RSS   AIC
<none>                                    109 -28.7
+ topicKnowledge            1      1.384 108 -28.5
+ addonIntuitive            1      0.809 108 -27.7
+ ISEQGen                   1      0.795 108 -27.7
+ age                       1      0.598 109 -27.5
+ GPA                       1      0.488 109 -27.3
+ searchExperience          1      0.413 109 -27.2
+ partnerAgreement          1      0.335 109 -27.1
- partnerFamiliarity        1      2.964 112 -26.9
- taskSatisfying            1      2.981 112 -26.9
+ ISEQJust                  1      0.055 109 -26.8
+ collaborativeSatisfying   1      0.000 109 -26.7
Start:  AIC=-97.8
SynthScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerAgreement + partnerFamiliarity


                          Df Sum of Sq  RSS    AIC
- partnerFamiliarity        1      0.026 58.7 -99.7
- taskSatisfying            1      0.030 58.7 -99.7
- GPA                       1      0.048 58.7 -99.7
- addonIntuitive            1      0.051 58.7 -99.7
- ISEQJust                  1      0.218 58.9 -99.3
- collaborativeSatisfying   1      0.244 58.9 -99.2
- ISEQGen                   1      0.336 59.0 -99.0
- age                       1      0.357 59.0 -99.0
- partnerAgreement          1      0.415 59.1 -98.8
- searchExperience          1      0.746 59.4 -98.0
<none>                                  58.7 -97.8
- topicKnowledge            1      3.040 61.7 -92.7


Step:  AIC=-99.7
SynthScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerAgreement


                          Df Sum of Sq  RSS    AIC
- taskSatisfying            1      0.032 58.7 -101.7
- GPA                       1      0.047 58.7 -101.6
- addonIntuitive            1      0.056 58.7 -101.6
- ISEQJust                  1      0.231 58.9 -101.2
- collaborativeSatisfying   1      0.243 58.9 -101.2
- age                       1      0.356 59.0 -100.9
- ISEQGen                   1      0.385 59.1 -100.8
- partnerAgreement          1      0.403 59.1 -100.8
- searchExperience          1      0.741 59.4 -100.0
<none>                                  58.7  -99.7
+ partnerFamiliarity        1      0.026 58.7  -97.8
```

```
- topicKnowledge             1     3.015 61.7  -94.7

Step:  AIC=-102
SynthScore ~ age + searchExperience + collaborativeSatisfying +
    addonIntuitive + GPA + ISEQGen + ISEQJust + topicKnowledge +
    partnerAgreement

                             Df Sum of Sq  RSS    AIC
- GPA                        1     0.060 58.8 -103.5
- addonIntuitive             1     0.109 58.8 -103.4
- collaborativeSatisfying    1     0.231 58.9 -103.1
- ISEQJust                   1     0.238 58.9 -103.1
- age                        1     0.340 59.0 -102.9
- ISEQGen                    1     0.394 59.1 -102.7
- partnerAgreement           1     0.395 59.1 -102.7
- searchExperience           1     0.721 59.4 -102.0
<none>                                    58.7 -101.7
+ taskSatisfying             1     0.032 58.7  -99.7
+ partnerFamiliarity         1     0.028 58.7  -99.7
- topicKnowledge             1     3.040 61.7  -96.6

Step:  AIC=-104
SynthScore ~ age + searchExperience + collaborativeSatisfying +
    addonIntuitive + ISEQGen + ISEQJust + topicKnowledge + partnerAgreement

                             Df Sum of Sq  RSS    AIC
- addonIntuitive             1     0.098 58.9 -105.3
- ISEQJust                   1     0.209 59.0 -105.0
- collaborativeSatisfying    1     0.219 59.0 -105.0
- age                        1     0.324 59.1 -104.8
- ISEQGen                    1     0.391 59.2 -104.6
- partnerAgreement           1     0.408 59.2 -104.6
- searchExperience           1     0.685 59.5 -103.9
<none>                                    58.8 -103.5
+ GPA                        1     0.060 58.7 -101.7
+ taskSatisfying             1     0.045 58.7 -101.6
+ partnerFamiliarity         1     0.027 58.7 -101.6
- topicKnowledge             1     3.000 61.8  -98.6

Step:  AIC=-105
SynthScore ~ age + searchExperience + collaborativeSatisfying +
    ISEQGen + ISEQJust + topicKnowledge + partnerAgreement

                             Df Sum of Sq  RSS  AIC
- collaborativeSatisfying    1     0.138 59.0 -107
- ISEQJust                   1     0.223 59.1 -107
- age                        1     0.303 59.2 -107
- partnerAgreement           1     0.362 59.2 -106
- ISEQGen                    1     0.376 59.2 -106
- searchExperience           1     0.805 59.7 -105
<none>                                    58.9 -105
+ addonIntuitive             1     0.098 58.8 -104
+ taskSatisfying             1     0.098 58.8 -104
+ GPA                        1     0.049 58.8 -103
+ partnerFamiliarity         1     0.037 58.8 -103
- topicKnowledge             1     3.025 61.9 -100

Step:  AIC=-107
SynthScore ~ age + searchExperience + ISEQGen + ISEQJust + topicKnowledge +
```

```
    partnerAgreement

                        Df Sum of Sq  RSS  AIC
- ISEQJust              1      0.233 59.2 -108
- partnerAgreement      1      0.251 59.3 -108
- age                   1      0.289 59.3 -108
- ISEQGen               1      0.388 59.4 -108
- searchExperience      1      0.790 59.8 -107
<none>                              59.0 -107
+ collaborativeSatisfying 1    0.138 58.9 -105
+ GPA                   1      0.045 59.0 -105
+ taskSatisfying        1      0.044 59.0 -105
+ partnerFamiliarity    1      0.031 59.0 -105
+ addonIntuitive        1      0.016 59.0 -105
- topicKnowledge        1      3.137 62.1 -102

Step:  AIC=-108
SynthScore ~ age + searchExperience + ISEQGen + topicKnowledge +
    partnerAgreement

                        Df Sum of Sq  RSS  AIC
- partnerAgreement      1      0.279 59.5 -110
- ISEQGen               1      0.286 59.5 -110
- age                   1      0.286 59.5 -110
- searchExperience      1      0.603 59.8 -109
<none>                              59.2 -108
+ ISEQJust              1      0.233 59.0 -107
+ collaborativeSatisfying 1    0.148 59.1 -107
+ taskSatisfying        1      0.049 59.2 -106
+ partnerFamiliarity    1      0.046 59.2 -106
+ addonIntuitive        1      0.020 59.2 -106
+ GPA                   1      0.020 59.2 -106
- topicKnowledge        1      3.102 62.3 -103

Step:  AIC=-110
SynthScore ~ age + searchExperience + ISEQGen + topicKnowledge

                        Df Sum of Sq  RSS  AIC
- age                   1      0.315 59.8 -111
- ISEQGen               1      0.327 59.8 -111
- searchExperience      1      0.687 60.2 -110
<none>                              59.5 -110
+ partnerAgreement      1      0.279 59.2 -108
+ ISEQJust              1      0.261 59.3 -108
+ taskSatisfying        1      0.053 59.5 -108
+ partnerFamiliarity    1      0.033 59.5 -108
+ GPA                   1      0.028 59.5 -108
+ collaborativeSatisfying 1    0.028 59.5 -108
+ addonIntuitive        1      0.025 59.5 -108
- topicKnowledge        1      2.997 62.5 -105

Step:  AIC=-111
SynthScore ~ searchExperience + ISEQGen + topicKnowledge

                        Df Sum of Sq  RSS  AIC
- ISEQGen               1      0.276 60.1 -112
- searchExperience      1      0.768 60.6 -111
<none>                              59.8 -111
+ age                   1      0.315 59.5 -110
```

```
+ partnerAgreement          1      0.309 59.5 -110
+ ISEQJust                  1      0.260 59.6 -110
+ partnerFamiliarity        1      0.030 59.8 -109
+ taskSatisfying            1      0.030 59.8 -109
+ GPA                       1      0.020 59.8 -109
+ collaborativeSatisfying   1      0.019 59.8 -109
+ addonIntuitive            1      0.018 59.8 -109
- topicKnowledge            1      2.927 62.8 -106

Step:  AIC=-112
SynthScore ~ searchExperience + topicKnowledge

                            Df Sum of Sq  RSS   AIC
<none>                                   60.1 -112
- searchExperience          1      1.012 61.1 -112
+ partnerAgreement          1      0.346 59.8 -111
+ ISEQGen                   1      0.276 59.8 -111
+ age                       1      0.264 59.8 -111
+ ISEQJust                  1      0.153 60.0 -111
+ partnerFamiliarity        1      0.067 60.0 -110
+ taskSatisfying            1      0.033 60.1 -110
+ GPA                       1      0.024 60.1 -110
+ collaborativeSatisfying   1      0.019 60.1 -110
+ addonIntuitive            1      0.012 60.1 -110
- topicKnowledge            1      2.836 62.9 -108
Start:  AIC=-43
SourDivScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerAgreement + partnerFamiliarity

                            Df Sum of Sq  RSS   AIC
- GPA                       1      0.000 86.7 -45.1
- collaborativeSatisfying   1      0.002 86.7 -45.0
- taskSatisfying            1      0.010 86.7 -45.0
- ISEQJust                  1      0.012 86.7 -45.0
- searchExperience          1      0.083 86.8 -44.9
- partnerAgreement          1      0.120 86.8 -44.9
- topicKnowledge            1      0.175 86.9 -44.8
- partnerFamiliarity        1      0.222 86.9 -44.7
- ISEQGen                   1      0.313 87.0 -44.5
- age                       1      0.652 87.4 -44.0
- addonIntuitive            1      1.053 87.8 -43.4
<none>                                   86.7 -43.1

Step:  AIC=-45
SourDivScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + ISEQGen + ISEQJust + topicKnowledge +
    partnerAgreement + partnerFamiliarity

                            Df Sum of Sq  RSS   AIC
- collaborativeSatisfying   1      0.002 86.7 -47.0
- taskSatisfying            1      0.011 86.7 -47.0
- ISEQJust                  1      0.012 86.7 -47.0
- searchExperience          1      0.084 86.8 -46.9
- partnerAgreement          1      0.121 86.8 -46.9
- topicKnowledge            1      0.176 86.9 -46.8
- partnerFamiliarity        1      0.222 86.9 -46.7
- ISEQGen                   1      0.313 87.0 -46.5
- age                       1      0.654 87.4 -46.0
```

```
- addonIntuitive            1      1.070 87.8 -45.3
<none>                                    86.7 -45.1
+ GPA                       1      0.000 86.7 -43.1

Step:  AIC=-47
SourDivScore ~ age + searchExperience + taskSatisfying + addonIntuitive +
    ISEQGen + ISEQJust + topicKnowledge + partnerAgreement +
    partnerFamiliarity

                          Df Sum of Sq  RSS    AIC
- ISEQJust                 1      0.011 86.7 -49.0
- taskSatisfying           1      0.012 86.7 -49.0
- searchExperience         1      0.083 86.8 -48.9
- partnerAgreement         1      0.155 86.9 -48.8
- topicKnowledge           1      0.180 86.9 -48.8
- partnerFamiliarity       1      0.222 86.9 -48.7
- ISEQGen                  1      0.313 87.0 -48.5
- age                      1      0.652 87.4 -48.0
- addonIntuitive           1      1.189 87.9 -47.1
<none>                                    86.7 -47.0
+ collaborativeSatisfying  1      0.002 86.7 -45.1
+ GPA                      1      0.000 86.7 -45.0

Step:  AIC=-49
SourDivScore ~ age + searchExperience + taskSatisfying + addonIntuitive +
    ISEQGen + topicKnowledge + partnerAgreement + partnerFamiliarity

                          Df Sum of Sq  RSS    AIC
- taskSatisfying           1      0.012 86.7 -51.0
- searchExperience         1      0.072 86.8 -50.9
- partnerAgreement         1      0.161 86.9 -50.8
- topicKnowledge           1      0.181 86.9 -50.7
- partnerFamiliarity       1      0.232 87.0 -50.7
- ISEQGen                  1      0.356 87.1 -50.5
- age                      1      0.652 87.4 -50.0
- addonIntuitive           1      1.186 87.9 -49.1
<none>                                    86.7 -49.0
+ ISEQJust                 1      0.011 86.7 -47.0
+ collaborativeSatisfying  1      0.002 86.7 -47.0
+ GPA                      1      0.000 86.7 -47.0

Step:  AIC=-51
SourDivScore ~ age + searchExperience + addonIntuitive + ISEQGen +
    topicKnowledge + partnerAgreement + partnerFamiliarity

                          Df Sum of Sq  RSS    AIC
- searchExperience         1      0.066 86.8 -52.9
- partnerAgreement         1      0.161 86.9 -52.8
- topicKnowledge           1      0.179 86.9 -52.7
- partnerFamiliarity       1      0.235 87.0 -52.6
- ISEQGen                  1      0.353 87.1 -52.4
- age                      1      0.673 87.4 -51.9
<none>                                    86.7 -51.0
- addonIntuitive           1      1.329 88.1 -50.9
+ taskSatisfying           1      0.012 86.7 -49.0
+ ISEQJust                 1      0.012 86.7 -49.0
+ collaborativeSatisfying  1      0.002 86.7 -49.0
+ GPA                      1      0.000 86.7 -49.0
```

```
Step:  AIC=-52.9
SourDivScore ~ age + addonIntuitive + ISEQGen + topicKnowledge +
    partnerAgreement + partnerFamiliarity

                          Df Sum of Sq  RSS   AIC
- partnerAgreement        1     0.180 87.0 -54.6
- topicKnowledge          1     0.186 87.0 -54.6
- partnerFamiliarity      1     0.224 87.0 -54.5
- ISEQGen                 1     0.305 87.1 -54.4
- age                     1     0.644 87.5 -53.9
<none>                                  86.8 -52.9
- addonIntuitive          1     1.266 88.1 -52.9
+ searchExperience        1     0.066 86.7 -51.0
+ taskSatisfying          1     0.006 86.8 -50.9
+ collaborativeSatisfying 1     0.001 86.8 -50.9
+ ISEQJust                1     0.001 86.8 -50.9
+ GPA                     1     0.000 86.8 -50.9

Step:  AIC=-54.6
SourDivScore ~ age + addonIntuitive + ISEQGen + topicKnowledge +
    partnerFamiliarity

                          Df Sum of Sq  RSS   AIC
- partnerFamiliarity      1     0.198 87.2 -56.3
- topicKnowledge          1     0.214 87.2 -56.3
- ISEQGen                 1     0.262 87.3 -56.2
- age                     1     0.608 87.6 -55.6
- addonIntuitive          1     1.220 88.2 -54.7
<none>                                  87.0 -54.6
+ partnerAgreement        1     0.180 86.8 -52.9
+ searchExperience        1     0.085 86.9 -52.8
+ collaborativeSatisfying 1     0.038 87.0 -52.7
+ taskSatisfying          1     0.006 87.0 -52.6
+ ISEQJust                1     0.001 87.0 -52.6
+ GPA                     1     0.000 87.0 -52.6

Step:  AIC=-56.3
SourDivScore ~ age + addonIntuitive + ISEQGen + topicKnowledge

                          Df Sum of Sq  RSS   AIC
- ISEQGen                 1     0.199 87.4 -58.0
- topicKnowledge          1     0.244 87.4 -57.9
- age                     1     0.619 87.8 -57.3
- addonIntuitive          1     1.140 88.3 -56.5
<none>                                  87.2 -56.3
+ partnerFamiliarity      1     0.198 87.0 -54.6
+ partnerAgreement        1     0.154 87.0 -54.5
+ searchExperience        1     0.071 87.1 -54.4
+ collaborativeSatisfying 1     0.034 87.2 -54.4
+ taskSatisfying          1     0.009 87.2 -54.3
+ ISEQJust                1     0.006 87.2 -54.3
+ GPA                     1     0.000 87.2 -54.3

Step:  AIC=-58
SourDivScore ~ age + addonIntuitive + topicKnowledge

                          Df Sum of Sq  RSS   AIC
- topicKnowledge          1     0.216 87.6 -59.6
- age                     1     0.573 88.0 -59.1
```

```
- addonIntuitive             1      1.127 88.5 -58.2
<none>                              87.4 -58.0
+ ISEQGen                    1      0.199 87.2 -56.3
+ partnerFamiliarity         1      0.134 87.3 -56.2
+ partnerAgreement           1      0.123 87.3 -56.2
+ ISEQJust                   1      0.036 87.4 -56.0
+ collaborativeSatisfying    1      0.030 87.4 -56.0
+ searchExperience           1      0.027 87.4 -56.0
+ taskSatisfying             1      0.008 87.4 -56.0
+ GPA                        1      0.000 87.4 -56.0

Step:  AIC=-59.6
SourDivScore ~ age + addonIntuitive

                            Df Sum of Sq  RSS    AIC
- age                        1      0.548 88.2 -60.8
- addonIntuitive             1      1.108 88.7 -59.9
<none>                              87.6 -59.6
+ topicKnowledge             1      0.216 87.4 -58.0
+ ISEQGen                    1      0.171 87.4 -57.9
+ partnerFamiliarity         1      0.162 87.4 -57.9
+ partnerAgreement           1      0.147 87.5 -57.9
+ collaborativeSatisfying    1      0.046 87.6 -57.7
+ ISEQJust                   1      0.035 87.6 -57.7
+ searchExperience           1      0.035 87.6 -57.7
+ taskSatisfying             1      0.006 87.6 -57.6
+ GPA                        1      0.001 87.6 -57.6

Step:  AIC=-60.8
SourDivScore ~ addonIntuitive

                            Df Sum of Sq  RSS    AIC
- addonIntuitive             1      1.068 89.2 -61.1
<none>                              88.2 -60.8
+ age                        1      0.548 87.6 -59.6
+ topicKnowledge             1      0.191 88.0 -59.1
+ partnerFamiliarity         1      0.177 88.0 -59.0
+ ISEQGen                    1      0.130 88.0 -59.0
+ partnerAgreement           1      0.118 88.0 -58.9
+ ISEQJust                   1      0.038 88.1 -58.8
+ collaborativeSatisfying    1      0.025 88.1 -58.8
+ taskSatisfying             1      0.024 88.1 -58.8
+ searchExperience           1      0.018 88.1 -58.8
+ GPA                        1      0.005 88.1 -58.8

Step:  AIC=-61.1
SourDivScore ~ 1

                            Df Sum of Sq  RSS    AIC
<none>                              89.2 -61.1
+ addonIntuitive             1      1.068 88.2 -60.8
+ age                        1      0.508 88.7 -59.9
+ topicKnowledge             1      0.175 89.0 -59.3
+ ISEQGen                    1      0.122 89.1 -59.3
+ partnerFamiliarity         1      0.105 89.1 -59.2
+ partnerAgreement           1      0.089 89.1 -59.2
+ taskSatisfying             1      0.086 89.1 -59.2
+ collaborativeSatisfying    1      0.074 89.1 -59.2
+ ISEQJust                   1      0.046 89.2 -59.1
```

```
+ GPA                          1      0.019 89.2 -59.1
+ searchExperience             1      0.001 89.2 -59.1
Start:  AIC=-51.7
SourQualScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerAgreement + partnerFamiliarity

                             Df Sum of Sq  RSS    AIC
- ISEQGen                     1      0.001 81.5 -53.7
- taskSatisfying              1      0.013 81.5 -53.7
- partnerFamiliarity          1      0.033 81.5 -53.7
- ISEQJust                    1      0.096 81.6 -53.6
- GPA                         1      0.322 81.8 -53.2
- age                         1      0.458 82.0 -53.0
- addonIntuitive              1      0.676 82.2 -52.6
- collaborativeSatisfying     1      0.757 82.3 -52.4
<none>                                     81.5 -51.7
- searchExperience            1      1.471 83.0 -51.2
- partnerAgreement            1      2.158 83.7 -50.1
- topicKnowledge              1      2.376 83.9 -49.7

Step:  AIC=-53.7
SourQualScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQJust + topicKnowledge +
    partnerAgreement + partnerFamiliarity

                             Df Sum of Sq  RSS    AIC
- taskSatisfying              1      0.013 81.5 -55.7
- partnerFamiliarity          1      0.036 81.5 -55.7
- ISEQJust                    1      0.104 81.6 -55.6
- GPA                         1      0.323 81.8 -55.2
- age                         1      0.464 82.0 -54.9
- addonIntuitive              1      0.681 82.2 -54.6
- collaborativeSatisfying     1      0.757 82.3 -54.4
<none>                                     81.5 -53.7
- searchExperience            1      1.489 83.0 -53.2
- partnerAgreement            1      2.182 83.7 -52.0
+ ISEQGen                     1      0.001 81.5 -51.7
- topicKnowledge              1      2.384 83.9 -51.7

Step:  AIC=-55.7
SourQualScore ~ age + searchExperience + collaborativeSatisfying +
    addonIntuitive + GPA + ISEQJust + topicKnowledge + partnerAgreement +
    partnerFamiliarity

                             Df Sum of Sq  RSS    AIC
- partnerFamiliarity          1      0.037 81.6 -57.6
- ISEQJust                    1      0.102 81.6 -57.5
- GPA                         1      0.311 81.8 -57.2
- age                         1      0.482 82.0 -56.9
- addonIntuitive              1      0.712 82.2 -56.5
- collaborativeSatisfying     1      0.777 82.3 -56.4
<none>                                     81.5 -55.7
- searchExperience            1      1.520 83.0 -55.1
- partnerAgreement            1      2.173 83.7 -54.0
+ taskSatisfying              1      0.013 81.5 -53.7
+ ISEQGen                     1      0.001 81.5 -53.7
- topicKnowledge              1      2.375 83.9 -53.7
```

```
Step:  AIC=-57.6
SourQualScore ~ age + searchExperience + collaborativeSatisfying +
    addonIntuitive + GPA + ISEQJust + topicKnowledge + partnerAgreement

                           Df Sum of Sq  RSS    AIC
- ISEQJust                 1     0.097  81.7  -59.5
- GPA                      1     0.314  81.9  -59.1
- age                      1     0.489  82.0  -58.8
- addonIntuitive           1     0.691  82.2  -58.5
- collaborativeSatisfying  1     0.778  82.3  -58.3
<none>                                   81.6  -57.6
- searchExperience         1     1.519  83.1  -57.1
- partnerAgreement         1     2.150  83.7  -56.0
+ partnerFamiliarity       1     0.037  81.5  -55.7
+ taskSatisfying           1     0.015  81.5  -55.7
+ ISEQGen                  1     0.004  81.6  -55.7
- topicKnowledge           1     2.442  84.0  -55.5

Step:  AIC=-59.5
SourQualScore ~ age + searchExperience + collaborativeSatisfying +
    addonIntuitive + GPA + topicKnowledge + partnerAgreement

                           Df Sum of Sq  RSS    AIC
- GPA                      1     0.271  81.9  -61.0
- age                      1     0.493  82.1  -60.6
- addonIntuitive           1     0.715  82.4  -60.3
- collaborativeSatisfying  1     0.799  82.5  -60.1
<none>                                   81.7  -59.5
- searchExperience         1     1.442  83.1  -59.0
- partnerAgreement         1     2.105  83.8  -57.9
+ ISEQJust                 1     0.097  81.6  -57.6
+ partnerFamiliarity       1     0.032  81.6  -57.5
+ ISEQGen                  1     0.016  81.6  -57.5
+ taskSatisfying           1     0.013  81.6  -57.5
- topicKnowledge           1     2.422  84.1  -57.4

Step:  AIC=-61
SourQualScore ~ age + searchExperience + collaborativeSatisfying +
    addonIntuitive + topicKnowledge + partnerAgreement

                           Df Sum of Sq  RSS    AIC
- age                      1     0.452  82.4  -62.2
- addonIntuitive           1     0.651  82.6  -61.9
- collaborativeSatisfying  1     0.746  82.7  -61.7
<none>                                   81.9  -61.0
- searchExperience         1     1.374  83.3  -60.7
- partnerAgreement         1     2.062  84.0  -59.5
+ GPA                      1     0.271  81.7  -59.5
+ ISEQJust                 1     0.055  81.9  -59.1
+ partnerFamiliarity       1     0.036  81.9  -59.1
- topicKnowledge           1     2.340  84.3  -59.1
+ ISEQGen                  1     0.014  81.9  -59.0
+ taskSatisfying           1     0.001  81.9  -59.0

Step:  AIC=-62.2
SourQualScore ~ searchExperience + collaborativeSatisfying +
    addonIntuitive + topicKnowledge + partnerAgreement

                           Df Sum of Sq  RSS    AIC
```

```
- addonIntuitive              1        0.589 83.0 -63.3
- collaborativeSatisfying  1        0.681 83.1 -63.1
<none>                                       82.4 -62.2
- searchExperience          1        1.485 83.9 -61.7
+ age                          1        0.452 81.9 -61.0
- partnerAgreement          1        2.027 84.4 -60.8
+ GPA                          1        0.230 82.1 -60.6
- topicKnowledge            1        2.287 84.7 -60.4
+ ISEQJust                   1        0.060 82.3 -60.4
+ partnerFamiliarity        1        0.042 82.3 -60.3
+ ISEQGen                    1        0.032 82.3 -60.3
+ taskSatisfying            1        0.011 82.4 -60.3


Step:  AIC=-63.2
SourQualScore ~ searchExperience + collaborativeSatisfying +
    topicKnowledge + partnerAgreement


                               Df Sum of Sq  RSS    AIC
- collaborativeSatisfying  1        0.308 83.3 -64.7
<none>                                       83.0 -63.3
+ addonIntuitive            1        0.589 82.4 -62.2
- searchExperience          1        1.851 84.8 -62.2
+ age                          1        0.390 82.6 -61.9
+ GPA                          1        0.177 82.8 -61.5
+ ISEQJust                   1        0.082 82.9 -61.4
- topicKnowledge            1        2.346 85.3 -61.3
+ ISEQGen                    1        0.045 82.9 -61.3
+ taskSatisfying            1        0.036 82.9 -61.3
+ partnerFamiliarity        1        0.019 82.9 -61.3
- partnerAgreement          1        2.392 85.4 -61.3


Step:  AIC=-64.7
SourQualScore ~ searchExperience + topicKnowledge + partnerAgreement


                               Df Sum of Sq  RSS    AIC
<none>                                       83.3 -64.7
- searchExperience          1        1.80 85.1 -63.7
+ age                          1        0.36 82.9 -63.3
+ collaborativeSatisfying  1        0.31 83.0 -63.3
+ addonIntuitive            1        0.22 83.1 -63.1
+ GPA                          1        0.16 83.1 -63.0
+ ISEQJust                   1        0.09 83.2 -62.9
+ ISEQGen                    1        0.04 83.2 -62.8
+ partnerFamiliarity        1        0.03 83.2 -62.8
+ taskSatisfying            1        0.00 83.3 -62.7
- topicKnowledge            1        2.48 85.8 -62.6
- partnerAgreement          1        3.51 86.8 -60.9
Start:  AIC=235
totalhand ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerAgreement + partnerFamiliarity


                               Df Sum of Sq RSS AIC
- searchExperience          1        0.78 635 234
- taskSatisfying            1        1.02 635 234
- ISEQJust                   1        1.03 635 234
- GPA                          1        2.51 636 234
- collaborativeSatisfying  1        2.96 637 234
- partnerAgreement          1        3.26 637 234
```

```
- ISEQGen                          1        5.51 639 235
- partnerFamiliarity               1        6.56 640 235
- topicKnowledge                   1        6.64 640 235
<none>                                           634 235
- age                              1        9.17 643 235
- addonIntuitive                   1       10.97 645 236

Step:  AIC=234
totalhand ~ age + collaborativeSatisfying + taskSatisfying +
    addonIntuitive + GPA + ISEQGen + ISEQJust + topicKnowledge +
    partnerAgreement + partnerFamiliarity

                          Df Sum of Sq RSS AIC
- ISEQJust                 1        0.61 635 232
- taskSatisfying           1        0.88 635 232
- GPA                      1        2.28 637 232
- partnerAgreement         1        2.96 637 232
- collaborativeSatisfying  1        3.09 638 232
- ISEQGen                  1        6.18 641 233
- topicKnowledge           1        6.47 641 233
- partnerFamiliarity       1        6.65 641 233
<none>                                    635 234
- age                      1        9.56 644 234
- addonIntuitive           1       12.53 647 234
+ searchExperience         1        0.78 634 235

Step:  AIC=232
totalhand ~ age + collaborativeSatisfying + taskSatisfying +
    addonIntuitive + GPA + ISEQGen + topicKnowledge + partnerAgreement +
    partnerFamiliarity

                          Df Sum of Sq RSS AIC
- taskSatisfying           1        0.98 636 230
- GPA                      1        2.03 637 230
- partnerAgreement         1        2.84 638 230
- collaborativeSatisfying  1        3.16 638 230
- ISEQGen                  1        5.58 641 231
- partnerFamiliarity       1        6.34 641 231
- topicKnowledge           1        6.41 642 231
<none>                                    635 232
- age                      1        9.43 645 232
- addonIntuitive           1       12.28 647 232
+ ISEQJust                 1        0.61 635 234
+ searchExperience         1        0.35 635 234

Step:  AIC=230
totalhand ~ age + collaborativeSatisfying + addonIntuitive +
    GPA + ISEQGen + topicKnowledge + partnerAgreement + partnerFamiliarity

                          Df Sum of Sq RSS AIC
- GPA                      1        2.52 639 228
- collaborativeSatisfying  1        2.91 639 229
- partnerAgreement         1        2.99 639 229
- ISEQGen                  1        5.59 642 229
- partnerFamiliarity       1        6.16 642 229
- topicKnowledge           1        6.62 643 229
- age                      1        8.94 645 230
<none>                                    636 230
- addonIntuitive           1       17.29 653 232
```

```
+ taskSatisfying             1       0.98 635 232
+ ISEQJust                   1       0.71 635 232
+ searchExperience           1       0.24 636 232

Step:  AIC=228
totalhand ~ age + collaborativeSatisfying + addonIntuitive +
    ISEQGen + topicKnowledge + partnerAgreement + partnerFamiliarity

                             Df Sum of Sq RSS AIC
- collaborativeSatisfying    1       2.59 641 227
- partnerAgreement           1       2.87 642 227
- ISEQGen                    1       5.66 644 228
- topicKnowledge             1       6.20 645 228
- partnerFamiliarity         1       6.32 645 228
- age                        1       8.38 647 228
<none>                                     639 228
+ GPA                        1       2.52 636 230
- addonIntuitive             1      16.27 655 230
+ taskSatisfying             1       1.47 637 230
+ ISEQJust                   1       0.43 638 230
+ searchExperience           1       0.15 638 230

Step:  AIC=227
totalhand ~ age + addonIntuitive + ISEQGen + topicKnowledge +
    partnerAgreement + partnerFamiliarity

                             Df Sum of Sq RSS AIC
- ISEQGen                    1       5.71 647 226
- partnerAgreement           1       6.20 647 226
- partnerFamiliarity         1       6.33 648 226
- topicKnowledge             1       6.84 648 226
- age                        1       7.89 649 227
<none>                                     641 227
- addonIntuitive             1      13.69 655 228
+ collaborativeSatisfying    1       2.59 639 228
+ GPA                        1       2.20 639 229
+ taskSatisfying             1       1.14 640 229
+ ISEQJust                   1       0.49 641 229
+ searchExperience           1       0.21 641 229

Step:  AIC=226
totalhand ~ age + addonIntuitive + topicKnowledge + partnerAgreement +
    partnerFamiliarity

                             Df Sum of Sq RSS AIC
- partnerFamiliarity         1       4.60 652 225
- partnerAgreement           1       5.11 652 225
- topicKnowledge             1       6.26 653 226
- age                        1       6.97 654 226
<none>                                     647 226
+ ISEQGen                    1       5.71 641 227
- addonIntuitive             1      13.10 660 227
+ collaborativeSatisfying    1       2.63 644 228
+ GPA                        1       2.26 645 228
+ taskSatisfying             1       1.15 646 228
+ searchExperience           1       0.91 646 228
+ ISEQJust                   1       0.00 647 228

Step:  AIC=225
```

```
totalhand ~ age + addonIntuitive + topicKnowledge + partnerAgreement

                          Df Sum of Sq RSS AIC
- partnerAgreement         1      4.62 656 224
- topicKnowledge           1      7.24 659 225
- age                      1      7.26 659 225
<none>                                652 225
- addonIntuitive           1     11.82 663 226
+ partnerFamiliarity       1      4.60 647 226
+ ISEQGen                  1      3.98 648 226
+ collaborativeSatisfying  1      2.64 649 227
+ GPA                      1      2.38 649 227
+ taskSatisfying           1      0.99 651 227
+ searchExperience         1      0.96 651 227
+ ISEQJust                 1      0.01 652 227

Step:   AIC=224
totalhand ~ age + addonIntuitive + topicKnowledge

                          Df Sum of Sq RSS AIC
- age                      1      6.70 663 224
- topicKnowledge           1      8.13 664 224
<none>                                656 224
- addonIntuitive           1     11.20 667 225
+ collaborativeSatisfying  1      5.38 651 225
+ partnerAgreement         1      4.62 652 225
+ partnerFamiliarity       1      4.11 652 225
+ ISEQGen                  1      3.21 653 226
+ GPA                      1      2.05 654 226
+ taskSatisfying           1      1.01 655 226
+ searchExperience         1      0.57 656 226
+ ISEQJust                 1      0.01 656 226

Step:   AIC=224
totalhand ~ addonIntuitive + topicKnowledge

                          Df Sum of Sq RSS AIC
- topicKnowledge           1      7.60 670 223
<none>                                663 224
- addonIntuitive           1     10.75 674 224
+ age                      1      6.70 656 224
+ collaborativeSatisfying  1      4.52 658 225
+ partnerFamiliarity       1      4.40 658 225
+ partnerAgreement         1      4.06 659 225
+ ISEQGen                  1      2.57 660 225
+ GPA                      1      1.64 661 225
+ searchExperience         1      0.86 662 226
+ taskSatisfying           1      0.55 662 226
+ ISEQJust                 1      0.01 663 226

Step:   AIC=223
totalhand ~ addonIntuitive

                          Df Sum of Sq RSS AIC
<none>                                670 223
- addonIntuitive           1     10.43 681 223
+ topicKnowledge           1      7.60 663 224
+ age                      1      6.16 664 224
+ collaborativeSatisfying  1      5.68 665 224
```

```
+ partnerFamiliarity        1        5.33 665 224
+ partnerAgreement          1        4.89 666 224
+ ISEQGen                   1        2.02 668 225
+ GPA                       1        1.24 669 225
+ taskSatisfying            1        0.66 670 225
+ searchExperience          1        0.63 670 225
+ ISEQJust                  1        0.01 670 225
[[1]]

Call:
lm(formula = TopicScore ~ taskSatisfying + partnerFamiliarity,
    data = no.na.data)

Coefficients:
       (Intercept)      taskSatisfying   partnerFamiliarity
            1.1203              0.1061               0.0556


[[2]]

Call:
lm(formula = SynthScore ~ searchExperience + topicKnowledge,
    data = no.na.data)

Coefficients:
      (Intercept)   searchExperience      topicKnowledge
            1.764              0.143               0.253


[[3]]

Call:
lm(formula = SourDivScore ~ 1, data = no.na.data)

Coefficients:
(Intercept)
       2.26


[[4]]

Call:
lm(formula = SourQualScore ~ searchExperience + topicKnowledge +
    partnerAgreement, data = no.na.data)

Coefficients:
     (Intercept)   searchExperience       topicKnowledge   partnerAgreement
           2.185              0.192                0.238             -0.118


[[5]]

Call:
lm(formula = totalhand ~ addonIntuitive, data = no.na.data)

Coefficients:
   (Intercept)  addonIntuitive
         7.303           0.207
```

MDP

Stepwise multiple linear

```
> forms <-
lapply(c(paste(group_vars[4:8],"~age+searchExperience+collaborativeSatisfyin
g+taskSatisfying+addonIntuitive+GPA+ISEQGen+ISEQJust+topicKnowledge+partnerA
greement+partnerFamiliarity",sep="")),FUN=formula) #this creates a list of
formulae
> predictors <-
c("ISEQGen","ISEQJust","age","searchExperience","collaborativeSatisfying","t
askSatisfying","addonIntuitive","GPA","topicKnowledge","partnerAgreement","p
artnerFamiliarity")
> no.na.data <- na.omit(y[c(predictors, group_vars[4:8])])
> no.na.data[c("age","topicKnowledge")] <-
c(log(no.na.data$age),log(no.na.data$topicKnowledge))
> forms <-
lapply(c(paste(group_vars[4:8],"~age+searchExperience+collaborativeSatisfyin
g+taskSatisfying+addonIntuitive+GPA+ISEQGen+ISEQJust+topicKnowledge+partnerA
greement+partnerFamiliarity",sep="")),FUN=formula)
> model <- lapply(forms,function(dep){lm(formula=dep,
+             no.na.data)})
> lapply(model, function (dv) step(dv,direction="both"))
Start:  AIC=-59.8
TopicScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerAgreement + partnerFamiliarity


                           Df Sum of Sq  RSS    AIC
- partnerAgreement          1      0.00 44.9 -61.8
- addonIntuitive            1      0.00 44.9 -61.8
- taskSatisfying            1      0.00 44.9 -61.8
- collaborativeSatisfying   1      0.07 44.9 -61.6
- GPA                       1      0.17 45.0 -61.4
- ISEQGen                   1      0.25 45.1 -61.2
- partnerFamiliarity        1      0.42 45.3 -60.9
- searchExperience          1      0.83 45.7 -59.9
- age                       1      0.83 45.7 -59.9
<none>                                  44.9 -59.8
- topicKnowledge            1      0.97 45.8 -59.6
- ISEQJust                  1      3.71 48.6 -53.7

Step:  AIC=-61.8
TopicScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerFamiliarity


                           Df Sum of Sq  RSS    AIC
- addonIntuitive            1      0.00 44.9 -63.8
- taskSatisfying            1      0.00 44.9 -63.8
- collaborativeSatisfying   1      0.08 44.9 -63.6
- GPA                       1      0.17 45.0 -63.4
- ISEQGen                   1      0.26 45.1 -63.2
- partnerFamiliarity        1      0.43 45.3 -62.8
- age                       1      0.83 45.7 -61.9
- searchExperience          1      0.88 45.7 -61.8
<none>                                  44.9 -61.8
- topicKnowledge            1      1.01 45.9 -61.5
+ partnerAgreement          1      0.00 44.9 -59.8
- ISEQJust                  1      3.71 48.6 -55.7
```

```
Step:  AIC=-63.8
TopicScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + GPA + ISEQGen + ISEQJust + topicKnowledge +
    partnerFamiliarity

                            Df Sum of Sq  RSS   AIC
- taskSatisfying            1      0.00 44.9 -65.8
- collaborativeSatisfying   1      0.09 44.9 -65.6
- GPA                       1      0.17 45.0 -65.4
- ISEQGen                   1      0.26 45.1 -65.2
- partnerFamiliarity        1      0.43 45.3 -64.8
- age                       1      0.85 45.7 -63.9
<none>                                  44.9 -63.8
- searchExperience          1      0.89 45.7 -63.8
- topicKnowledge            1      1.04 45.9 -63.5
+ addonIntuitive            1      0.00 44.9 -61.8
+ partnerAgreement          1      0.00 44.9 -61.8
- ISEQJust                  1      3.74 48.6 -57.6

Step:  AIC=-65.8
TopicScore ~ age + searchExperience + collaborativeSatisfying +
    GPA + ISEQGen + ISEQJust + topicKnowledge + partnerFamiliarity

                            Df Sum of Sq  RSS   AIC
- collaborativeSatisfying   1      0.09 44.9 -67.6
- GPA                       1      0.17 45.0 -67.4
- ISEQGen                   1      0.26 45.1 -67.2
- partnerFamiliarity        1      0.43 45.3 -66.8
- age                       1      0.86 45.7 -65.9
- searchExperience          1      0.89 45.7 -65.8
<none>                                  44.9 -65.8
- topicKnowledge            1      1.08 45.9 -65.4
+ taskSatisfying            1      0.00 44.9 -63.8
+ partnerAgreement          1      0.00 44.9 -63.8
+ addonIntuitive            1      0.00 44.9 -63.8
- ISEQJust                  1      3.80 48.7 -59.5

Step:  AIC=-67.6
TopicScore ~ age + searchExperience + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerFamiliarity

                            Df Sum of Sq  RSS   AIC
- GPA                       1      0.23 45.2 -69.1
- ISEQGen                   1      0.28 45.2 -69.0
- partnerFamiliarity        1      0.46 45.4 -68.6
- age                       1      0.82 45.8 -67.7
- searchExperience          1      0.88 45.8 -67.6
<none>                                  44.9 -67.6
- topicKnowledge            1      1.11 46.1 -67.1
+ collaborativeSatisfying   1      0.09 44.9 -65.8
+ addonIntuitive            1      0.02 44.9 -65.6
+ taskSatisfying            1      0.00 44.9 -65.6
+ partnerAgreement          1      0.00 44.9 -65.6
- ISEQJust                  1      3.84 48.8 -61.2

Step:  AIC=-69.1
TopicScore ~ age + searchExperience + ISEQGen + ISEQJust + topicKnowledge +
    partnerFamiliarity
```

```
                           Df Sum of Sq  RSS   AIC
- ISEQGen                  1      0.36  45.5 -70.3
- partnerFamiliarity       1      0.48  45.7 -70.0
- age                      1      0.72  45.9 -69.4
- searchExperience         1      0.78  46.0 -69.3
<none>                                  45.2 -69.1
- topicKnowledge           1      1.02  46.2 -68.8
+ GPA                      1      0.23  44.9 -67.6
+ collaborativeSatisfying  1      0.16  45.0 -67.4
+ taskSatisfying           1      0.02  45.2 -67.1
+ partnerAgreement         1      0.02  45.2 -67.1
+ addonIntuitive           1      0.01  45.2 -67.1
- ISEQJust                 1      4.02  49.2 -62.4

Step:  AIC=-70.3
TopicScore ~ age + searchExperience + ISEQJust + topicKnowledge +
    partnerFamiliarity

                           Df Sum of Sq  RSS   AIC
- partnerFamiliarity       1      0.53  46.1 -71.1
- age                      1      0.62  46.2 -70.9
- searchExperience         1      0.73  46.3 -70.6
<none>                                  45.5 -70.3
- topicKnowledge           1      1.06  46.6 -69.9
+ ISEQGen                  1      0.36  45.2 -69.1
+ GPA                      1      0.31  45.2 -69.0
+ collaborativeSatisfying  1      0.20  45.3 -68.7
+ taskSatisfying           1      0.04  45.5 -68.3
+ addonIntuitive           1      0.03  45.5 -68.3
+ partnerAgreement         1      0.02  45.5 -68.3
- ISEQJust                 1      3.67  49.2 -64.4

Step:  AIC=-71.1
TopicScore ~ age + searchExperience + ISEQJust + topicKnowledge

                           Df Sum of Sq  RSS   AIC
- age                      1      0.52  46.6 -71.9
- searchExperience         1      0.75  46.8 -71.4
- topicKnowledge           1      0.85  46.9 -71.2
<none>                                  46.1 -71.1
+ partnerFamiliarity       1      0.53  45.5 -70.3
+ ISEQGen                  1      0.41  45.7 -70.0
+ GPA                      1      0.34  45.7 -69.8
+ collaborativeSatisfying  1      0.26  45.8 -69.7
+ taskSatisfying           1      0.09  46.0 -69.3
+ partnerAgreement         1      0.08  46.0 -69.2
+ addonIntuitive           1      0.06  46.0 -69.2
- ISEQJust                 1      3.68  49.8 -65.2

Step:  AIC=-71.9
TopicScore ~ searchExperience + ISEQJust + topicKnowledge

                           Df Sum of Sq  RSS   AIC
- searchExperience         1      0.61  47.2 -72.6
- topicKnowledge           1      0.85  47.4 -72.1
<none>                                  46.6 -71.9
+ age                      1      0.52  46.1 -71.1
+ partnerFamiliarity       1      0.44  46.2 -70.9
```

```
+ ISEQGen                        1       0.30 46.3 -70.6
+ GPA                            1       0.22 46.4 -70.4
+ collaborativeSatisfying        1       0.18 46.4 -70.3
+ partnerAgreement               1       0.05 46.5 -70.0
+ taskSatisfying                 1       0.04 46.5 -70.0
+ addonIntuitive                 1       0.01 46.6 -69.9
- ISEQJust                       1       3.60 50.2 -66.3


Step:  AIC=-72.6
TopicScore ~ ISEQJust + topicKnowledge

                                Df Sum of Sq  RSS    AIC
- topicKnowledge                 1      0.678 47.9 -73.1
<none>                                       47.2 -72.6
+ searchExperience               1      0.611 46.6 -71.9
+ partnerFamiliarity             1      0.466 46.7 -71.6
+ age                            1      0.384 46.8 -71.4
+ ISEQGen                        1      0.266 46.9 -71.2
+ partnerAgreement               1      0.164 47.0 -70.9
+ collaborativeSatisfying        1      0.163 47.0 -70.9
+ GPA                            1      0.148 47.1 -70.9
+ taskSatisfying                 1      0.055 47.1 -70.7
+ addonIntuitive                 1      0.003 47.2 -70.6
- ISEQJust                       1      3.089 50.3 -68.1


Step:  AIC=-73.1
TopicScore ~ ISEQJust

                                Df Sum of Sq  RSS    AIC
<none>                                       47.9 -73.1
+ topicKnowledge                 1      0.678 47.2 -72.6
+ searchExperience               1      0.439 47.4 -72.1
+ age                            1      0.400 47.5 -72.0
+ ISEQGen                        1      0.288 47.6 -71.8
+ partnerFamiliarity             1      0.286 47.6 -71.7
+ partnerAgreement               1      0.269 47.6 -71.7
+ collaborativeSatisfying        1      0.172 47.7 -71.5
+ GPA                            1      0.097 47.8 -71.3
+ addonIntuitive                 1      0.018 47.9 -71.2
+ taskSatisfying                 1      0.008 47.9 -71.2
- ISEQJust                       1      2.991 50.9 -69.0
Start:  AIC=-37.8
SynthScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerAgreement + partnerFamiliarity

                                Df Sum of Sq  RSS    AIC
- collaborativeSatisfying        1      0.003 55.7 -39.8
- partnerFamiliarity             1      0.180 55.8 -39.5
- taskSatisfying                 1      0.216 55.9 -39.4
- GPA                            1      0.309 56.0 -39.2
- topicKnowledge                 1      0.515 56.2 -38.9
- addonIntuitive                 1      0.586 56.2 -38.7
- ISEQJust                       1      0.735 56.4 -38.5
- partnerAgreement               1      0.760 56.4 -38.4
- searchExperience               1      1.037 56.7 -37.9
<none>                                       55.7 -37.8
- ISEQGen                        1      1.190 56.8 -37.6
- age                            1      1.400 57.1 -37.3
```

```
Step:  AIC=-39.8
SynthScore ~ age + searchExperience + taskSatisfying + addonIntuitive +
    GPA + ISEQGen + ISEQJust + topicKnowledge + partnerAgreement +
    partnerFamiliarity

                          Df Sum of Sq  RSS    AIC
- partnerFamiliarity       1     0.180 55.8 -41.5
- taskSatisfying           1     0.213 55.9 -41.4
- GPA                      1     0.332 56.0 -41.2
- topicKnowledge           1     0.515 56.2 -40.8
- addonIntuitive           1     0.668 56.3 -40.6
- ISEQJust                 1     0.736 56.4 -40.4
- partnerAgreement         1     0.829 56.5 -40.3
- searchExperience         1     1.034 56.7 -39.9
<none>                                 55.7 -39.8
- ISEQGen                  1     1.187 56.8 -39.6
- age                      1     1.401 57.1 -39.3
+ collaborativeSatisfying  1     0.003 55.7 -37.8

Step:  AIC=-41.5
SynthScore ~ age + searchExperience + taskSatisfying + addonIntuitive +
    GPA + ISEQGen + ISEQJust + topicKnowledge + partnerAgreement

                          Df Sum of Sq  RSS    AIC
- taskSatisfying           1     0.262 56.1 -43.0
- GPA                      1     0.339 56.2 -42.8
- topicKnowledge           1     0.430 56.3 -42.7
- addonIntuitive           1     0.678 56.5 -42.2
- partnerAgreement         1     0.728 56.6 -42.1
- ISEQJust                 1     0.737 56.6 -42.1
- searchExperience         1     1.060 56.9 -41.5
<none>                                 55.8 -41.5
- ISEQGen                  1     1.236 57.1 -41.2
- age                      1     1.468 57.3 -40.8
+ partnerFamiliarity       1     0.180 55.7 -39.8
+ collaborativeSatisfying  1     0.003 55.8 -39.5

Step:  AIC=-43
SynthScore ~ age + searchExperience + addonIntuitive + GPA +
    ISEQGen + ISEQJust + topicKnowledge + partnerAgreement

                          Df Sum of Sq  RSS    AIC
- GPA                      1     0.245 56.3 -44.5
- topicKnowledge           1     0.305 56.4 -44.4
- addonIntuitive           1     0.429 56.5 -44.2
- ISEQJust                 1     0.606 56.7 -43.9
- partnerAgreement         1     0.675 56.8 -43.8
- searchExperience         1     0.998 57.1 -43.2
<none>                                 56.1 -43.0
- ISEQGen                  1     1.235 57.3 -42.8
- age                      1     1.441 57.5 -42.4
+ taskSatisfying           1     0.262 55.8 -41.5
+ partnerFamiliarity       1     0.229 55.9 -41.4
+ collaborativeSatisfying  1     0.001 56.1 -41.0

Step:  AIC=-44.5
SynthScore ~ age + searchExperience + addonIntuitive + ISEQGen +
    ISEQJust + topicKnowledge + partnerAgreement
```

```
                              Df Sum of Sq  RSS    AIC
- topicKnowledge            1     0.388 56.7 -45.8
- addonIntuitive            1     0.402 56.7 -45.8
- ISEQJust                  1     0.671 57.0 -45.3
- partnerAgreement          1     0.834 57.2 -45.0
- searchExperience          1     0.869 57.2 -45.0
- ISEQGen                   1     1.113 57.5 -44.5
<none>                                  56.3 -44.5
- age                       1     1.314 57.7 -44.2
+ GPA                       1     0.245 56.1 -43.0
+ partnerFamiliarity        1     0.226 56.1 -42.9
+ taskSatisfying            1     0.168 56.2 -42.8
+ collaborativeSatisfying   1     0.006 56.3 -42.5

Step:  AIC=-45.8
SynthScore ~ age + searchExperience + addonIntuitive + ISEQGen +
    ISEQJust + partnerAgreement

                              Df Sum of Sq  RSS    AIC
- addonIntuitive            1     0.355 57.1 -47.2
- partnerAgreement          1     0.671 57.4 -46.6
- ISEQJust                  1     0.738 57.5 -46.5
- searchExperience          1     1.100 57.8 -45.9
- ISEQGen                   1     1.123 57.9 -45.8
<none>                                  56.7 -45.8
- age                       1     1.276 58.0 -45.6
+ topicKnowledge            1     0.388 56.3 -44.5
+ GPA                       1     0.328 56.4 -44.4
+ partnerFamiliarity        1     0.117 56.6 -44.0
+ taskSatisfying            1     0.057 56.7 -43.9
+ collaborativeSatisfying   1     0.015 56.7 -43.9

Step:  AIC=-47.2
SynthScore ~ age + searchExperience + ISEQGen + ISEQJust + partnerAgreement

                              Df Sum of Sq  RSS    AIC
- ISEQJust                  1     0.720 57.8 -47.9
- partnerAgreement          1     0.833 57.9 -47.7
- searchExperience          1     0.987 58.1 -47.5
- ISEQGen                   1     0.991 58.1 -47.4
- age                       1     1.083 58.2 -47.3
<none>                                  57.1 -47.2
+ addonIntuitive            1     0.355 56.7 -45.8
+ topicKnowledge            1     0.340 56.7 -45.8
+ GPA                       1     0.293 56.8 -45.7
+ collaborativeSatisfying   1     0.120 57.0 -45.4
+ partnerFamiliarity        1     0.104 57.0 -45.4
+ taskSatisfying            1     0.004 57.1 -45.2

Step:  AIC=-47.9
SynthScore ~ age + searchExperience + ISEQGen + partnerAgreement

                              Df Sum of Sq  RSS    AIC
- searchExperience          1     0.656 58.5 -48.8
- partnerAgreement          1     0.848 58.7 -48.4
- age                       1     0.976 58.8 -48.2
<none>                                  57.8 -47.9
+ ISEQJust                  1     0.720 57.1 -47.2
```

```
- ISEQGen                        1      1.675 59.5 -47.0
+ topicKnowledge                 1      0.403 57.4 -46.6
+ GPA                            1      0.375 57.4 -46.6
+ addonIntuitive                 1      0.337 57.5 -46.5
+ collaborativeSatisfying        1      0.153 57.7 -46.2
+ partnerFamiliarity             1      0.088 57.7 -46.1
+ taskSatisfying                 1      0.044 57.8 -46.0

Step:  AIC=-48.8
SynthScore ~ age + ISEQGen + partnerAgreement

                              Df Sum of Sq  RSS    AIC
- age                          1      0.834 59.3 -49.3
<none>                                      58.5 -48.8
- partnerAgreement             1      1.280 59.7 -48.6
- ISEQGen                      1      1.653 60.1 -47.9
+ searchExperience             1      0.656 57.8 -47.9
+ topicKnowledge               1      0.572 57.9 -47.8
+ ISEQJust                     1      0.390 58.1 -47.5
+ addonIntuitive               1      0.247 58.2 -47.2
+ GPA                          1      0.228 58.2 -47.2
+ partnerFamiliarity           1      0.088 58.4 -46.9
+ collaborativeSatisfying      1      0.083 58.4 -46.9
+ taskSatisfying               1      0.037 58.4 -46.8

Step:  AIC=-49.3
SynthScore ~ ISEQGen + partnerAgreement

                              Df Sum of Sq  RSS    AIC
- partnerAgreement             1      1.085 60.4 -49.5
<none>                                      59.3 -49.3
+ age                          1      0.834 58.5 -48.8
+ topicKnowledge               1      0.523 58.8 -48.2
+ searchExperience             1      0.514 58.8 -48.2
- ISEQGen                      1      1.941 61.2 -48.0
+ ISEQJust                     1      0.346 59.0 -47.9
+ GPA                          1      0.138 59.2 -47.6
+ partnerFamiliarity           1      0.131 59.2 -47.5
+ addonIntuitive               1      0.115 59.2 -47.5
+ collaborativeSatisfying      1      0.047 59.3 -47.4
+ taskSatisfying               1      0.010 59.3 -47.3

Step:  AIC=-49.5
SynthScore ~ ISEQGen

                              Df Sum of Sq  RSS    AIC
<none>                                      60.4 -49.5
+ partnerAgreement             1      1.085 59.3 -49.3
+ searchExperience             1      0.884 59.5 -49.0
+ age                          1      0.638 59.7 -48.6
- ISEQGen                      1      1.823 62.2 -48.4
+ topicKnowledge               1      0.322 60.1 -48.0
+ ISEQJust                     1      0.295 60.1 -48.0
+ collaborativeSatisfying      1      0.253 60.1 -47.9
+ GPA                          1      0.235 60.1 -47.9
+ addonIntuitive               1      0.229 60.2 -47.9
+ taskSatisfying               1      0.049 60.3 -47.6
+ partnerFamiliarity           1      0.039 60.3 -47.5
Start:  AIC=-117
```

```
SourDivScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerAgreement + partnerFamiliarity

                            Df Sum of Sq  RSS  AIC
- partnerAgreement           1      0.086 25.7 -118
- addonIntuitive             1      0.198 25.8 -118
- taskSatisfying             1      0.214 25.9 -118
- collaborativeSatisfying    1      0.218 25.9 -118
- partnerFamiliarity         1      0.227 25.9 -118
- ISEQGen                    1      0.300 25.9 -118
- GPA                        1      0.340 26.0 -118
- searchExperience           1      0.487 26.1 -117
<none>                                   25.6 -117
- topicKnowledge             1      0.542 26.2 -117
- age                        1      1.364 27.0 -114
- ISEQJust                   1      1.768 27.4 -112

Step:  AIC=-119
SourDivScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerFamiliarity

                            Df Sum of Sq  RSS  AIC
- collaborativeSatisfying    1      0.168 25.9 -120
- partnerFamiliarity         1      0.190 25.9 -120
- addonIntuitive             1      0.195 25.9 -120
- taskSatisfying             1      0.215 25.9 -120
- ISEQGen                    1      0.308 26.0 -119
- GPA                        1      0.311 26.0 -119
- searchExperience           1      0.412 26.1 -119
- topicKnowledge             1      0.476 26.2 -119
<none>                                   25.7 -118
+ partnerAgreement           1      0.086 25.6 -117
- age                        1      1.383 27.1 -115
- ISEQJust                   1      1.772 27.5 -114

Step:  AIC=-120
SourDivScore ~ age + searchExperience + taskSatisfying + addonIntuitive +
    GPA + ISEQGen + ISEQJust + topicKnowledge + partnerFamiliarity

                            Df Sum of Sq  RSS  AIC
- taskSatisfying             1      0.156 26.1 -121
- partnerFamiliarity         1      0.201 26.1 -121
- ISEQGen                    1      0.316 26.2 -121
- addonIntuitive             1      0.351 26.2 -120
- searchExperience           1      0.412 26.3 -120
- GPA                        1      0.437 26.3 -120
- topicKnowledge             1      0.500 26.4 -120
<none>                                   25.9 -120
+ collaborativeSatisfying    1      0.168 25.7 -118
+ partnerAgreement           1      0.036 25.9 -118
- age                        1      1.386 27.3 -116
- ISEQJust                   1      1.782 27.7 -115

Step:  AIC=-121
SourDivScore ~ age + searchExperience + addonIntuitive + GPA +
    ISEQGen + ISEQJust + topicKnowledge + partnerFamiliarity
```

```
                             Df Sum of Sq  RSS  AIC
- partnerFamiliarity        1      0.164 26.2 -123
- addonIntuitive            1      0.208 26.2 -122
- ISEQGen                   1      0.320 26.4 -122
- GPA                       1      0.355 26.4 -122
- searchExperience          1      0.374 26.4 -122
<none>                                     26.1 -121
- topicKnowledge            1      0.647 26.7 -121
+ taskSatisfying            1      0.156 25.9 -120
+ collaborativeSatisfying   1      0.109 25.9 -120
+ partnerAgreement          1      0.043 26.0 -119
- age                       1      1.361 27.4 -118
- ISEQJust                  1      1.650 27.7 -117

Step:  AIC=-123
SourDivScore ~ age + searchExperience + addonIntuitive + GPA +
    ISEQGen + ISEQJust + topicKnowledge


                             Df Sum of Sq  RSS  AIC
- addonIntuitive            1      0.235 26.4 -124
- ISEQGen                   1      0.342 26.6 -123
- GPA                       1      0.371 26.6 -123
- searchExperience          1      0.387 26.6 -123
<none>                                     26.2 -123
- topicKnowledge            1      0.559 26.8 -122
+ partnerFamiliarity        1      0.164 26.1 -121
+ collaborativeSatisfying   1      0.123 26.1 -121
+ taskSatisfying            1      0.118 26.1 -121
+ partnerAgreement          1      0.018 26.2 -121
- age                       1      1.307 27.5 -120
- ISEQJust                  1      1.672 27.9 -118

Step:  AIC=-124
SourDivScore ~ age + searchExperience + GPA + ISEQGen + ISEQJust +
    topicKnowledge


                             Df Sum of Sq  RSS  AIC
- searchExperience          1      0.352 26.8 -124
- GPA                       1      0.358 26.8 -124
- ISEQGen                   1      0.422 26.9 -124
<none>                                     26.4 -124
- topicKnowledge            1      0.627 27.1 -123
+ collaborativeSatisfying   1      0.275 26.2 -123
+ addonIntuitive            1      0.235 26.2 -123
+ partnerFamiliarity        1      0.191 26.2 -122
+ partnerAgreement          1      0.005 26.4 -122
+ taskSatisfying            1      0.002 26.4 -122
- age                       1      1.148 27.6 -121
- ISEQJust                  1      1.661 28.1 -120

Step:  AIC=-124
SourDivScore ~ age + GPA + ISEQGen + ISEQJust + topicKnowledge


                             Df Sum of Sq  RSS  AIC
- GPA                       1      0.283 27.1 -125
- ISEQGen                   1      0.395 27.2 -125
- topicKnowledge            1      0.509 27.3 -124
<none>                                     26.8 -124
+ searchExperience          1      0.352 26.4 -124
```

```
+ collaborativeSatisfying  1      0.267 26.5 -123
+ partnerFamiliarity       1      0.202 26.6 -123
+ addonIntuitive           1      0.200 26.6 -123
- age                      1      0.993 27.8 -123
+ partnerAgreement         1      0.008 26.8 -122
+ taskSatisfying           1      0.001 26.8 -122
- ISEQJust                 1      1.393 28.2 -121

Step:  AIC=-125
SourDivScore ~ age + ISEQGen + ISEQJust + topicKnowledge

                           Df Sum of Sq  RSS  AIC
- topicKnowledge           1      0.446 27.5 -126
- ISEQGen                  1      0.501 27.6 -125
<none>                                   27.1 -125
+ collaborativeSatisfying  1      0.387 26.7 -125
+ GPA                      1      0.283 26.8 -124
+ searchExperience         1      0.276 26.8 -124
+ partnerFamiliarity       1      0.216 26.9 -124
+ addonIntuitive           1      0.193 26.9 -124
- age                      1      0.883 28.0 -124
+ partnerAgreement         1      0.023 27.1 -123
+ taskSatisfying           1      0.004 27.1 -123
- ISEQJust                 1      1.560 28.6 -122

Step:  AIC=-126
SourDivScore ~ age + ISEQGen + ISEQJust

                           Df Sum of Sq  RSS  AIC
- ISEQGen                  1      0.528 28.1 -126
<none>                                   27.5 -126
+ topicKnowledge           1      0.446 27.1 -125
+ collaborativeSatisfying  1      0.397 27.1 -125
+ addonIntuitive           1      0.252 27.3 -124
+ GPA                      1      0.220 27.3 -124
+ searchExperience         1      0.187 27.3 -124
- age                      1      0.909 28.4 -124
+ partnerFamiliarity       1      0.119 27.4 -124
+ partnerAgreement         1      0.061 27.5 -124
+ taskSatisfying           1      0.003 27.5 -124
- ISEQJust                 1      1.517 29.0 -122

Step:  AIC=-126
SourDivScore ~ age + ISEQJust

                           Df Sum of Sq  RSS  AIC
<none>                                   28.1 -126
+ ISEQGen                  1      0.528 27.5 -126
+ collaborativeSatisfying  1      0.490 27.6 -126
+ topicKnowledge           1      0.473 27.6 -125
+ addonIntuitive           1      0.347 27.7 -125
- age                      1      0.764 28.8 -125
+ GPA                      1      0.316 27.7 -125
+ searchExperience         1      0.153 27.9 -124
+ partnerFamiliarity       1      0.147 27.9 -124
+ partnerAgreement         1      0.067 28.0 -124
+ taskSatisfying           1      0.000 28.1 -124
- ISEQJust                 1      1.145 29.2 -124
Start:  AIC=-36.5
```

```
SourQualScore ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerAgreement + partnerFamiliarity

                          Df Sum of Sq  RSS    AIC
- collaborativeSatisfying  1     0.013 56.4 -38.4
- ISEQJust                 1     0.052 56.4 -38.4
- taskSatisfying           1     0.097 56.5 -38.3
- topicKnowledge           1     0.212 56.6 -38.1
- searchExperience         1     0.529 56.9 -37.5
- age                      1     0.539 56.9 -37.5
- ISEQGen                  1     0.604 57.0 -37.4
- GPA                      1     0.679 57.1 -37.2
<none>                                 56.4 -36.5
- partnerFamiliarity       1     1.158 57.5 -36.4
- addonIntuitive           1     1.188 57.6 -36.3
- partnerAgreement         1     1.708 58.1 -35.4

Step:  AIC=-38.4
SourQualScore ~ age + searchExperience + taskSatisfying + addonIntuitive +
    GPA + ISEQGen + ISEQJust + topicKnowledge + partnerAgreement +
    partnerFamiliarity

                          Df Sum of Sq  RSS    AIC
- ISEQJust                 1     0.053 56.4 -40.3
- taskSatisfying           1     0.087 56.5 -40.3
- topicKnowledge           1     0.214 56.6 -40.1
- searchExperience         1     0.520 56.9 -39.5
- age                      1     0.540 56.9 -39.5
- ISEQGen                  1     0.601 57.0 -39.4
- GPA                      1     0.741 57.1 -39.1
<none>                                 56.4 -38.4
- partnerFamiliarity       1     1.160 57.6 -38.4
- addonIntuitive           1     1.379 57.8 -38.0
- partnerAgreement         1     1.883 58.3 -37.1
+ collaborativeSatisfying  1     0.013 56.4 -36.5

Step:  AIC=-40.4
SourQualScore ~ age + searchExperience + taskSatisfying + addonIntuitive +
    GPA + ISEQGen + topicKnowledge + partnerAgreement + partnerFamiliarity

                          Df Sum of Sq  RSS    AIC
- taskSatisfying           1     0.066 56.5 -42.2
- topicKnowledge           1     0.215 56.7 -42.0
- searchExperience         1     0.471 56.9 -41.5
- age                      1     0.523 57.0 -41.4
- GPA                      1     0.758 57.2 -41.0
- ISEQGen                  1     0.768 57.2 -41.0
<none>                                 56.4 -40.3
- partnerFamiliarity       1     1.160 57.6 -40.3
- addonIntuitive           1     1.339 57.8 -40.0
- partnerAgreement         1     1.882 58.3 -39.0
+ ISEQJust                 1     0.053 56.4 -38.4
+ collaborativeSatisfying  1     0.014 56.4 -38.4

Step:  AIC=-42.2
SourQualScore ~ age + searchExperience + addonIntuitive + GPA +
    ISEQGen + topicKnowledge + partnerAgreement + partnerFamiliarity
```

```
                               Df Sum of Sq  RSS    AIC
- topicKnowledge               1      0.175 56.7  -43.9
- searchExperience             1      0.463 57.0  -43.4
- age                          1      0.516 57.0  -43.3
- GPA                          1      0.700 57.2  -43.0
- ISEQGen                      1      0.743 57.3  -42.9
<none>                                      56.5  -42.2
- partnerFamiliarity           1      1.238 57.8  -42.0
- addonIntuitive               1      1.401 57.9  -41.7
- partnerAgreement             1      1.853 58.4  -40.9
+ taskSatisfying               1      0.066 56.4  -40.3
+ ISEQJust                     1      0.032 56.5  -40.3
+ collaborativeSatisfying      1      0.005 56.5  -40.2

Step:  AIC=-43.9
SourQualScore ~ age + searchExperience + addonIntuitive + GPA +
    ISEQGen + partnerAgreement + partnerFamiliarity


                               Df Sum of Sq  RSS    AIC
- age                          1      0.515 57.2  -45.0
- searchExperience             1      0.587 57.3  -44.9
- ISEQGen                      1      0.781 57.5  -44.5
- GPA                          1      0.808 57.5  -44.5
- partnerFamiliarity           1      1.104 57.8  -43.9
<none>                                      56.7  -43.9
- addonIntuitive               1      1.342 58.0  -43.5
- partnerAgreement             1      1.696 58.4  -42.9
+ topicKnowledge               1      0.175 56.5  -42.2
+ ISEQJust                     1      0.039 56.7  -42.0
+ taskSatisfying               1      0.026 56.7  -42.0
+ collaborativeSatisfying      1      0.008 56.7  -41.9

Step:  AIC=-45
SourQualScore ~ searchExperience + addonIntuitive + GPA + ISEQGen +
    partnerAgreement + partnerFamiliarity


                               Df Sum of Sq  RSS    AIC
- searchExperience             1      0.458 57.7  -46.2
- GPA                          1      0.657 57.9  -45.8
- ISEQGen                      1      0.866 58.1  -45.5
- addonIntuitive               1      1.097 58.3  -45.0
<none>                                      57.2  -45.0
- partnerFamiliarity           1      1.201 58.4  -44.9
- partnerAgreement             1      1.651 58.9  -44.1
+ age                          1      0.515 56.7  -43.9
+ topicKnowledge               1      0.174 57.0  -43.3
+ ISEQJust                     1      0.025 57.2  -43.0
+ taskSatisfying               1      0.022 57.2  -43.0
+ collaborativeSatisfying      1      0.010 57.2  -43.0

Step:  AIC=-46.2
SourQualScore ~ addonIntuitive + GPA + ISEQGen + partnerAgreement +
    partnerFamiliarity


                               Df Sum of Sq  RSS    AIC
- GPA                          1      0.516 58.2  -47.3
- ISEQGen                      1      0.807 58.5  -46.8
- addonIntuitive               1      0.986 58.7  -46.4
<none>                                      57.7  -46.2
```

```
- partnerFamiliarity        1      1.177 58.8 -46.1
+ searchExperience          1      0.458 57.2 -45.0
+ age                       1      0.387 57.3 -44.9
+ topicKnowledge            1      0.282 57.4 -44.7
- partnerAgreement          1      2.271 59.9 -44.2
+ taskSatisfying            1      0.012 57.7 -44.2
+ collaborativeSatisfying   1      0.003 57.7 -44.2
+ ISEQJust                  1      0.000 57.7 -44.2

Step:  AIC=-47.3
SourQualScore ~ addonIntuitive + ISEQGen + partnerAgreement +
    partnerFamiliarity

                           Df Sum of Sq  RSS   AIC
- ISEQGen                   1      0.675 58.9 -48.1
- addonIntuitive            1      0.961 59.1 -47.6
- partnerFamiliarity        1      1.117 59.3 -47.3
<none>                                   58.2 -47.3
+ GPA                       1      0.516 57.7 -46.2
+ topicKnowledge            1      0.365 57.8 -45.9
+ searchExperience          1      0.318 57.9 -45.8
+ age                       1      0.286 57.9 -45.8
+ collaborativeSatisfying   1      0.053 58.1 -45.4
+ ISEQJust                  1      0.008 58.2 -45.3
+ taskSatisfying            1      0.002 58.2 -45.3
- partnerAgreement          1      2.535 60.7 -44.9

Step:  AIC=-48.1
SourQualScore ~ addonIntuitive + partnerAgreement + partnerFamiliarity

                           Df Sum of Sq  RSS   AIC
- addonIntuitive            1      0.760 59.6 -48.8
<none>                                   58.9 -48.1
- partnerFamiliarity        1      1.201 60.1 -48.0
+ ISEQGen                   1      0.675 58.2 -47.3
+ topicKnowledge            1      0.393 58.5 -46.8
+ GPA                       1      0.385 58.5 -46.8
+ age                       1      0.361 58.5 -46.7
+ searchExperience          1      0.288 58.6 -46.6
+ ISEQJust                  1      0.101 58.8 -46.3
+ collaborativeSatisfying   1      0.042 58.8 -46.2
+ taskSatisfying            1      0.007 58.8 -46.1
- partnerAgreement          1      2.497 61.4 -45.8

Step:  AIC=-48.8
SourQualScore ~ partnerAgreement + partnerFamiliarity

                           Df Sum of Sq  RSS   AIC
- partnerFamiliarity        1      1.108 60.7 -48.9
<none>                                   59.6 -48.8
+ addonIntuitive            1      0.760 58.9 -48.1
+ ISEQGen                   1      0.475 59.1 -47.6
+ GPA                       1      0.383 59.2 -47.4
+ topicKnowledge            1      0.300 59.3 -47.3
+ collaborativeSatisfying   1      0.293 59.3 -47.3
+ taskSatisfying            1      0.229 59.4 -47.2
+ searchExperience          1      0.218 59.4 -47.2
+ age                       1      0.192 59.4 -47.1
+ ISEQJust                  1      0.086 59.5 -46.9
```

```
- partnerAgreement            1    2.929 62.5 -45.9

Step:  AIC=-48.9
SourQualScore ~ partnerAgreement

                            Df Sum of Sq  RSS    AIC
<none>                                    60.7 -48.9
+ partnerFamiliarity         1    1.108 59.6 -48.8
+ addonIntuitive             1    0.667 60.1 -48.0
+ ISEQGen                    1    0.553 60.2 -47.8
+ GPA                        1    0.325 60.4 -47.4
+ age                        1    0.269 60.5 -47.4
+ collaborativeSatisfying    1    0.238 60.5 -47.3
+ searchExperience           1    0.210 60.5 -47.3
+ topicKnowledge             1    0.111 60.6 -47.1
+ taskSatisfying             1    0.091 60.6 -47.1
+ ISEQJust                   1    0.076 60.6 -47.0
- partnerAgreement           1    2.452 63.2 -46.9
Start:  AIC=127
totalhand ~ age + searchExperience + collaborativeSatisfying +
    taskSatisfying + addonIntuitive + GPA + ISEQGen + ISEQJust +
    topicKnowledge + partnerAgreement + partnerFamiliarity

                            Df Sum of Sq RSS AIC
- searchExperience           1     0.02 280 125
- age                        1     0.03 280 125
- taskSatisfying             1     0.07 280 125
- GPA                        1     0.15 280 125
- collaborativeSatisfying    1     0.32 280 125
- addonIntuitive             1     1.94 282 126
- ISEQJust                   1     4.71 284 127
<none>                                  280 127
- partnerAgreement           1     6.19 286 127
- partnerFamiliarity         1     6.89 286 127
- topicKnowledge             1     8.40 288 128
- ISEQGen                    1     8.53 288 128

Step:  AIC=125
totalhand ~ age + collaborativeSatisfying + taskSatisfying +
    addonIntuitive + GPA + ISEQGen + ISEQJust + topicKnowledge +
    partnerAgreement + partnerFamiliarity

                            Df Sum of Sq RSS AIC
- age                        1     0.03 280 123
- taskSatisfying             1     0.07 280 123
- GPA                        1     0.14 280 123
- collaborativeSatisfying    1     0.33 280 123
- addonIntuitive             1     1.92 282 124
- ISEQJust                   1     5.15 285 125
<none>                                  280 125
- partnerAgreement           1     6.89 287 125
- partnerFamiliarity         1     6.91 287 125
- ISEQGen                    1     8.55 288 126
- topicKnowledge             1     8.77 288 126
+ searchExperience           1     0.02 280 127

Step:  AIC=123
totalhand ~ collaborativeSatisfying + taskSatisfying + addonIntuitive +
    GPA + ISEQGen + ISEQJust + topicKnowledge + partnerAgreement +
```

```
      partnerFamiliarity

                          Df Sum of Sq RSS AIC
- taskSatisfying          1      0.07 280 121
- GPA                     1      0.16 280 121
- collaborativeSatisfying 1      0.34 280 121
- addonIntuitive          1      2.04 282 122
- ISEQJust                1      5.14 285 123
<none>                                280 123
- partnerFamiliarity      1      6.88 287 123
- partnerAgreement        1      6.99 287 123
- ISEQGen                 1      8.52 288 124
- topicKnowledge          1      8.81 288 124
+ age                     1      0.03 280 125
+ searchExperience        1      0.03 280 125

Step:  AIC=121
totalhand ~ collaborativeSatisfying + addonIntuitive + GPA +
    ISEQGen + ISEQJust + topicKnowledge + partnerAgreement +
    partnerFamiliarity

                          Df Sum of Sq RSS AIC
- GPA                     1      0.13 280 119
- collaborativeSatisfying 1      0.41 280 119
- addonIntuitive          1      2.11 282 120
- ISEQJust                1      5.47 285 121
<none>                                280 121
- partnerAgreement        1      6.97 287 121
- partnerFamiliarity      1      7.11 287 122
- ISEQGen                 1      8.50 288 122
- topicKnowledge          1      8.90 289 122
+ taskSatisfying          1      0.07 280 123
+ age                     1      0.04 280 123
+ searchExperience        1      0.02 280 123

Step:  AIC=119
totalhand ~ collaborativeSatisfying + addonIntuitive + ISEQGen +
    ISEQJust + topicKnowledge + partnerAgreement + partnerFamiliarity

                          Df Sum of Sq RSS AIC
- collaborativeSatisfying 1      0.32 280 117
- addonIntuitive          1      2.03 282 118
- ISEQJust                1      5.35 285 119
<none>                                280 119
- partnerFamiliarity      1      7.09 287 120
- partnerAgreement        1      7.13 287 120
- ISEQGen                 1      8.37 288 120
- topicKnowledge          1      9.19 289 120
+ GPA                     1      0.13 280 121
+ age                     1      0.05 280 121
+ taskSatisfying          1      0.04 280 121
+ searchExperience        1      0.01 280 121

Step:  AIC=117
totalhand ~ addonIntuitive + ISEQGen + ISEQJust + topicKnowledge +
    partnerAgreement + partnerFamiliarity

                          Df Sum of Sq RSS AIC
- addonIntuitive          1      1.70 282 116
```

```
<none>                             280 117
- ISEQJust              1     5.61 286 117
- partnerAgreement      1     6.82 287 118
- partnerFamiliarity    1     7.13 287 118
- ISEQGen               1     8.57 289 118
- topicKnowledge        1     9.01 289 118
+ collaborativeSatisfying 1   0.32 280 119
+ taskSatisfying        1     0.10 280 119
+ GPA                   1     0.05 280 119
+ age                   1     0.05 280 119
+ searchExperience      1     0.02 280 119

Step:  AIC=116
totalhand ~ ISEQGen + ISEQJust + topicKnowledge + partnerAgreement +
    partnerFamiliarity

                        Df Sum of Sq RSS AIC
- ISEQJust              1     5.53 288 116
<none>                             282 116
- partnerFamiliarity    1     6.75 289 116
- ISEQGen               1     7.70 290 116
- partnerAgreement      1     7.72 290 116
- topicKnowledge        1     8.43 290 117
+ addonIntuitive        1     1.70 280 117
+ age                   1     0.19 282 118
+ taskSatisfying        1     0.13 282 118
+ GPA                   1     0.04 282 118
+ searchExperience      1     0.00 282 118
+ collaborativeSatisfying 1   0.00 282 118

Step:  AIC=116
totalhand ~ ISEQGen + topicKnowledge + partnerAgreement + partnerFamiliarity

                        Df Sum of Sq RSS AIC
- ISEQGen               1     4.86 292 115
<none>                             288 116
+ ISEQJust              1     5.53 282 116
- partnerFamiliarity    1     7.11 295 116
- topicKnowledge        1     8.22 296 117
- partnerAgreement      1     8.33 296 117
+ addonIntuitive        1     1.62 286 117
+ searchExperience      1     0.40 287 118
+ age                   1     0.12 287 118
+ collaborativeSatisfying 1   0.03 287 118
+ GPA                   1     0.01 288 118
+ taskSatisfying        1     0.00 288 118

Step:  AIC=115
totalhand ~ topicKnowledge + partnerAgreement + partnerFamiliarity

                        Df Sum of Sq RSS AIC
<none>                             292 115
+ ISEQGen               1     4.86 288 116
- partnerFamiliarity    1     7.86 300 116
- partnerAgreement      1     8.00 300 116
- topicKnowledge        1     8.75 301 116
+ ISEQJust              1     2.69 290 116
+ addonIntuitive        1     0.91 291 117
+ searchExperience      1     0.33 292 117
```

```
+ collaborativeSatisfying  1      0.14 292 117
+ GPA                      1      0.11 292 117
+ taskSatisfying           1      0.02 292 117
+ age                      1      0.01 292 117
[[1]]

Call:
lm(formula = TopicScore ~ ISEQJust, data = no.na.data)

Coefficients:
(Intercept)      ISEQJust
       3.29         -0.31


[[2]]

Call:
lm(formula = SynthScore ~ ISEQGen, data = no.na.data)

Coefficients:
(Intercept)      ISEQGen
      2.939        -0.249


[[3]]

Call:
lm(formula = SourDivScore ~ age + ISEQJust, data = no.na.data)

Coefficients:
(Intercept)          age     ISEQJust
      8.854       -1.916       -0.192


[[4]]

Call:
lm(formula = SourQualScore ~ partnerAgreement, data = no.na.data)

Coefficients:
    (Intercept)   partnerAgreement
          0.780              0.119


[[5]]

Call:
lm(formula = totalhand ~ topicKnowledge + partnerAgreement +
    partnerFamiliarity, data = no.na.data)

Coefficients:
      (Intercept)        topicKnowledge        partnerAgreement
partnerFamiliarity
            6.952                 0.536                   0.221                    -
0.110
```

Appendix 16 – Raw Stepwise Regression for Trace and Survey Data Predictive of Outcomes

Cis

topicscore

Step: AIC=-67.4

TopicScore ~ contrib_symmetry + nchar + ChatTaskTotal + synthesis +

   query_depth + searchExperience + taskSatisfying + topicKnowledge +

   partnerAgreement + partnerFamiliarity

Step: AIC=-68.3

TopicScore ~ contrib_symmetry + nchar + ChatTaskTotal + synthesis +

   query_depth + searchExperience + taskSatisfying + topicKnowledge +

   partnerFamiliarity

Step: AIC=-68.5

TopicScore ~ contrib_symmetry + nchar + ChatTaskTotal + synthesis +

   searchExperience + taskSatisfying + topicKnowledge + partnerFamiliarity

Step: AIC=-68.9

TopicScore ~ contrib_symmetry + nchar + ChatTaskTotal + synthesis +

   taskSatisfying + topicKnowledge + partnerFamiliarity

Step: AIC=-69.1

TopicScore ~ contrib_symmetry + nchar + ChatTaskTotal + synthesis +

   topicKnowledge + partnerFamiliarity

synthscore

Step: AIC=-103

SynthScore ~ f + nchar + explor + Topic + synthesis + Query_vocabulary_richness +

   searchExperience + collaborativeSatisfying + topicKnowledge +

   partnerAgreement

Step: AIC=-104

SynthScore ~ f + nchar + explor + Topic + synthesis + Query_vocabulary_richness +

   collaborativeSatisfying + topicKnowledge + partnerAgreement

Step: AIC=-105

SynthScore ~ f + explor + Topic + synthesis + Query_vocabulary_richness +

   collaborativeSatisfying + topicKnowledge + partnerAgreement

Step: AIC=-106

SynthScore ~ f + explor + Topic + synthesis + Query_vocabulary_richness +

   topicKnowledge + partnerAgreement

Step: AIC=-107

SynthScore ~ f + explor + Topic + synthesis + topicKnowledge +

   partnerAgreement

Step: AIC=-108

SynthScore ~ f + explor + Topic + topicKnowledge + partnerAgreement

Step: AIC=-108

SynthScore ~ explor + Topic + topicKnowledge + partnerAgreement

sourdivscore
Step: AIC=-98

SourDivScore ~ ISEQGen + ISEQJust + f + nchar + n_touch_points.x +

   explor + ChatTaskTotal + topicKnowledge

Step: AIC=-99.2

SourDivScore ~ ISEQGen + ISEQJust + f + nchar + explor + ChatTaskTotal +

   topicKnowledge

Step: AIC=-101

SourDivScore ~ ISEQGen + ISEQJust + f + nchar + explor + ChatTaskTotal

Step: AIC=-101

SourDivScore ~ ISEQJust + f + nchar + explor + ChatTaskTotal

Step: AIC=-102

SourDivScore ~ f + nchar + explor + ChatTaskTotal

sourqualscore
Step: AIC=-83.8

SourQualScore ~ ISEQGen + ISEQJust + nchar + explor + sourQual +

   Topic + searchExperience + collaborativeSatisfying + taskSatisfying +

   topicKnowledge + partnerAgreement

Step: AIC=-85.4

SourQualScore ~ ISEQGen + nchar + explor + sourQual + Topic +

   searchExperience + collaborativeSatisfying + taskSatisfying +

   topicKnowledge + partnerAgreement

Step: AIC=-87

SourQualScore ~ ISEQGen + nchar + explor + sourQual + Topic +

  searchExperience + collaborativeSatisfying + topicKnowledge +

  partnerAgreement

Step: AIC=-88.3

SourQualScore ~ ISEQGen + nchar + explor + sourQual + Topic +

  searchExperience + collaborativeSatisfying + topicKnowledge

Step: AIC=-88.8

SourQualScore ~ ISEQGen + nchar + explor + Topic + searchExperience +

  collaborativeSatisfying + topicKnowledge

totalhand
Step: AIC=168

totalhand ~ ISEQGen + ISEQJust + f + nchar + n_touch_points.x +

  explor + ChatTaskTotal + Topic + collaborativeSatisfying

Step: AIC=167

totalhand ~ ISEQGen + ISEQJust + f + nchar + explor + ChatTaskTotal +

  Topic + collaborativeSatisfying

Step: AIC=166

totalhand ~ ISEQGen + ISEQJust + f + nchar + explor + ChatTaskTotal +

  Topic

Step: AIC=165

totalhand ~ ISEQGen + f + nchar + explor + ChatTaskTotal + Topic

Step: AIC=163

totalhand ~ f + nchar + explor + ChatTaskTotal + Topic

Step: AIC=163

totalhand ~ f + nchar + explor + ChatTaskTotal

Mdp

Topicscore
Step: AIC=-91.3

TopicScore ~ ISEQJust + f + nchar + ChatTaskTotal + sourQual +

  searchExperience + collaborativeSatisfying + addonIntuitive

Step: AIC=-92.4

TopicScore ~ ISEQJust + f + nchar + ChatTaskTotal + searchExperience +

   collaborativeSatisfying + addonIntuitive

Step: AIC=-93.4

TopicScore ~ ISEQJust + f + nchar + ChatTaskTotal + collaborativeSatisfying +

   addonIntuitive

Step: AIC=-94.4

TopicScore ~ ISEQJust + f + nchar + ChatTaskTotal + addonIntuitive

Step: AIC=-95.6

TopicScore ~ ISEQJust + f + nchar + ChatTaskTotal

Step: AIC=-95.8

TopicScore ~ f + nchar + ChatTaskTotal

synthscore
Step: AIC=-47.8

SynthScore ~ ISEQGen + contrib_symmetry + nchar + sourQual +

   synthesis + topicKnowledge + partnerAgreement + partnerFamiliarity

Step: AIC=-49.3

SynthScore ~ ISEQGen + contrib_symmetry + nchar + sourQual +

   topicKnowledge + partnerAgreement + partnerFamiliarity

Step: AIC=-50.7

SynthScore ~ ISEQGen + contrib_symmetry + nchar + sourQual +

   topicKnowledge + partnerAgreement

Step: AIC=-51.5

SynthScore ~ ISEQGen + contrib_symmetry + nchar + topicKnowledge +

   partnerAgreement

Step: AIC=-52.2

SynthScore ~ ISEQGen + contrib_symmetry + nchar + partnerAgreement

sourdivscore
Step: AIC=-156

SourDivScore ~ ISEQJust + f + nchar + explor + collaborativeSatisfying +

   taskSatisfying + addonIntuitive + topicKnowledge + partnerFamiliarity

Step: AIC=-158

SourDivScore ~ ISEQJust + f + nchar + explor + collaborativeSatisfying +

taskSatisfying + addonIntuitive + partnerFamiliarity

Step: AIC=-159

SourDivScore ~ ISEQJust + f + nchar + explor + collaborativeSatisfying +

   taskSatisfying + addonIntuitive

Step: AIC=-160

SourDivScore ~ ISEQJust + f + nchar + explor + collaborativeSatisfying +

   addonIntuitive

Step: AIC=-161

SourDivScore ~ f + nchar + explor + collaborativeSatisfying +

   addonIntuitive

sourqualscore

Step: AIC=-42.8

SourQualScore ~ ISEQJust + contrib_symmetry + nchar + sourQual +

   synthesis + searchExperience + addonIntuitive + partnerAgreement +

   partnerFamiliarity

Step: AIC=-43.5

SourQualScore ~ contrib_symmetry + nchar + sourQual + synthesis +

   searchExperience + addonIntuitive + partnerAgreement + partnerFamiliarity

Step: AIC=-44.4

SourQualScore ~ contrib_symmetry + sourQual + synthesis + searchExperience +

   addonIntuitive + partnerAgreement + partnerFamiliarity

Step: AIC=-45.5

SourQualScore ~ contrib_symmetry + sourQual + synthesis + addonIntuitive +

   partnerAgreement + partnerFamiliarity

Step: AIC=-46.5

SourQualScore ~ contrib_symmetry + sourQual + synthesis + partnerAgreement +

   partnerFamiliarity

Step: AIC=-47

SourQualScore ~ contrib_symmetry + sourQual + synthesis + partnerAgreement

Step: AIC=-47.3

SourQualScore ~ contrib_symmetry + sourQual + partnerAgreement

Step: AIC=-47.6

SourQualScore ~ contrib_symmetry + partnerAgreement

totalhand

Step: AIC=116

totalhand ~ ISEQGen + f + contrib_symmetry + nchar + ChatTaskTotal +

   sourQual + collaborativeSatisfying + addonIntuitive + topicKnowledge +

   partnerAgreement + partnerFamiliarity

Step: AIC=115

totalhand ~ ISEQGen + f + contrib_symmetry + nchar + ChatTaskTotal +

   collaborativeSatisfying + addonIntuitive + topicKnowledge +

   partnerAgreement + partnerFamiliarity

Step: AIC=113

totalhand ~ ISEQGen + f + contrib_symmetry + nchar + ChatTaskTotal +

   addonIntuitive + topicKnowledge + partnerAgreement + partnerFamiliarity

Step: AIC=113

totalhand ~ ISEQGen + f + contrib_symmetry + nchar + addonIntuitive +

   topicKnowledge + partnerAgreement + partnerFamiliarity

Step: AIC=112

totalhand ~ f + contrib_symmetry + nchar + addonIntuitive + topicKnowledge +

   partnerAgreement + partnerFamiliarity

Step: AIC=111

totalhand ~ f + contrib_symmetry + nchar + topicKnowledge + partnerAgreement +

   partnerFamiliarity

Step: AIC=111

totalhand ~ f + contrib_symmetry + nchar + topicKnowledge + partnerAgreement

Appendix 17 – Least and Most Trusted URLs Raw Output List

> unique(trustratingsCIS$leastTrustworthyURL)

[1] ""

[2] "http://www.webmd.com/vitamins-supplements/ingredientmono-925-
red%20yeast%20rice%20%28red%20yeast.aspx?activeingredientid=925&activeingredientname=r
ed%20yeast%20rice%20%28red%20yeast"

[3] "http://www.medicinenet.com/red_yeast_rice_and_cholesterol/article.htm"

[4] "http://www.lexology.com/library/detail.aspx?g=10fb65d0-6d6d-4651-b3b1-20d109168671"

[5] "http://altmedicine.about.com/od/herbsupplementguide/a/redyeastrice.htm"

[6] "http://www.webmd.com/cholesterol-management/red-yeast-rice"

[7] "http://sciencebasedpharmacy.wordpress.com/2009/08/06/red-yeast-rice/"

[8] "http://www.drugs.com/mtm/red-yeast-rice.html"

[9] "Medicine.com (not sure which one it was"

[10] "altmedicine.about.com"

[11]
"http://www.medicinenet.com/red_yeast_rice_and_cholesterol/page4.htm#how_safe_are_red_
yeast_rice_products"

[12] "wikipedia.org"

[13] "wikipedia"

[14] "http://www.emedicinehealth.com/drug-red_yeast_rice/article_em.htm#sideeffects"

[15] "http://www.drugs.com/sfx/lovastatin-side-effects.html"

[16] "www.wikipedia.com"

[17] "http://www.mayoclinic.org/drugs-supplements/red-yeast-rice/safety/hrb-
20059910\r\nDespite the growing interest in dietary supplements, red yeast rice (M. purpureus)
is not recommended for patients with hypercholesterolemia. A lack of uniformity among
products, the possibility of contamination, and the risk of severe adverse reactions pose a threat
to individuals using this product. Overall, red yeast rice has not been shown to be a safe
alternative to statins for patients with hyperlipidemia despite its demonstrated efficacy in
controlled clinical trials. Physicians should be aware of its popularity as a "natural" way to lower
serum cholesterol, and they should discuss the risks and benefits of this supplement with their
patients.\r\nhttp://www.medicinenet.com/red_yeast_rice_and_cholesterol/page4.htm\r\ncheck
this out conclusion \r\n\r\nhttp://en.wikipedia.org/wiki/Red_yeast_rice#Safety"

[18] "www.webmd.com"

[19] "www.wikipedia.org"

[20] "http://en.wikipedia.org/wiki/Red_yeast_rice"

[21] "http://www.mayoclinic.org/drugs-supplements/red-yeast-rice/safety/hrb-20059910"

[22] "http://articles.mercola.com/sites/articles/archive/2009/09/10/why-you-should-avoid-red-
rice-yeast.aspx"

[23] " "

[24] "ww.webmd.com"

> unique(trustratingsCIS$mostTrustworthyURL)

[1] "http://umm.edu/health/medical/altmed/supplement/red-yeast-rice"

[2] ""

[3] "http://articles.mercola.com/sites/articles/archive/2009/09/10/why-you-should-avoid-red-
rice-yeast.aspx"

[4] "used the quciksearch from the universities online libary...."

[5]
"https://www.consumerlab.com/news/red_yeast_rice_reviewed_by_ConsumerLab/08_20_2014/
\r\nhttp://www.reuters.com/article/2008/07/09/us-contamination-common-
idUSCOL97022820080709\r\nhttp://www.linkedin.com/pub/anne-harding/6/a39/40"

[6]
"http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2012/sklm_rotschimm
elreis_121218.pdf"

[7] "http://www.emedicinehealth.com/drug-red_yeast_rice/article_em.htm"

[8] "http://www.ncbi.nlm.nih.gov/pmc/articles/PMC317299/"

[9] "http://clinicaltrials.gov/show/NCT00639223"

[10] "https://www.anses.fr/en/content/food-supplements-containing-red-yeast-rice-consumption-ask-healthcare-professional"

[11] "http://nccam.nih.gov/health/redyeastrice"

[12] "http://www.efsa.europa.eu/en/efsajournal/pub/2304.htm"

[13] "http://nccam.nih.gov/health/redyeastrice\r\nhttp://umm.edu/health/medical/altmed/supplement/red-yeast-rice"

[14] "FDA "

[15] "nccam.nih.gov"

[16] "webpage of ANSES"

[17] "http://en.wikipedia.org/wiki/Red_yeast_rice"

[18] "NCCAM website"

[19] "http://www.fda.gov"

[20] "http://www.sciencebasedmedicine.org/red-yeast-rice-and-cholesterol/"

[21] "http://altmedicine.about.com/od/herbsupplementguide/a/redyeastrice.htm"

[22] "European Food safety autority, I guess"

[23] "http://www.reuters.com/article/2008/07/09/us-contamination-common-idUSCOL97022820080709"

[24] "http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2697909/"

[25] "http://www.mayoclinic.org/drugs-supplements/red-yeast-rice/safety/hrb-20059910\r\nDespite the growing interest in dietary supplements, red yeast rice (M. purpureus) is not recommended for patients with hypercholesterolemia. A lack of uniformity among products, the possibility of contamination, and the risk of severe adverse reactions pose a threat to individuals using this product. Overall, red yeast rice has not been shown to be a safe alternative to statins for patients with hyperlipidemia despite its demonstrated efficacy in controlled clinical trials. Physicians should be aware of its popularity as a "natural" way to lower serum cholesterol, and they should discuss the risks and benefits of this supplement with their patients.\r\nhttp://www.medicinenet.com/red_yeast_rice_and_cholesterol/page4.htm\r\ncheck this out conclusion \r\nhttp://www.ncbi.nlm.nih.gov/pmc/articles/PMC2697909/\r\nyes looks good"

[26] "vsearch.nlm.nih.gov/"

[27] "reuters"

[28] "reuters is basically very reliable"

[29] "http://www.nlm.nih.gov/medlineplus/druginfo/natural/925.html"

[30] "http://www.diabecinn.com/files/8513/1012/3780/Review_of_the_studies_on_the_red_yeast....pdf"

[31] "fda.com"

[32] "http://www.medicinenet.com/red_yeast_rice_and_cholesterol/article.htm"

[33] "http://www.medicinenet.com/red_yeast_rice_and_cholesterol/page4.htm"

[34] "http://www.medicinenet.com/red_yeast_rice_and_cholesterol/page5.htm"

[35] "http://www.healthyheartexperts.com/red-rice-yeast-and-cholesterol/"

[36] "www.UniversityofMaryland.com"

[37] "http://pennstatehershey.adam.com/content.aspx?productId=107&pid=33&gid=000323"

Appendix 18 – Least and Most Trustworthy Rated URLs

Domains given in the most and least trustworthy text box were compiled and tallied for both types. As various sub-pages were given, these raw URLs are largely condensed to top-level domains. However to make it clear which area of the website is generally being referred to some of the URLs below refer to specific pages or areas on the website (e.g. http://umm.edu/health/medical/altmed/ not http://umm.edu/ is listed). A number of pages are listed by different participants as both the most, and least, trustworthy page encountered.

Most count	Least count	domain
215	235	[blank]
14	0	http://umm.edu/health/medical/altmed/
3	7	http://www.medicinenet.com
8	0	http://nccam.nih.gov/health/
1	7	http://en.wikipedia.org/wiki/
0	6	http://www.webmd.com/vitamins-supplements/
5	0	http://www.reuters.com/
3	1	http://articles.mercola.com/
4	0	http://www.nlm.nih.gov/medlineplus/druginfo/
0	4	http://www.mayoclinic.org/drugs-supplements/
1	2	http://altmedicine.about.com/
3	0	http://www.fda.gov
2	0	http://www.efsa.europa.eu/
2	0	https://www.anses.fr/en/content/
0	2	http://www.drugs.com/mtm/red-yeast-rice.html
1	0	http://clinicaltrials.gov/show/NCT00639223
1	0	http://pennstatehershey.adam.com/content.aspx?productId=107&pid=33&gid=000323
1	0	http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2012/sklm_rotschimmelreis_121218.pdf
1	0	http://www.diabecinn.com/files/8513/1012/3780/Review_of_the_studies_on_the_red_yeast....pdf
1	0	http://www.emedicinehealth.com/drug-red_yeast_rice/article_em.htm
1	0	http://www.healthyheartexperts.com/
1	0	http://www.sciencebasedmedicine.org/
1	0	https://www.consumerlab.com/news/red_yeast_rice_reviewed_by_ConsumerLab/08_20_2014/
1	0	used the quciksearch from the universities online libary....
0	1	http://sciencebasedpharmacy.wordpress.com/2009/08/06/red-yeast-rice/
0	1	http://www.emedicinehealth.com/drug-red_yeast_rice/article_em.htm#sideeffects
0	1	http://www.lexology.com/library/detail.aspx?g=10fb65d0-6d6d-4651-b3b1-20d109168671

Appendix 19 – Tabulated Key Relationships Between Thesis Findings and Prior Research Evidence

Data source	Finding from this research	Evidence provided by source	Citation
N communication	Weak, but significant, negative correlations between topic-chat and ISEQ justification factor for CIS but not MDP tasks	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.	(Strømsø & Bråten, 2010)
URLs used	Weak but significant negative correlations between n of pages viewed and the ISEQ Justification factor for MDP but not CIS tasks	Students believing facts needed checking (and reasoning) were more likely to report engaging in self-regulatory strategies like planning.	
URLs used, source quality, topic coverage		Students with doubts about the need to check sources were more likely to have a one-sided representation	(Kammerer et al., 2013)
Source quality chat & scores	No correlation observed between ISEQ scores and source quality chat or output scores.	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.	
Trust ratings and ISEQ scores	No clear patterns in the MDP data. Small relationships between ISEQ scores and CIS trustworthiness scores.	"readers who believe that knowledge claims should be critically evaluated through logic and rules rated the science text as more trustworthy…. These effects hold true while controlling for readers' prior knowledge and text comprehensibility". (pt.abstract)	(Strømsø et al., 2011)
MDP Trust ratings	No clear preference patterns in the MDP data	Participants given 2 texts of varying quality are more likely to favour high quality sources,	Kobayashi (2014)
Source quality chat	Confirmatory findings, source chat is infrequent and few source features are directly referred to in this collaborative context. The outputs have a range of evaluation types in them, with many students failing to engage in any evaluation (including source evaluation) at all.	they make little reference to source features (on average only 1.85 out of 10 features); and rarely (<6% of the 154 participants) explicitly use source information for justifying their evaluation of the text's explanation, that is, they do not make connections between source metadata and their evaluative stance.	
		only 10% of the participants made remarks about the author or their credentials, and none actually verified those credentials, while superficial cues such as search rank and domain names (for example, .edu, .gov) were seen as credibility indicators.	(Hargittai et al., 2010, pt. abstract).
Synthesis MDP Trust Ratings	No patterns in MDP trust ratings which might indicate balancing of perspectives or selectivity of sources. No clear relationships between	students with more sophisticated (evaluativist) epistemic-perspectives have higher comprehension of conflicting sources.	Barzilai and Eshet-Alkalai (2015),

	ISEQ and synthesis scores.		
Page use Query depth Query richness	Delving beyond the first SERP is rare (query depth), there are differences in pageviews and pageuse	dwell time, links opened, and bookmarked pages while reading for a task on a socio-scientific issue indicated that students tended to employ a 'top link' heuristic – trusting search engine results page ranking as an indicator of credibility – while searching web pages for a socio-scientific topic	(Salmerón et al., 2013)
Trust ratings Source chat Evaluation	The lack of pattern in MDP trust ratings, the low incidence of source chat, and range in the output source/evaluation indicates that conflicts, and absence of claims in some of the stronger works (e.g. the review papers, regarding health issues) were not accounted for	In other recent work exploring the ways search engine and searcher bias relate, log analysis indicated in health based search, participants favour positive (i.e., affirming a claim) over negative (i.e. denying a claim) results, and that search engines are more likely to display such positive results, despite the fact this meant around half of answers searchers settled on were incorrect	
Page use Query richness Source diversity	Few ISEQ correlations with limited confirmatory indications of relationship between ISEQ and self-report search expertise and page-viewing behaviours.	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)
Pageuse interaction with source quality chat Trustworthiness ratings	No obvious interaction between pageuse and incidence of source quality chat (suggesting selectivity in page use). We might expect higher trust – fewer pages viewed or used (in the CIS task at least), but this doesn't seem to be the case. No clear links to trust ratings.	"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, p. 246)	(Dimopoulos & Asimakopoulos, 2010)
Variety	No observed relationships between ISEQ scores and searching behaviours (except self-report). Regressions indicate that variance in output scores can be explained by differences in trace behaviours.	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.	(C.-C. Lin & Tsai, 2007, p. 689)

		Lin & Tsai, 2007, p. 692).	
Query richness Query depth Sequence analysis	In CIS, those who visit more pages use more query terms, but query less (as reflected in query-vocabulary richness metrics). Small to moderate relationships between page views and querying. No clear relationship to query depth. Sequence analysis not conducted in this work, but will form part of secondary data analysis.	High epistemic belief students more likely to: show bi-directional sequences of 'query-results browsing' (e.g. query->page->query); and more likely to go deeper into query results pages. They were also more likely to use the 'back' button to revisit earlier information.	Hsu, Tsai, Hou, and Tsai (2013)